



Universidade Federal do Piauí
Centro de Ciências da Natureza
Programa de Pós-Graduação em Ciência da Computação

Classificação de Imagens Dermatoscópicas utilizando Aprendizado Profundo

Alan Rafael Ferreira dos Santos

Teresina-PI, Setembro de 2017

Alan Rafael Ferreira dos Santos

Classificação de Imagens Dermatoscópicas utilizando Aprendizado Profundo

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Prof. Dr. Kelson Rômulo Teixeira Aires

Teresina-PI

Setembro de 2017

S237c Santos, Alan Rafael Ferreira dos.

Classificação de imagens dermatoscópicas utilizando aprendizado profundo / Alan Rafael Ferreira dos Santos. – Teresina, 2017.

80f.: il.:color.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Ciência da Computação, 2017.

Orientador: Prof. Dr. Kelson Rômulo Teixeira Aires.

1. Sistemas de Computação. 2. Redes Neurais Convolucionais. 3. Aprendizado Profundo. Título.

CDD 004.21

“Classificação de Imagens Dermatoscópicas utilizando Aprendizado Profundo”

ALAN RAFAEL FERREIRA DOS SANTOS

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Natureza da Universidade Federal do Piauí, como parte integrante dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

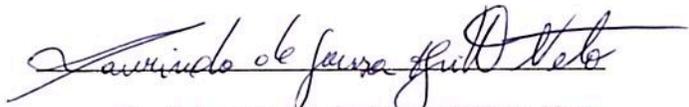
Aprovada por:



Prof. Kelson Rômulo Teixeira Aires
(Presidente da Banca Examinadora)



Prof. Geraldo Braz Júnior
(Examinador Externo)



Prof. Laurindo de Sousa Britto Neto
(Examinador Interno)



Prof. Rodrigo de Melo Souza Vêras
(Examinador Interno)

Agradecimentos

A conclusão desse trabalho é um marco na minha vida acadêmica e profissional, sendo fruto de muito esforço e dedicação. Agradecer é a melhor forma de expressar a minha gratidão.

Inicio meus agradecimentos a todos os profissionais que me guiaram por esta jornada. Em especial, a meu orientador, professor Kelson Rômulo Teixeira Aires, que proporcionou todos os ensinamentos necessários. Sou grato por sua dedicação confiança e pelas oportunidades dadas.

Agradeço a meu pai Antônio Leal dos Santos e Maria Irene Ferreira dos Santos por todo amor e carinho dedicados neste período, e a meu irmão Antônio Carlos Ferreira dos Santos pelo apoio incondicional em todas as situações de dificuldades.

A agradeço à minha companheira de todos os dias, Maria das Graças Oliveira Araújo por todo amor, afeto e atenção em todas as horas, apoiando e incentivando todas as etapas deste trabalho.

Agradeço ainda a todos meus amigos, em especial, Ivenilton Alexandre, Dennis Sávio, Leonardo Pereira, Francisco Imperes, Ismael Holanda, Romuere Rodrigues e Kalyf Abdalla.

Por fim, não poderia deixar de agradecer a Deus. As dificuldades e obstáculos encontrados nessa jornada me ensinaram o valor da vida. Obrigado por estar sempre me conduzindo.

Resumo

O melanoma é considerado o câncer de pele de maior gravidade e a sua descoberta em estágios avançados pode levar o paciente a óbito. O diagnóstico precoce ainda é a melhor forma de prescrever um tratamento adequado. Nesse contexto, o presente trabalho apresenta duas abordagens de classificação de imagens dermatoscópicas por meio da utilização de aprendizado profundo com as Redes Neurais Convolucionais. A primeira delas é baseada nos princípios tradicionais de detecção do melanoma, utilizando os conceitos da regra ABCD. Para tal, essa proposta de abordagem desenvolve as etapas de pré-processamento com filtros morfológicos, segmentação com o método *Fuzzy K-means*, extração de características de assimetria com o descritor de geometria, bordas com o método de *Histogram Oriented Gradient*, cor com o descritor Estatístico em Canais de Cores e os métodos de *Haralick* e de *Gabor Bank Filter* para obter os dados de texturas. Por fim, a classificação desses dados é realizada com uma arquitetura convolucional. A segunda abordagem sugerida neste trabalho utiliza as arquiteturas de redes convolucionais AlexNet e VGG-F. Antes de treinar essas redes, o conjunto de imagens é ajustado por meio do método de *data augmentation*. Após o treino, os modelos convolucionais são utilizados como descritores de características por intermédio das camadas FC₆ e FC₇. Em ambas as abordagens propostas, são utilizados os métodos de classificação *Support Vector Machine*, *K-Nearest Neighbor* e *MultiLayer Perceptron*. Esses classificadores são implementados na maioria das formas de diagnóstico de lesões cancerígenas. Os testes realizados neste trabalho utilizaram as bases de imagens PH² e ISIC. A abordagem de classificação com as redes convolucionais obteve uma acurácia de 93,1% na classificação correta do melanoma. Já a abordagem com modelos convolucionais treinados obteve, em seu melhor resultado, uma acurácia de 91,5% na classificação das lesões cancerígenas com a descrição da arquitetura AlexNet, utilizando um treinamento com o método de *batch normalization*.

Palavras-chaves: melanoma, aprendizado profundo, redes neurais convolucionais.

Abstract

Melanoma is considered the worst kind of skin cancer, and its discovery at advanced stages can bring the patient to death. The early diagnosis is the best way to prescribe an adequate treatment. In this context, the present work presents two approaches of dermoscopic image classification, through the use of deep learning with convolutional neural networks. The first is based on the traditional principles of melanoma detection, using the concepts of the ABCD rule. In this regard, this approach proposal develops the pre-processing steps with morphological filters, segmentation with the Fuzzy K-means method, asymmetry characteristic extraction with the geometry descriptor, borders with the Histogram Oriented Gradient method, color with the color channel statistic descriptor and the Haralic and Gabor Bank Filter methods to obtain the texture data. Finally, the classification of these data is performed with a convolutional architecture. The second approach suggested in this work uses the AlexNet and VGG-F convolutional network architectures. Before training these networks, the image set is adjusted through the data augmentation method. After the training, the convolutional models are used as characteristic descriptors through the FC₆ and FC₇ layers. In both approaches, are used the Support Vector Machine, K-Nearest Neighbor and MultiLayer Perceptron classification methods. These classifiers are implemented in most of the cancerous lesions diagnosis means. The tests performed in this work used the PH 2 and ISIC image databases. The classification approach with the convolutional networks obtained a 93,1% accuracy on classifying melanoma. The approach with trained convolutional models obtained, at its best result, a 91,5% accuracy on classifying cancerous lesions with the description of the AlexNet architecture, using a training with the method of bach normalization.

Keywords: melanoma, deep learning, convolutional neural network.

Lista de ilustrações

Figura 1 – As camadas da Pele. Adaptado de (MARIEB; HOEHN, 2009).	9
Figura 2 – Estrutura da epiderme em camadas. Adaptado de (DÂNGELO; FATTINI, 1998).	10
Figura 3 – Pigmentos de melanina. Adaptado de (MARIEB; HOEHN, 2009).	11
Figura 4 – Estrutura da derme. Adaptado de (MARIEB; HOEHN, 2009).	11
Figura 5 – Estrutura da hipoderme. Adaptado de (DÂNGELO; FATTINI, 1998).	12
Figura 6 – Carcinoma basocelular. (a) estrutura do carcinoma basocelular. (b) aparência do carcinoma basocelular. Adaptado de (SOCIETY, 2016).	14
Figura 7 – Carcinoma espinocelular. (a) estrutura do carcinoma espinocelular. (b) aparência do carcinoma espinocelular. Adaptado de (SOCIETY, 2016).	14
Figura 8 – Melanoma. (a) estrutura do melanoma. (b) aparência do melanoma. Adaptado de (SOCIETY, 2016).	15
Figura 9 – Padrão ABCD. Adaptado de (SBD, 2016).	16
Figura 10 – Estágios do melanoma. (a) estágios iniciais. (b) metástase. Adaptado de (SOCIETY, 2016).	16
Figura 11 – Classificação de dados com o KNN.	36
Figura 12 – Modelo de neurônio artificial. Adaptado de (HAYKIN, 2001)	36
Figura 13 – Exemplo de classificação com SVM.	38
Figura 14 – Hiperplano ótimo.	38
Figura 15 – Convolução.	41
Figura 16 – Reticulado de neurônios convolucionais.	42
Figura 17 – Compartilhamento de pesos de campos receptivos em três camadas.	43
Figura 18 – <i>Pooling</i> em mapa de características.	43
Figura 19 – Comportamento das Funções de Ativação em RNCs.	47
Figura 20 – Comportamento do <i>Dropout</i>	48
Figura 21 – Exemplos de <i>Data Augmentation</i> para imagens dermatoscópicas.	49
Figura 22 – Proposta de reconhecimento do melanoma com a classificação por RNCs.	52
Figura 23 – Remoção de pelos pelo método de Lee et al. (1997). (a) Imagem de entrada. (b) Máscara dos pêlos (c) Imagem corrigida. Adaptado de (MENDONÇA et al., 2013)	52
Figura 24 – Segmentação da lesão com método <i>Fuzzy K-means</i> . (a) Imagem de entrada (b) Segmentação. (c) Máscara. (d) Imagem de saída. Adaptado de (MENDONÇA et al., 2013).	54
Figura 25 – Proposta de reconhecimento do melanoma pela descrição com RNCs.	56
Figura 26 – Arquitetura da RNC AlexNet.	57
Figura 27 – Arquitetura da RNC VGG-F.	58

Lista de tabelas

Tabela 1 – Trabalhos relacionados a abordagem convencional de identificação do melanoma	22
Tabela 2 – Trabalhos relacionados ao aprendizado profundo	24
Tabela 3 – Nível de exatidão de classificação de dados com o índice kappa.	63
Tabela 4 – Nível de precisão por meio da AUC.	64
Tabela 5 – Resultados da classificação com o SVM.	66
Tabela 6 – Resultados da classificação com o KNN.	66
Tabela 7 – Resultados da classificação com o MLP.	67
Tabela 8 – Resultados da classificação com a RNC.	67
Tabela 9 – Resultados da classificação mediante a descrição de RNCs no SVM.	68
Tabela 10 – Resultados da classificação mediante a descrição de RNCs no KNN.	68
Tabela 11 – Resultados da classificação mediante a descrição de RNCs no MLP.	69
Tabela 12 – Melhores resultados para a classificação com RNCs.	69
Tabela 13 – Melhores resultados para classificação mediante a descrição de RNCs.	70
Tabela 14 – Comparações com a Abordagem de classificação utilizando RNC.	70
Tabela 15 – Comparações com a Abordagem de Descrição Utilizando Modelos de RNCs.	70

Lista de abreviaturas e siglas

ABCD	Assimetria, Borda, Cor e Diâmetro
AUC	<i>Area Under Curve</i>
HSV	<i>Hue, Saturation and Value</i>
HOG	<i>Histogram of Oriented dos Gradients</i>
FP	Falso Positivo
KNN	<i>K-Nearest Neighbor</i>
MLP	<i>MultiLayer Perceptron</i>
RNAs	Redes Neurais Artificiais
RNCs	Redes Neurais Convolucionais
ROC	<i>Receiver Operating Characteristic</i>
RGB	<i>Red, Green, and Blue</i>
SVM	<i>Support Vector Machine</i>
UV	Ultravioleta
VP	Verdadeiro Positivo
FP	Falso Positivo

Lista de símbolos

σ	Letra Grega Sigma
θ	Letra Grega Teta
ρ	Letra Grega Ró
η	Letra Grega Eta
δ	Letra Grega Delta
π	Letra Grega Pi
υ	Letra Grega Upsilon
ψ	Letra Grega Psi
ϵ	Letra Grega Epsilon
λ	Letra Grega Lambdax

Sumário

1	INTRODUÇÃO	1
1.1	Caracterização do Problema	2
1.2	Visão Geral da Proposta	3
1.3	Objetivos	4
1.4	Justificativa	5
1.5	Produções Bibliográficas	5
1.6	Contribuições	6
1.7	Estrutura do Trabalho	6
2	CÂNCER DE PELE	9
2.1	Anatomia e Fisiologia da Pele	9
2.2	Câncer de Pele	12
2.3	Tipos de Câncer de Pele	13
2.4	Diagnóstico e Tratamento do Melanoma	15
2.5	Considerações Finais	17
3	TRABALHOS RELACIONADOS	19
3.1	Formas Convencionais de Identificação de Lesões de Pele	19
3.2	Pesquisas em Aprendizado Profundo	22
3.3	Considerações Finais	24
4	MATERIAIS E MÉTODOS	25
4.1	Estrutura Tradicional de Reconhecimento do Melanoma	25
4.1.1	Pré-Processamento	26
4.1.2	Segmentação	27
4.1.3	Extração de Características	29
4.1.3.1	Descritor de Geometria	29
4.1.3.2	Descritor <i>Histogram of Oriented Gradients</i>	30
4.1.3.3	Descritor de Estatística em Canais de Cores	31
4.1.3.4	Descritor com Banco de Filtros de Gabor	32
4.1.3.5	Descritor <i>Local Binary Pattern</i>	33
4.1.3.6	Descritor de <i>Haralick</i>	34
4.1.4	Classificação	34
4.1.4.1	<i>K-Nearest Neighbor</i>	35
4.1.4.2	<i>MultiLayer Perceptron</i>	35
4.1.4.3	<i>Support Vector Machine</i>	37

4.2	Métodos de Aprendizado Profundo	39
4.2.1	Redes Neurais Convolucionais	40
4.2.2	Camada de convolução	40
4.2.3	Camada de <i>Pooling</i>	43
4.2.4	Camada Totalmente Conectada	44
4.3	Treinamento de Redes Neurais Convolucionais	45
4.3.1	Ativação	46
4.3.2	<i>Dropout</i>	47
4.3.2.1	<i>Data Augmentation</i>	48
4.3.3	<i>Batch Normalization</i>	49
4.4	Considerações Finais	50
5	CLASSIFICAÇÃO COM APRENDIZADO PROFUNDO	51
5.1	Reconhecimento do Melanoma pela Classificação com RNCs	51
5.1.1	Pré-processamento	52
5.1.2	Segmentação	53
5.1.3	Extração de Características	53
5.1.4	Classificação	55
5.2	Reconhecimento do Melanoma pela Descrição com RNCs	56
5.2.1	Arquitetura Convolutacional AlexNet	57
5.2.2	Arquitetura Convolutacional VGG - F	58
5.2.3	Definição do Aumento de Dados	58
5.2.4	Extração de Características	59
5.2.5	Classificação	60
5.3	Considerações Finais	60
6	EXPERIMENTOS E DISCUSSÕES	61
6.1	Bases de Imagens	61
6.2	Avaliação de Resultados	61
6.2.1	Definição dos Conjuntos de Treino, Validação e Teste	63
6.3	Configuração dos Experimentos	64
6.4	Resultados do Reconhecimento pela Classificação com RNCs	66
6.5	Resultados do Reconhecimento pela Descrição de RNCs	68
6.6	Discussões e Comparações	69
6.7	Considerações Finais	70
7	CONCLUSÕES E TRABALHOS FUTUROS	73
	REFERÊNCIAS	75

1 Introdução

No Brasil, o câncer de pele é o mais incidente na população devido ao comportamento das pessoas em relação à exposição solar. Estimativas mostram que 9% da sociedade brasileira possui um prognóstico desse câncer, dos quais, 25% dos casos são considerados malignos (SBD, 2016). Dentre os tipos de câncer de pele, o melanoma é o de menor incidência no país, entretanto, a sua forma é a mais perigosa e letal. Esse tipo se desenvolve na pele a partir de lesões anormais, que surgem por meio da exposição solar excessiva (INCA, 2016).

Em estágio avançado, o melanoma pode invadir as camadas profundas da pele, espalhando células cancerígenas para todo o organismo, o que acarreta na morte do paciente em um curto período de tempo. O diagnóstico precoce ainda é a melhor solução para essa doença, pois aumenta as chances de cura e evita o óbito (SBD, 2016).

O meio mais utilizado para a detecção do melanoma é o exame dermatoscópico, que consiste em uma técnica que produz imagens digitais de alta qualidade das lesões de pele para serem analisadas e diagnosticadas visualmente por um médico dermatologista (FRANGE; ARRUDA; DALDON, 2012). A identificação do melanoma segue um padrão com base nas seguintes características: Assimetria, Bordas irregulares, aspectos de Cor variável e Diâmetro da lesão acima de 5 mm (regra de detecção ABCD). Em geral, as lesões que possuem essas características têm fortes indícios de serem consideradas câncer melanômico nas imagens dermatoscópicas (MAGLOGIANNIS; DOUKAS, 2009).

A identificação precoce do melanoma ainda é uma tarefa complexa que requer grande experiência do dermatologista. Um dos principais obstáculos encontrados no diagnóstico dessa doença é a grande quantidade de dados que precisam ser analisados. Essa questão impulsiona o desenvolvimento de sistemas automáticos que possibilitam uma análise eficaz, favorecendo um diagnóstico e uma segunda opinião ao especialista (SOARES, 2008).

Na literatura diversos sistemas já foram propostos para auxiliar os dermatologistas. A maioria dessas ferramentas são inspiradas na regra de detecção ABCD, contudo, conceitos de texturas também são utilizados para favorecer um melhor reconhecimento. A classificação é a etapa mais pesquisada nesses sistemas, pois os métodos empregados definem a qualidade do diagnóstico das lesões. Nesse contexto, novas formas poderiam ser utilizadas, com o objetivo de promover abordagens eficazes no reconhecimento de imagens dermatoscópicas (BARATA et al., 2014a; CELEBI; MENDONCA; MARQUES, 2015; KOROTKOV; GARCIA, 2012).

Os conceitos de aprendizado profundo utilizando Redes Neurais Convolucionais (RNCs) vêm obtendo excelentes resultados para uma diversidade de problemas que envolvem

reconhecimento, principalmente para a classificação de imagens médicas. Essas redes são capazes de transformar estruturas de dados em conjuntos invariantes de características que são utilizadas para definir as probabilidades de ocorrência em classes. A partir desse conceito, o presente trabalho sugere a sua utilização na classificação do melanoma em imagens dermatoscópicas (SHI; HE, 2011; HUYNH; TRAN; KIM, 2016). Essas redes poderiam proporcionar melhores formas de reconhecimento dessa doença, favorecendo novos sistemas que são voltados para a busca de um diagnóstico mais preciso.

Algumas pesquisas recentes já sugeriram o aprendizado profundo com Redes Neurais Convolucionais para classificação de imagens e sons. Esses trabalhos passaram a ser notórios nos últimos anos por conta dos seus excelentes resultados. Ainda não é possível definir em que áreas específicas eles estão, entretanto, é possível determinar duas linhas semelhantes:

- **Processamento Digital de Imagens (PDI):** Detecção do melanoma em imagens dermatoscópicas (CODELLA et al., 2015), diagnóstico em tomografia computadorizada (SHEN et al., 2015), classificação de nódulos pulmonares (ANTHIMOPOULOS et al., 2016), reconhecimento facial (HONG et al., 2015), reconhecimento de expressões faciais (HUYNH; TRAN; KIM, 2016), identificação de caracteres (SERMANET; CHINTALA; LECUN, 2012), classificação de imagens geográficas (ZHAO; DU, 2016);
- **Processamento Digital de Sons (PDS):** Classificação de vozes (PRICE; ISO; SHINODA, 2016), reconhecimentos de conversas (SIOHAN; RYBACH, 2015), identificação de idiomas (LOZANO-DIEZ et al., 2015), classificação de atores em diálogos (UZAN; WOLF, 2015).

A partir desses trabalhos, podemos afirmar que os conceitos de aprendizado profundo são emergentes na literatura. Os estudos que utilizam essas teorias sugerem pesquisas mais aprofundadas, juntamente com a sua aplicação em uma diversidade de outros problemas, principalmente as Redes Neurais Convolucionais. Esse método poderia trazer excelentes resultados em novas abordagens que envolvem reconhecimento de imagens.

A utilização de Redes Neurais Convolucionais para o reconhecimento do melanoma neste trabalho surgiu da necessidade de diversificar o emprego do aprendizado profundo, a partir dos excelentes resultados de algumas pesquisas que utilizam imagens médicas. Nesse sentido, mais trabalhos relacionados serão discutidos ao longo do texto para justificar o propósito na forma de utilizar essa técnica para diagnosticar imagens dermatoscópicas.

1.1 Caracterização do Problema

As abordagens de detecção do melanoma em imagens dermatoscópicas são desenvolvidas por meio de uma estrutura convencional, baseada nas etapas de pré-processamento,

segmentação, extração de características e classificação. Em cada uma dessas etapas são traçados objetivos que resultam no diagnóstico das lesões (KOROTKOV; GARCIA, 2012).

A etapa de pré-processamento é responsável por melhorar a aparência das imagens. As operações realizadas nessa fase visam remover elementos indesejados, sem diminuir significativamente a qualidade dessas imagens. Na etapa de segmentação, as lesões são separadas do fundo das imagens, com o objetivo de manter uma maior riqueza de detalhes dos seus contornos. Já a extração de características tem o intuito de calcular uma definição matemática que obtém os atributos numéricos que vão representar as lesões. Por fim, a classificação recebe esses dados e aplica um modelo de predição para determinar o diagnóstico final das lesões em melanoma ou não melanoma (MAGLOGIANNIS; DOUKAS, 2009).

A composição da estrutura de uma abordagem influencia diretamente na qualidade do reconhecimento. A maioria das abordagens desenvolvidas para o diagnóstico do melanoma realiza essa composição de maneira empírica. Isso pode comprometer a sua eficácia por priorizar somente aspectos experimentais (ARGENZIANO et al., 2003). Além disso, a validação necessita comparar os resultados obtidos por intermédio de boas medidas de avaliação, em que outras abordagens similares são confrontadas. Uma proposta de reconhecimento só é considerada satisfatória se todos esses requisitos forem atingidos.

Uma boa parte das abordagens de reconhecimento do melanoma desenvolvidas atualmente priorizam a estrutura convencional em que a etapa de classificação, na maioria dos casos, utiliza métodos tradicionais. A literatura defende que, para essa etapa, os métodos Máquina de Vetor de Suporte (*Support Vector Machine* – SVM), K-Vizinhos mais Próximos (*K - Nearest Neighbor* – KNN) e Perceptron Multicamadas (*MultiLayer Perceptron* – MLP) como sendo algumas das melhores formas de alcançar bons resultados. Entretanto, novas pesquisas estão sugerindo a substituição desses classificadores por métodos de aprendizado profundo com Redes Neurais Convolucionais, tanto na etapa de classificação quanto na extração de características (SHI; HE, 2011; BARATA et al., 2014a; CELEBI; MENDONCA; MARQUES, 2015; KOROTKOV; GARCIA, 2012).

A partir do bom desempenho apresentado em trabalhos anteriores, as Redes Neurais Convolucionais podem ser utilizadas de formas diferentes em abordagens de detecção do melanoma, considerando que elas seriam de grande valia para promover diagnósticos mais eficazes das lesões, além de proporcionar novos estudos nesse sentido.

1.2 Visão Geral da Proposta

A partir da arquitetura convencional dos sistemas de reconhecimento do melanoma, o presente trabalho apresenta duas formas de reconhecer o câncer de pele para imagens

dermatoscópicas. A primeira forma é baseada em sistemas convencionais, entretanto, o seu diferencial está na utilização do aprendizado profundo na etapa de classificação. Já a segunda forma trata o aprendizado na etapa de extração de características.

A forma de reconhecer o melanoma por intermédio da classificação pelas Redes Neurais Convolucionais desenvolvida neste trabalho sugere o uso de operações morfológicas, interpolação bilinear e suavização para remover os artefatos na etapa de pré-processamento. A segmentação propõe o método *Fuzzy K-means* (BEZDEK; EHRlich; FULL, 1984) para localizar e separar as lesões. Já a extração de características sugere o uso dos descritores de *Geometria* (GANSTER et al., 2001) para dados de assimetria, *Histogram of Oriented Gradients* – HOG (DALAL; TRIGGS, 2005) para dados de borda, Estatística em Canais de Cores (SCHMID-SAUGEONA; GUILLODB; THIRANA, 2003) para variações de dados na coloração, *Gabor Bank Filter* – GBF (FOGEL; SAGI, 1989), *Local Binary Pattern* – LBP (MU et al., 2008) e *Haralick features* (HARALICK, 1979) para dados de textura. Por fim, a etapa de classificação utiliza uma arquitetura de Rede Neurais Convolutiva na predição dessas características.

A segunda forma desenvolvida neste trabalho não é baseada na regra ABCD. A sua estrutura não utiliza nenhum método de pré-processamento ou segmentação, apenas arquiteturas de Redes Neurais Convolucionais treinadas com imagens dermatoscópicas, que servem como modelos para extrair características espaciais relevantes das lesões. Por fim, os dados extraídos são utilizados para realizar a predição.

As abordagens propostas tiveram seus dados classificados com os métodos tradicionais SVM, KNN e MLP. Essas classificações têm o objetivo de avaliar as propostas em relação aos seus resultados. Para isso, a análise desses resultados leva em consideração as medidas de acurácia, precisão, *recall*, *f-measure* e índice *kappa*, obtidas com a matriz de confusão. Outra medida utilizada é a Área sob a Curva (*Area Under Curve* – AUC) de Característica de Operação do Receptor (*Receiver Operating Characteristic* – ROC).

1.3 Objetivos

O objetivo geral desse trabalho é desenvolver formas de diagnosticar o melanoma em imagens dermatoscópicas, variando os princípios de aprendizado profundo com as Redes Neurais Convolucionais tanto no processo de extração de características como na classificação de lesões de pele. Essa proposta visa desenvolver estruturas que comportem a utilização dessas redes de duas formas distintas na obtenção dos dados necessários para promover um diagnóstico mais preciso. Para isso, a execução das formas sugeridas traçaram alguns objetivos específicos, como:

- estudar conceitos que envolvem a detecção do câncer de pele;

- investigar as formas automáticas de identificação do melanoma;
- pesquisar os principais métodos de pré-processamento, segmentação, extração de características e classificação de imagens dermatoscópicas;
- analisar os conceitos de aprendizado profundo utilizando as redes neurais por convolução;
- estudar a utilização das arquiteturas convolucionais na extração de características e na classificação de dados;
- pesquisar melhores arquiteturas de convolução que se adequam ao problema;
- investigar metodologias de avaliação de dados, e;
- entender as vantagens e limitações das propostas de diagnóstico sugeridas.

1.4 Justificativa

As abordagens que identificam o melanoma em imagens dermatoscópicas possibilitam um auxílio ao dermatologista no diagnóstico das lesões de pele. Nesse sentido, várias propostas de classificar essa doença já foram desenvolvidas na bibliografia, entretanto, percebe-se que ainda há diversas lacunas a serem preenchidas (LEE; CHEN, 2015). Uma delas é a precisão nesse diagnóstico (SHI; HE, 2011). Percebendo isso, esse trabalho propõe duas formas de identificar e classificar lesões em imagens dermatoscópicas, introduzindo o aprendizado profundo como mais uma opção na busca por diagnóstico eficaz do melanoma.

As abordagens de detecção do melanoma propostas nessa pesquisa são estruturadas com as Redes Neurais Convolucionais, utilizando tanto a arquitetura convencional dos sistemas de detecção dessa doença como novas formas reconhecimento. Elas permitem o uso diferenciado das redes convolucionais, com o objetivo de encontrar a combinação que proporcione a maior acurácia no diagnóstico. Por fim, é importante destacar que as abordagens sugeridas possuem algumas particularidades em relação as demais, uma delas é o treinamento de arquiteturas profundas com imagens específicas para criar modelos de extração de características. Isso pode trazer benefícios significativos em novas pesquisas.

1.5 Produções Bibliográficas

As principais contribuições relacionadas às abordagens propostas neste trabalho são:

- FERREIRA SANTOS, A. R.; TEIXEIRA AIRES, K. R.; SOUZA VERAS, R. M. Classificação de Imagens Dermatoscópicas Utilizando Aprendizado Profundo para Características de Borda, Cor e Textura. In: Congresso Brasileiro de Automática (CBA), 2016, Vitória/ES. XXI CBA.

- FERREIRA SANTOS, A. R.; TEIXEIRA AIRES, K. R.; SOUZA VERAS, R. M.; UCHÔA, V. S.; TEIXEIRA SANTOS, L. G. Uma Abordagem de Classificação de Imagens Dermatoscópicas Utilizando Aprendizado Profundo com Redes Neurais Convolucionais. In: *Workshop de Informática Médica (WIM)*, 2017, São Paulo/SP. XVII WIN.

1.6 Contribuições

O presente trabalho contribui propondo formas de reconhecimento do melanoma para imagens dermatoscópicas por meio de duas estruturas distintas que utilizam aprendizado profundo com Redes Neurais Convolucionais. Atualmente existe uma diversidade de trabalhos nesse sentido, entretanto, a maioria considera apenas as estruturas convencionais e muitas vezes a etapa de classificação emprega sempre os mesmos métodos para classificar as lesões de pele. Neste trabalho, realizamos uma mudança desse cenário com duas arquiteturas de detecção que diversificam a literatura. A primeira utiliza as formas convencionais, contudo, as operações de convolução são utilizadas para classificar as características das lesões. Já na segunda é desenvolvida uma nova forma de classificar as imagens dermatoscópicas, por meio do treinamento de arquiteturas de Redes Neurais Convolucionais, que servem como modelos para extração de características.

Além disso, também foi possível avaliar o poder do aprendizado profundo com as redes de convolução, por meio da comparação com os principais métodos tradicionais de classificação que foram sugeridos pela literatura, visto que poucos trabalhos realizam essa tarefa. Isso enfatiza a importância deste trabalho e sugere estudos mais aprofundados para reconhecimento do melanoma com as Redes Neurais Convolucionais.

1.7 Estrutura do Trabalho

A estrutura deste trabalho é dividida a partir dos seguintes itens:

- o Capítulo 2 apresenta os principais conceitos sobre anatomia e fisiologia da pele, os fatores de risco para o câncer de pele, os seus tipos e o diagnóstico e tratamento do melanoma;
- o Capítulo 3 apresenta alguns dos principais trabalhos desenvolvidos para classificação do melanoma em imagens dermatoscópicas, além de novos estudos que desenvolvem as Redes Neurais Convolucionais de aprendizado profundo para problemas semelhantes;
- o Capítulo 4 descreve os principais matérias e métodos utilizados para o desenvolvimento da proposta deste trabalho;

-
- no Capítulo 5 são apresentadas as sugestões de abordagens com aprendizado profundo, que utilizam as Redes Neurais Convolucionais tanto na forma tradicional de detecção do melanoma quanto em novas formas de identificação dessa doença;
 - já o Capítulo 6 apresenta todas as configurações utilizadas nas abordagens desenvolvidas, os resultados obtidos nos experimentos e algumas discussões;
 - por fim, o Capítulo 7 apresenta as conclusões referentes as abordagens desenvolvidas, juntamente com os aspectos que justificam as suas investigações e continuidade.

2 Câncer de Pele

Este capítulo apresenta um embasamento teórico sobre a estrutura da pele humana, os fatores de riscos do câncer de pele, os seus tipos de manifestações e os aspectos para o diagnóstico e tratamento do melanoma. O conhecimento da maior parte deste conteúdo permite a estruturação e o desenvolvimento de abordagens automáticas de detecção do câncer de pele.

2.1 Anatomia e Fisiologia da Pele

A pele é considerada o maior órgão do corpo humano, correspondendo a 15% da massa corpórea de um indivíduo. A sua estrutura faz o revestimento de todo o organismo, com a função de proteger todas as camadas subjacentes contra uma diversidade de agentes físicos, químicos e biológicos (SEELEY, 2001).

As principais características da pele são as suas variações de espessura, formato e aparência ao longo do corpo. Essas diferenças são visivelmente notadas nas palmas das mãos e nas plantas dos pés. Outro aspecto relevante é a variedade dos tons de pele por indivíduo, essa característica é determinada por fatores genéticos, regionais ou sexuais (DÂNGELO; FATTINI, 1998).

A superfície da pele é dividida em epiderme, derme e hipoderme, sendo a epiderme a parte mais externa da pele, a derme a parte intermediária e a hipoderme a parte mais profunda (ELDER; ELENITSAS; JR, 2000). A Figura 1 mostra essa divisão.

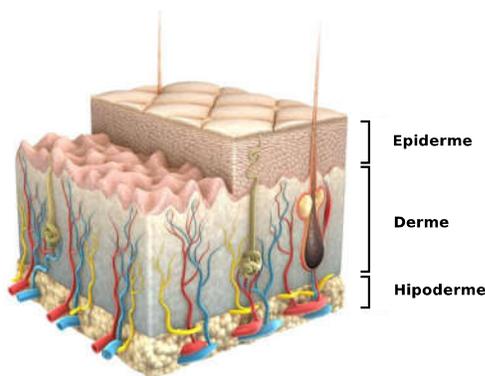


Figura 1 – As camadas da Pele. Adaptado de (MARIEB; HOEHN, 2009).

A epiderme é a parte da pele que tem contato direto com o exterior. A sua espessura pode variar conforme a localização no corpo, podendo atingir até 1,5 mm nas plantas dos pés. A mesma possui células próprias, conhecidas como: queratinócitos, melanócitos

e *Langerhans*. Essas células se encontram distribuídas em camadas: córnea, granulosa, espinhosa e basal (MARIEB; HOEHN, 2009). A Figura 2 ilustra a localização dessas células em relação às camadas apresentadas.

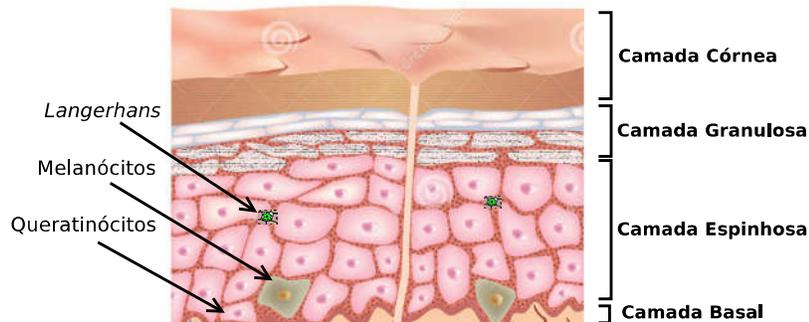


Figura 2 – Estrutura da epiderme em camadas. Adaptado de (DÂNGELO; FATTINI, 1998).

Os queratinócitos são as células em maior disposição na epiderme, estando presentes em todas as camadas. Essas células mantêm ciclos contantes de renovações por atividades mitóticas que são realizados entre as partes posteriores da pele, acarretando na morte de algumas células e no surgimento de outras novas (DÂNGELO; FATTINI, 1998).

A camada córnea é constituída por queratinócitos mortos, sem núcleos e achatados. Essas células são continuamente empurradas pelos queratinócitos vivos até a superfície da pele, gerando os ciclos de renovações. Além disso, essa camada é responsável pela proteção direta da pele e a sua estrutura é semelhante a um conjunto de laminas rígidas e sobrepostas (MARIEB; HOEHN, 2009).

A camada granulosa é formada por queratinócitos vivos, em forma de polígonos achatados, com queratina em seu citoplasma. Essa camada é responsável por produzir proteínas que impedem a penetração de substâncias líquidas, tornando a pele impermeável e impedindo a desidratação do organismo (ELDER; ELENITSAS; JR, 2000).

Já na camada espinhosa, os queratinócitos tem forma de cuboide, com pouco achatamento. Essas células possuem expansões próximas de seus citoplasmas, proporcionando um aspecto espinhoso. A imunovigilância dessa camada é realizada pelas células *Langerhans*. Elas são responsáveis por eliminar partículas estranhas que invadem a camada (DÂNGELO; FATTINI, 1998).

A camada basal é a mais profunda da epiderme. A sua estrutura é formada por queratinócitos em forma de prismas. Nessa camada ocorre uma intensa produção de queratinócitos por meio das divisões celulares por contato com a derme (DIO, 1998).

Já os melanócitos são as células mais importantes da epiderme, pois elas são responsáveis pela produção de grânulos de melanina. Eles se espalham entre as reentrâncias dos queratinócitos, com o objetivo de criar uma barreira de proteção contra os raios

ultravioleta (UV). A melanina proporciona a coloração da pele, a sua menor quantidade determina tons mais claros e a sua maior tons mais escuros (DIO, 1998). A Figura 3 mostra a localização desses pigmentos na pele.

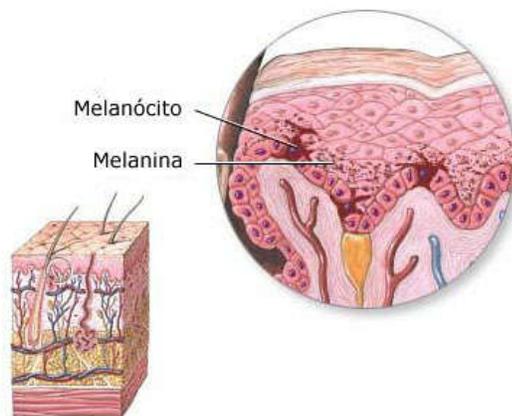


Figura 3 – Pigmentos de melanina. Adaptado de (MARIEB; HOEHN, 2009).

A exposição solar estimula a produção dos queratinócitos. No entanto, quando a mesma se torna excessiva, os melanócitos produzem mais melanina, resultando no escurecimento da pele (bronzamento) (KLIGERMAN, 2000).

A derme é a parte da pele formada por tecidos conjuntivos que interligam a epiderme com o hipoderme. Essa camada é responsável por 90% da espessura da pele e sua composição é feita por colágeno e elastina (HARGIS, 1990). A Figura 4 mostra a estrutura dessa parte da pele.

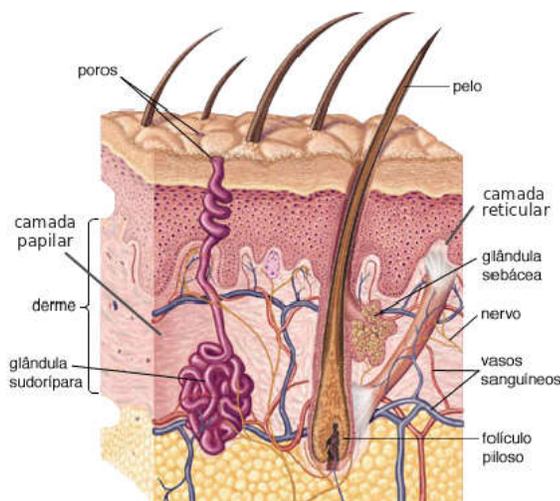


Figura 4 – Estrutura da derme. Adaptado de (MARIEB; HOEHN, 2009).

A derme é formada por duas camadas pouco visíveis: a papilar e a reticular. A camada papilar é formada pelas fibrilas de colágeno que ajudam a derme a se firmar na epiderme, facilitando a nutrição das células pelos vasos sanguíneos. A camada reticular é

formada por um tecido conjuntivo mais denso e rico em fibras. Ambas as camadas são responsáveis pela elasticidade da pele (DÂNGELO; FATTINI, 1998).

Na derme se encontram os pelos, vasos sanguíneos, nervos, folículos pilosos, glândulas sudoríparas e sebáceas. Além disso, essa camada possui receptores nervosos que fazem parte do sistema sensorial. Eles são responsáveis por captar as sensações de frio, calor e dor (SEELEY, 2001).

Os pelos são considerados filamentos resistentes que estão distribuídos por toda a pele, exceto nas extremidades das mãos e nas plantas dos pés. A sua composição é feita por uma estrutura delgada de queratina, que emerge do folículo piloso até a epiderme (DIO, 1998).

A hipoderme é conhecida como a camada mais profunda da pele. A sua composição é feita por tecidos adiposos que são compostos por células que armazenam gordura. Essa camada possui um tecido conjuntivo frouxo, que une de maneira pouco firme a derme à hipoderme, permitindo que a pele tenha um deslizamento não uniforme de acordo com a sua localização. O objetivo dessa camada é armazenar energia, manter a temperatura corpórea, modelar a superfície e realizar a absorção de choques (DIO, 1998). A Figura 9 ilustra a estrutura da hipoderme.

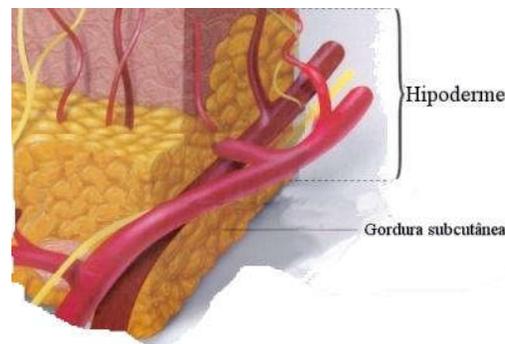


Figura 5 – Estrutura da hipoderme. Adaptado de (DÂNGELO; FATTINI, 1998).

A saúde da pele é essencial para o bem estar do ser humano. Uma pele saudável atua como uma barreira contra o mundo exterior, sendo a melhor defesa que o organismo possui. Por fim, quando a estrutura da pele sofre alterações, sua função protetora e aparência ficam comprometidas, propiciando o aparecimento de doenças ou anomalias (MARIEB; HOEHN, 2009).

2.2 Câncer de Pele

O câncer de pele é considerado um crescimento anormal e descontrolado dos queratinócitos, produzindo as conhecidas lesões cutâneas ou nevos. Existem vários fatores

de riscos que causam o surgimento dessa doença, dentre os quais podemos destacar (SAMPAIO; RIVITTI, 2001):

- **exposição solar:** esse é considerado o principal fator de risco de qualquer câncer de pele. Os raios UV estimulam a produção de melanócitos, fazendo com que a pele fique com um aspecto mais escuro. Uma longa exposição pode gerar queimaduras solares, manchas ou lesões, que podem evoluir para um câncer de pele;
- **pele clara:** indivíduos que possuem tons de pele leitosos são propensos a desenvolver câncer, pois possuem menos melanina e, conseqüentemente, tem menor proteção contra os raios UV;
- **idade e sexo:** pessoas com mais de 50 anos são propensas ao surgimento de lesões cancerígenas causadas pela exposição solar acumulada ao longo da sua vida. Contudo, muitos casos desse câncer vêm aparecendo em indivíduos mais jovens. Os homens tem o dobro da probabilidade de desenvolver o câncer de pele, devido à maior tendência à exposição solar;
- **histórico familiar:** pessoas que possuem casos de câncer de pele na família também estão mais propensas a desenvolver a doença; e,
- **imunodeficiência:** problemas no sistema imunológico, causados por doenças como leucemia e linfoma, assim como o uso de algum medicamento que suprima o sistema imunológico, podem aumentar a chance de ocorrência desse câncer.

2.3 Tipos de Câncer de Pele

A literatura costuma classificar o câncer de pele em três tipos: carcinoma basocelular, carcinoma espinocelular e melanoma (KLIGERMAN, 2000). O carcinoma basocelular surge nas células mais profundas da epiderme e é considerado o mais comum dentre os tipos. Geralmente, surge em partes do corpo que sofrem exposição solar, contudo podem aparecer em outras regiões. Algumas manifestações desse tipo de câncer se assemelham a eczemas e psoríase (SBD, 2016). A Figura 6 mostra a estrutura e aparência dessas lesões.

O carcinoma espinocelular é o segundo tipo de câncer mais comum. Ele se manifesta nas células da camada basal e o seu aparecimento também está relacionado com a exposição solar. As principais características dessas lesões são: bordas elevadas, coloração variada, sangramento e falta de elasticidade no local. Normalmente, o tipo espinocelular tem aparência avermelhada em forma de machucado, feridas descamativas ou verrugas. Em alguns casos, ele é associado a cicatrizes, feridas crônicas, exposição a substâncias químicas ou radiação (SBD, 2016). A Figura 7 apresenta lesões desse tipo.

O carcinoma basocelular e espinocelular são os tipos mais incidentes na população brasileira, correspondendo a 30% de todos os casos de lesões malignas. Apesar da incidência,

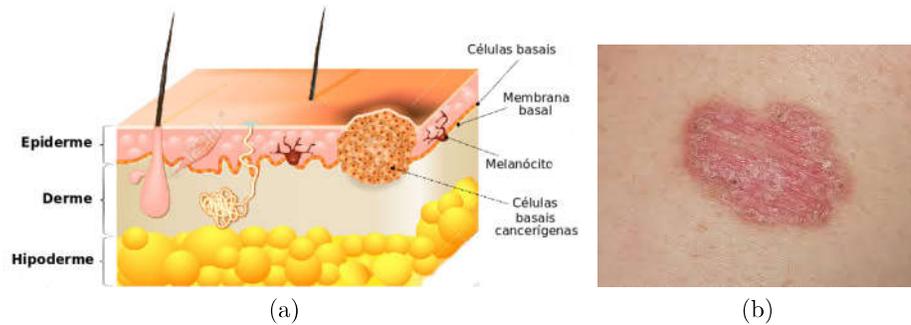


Figura 6 – Carcinoma basocelular. (a) estrutura do carcinoma basocelular. (b) aparência do carcinoma basocelular. Adaptado de (SOCIETY, 2016).

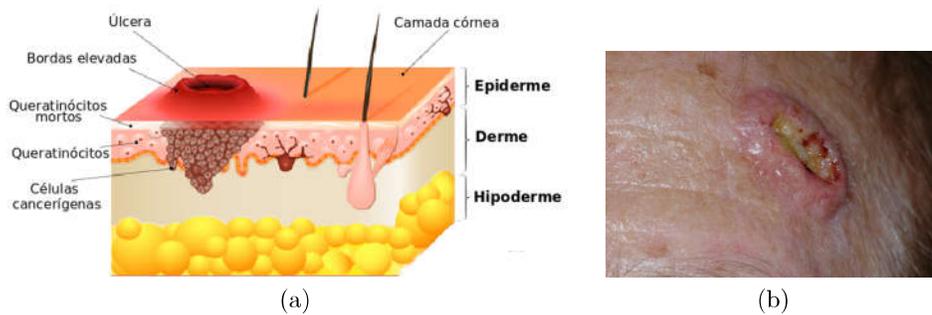


Figura 7 – Carcinoma espinocelular. (a) estrutura do carcinoma espinocelular. (b) aparência do carcinoma espinocelular. Adaptado de (SOCIETY, 2016).

apresentam altos percentuais de cura no caso de detecção precoce, pois geralmente não evoluem para estados que possam provocar mortes (SBD, 2016).

Apesar de ser o tipo menos frequente, o melanoma é o mais perigoso por apresentar o pior prognóstico. No Brasil, as lesões melanômicas correspondem a 3% dos casos registrados de câncer de pele. Embora essa porcentagem seja pequena, essa doença em estágios avançados pode levar o paciente a óbito (SBD, 2016). Segundo o INCA (2016) são esperados 5.670 óbitos para o ano de 2016 no país. No ano de 2013 o melanoma levou a óbito 1.547 pessoas.

As lesões melanômicas surgem a partir de manchas escuras, com bordas irregulares que podem causar coceira, descamação e sangramento (CELEBI; MENDONCA; MARQUES, 2015). Na maioria dos casos, essa doença não tem sintomas (SBD, 2016). Quando esse tipo de câncer surge de uma lesão já existente, observa-se que ela evolui de tamanho, muda de coloração e passa a apresentar bordas irregulares. Nessa situação, é necessária a avaliação de um médico especialista (ELDER; ELENITSAS; JR, 2000). A Figura 8 mostra essa lesão.

Na fase de crescimento da lesão, a epiderme e a derme podem ser invadidas por uma massa de células cancerígenas, formando nódulos visíveis e palpáveis. Essas lesões tem origem nos melanócitos, as células que produzem melanina. Normalmente, o surgimento

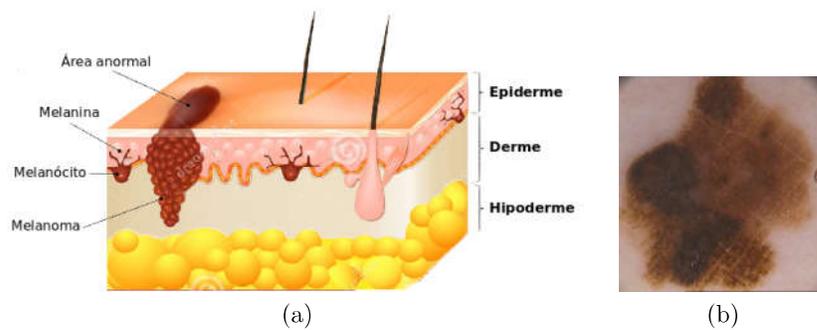


Figura 8 – Melanoma. (a) estrutura do melanoma. (b) aparência do melanoma. Adaptado de (SOCIETY, 2016).

dessa doença também está diretamente relacionado a exposição solar (ARGENZIANO et al., 2003).

2.4 Diagnóstico e Tratamento do Melanoma

O médico especialista na área da dermatologia é responsável por analisar as lesões de pele, diagnosticar o tipo de câncer e prescrever o seu tratamento. Em caso de dúvidas no diagnóstico visual, o paciente é submetido a uma biópsia. Esse procedimento é invasivo e tem o objetivo de coletar uma parte da lesão, com o intuito de favorecer uma análise laboratorial. A biópsia é a única forma efetiva para o diagnóstico definitivo do melanoma, pois ela apresenta quais camadas da pele já foram afetadas pela lesão maligna. No entanto, a prescrição desse método só deve ser sugerida em último caso (FRANGE; ARRUDA; DALDON, 2012).

O diagnóstico do melanoma também pode ser realizado por meio do exame dermatoscópico. Essa técnica captura imagens digitais ampliadas das lesões de pele, com alta resolução. Os especialistas analisam as características das lesões nessas imagens tomando por base os padrões visuais sugeridos pelos atlas da dermatologia. Esses são descrições de métodos para a identificação de doenças. Os principais métodos utilizados para o diagnóstico das lesões melanômicas em imagens dermatoscópicas são: regra ABCD (Assimetria, Borda, Cor e Diâmetro) (NACHBAR et al., 1994), *menzies* (MENZIES; CROTTY; MCCARTHY, 1995), lista de sete pontos (ARGENZIANO et al., 1998) e regra de três pontos (SOYER et al., 2001). A aplicação do método ABCD pode ser descrito pela Figura ??, em que cada componente é descrito.

Em estágios iniciais, o melanoma se desenvolve apenas na camada mais superficial da pele. Já nos estágios mais avançados, a lesão é mais profunda e espessa, o que aumenta a chance de metástase ¹ para outros órgãos (SOYER et al., 2001). A Figura 10 mostra os

¹ Disseminação de células cancerígenas para os outros órgão do corpo pela corrente sanguínea.



Figura 9 – Padrão ABCD. Adaptado de (SBD, 2016).

estágios do melanoma.

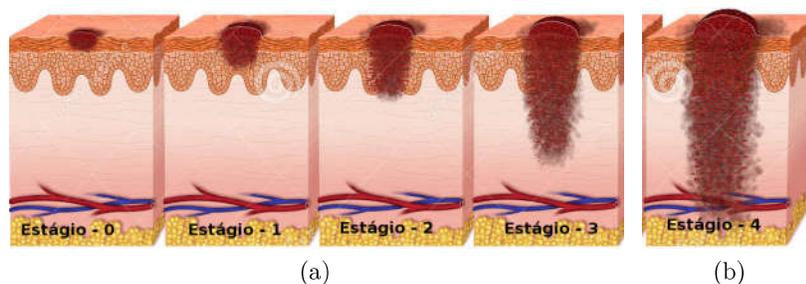


Figura 10 – Estágios do melanoma. (a) estágios iniciais. (b) metástase. Adaptado de (SOCIETY, 2016).

O tratamento do melanoma é realizado por meio de cirurgias de extração, radioterapias, imunoterapias e quimioterapias, de acordo com o estágio da lesão. Nos estágios 0 e 1 as lesões não ultrapassam a epiderme e seu tratamento é feito com uma incisão cirúrgica superficial. Já nos estágios 2 e 3 essas lesões invadem a derme. O tratamento é o mesmo das etapas anteriores, mas em alguns casos pode ser utilizada a radioterapia.

No estágio 4, essas lesões são difíceis de serem tratadas por ter ocorrido a metástase. Nessa situação, o tratamento é feito pela incisão, quimioterapia e radioterapia. Em muitos casos, esse tratamento serve apenas para proporcionar uma sobrevida ao paciente. Por isso, o diagnóstico precoce do melanoma é fundamental para evitar esse estágio final (SOYER et al., 2001).

2.5 Considerações Finais

Este capítulo apresentou conceitos relacionados à anatomia e fisiologia da pele humana, aos fatores de riscos relacionados à ocorrência de câncer de pele, os tipos de nevos cancerígenos e ao diagnóstico e tratamento dessas lesões. Os conteúdos descritos neste capítulo são utilizados, em partes, na criação e no desenvolvimento formas automáticas de detecção do melanoma. O próximo capítulo apresenta os principais trabalhos relacionados que proporcionaram o desenvolvimento da proposta deste trabalho

3 Trabalhos Relacionados

Uma boa parte das formas automáticas de reconhecimento do câncer de pele segue uma estrutura convencional, baseada em métodos bem conhecidos na literatura. Entretanto, novas pesquisas que utilizam os princípios do Aprendizado Profundo vêm ganhando uma ênfase maior na bibliografia para o diagnóstico do melanoma, por proporcionarem bons resultados. Nesse contexto, o presente capítulo apresenta alguns dos principais trabalhos relacionados a forma convencional de identificação do melanoma, além de algumas pesquisas que desenvolvem o Aprendizado Profundo em circunstâncias semelhantes para imagens médicas.

3.1 Formas Convencionais de Identificação de Lesões de Pele

Como dito antes, as abordagens convencionais de diagnóstico de lesões de pele desenvolvem quatro etapas básicas: pré-processamento, segmentação, extração de características e classificação. Os métodos utilizados em cada uma dessas etapas possuem objetivos específicos que, ao fim da sua execução, realizam transformações nas imagens, contribuindo significativamente para atingir o diagnóstico do melanoma. Para esse cenário, várias propostas já desenvolveram o reconhecimento deste tipo de câncer. Entretanto, os próximos trabalhos que serão apresentados definem uma estrutura tradicional, que evidencia uma padronização no diagnóstico do melanoma para imagens dermatoscópicas.

Nos estudos de [Ercal et al. \(1994\)](#) é desenvolvida uma forma de diagnosticar o melanoma seguindo as etapas de extração de características e a classificação de dados. A primeira situação considera as descrições de coordenadas de cor, cromaticidade relativa, percentual de assimetria e indexação de irregularidades. Já a segunda classifica esses dados por intermédio de uma rede neural multicamadas. Nos testes realizados para 326 imagens de lesões de pele em duas classes (melanoma e não melanoma) obtiveram uma precisão de 80% na identificação das lesões melanômicas.

Já em [Ganster et al. \(2001\)](#) é proposto um sistema automático de identificação do melanoma, que na etapa de pré-processamento realiza operações morfológicas para remoção dos artefatos. Em seguida, na etapa de segmentação, utiliza a junção dos métodos de limiar adaptativo e agrupamento de cor para separar as lesões. Já a etapa de extração de características considera as descrições de área, perímetro, regiões polares e estatística em canais de cores. Por fim, a classificação desses dados é realizada com o método de KNN. Os resultados obtidos nesse trabalho para 5.393 imagens em três classes (normal, anormal e melanoma) atingiram uma sensibilidade de 87% e uma especificidade de 92%

no diagnóstico correto dessas lesões.

No trabalho de [Celebi et al. \(2007\)](#) é criada uma forma de identificar o melanoma semelhante ao proposto por [Ganster et al. \(2001\)](#). No entanto, a etapa de pré-processamento emprega o uso de suavização por mediana de cor para remover os pelos. A etapa de segmentação utiliza os métodos de quantização e agrupamento de cor para separar as lesões. Já na extração de características é sugerido o uso das descrições de textura que estão contidas em matrizes de co-ocorrência. O diferencial dessa abordagem é a utilização de uma etapa de seleção de características, com a combinação dos métodos *Relief*, mutualidade e correlação. A classificação desses dados selecionados é realizada com a SVM. Nos melhores resultados para 526 classificadas em duas classes (melanoma e não melanoma), obtiveram uma especificidade de 92,34% e sensibilidade de 93,33% na identificação correta das imagens.

Em [Rahman, Bhattacharya e Desai \(2008\)](#) é sugerido um sistema integrado de suporte a decisões, que realiza a segmentação, a extração de características e a classificação de imagens dermatoscópicas. A primeira etapa faz a separação das lesões por meio dos métodos de limiar adaptativo e *Fuzzy K-means*. Já na extração de dados, utilizam os aspectos de textura que são encontrados nas matrizes de co-ocorrência. Por fim, a identificação desses dados em classes (normal, anormal e melanoma) é realizada pela combinação dos métodos de SVM, KNN e MLP. Os resultados obtidos nessas configurações para 326 imagens obtiveram uma acurácia de 83,75% na identificação correta.

No trabalho de [Maglogiannis e Doukas \(2009\)](#) é apresentado um estudo das principais técnicas de detecção do melanoma para imagens dermatoscópicas. Os dados mostram que 36% das metodologias de classificação utilizam características de cor, 21% características de borda e 9% características de textura na descrição das lesões. Neste contexto, a maioria dos métodos citados fazem a classificação dos dados por intermédio das Redes Neurais, Aprendizado Estatístico e SVM. Alguns dos melhores resultados encontrados na literatura para esses descritores são dados pela SVM e o MLP, atingindo precisões de até 94% para uma classificação binária (melanoma e não melanoma).

Nas pesquisas de [Cavalcanti e Scharcanski \(2011\)](#) é proposta uma abordagem para identificar pigmentos em imagens dermatoscópicas de lesões melanômicas. A etapa de pré-processamento corrige as lesões com operações de fechamento morfológico no espaço de cor HSV. Na segmentação é sugerido o uso de um novo método que utiliza texturas e cores para identificar apenas a área das lesões. As operações realizadas nessa etapa fazem a identificação de luminosidade por meio de uma normalização de canal de cores com um limiar adaptativo. Na etapa de extração de características são priorizados os descritores de área, perímetro, diâmetro, magnitude, semelhança, gradientes, estatística e quantização de cores. Nesse trabalho, a classificação é realizada com os métodos KNN e Árvore de Decisão. Os testes dessa abordagem utilizaram 220 imagens e seus resultados atingiram

uma acurácia de 91,7% na identificação do melanoma.

Já em Ruiz et al. (2011) é proposto uma sistema sequencial e automatizado de análise e classificação de imagens dermatoscópicas. Ele possui o objetivo de extrair as lesões de pele, determinar suas características e definir um diagnóstico. Na etapa de pré-processamento são utilizadas operações morfológicas de fechamento em canais de cores para remover os pelos, juntamente com uma suavização. Na etapa de segmentação é realizado um limiar adaptativo, seguido da operação de binarização. Essa abordagem também apresenta uma etapa de seleção de características mediante a variação do método de KNN. Finalmente, a etapa de classificação é realizada por meio da combinação dos métodos KNN, MLP e *NaiveBayes*. Os testes dessa proposta utilizaram uma base com 98 imagens e os resultado da combinação desses classificadores atingiu uma acurácia de 87% na identificação das lesões melanômicas.

Em Korotkov e Garcia (2012) também é feito uma revisão literária nas formas de detecção do melanoma. No entanto, essa pesquisa levou em consideração apenas os melhores descritores, juntamente com as melhores formas de classificação. Nos dados apresentados, uma parte da literatura utiliza os descritores de *Haralick*, *Gabor* e para descrever texturas, o HOG para bordas e Estatísticas em Canais de Cores para coloração. As características extraídas por esses descritores conseguem descrever com riqueza de detalhes o câncer de pele, oferecendo dados mais precisos para sua identificação.

Já em Barata et al. (2014a) são comparadas duas abordagens de diagnóstico do melanoma. A primeira abordagem, denominada "global", utiliza uma segmentação com limiar adaptativo e uma extração de características baseada em cor e textura com histogramas. A classificação desses dados é realizada com algoritmos *Adaboost*, SVM e KNN. Já a segunda abordagem, denominada "local", identifica os melhores pontos das lesões por intermédio de histogramas. Esses dados são classificados com o método *Bag Of Features* (BOF). Nesse sentido, as abordagens propostas são testadas com 200 imagens dermatoscópicas e os resultados obtidos com a forma global atingiram uma sensibilidade de 93% e uma especificidade de 85% na identificação do melanoma, enquanto a forma local atingiu uma sensibilidade de 96% e uma especificidade de 80%.

A Tabela 1 apresenta um resumo de todos os trabalhos descritos. A sua análise mostra que os métodos utilizados para realizar o diagnóstico, na maioria das vezes, podem estar organizados em um modelo de arquitetura tradicional. Ainda, é possível perceber que os estudos realizados fazem uso de métodos de extração de características baseados na regra ABCD, entretanto, os dados de geometria, bordas, cores e texturas são priorizados no desenvolvimento da maioria dos diagnósticos desenvolvidos. Além disso, a etapa de classificação tende a usar formas convencionais, que são o SVM, o KNN e o MLP. Apesar de algumas das abordagens apresentadas apresentarem resultados aceitáveis, novas formas de tratar esse problema precisam ser investigadas.

Tabela 1 – Trabalhos relacionados a abordagem convencional de identificação do melanoma

Autores	Pré-processamento	Segmentação	Extração de Características	Classificação
Ercal et al. (1994)	*	*	- coordenadas de cor; - cromaticidade relativa; - percentual de assimetria, e; - indexação de irregularidades.	- MLP.
Gauster et al. (2001)	- operações morfológicas.	- limiar adaptativo, e; - agrupamento de cor.	- área e perímetro; - regiões polares, e; - estatística em cores.	- KNN.
Celcbi et al. (2007)	- suavização por mediana de cor.	- quantização, e; - agrupamento de cor.	- texturas com matrizes de co-ocorrências.	- SVM.
Rahman, Bhattacharya e Desai (2008)	*	- limiar adaptativo, e; - <i>Fuzzy K-means</i> .	- texturas com matrizes de co-ocorrência.	- SVM; - KNN, e; - MLP.
Maglogiannis e Doukas (2009)	*	*	- metodologias de texturas, e; - irregularidade de bordas.	- Redes Neurais, e; - Aprendizado estatístico.
Cavalcanti e Scharcanski (2011)	- fechamento morfológico	- dissimilaridade entre texturas e cores.	- área e perímetro; - diâmetro e magnitude; - semelhança e gradientes, e; - estatística e quantização de cores.	- KNN, e; - árvore de decisão
Ruiz et al. (2011)	- fechamento morfológico em canais, e; - suavização.	- limiar adaptativo.	- dissimilaridade de formas e cores.	- KNN; - MLP, e; - Naive Bayes
Korotkov e Garcia (2012)	*	*	- texturas com <i>Haralick</i> ; - texturas com Gabor; - estatísticas em cores, e; - irregularidades com HOG;	*
Barata et al. (2014a)	*	- limiar adaptativo	- dados de cor e textura por histogramas, e; - <i>Bag of Features</i> .	- Adaboost; - SVM, e; - KNN.

3.2 Pesquisas em Aprendizado Profundo

O Aprendizado Profundo com RNCs tem conduzido grandes avanços em muitas tarefas que envolvem descrição e classificação de imagens (KAWAHARA; BENTAIEB; HAMARNEH, 2016). Essas redes já foram bastante estudadas nos trabalhos de Cun et al. (1990) para identificação de dígitos numéricos, de Lawrence et al. (1997) para reconhecimento de faces, e de Sermanet e LeCun (2011) para reconhecimento de sinalizações de trânsito.

Recentemente, Shi e He (2011) sugeriram a utilização das redes RNCs para tratar problemas associados a imagens médicas, visando aproveitar as suas capacidades na identificação e classificação de enfermidades. Nesse sentido, descrevemos algumas abordagens que obtiveram êxito no reconhecimento do câncer maligno utilizando essas redes.

Em Cruz-Roa et al. (2013) foi desenvolvido um estudo avaliativo entre os descritores de *Bag of Features* (BOF) e *Discrete Cosine Transform* (DCT) (baseada em *Haar*) na identificação da metástase de células basais em imagens de histopatologias sanguíneas. A classificação desses dados utilizou aprendizado profundo por meio de um modelo de Rede Neural Convolutiva. Os testes realizados nesse trabalho utilizaram 1.417 imagens e os seus melhores índices foram atingidos pelos dados extraídos do descritor DCT. As

representações aprendidas pela rede convolucional proposta atingiram um precisão de 91,4% no diagnóstico correto da metástase.

No trabalho de [Carneiro, Nascimento e Bradley \(2015\)](#) é desenvolvida uma forma de reconhecer o câncer de mama com a utilização de imagens de mamografia. Essa proposta utiliza redes convolucionais na segmentação e na classificação dos nódulos cancerígenos em duas classes (benigna e maligna). O processo de segmentação é realizado pelo modelo pré-treinado ImageNet que define as classes de *pixels* que pertencem aos nódulos. Em seguida, o seu produto é utilizado para treinar outro modelo de rede convolucional. Por fim, a classificação dos nódulos é dada por meio das estimativas de riscos na saída da rede com a técnica *Breast Imaging-Reporting and Data System* (BI-RADS), que determina uma taxa de precisão no diagnóstico. Os testes realizados com essa proposta utilizou 1.090 imagens mamográficas e obteve uma acurácia de 90% na classificação correta do câncer maligno.

Em [Codella et al. \(2015\)](#) foi proposta uma abordagem de classificação de imagens dermatoscópicas utilizando Redes Neurais Convolucionais pré-treinadas. Esse método aborda uma combinação de Aprendizado Profundo, *sparce coding* e SVM para reconhecimento do melanoma. Os resultados obtidos nos testes com essas redes pré-treinadas para 2.624 imagens em três classes (normal, anormal e melanoma) atingiram uma acurácia de 72,3%. Essa abordagem considera apenas os recursos de treinamento das redes com imagens diversas, desconsiderando as características do melanoma pela regra ABCD.

Já em [Majtner, Yildirim-Yayilgan e Hardeberg \(2016\)](#) foi proposta uma abordagem de reconhecimento de lesões de pele com uma combinação de redes convolucionais e recursos convencionais de classificação. Na etapa de extração de características essa proposta sugere o uso da arquitetura de convolução AlexNet ([KRIZHEVSKY; SUTSKEVER; HINTON, 2012](#)) pré-treinada. Nesse trabalho, a estimativa de dados é dada pelas últimas camadas da arquitetura. A classificação dos dados é feita com SVM, que atinge uma acurácia de 80,5% na segunda fase da proposta para um total de 900 imagens em duas classes (melanoma e não melanoma). Essa abordagem também desconsidera o padrão ABCD em seu processo de classificação convolucional.

No trabalho de [Kawahara, BenTaieb e Hamarneh \(2016\)](#) é sugerida uma classificação do melanoma sem realizar nenhum tipo de pré-processamento ou segmentação de imagens dermatoscópicas. Essa abordagem também utiliza uma estrutura de rede convolucional pré-treinada para descrever características. O diferencial dessa abordagem é a aplicação de uma normalização em lotes de imagens (*batch normalization*) antes de extrair seus recursos. Os resultados obtidos nos testes com 1.300 imagens em dez classes atingiram uma precisão de 81,8% na identificação correta das lesões.

A Tabela 2 apresenta um breve resumo dos trabalhos relacionados. A sua análise prevê que ainda há muito o que ser explorado das Redes Neurais Convolucionais para

melhorar o reconhecimento em imagens. O treinamento dessas redes com elementos específicos pode promover melhores formas de descrever e classificar imagens. Apesar desse processo ser dispendioso, o seu produto final permite captar características e identificá-las para qualquer conjunto de imagens em poucos segundos.

Tabela 2 – Trabalhos relacionados ao aprendizado profundo

Autores	Segmentação	Extração de Características	Classificação
Cruz-Roa et al. (2013)	*	- <i>Bag of Features</i> , e; - <i>Discrete Cosine Transform</i> baseada em <i>Haar</i> .	- RNC.
Carneiro, Nascimento e Bradley (2015)	- RNC (ImageNet) pré-treinada.	*	- RNC com <i>Breast Reporting and Data System</i> .
Codella et al. (2015)	*	- RNC.	- SVM. e; - <i>sparse coding</i> .
Majtner, Yildirim-Yayilgan e Hardeberg (2016)	*	- RNC (AlexNet) pré-treinada.	- SVM.
Kawahara, BenTaieb e Hamarneh (2016)	*	*	- RNC pré-treinada com batch normalization.

3.3 Considerações Finais

Este capítulo descreveu alguns dos trabalhos relacionados que são voltados para a existência de uma estrutura convencional de detecção de lesões melanômicas, além de novas pesquisas que estão desenvolvidas utilizando aprendizado profundo com as Redes Neurais Convolucionais. O conteúdo apresentado é descrito por meio dos principais pontos de cada abordagem, priorizando os principais métodos desenvolvidos em cada uma das etapas do diagnóstico das imagens dermatoscópicas. O próximo capítulo descreve os materiais e métodos essenciais para o desenvolvimento da proposta deste trabalho.

4 Materiais e Métodos

A maioria das abordagens desenvolvidas para o reconhecimento do melanoma em imagens dermatoscópicas segue uma estrutura tradicional, que desenvolve uma diversidade de métodos no diagnóstico das lesões. Nesse contexto, este capítulo apresenta alguns dos principais métodos voltados para a estrutura convencional de classificação de imagens dermatoscópicas, além dos conceitos essenciais para o entendimento do aprendizado profundo, utilizando um dos seus principais métodos, as Redes Neurais Convolucionais. Esses materiais e métodos são os aparatos primordiais para a elaboração da proposta deste trabalho.

4.1 Estrutura Tradicional de Reconhecimento do Melanoma

Como descrito anteriormente, o processo de captura em qualquer abordagem voltada ao diagnóstico do melanoma sempre utiliza o dermatoscópio, que é uma câmera digital de alta resolução, capaz ampliar seu campo de captura em até dez vezes o tamanho normal de uma lesão de pele (FRANGE; ARRUDA; DALDON, 2012). Ela serve para produzir as imagens dermatoscópicas que são utilizadas nos sistemas de detecção para avaliar a presença ou ausência do melanoma por meio das características dos *pixels* (MAGLOGIANNIS; DOUKAS, 2009).

O objetivo das abordagens que utilizam imagens dermatoscópicas é fornecer um diagnóstico paralelo, além de uma segunda opinião no reconhecimento do melanoma. As ferramentas desenvolvidas nesse sentido, geralmente, são baseadas na regra ABCD. Ademais, os aspectos de rugosidade (texturas) também são interpretados em conjunto com essa regra. Essa característica garante uma maior efetividade no diagnóstico do melanoma (KOROTKOV; GARCIA, 2012).

Nas etapas básicas de pré-processamento, segmentação, extração de características e classificação são empregados métodos para cumprir os objetivos propostos, que fazem a correção, separação, representação e identificação das lesões. Uma seleção correta desses métodos pode propiciar ferramentas mais eficientes (BARATA et al., 2014b). Por outro lado, eles ainda precisam ser testados, avaliados e comparados a outras formas de detecção semelhantes. Essas fases têm o propósito de comprovar a efetividade do diagnóstico entre abordagens semelhantes (KOROTKOV; GARCIA, 2012).

Uma boa parte das abordagens tradicionais que já foram desenvolvidas para o problema de reconhecimento do melanoma sugerem a utilização de técnicas de correção baseadas em morfologia matemática e interpolação bilinear, separação de regiões com

fatores de dissimilaridade, extração de características seguindo a regra ABCD e classificação utilizando os métodos convencionais SVM, KNN e MLP (MAGLOGIANNIS; DOUKAS, 2009). Nesse sentido, as próximas seções e subseções apresentam uma descrição dos métodos utilizados neste trabalho para o desenvolvimento do diagnóstico do melanoma.

4.1.1 Pré-Processamento

Apesar do processo de aquisição favorecer os melhores meios de visualização das lesões de pele, as imagens produzidas podem conter elementos que influenciam negativamente na qualidade do diagnóstico do melanoma. A etapa de pré-processamento tem o objetivo de corrigir essas imperfeições, retirando ou pelo menos aliviando os seus efeitos (KOROTKOV; GARCIA, 2012).

O método selecionado para essa etapa na abordagem tradicional precisa rejeitar as estruturas espaciais indesejadas, reconstruindo ou atenuando as suas proporções nas imagens dermatoscópicas. Um dos elementos principais que podem comprometer qualquer sistema de reconhecimento são os pelos contidos nas lesões (BARATA et al., 2014a). Para essa correção, são sugeridas as operações de fechamento morfológico, interpolação bilinear e suavização. Apesar da simplicidade dessas operações, elas são amplamente utilizadas para amenizar ou remover esses artefatos indesejados e seus resultados foram comprovados por meio dos resultados obtidos nos trabalhos de Celebi et al. (2007), Nguyen, Lee e Atkins (2010), Abbas, Celebi e García (2011). A combinação dessas operações foi desenvolvida nas pesquisas de Lee et al. (1997).

A aplicação das operações de correção foi dividida em três fases na abordagem tradicional. A primeira localiza os *pixels* que representam os pelos nas imagens RGB (*Red*, *Green* e *Blue*), empregando a operação morfológica de fechamento, que é descrita por Gonzalez e Woods (2002). Essa operação é executada em um canal de cor por vez. Em seguida, cada subimagem de canal é rotacionada nos ângulos de 0°, 45° e 90° para localizar o sentido dos pelos. Em seguida, é calculada a maior resposta por canal de cada subimagem para cada um dos ângulos, que permite uma generalização da escala de cinza. Ao fim desta operação, uma máscara binária é produzida por intermédio da diferença entre as bandas de cores e a resposta máxima de cada uma das rotações. A Equação 4.1 apresenta um esboço dessa primeira fase com o canal de cor R,

$$G_R = |O_R - \max\{O_R \bullet S_0, O_R \bullet S_{45}, O_R \bullet S_{90}\}|, \quad (4.1)$$

em que “•” representa a operação de fechamento morfológico, G_R a máscara binária para o canal R, O_R o canal R original, e S_0 , S_{45} e S_{90} representam as operações de sentido. Essa função também é executada para os canais G e B. Uma máscara binária é computada com

base nas coordenadas da imagem, que é denotada por $M(x, y)$ na Equação 4.2. O valor de T representa um limiar pré definido.

$$M(x, y) = \begin{cases} 1 & \text{se } G_R(x, y) > T \\ 0 & \text{se outro valor} \end{cases} \quad (4.2)$$

Uma outra expressão que descreve essa máscara final é dada pela união das máscaras de cada canal. Essa operação é descrita na Equação 4.3,

$$M = M_R \cup M_G \cup M_B. \quad (4.3)$$

A máscara final serve para guiar a substituição dos *pixels* que representam os pelos. Porém, antes de realizar essa substituição, cada *pixel* da máscara é verificado, com o objetivo de assegurar que o mesmo esteja dentro de uma estrutura curvilínea.

A segunda fase desse processo é composta pela operação de interpolação bilinear, também descrita por [Gonzalez e Woods \(2002\)](#). Ela realiza a substituição dos segmentos *de pixels* que representam os pelos. A Equação 4.4 descreve essa operação,

$$I_n(x, y) = I_{i+j}(x, y) * \frac{D(I(x, y), I_{i+j}(x, y))}{D(I_{i+j}(x, y), I_{i+j}(x, y))} + I_{i+j}(x, y) * \frac{D(I(x, y), I_{i+j}(x, y))}{D(I_{i+j}(x, y), I_{i+j}(x, y))}, \quad (4.4)$$

em que o valor aproximado de $I_n(x, y)$ representa a interpolação de um *pixel* (n). Já o valor de D descreve a distância de um determinado *pixel* $I(x, y)$ e $I_{i+j}(x, y)$ os seus vizinhos mais próximos. A última fase desse processo de correção é definida por uma operação de suavização, utilizando o filtro da mediana, que reduz o desfoque e mantém as bordas.

4.1.2 Segmentação

Após a aplicação do método de correção, a etapa de segmentação realiza a separação da região de interesse do restante da imagem. Essa etapa tem o objetivo de localizar os *pixels* que correspondem as lesões e, em seguida, separá-los dos demais. O método selecionado nessa parte da abordagem tradicional deve ser robusto e sensível a detalhes, já que os aspectos de bordas e contornos são vitais para um melhor diagnóstico de lesões cancerígenas ([MAGLOGIANNIS; DOUKAS, 2009](#)).

A segmentação de imagens dermatoscópicas pode ser realizada por vários métodos distintos. Contudo, deve-se priorizar operações que mantêm um maior número de detalhes das lesões. Para essa etapa, sugeriu-se o método *Fuzzy K-means*. Ele obteve excelentes resultados nos trabalhos de [Schmid \(1999\)](#), [Zhou et al. \(2009\)](#) e [Xie e Bovik \(2013\)](#).

A forma de particionar uma imagem no *Fuzzy K-means* é derivada do método *K-means*. O diferencial entre ambos é a de taxa similaridade *Fuzzy*, que serve para determinar a pertinência de um *pixel* a um determinado grupo. Essa taxa é calculada com base em um parentesco (similaridade), dado pela distância entre os *pixels*. (NAZ; MAJEED; IRSHAD, 2010).

A aplicação do *Fuzzy K-means* considera uma imagem como um conjunto de *pixels* $P(P_1, P_2, P_3, \dots, P_n)$, em que os K centros desejados são escolhidos aleatoriamente. Em todos os valores de P são calculadas as distância entre os grupos $C_j (j = 1, 2, 3, \dots, C)$. Geralmente, esta etapa utiliza a distância euclidiana que é representada na Equação 4.5 (ALBERTI; ZIMMET, 1998),

$$D(x_i, x_j) = \|x_i, x_j\| = \sqrt{(x_i - x_j)^2 + (x_1 - x_j)^2}, \quad (4.5)$$

em que D representa a distância, i e j os valores das coordenadas de x e y de um *pixel*.

Os valores de K são atualizados por $D(x_i, x_j)$, que representa a distância média dos grupos de *pixels*. Os centroides devem ser deslocados até as coordenadas centrais dos grupos aos quais sejam pertinentes. Os cálculos de distância euclidiana e média dos centróides são executados até que os valores das coordenadas dos pontos centrais de cada um dos K grupos não sofram mais alterações (ALBERTI; ZIMMET, 1998).

A função de pertinência $U_{ik}(P) \rightarrow [1, 0]$, $i = 1, \dots, k$ indica o grau de parentesco de cada P_n com cada grupo K . A Equação 4.6 determina o cálculo dos centros dos grupos. Os valores de m são os coeficientes que vão ponderar o grau de pertinência por meio da distância euclidiana. O valor da pertinência de cada *pixel* é definido de acordo com a Equação 4.7. O valor de D é a distância euclidiana (JAIN; MURTY; FLYNN, 1999).

$$v_i = \frac{\sum_{k=1}^n (U_{ik})^m * P_n}{\sum_{k=1}^n (U_{ik})^m} \quad (4.6)$$

$$U_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{D_{xk}}{D_{yk}}\right)^{\frac{2}{(m-1)}}} \quad (4.7)$$

O método *Fuzzy K-means* encerra a sua execução quando um determinado número de iterações é definido ou quando U é menor que um δ limiar de convergência (JAIN; MURTY; FLYNN, 1999). A descrição do método *Fuzzy K-means* é dada no Algoritmo 1.

Por fim, o método *Fuzzy K-means* realiza várias combinações diferentes entre classes de *pixels* de uma imagem. Todavia, o agrupamento que determina a maior diferença entre o grau de similaridade das lesões e o fundo da imagem é sempre selecionado.

Algoritmo 1: *Fuzzy K-Means*

Entrada: Conjunto de *pixels* da imagem**Saída:** Distribuição do *pixels* P entre os K grupos**início**

1. Calcule os centros (w_1, \dots, w_k) , que representam os grupos
2. Calcule a distância euclidiana para cada *pixel* de entrada P
3. Calcule e atualize os valores de pertinência para U_{ik}
4. Repita os passos 2, 3 até que $U < \delta$

fim

4.1.3 Extração de Características

Na maioria dos sistemas, a etapa de extração de características tem o objetivo de representar as lesões das imagens dermatoscópicas por meio de uma estrutura matemática. Os descritores são os modelos utilizados para definir fronteiras ou regiões de *pixels*. A cada execução de um modelo específico, uma imagem é transformada em um vetor de características. Ele é uma descrição compacta dessa imagem que torna o processo de análise mais prático e simplificado (BARATA et al., 2014a).

Os descritores de características em qualquer abordagem de predição são definidos de acordo com o tratamento das imagens. Eles são escolhidos por meio da especificação do problema. Essa seleção pode proporcionar resultados mais eficazes no reconhecimento das lesões de pele. Para isso, essa pesquisa realizou um levantamento bibliográfico acerca dessa situação. As próximas subseções descrevem os principais descritores que obtiveram resultados satisfatórios na descrição de lesões de pele em imagens dermatoscópicas. Eles foram definidos por meio dos estudos de Maglogiannis e Doukas (2009) e Korotkov e Garcia (2012).

4.1.3.1 Descritor de Geometria

O descritor de geometria é baseado em operações de forma que caracterizam objetos em imagens. Os *pixels* que compõem esses objetos possuem distinções estruturais que podem ser aproveitadas por meio dos critérios de contorno (GONZALEZ; WOODS, 2002). Esse modelo descreve as linhas delimitadoras de um elemento. As operações desenvolvidas nesse descritor produzem dados que são capazes de detectar estruturas espaciais complexas das regiões de uma imagem.

As características espaciais são capazes de facilitar a identificação de dissimilaridades que podem ser utilizadas na classificação de assimetria de lesões (GANSTER et al., 2001). Nesse problema, a extração de características com o descritor de geometria atende os princípios de invariância a mudanças de orientação (rotação e translação), variação de escala e robustez a diversificação de ruídos. Esses parâmetros garantem que os dados das

estruturas de *pixels* semelhantes sejam diferenciados, independentemente das suas relações.

As operações utilizadas para descrever formas no descritor de geometria são descritas nos próximos itens e seguem os princípios dos estudos de Celebi et al. (2009) e Rangayyan, Ayres e Desautels (2007).

- área e perímetro;
- momentos invariantes e Geométricos;
- convexões de regiões (*Convex Hull*);
- delimitação de regiões (*Bounding Box*), e;
- centroide e orientação.

4.1.3.2 Descritor *Histogram of Oriented Gradients*

O descritor Histograma de Gradientes Orientados (*Histogram of Oriented Gradients* – HOG) é utilizado para identificar as características de bordas em imagens dermatoscópicas por intermédio do cálculo dos gradientes, considerando a sua orientação e magnitude. O funcionamento desse descritor consiste em contar ocorrências de orientações em porções de *pixels*, com o objetivo de codificar formas específicas que representam bordas. No seu processo de extração são desenvolvidas quatro etapas básicas que definem o vetor de atributos (DALAL; TRIGGS, 2005).

A primeira etapa consiste em realizar a aplicação das máscaras espaciais dadas por $[-1 \ 0 \ 1]$ e $[-1 \ 0 \ 1]^T$. Elas servem para ajustar a estrutura da imagem de acordo com a orientação e magnitude de cada *pixel*. Esse cálculo pode ser realizado em cada canal de cor de uma imagem, sendo considerada apenas a maior magnitude (MERY, 2011).

Na segunda etapa são designadas estruturas de células na imagem. Elas são utilizadas para criar histogramas que armazenam a direção dos gradientes. Geralmente, essas células possuem uma dimensão 8 x 8 e seus respectivos histogramas são classificados por meio da divisão em partes (*bins*), com valores entre 0° e 180° no eixo das abscissas e com pesos dos gradientes no eixo das coordenadas. A contribuição de cada *pixel* nessa fase é dada por intermédio da ponderação do valor de magnitude do gradiente em cada histograma de célula (DALAL; TRIGGS, 2005).

As mudanças de iluminação, contraste e gradiente precisam ser corrigidos nos histogramas das células. Para isso, a terceira etapa requer uma normalização local, com o agrupamento das células em blocos maiores para facilitar a correção. Geralmente, esses blocos possuem uma dimensão de 2 x 2, em que uma operação de normalização considera um conjunto de células v , $\|v\|_k$ a sua k -ésima norma e ε uma constante com valor baixo que evita divisões por zero. Essa operação é representada pela Equação 4.8, que faz com

que cada célula seja normalizada em relação a diferentes blocos (DALAL; TRIGGS, 2005).

$$v \rightarrow \frac{v}{\|v\|_2^2 + \varepsilon^2}. \quad (4.8)$$

Por fim, a quarta etapa realiza a concatenação dos blocos de uma imagem após a normalização, criando assim um vetor de blocos com histogramas de células. Esse vetor representa todas as características de bordas e arestas dos objetos presentes em uma imagem, mesmo sem o conhecimento exato da posição dos gradientes (MERY, 2011).

4.1.3.3 Descritor de Estatística em Canais de Cores

O descritor de Estatística em Canais de Cores – ECC é baseado em operações estatísticas simples em imagens, que definem características de cor. A sua implementação é baseada nos princípios de identificação de objetos, utilizando medidas que são aplicadas a cada canal de cor. Essas operações já foram utilizadas em uma diversidade de trabalhos que focam no reconhecimento de características de pele (MAGLOGIANNIS; DOUKAS, 2009).

A extração de características com o descritor de ECC é definida como um processo estatístico que determina dados referentes aos *pixels* combinados pelos canais de uma imagem, em que C_n representa um canal específico, O_x uma operação estatística x e R_n o seu resultado. Essa operação é apresentada pela Equação 4.9.

$$R_z = O_x(C_n) \quad (4.9)$$

De forma geral, o vetor de características v extraído pelo descritor ECC é representado pela combinação das operações estatísticas R_z para cada canal C_n de uma imagem. Esse vetor é representado na Equação 4.10.

$$v = \int(R_z) \quad (4.10)$$

As operações utilizadas na implementação do descritor ECC foram baseadas na promoção dos estudos de Vezhnevets, Sazonov e Andreeva (2003), no qual é possível perceber excelentes resultados de classificação por meio de características de estatísticas em cores. Além disso, esse desenvolvimento considera algumas medidas relacionadas ao reconhecimento de pele, descritas no estudo de Phung, Bouzerdoum e Chai (2005). Os itens a seguir apresentam todas as medidas implementadas nesse descritor.

- intensidades mínima, média, máxima, mediana e moda;
- desvio padrão unidimensional e bidimensional;

- energia;
- entropia;
- covariância mínima e máxima;
- correlação, e;
- *kurtosis* mínima e máxima.

4.1.3.4 Descritor com Banco de Filtros de Gabor

O descritor que utiliza o Banco de Filtros de *Gabor* (*Gabor Bank Filters* – GBF) é baseado em características que representam aspectos visuais de texturas em imagens. Esse método permite determinar relações entre dados por intermédio das frequências espaciais, aspectos de localização e orientações. As operações utilizadas nesse sentido são senóides complexas e bidimensionais, modeladas por uma função gaussiana (FOGEL; SAGI, 1989).

As *wavelets* de *Gabor* são bem localizadas em tempo e espaço, permitindo uma melhor representação de parâmetros por meio da Transformada Discreta de *Fourier*. A função de *Gabor* $g(x, y)$ e a sua respectiva transformada $G(u, v)$ podem ser definidas nas Equações 4.11 e 4.12,

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + i2\pi Wx \right], \quad (4.11)$$

$$G(u, v) = \exp \left[-\frac{1}{2} \left[\frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right], \quad (4.12)$$

em que $\sigma_u = \frac{1}{2\pi\sigma_x}$ e $\sigma_v = \frac{1}{2\pi\sigma_y}$. Nessa circunstância, a função $g(x, y)$ desenvolve um conjunto de características não ortogonais, que serve para expandir o sinal de propagação na função $G(u, v)$ (JONES; PALMER, 1987).

A criação de bancos de filtros que utiliza a função de *Gabor* pode ser definida no descritor GBF por meio da dilatação e rotação da imagem. Essas operações manipulam os parâmetros de frequência, orientação, excentricidade e simetria. A combinação dessas operações é definida por intermédio das Equações 4.13 e 4.14,

$$G(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{(x')^2 + (y')^2}{2\sigma^2}\right) \cos 2\pi \frac{x'}{\lambda} + \psi, \quad (4.13)$$

$$x' = (x \cos\theta + y \sen\theta) \quad y' = (-x \sen\theta + y \cos\theta), \quad (4.14)$$

no qual o valor de λ representa um comprimento de onda, θ os seus ângulos, σ o desvio padrão da distribuição que está relacionada a largura gaussiana de um filtro e, ψ o tamanho da janela de um núcleo γ . A Equação 4.13 produz uma função senoidal por meio

da modelagem gaussiana no qual a Equação 4.14 rotaciona a Equação 4.13, de acordo com o valor de θ (FOGEL; SAGI, 1989).

Finalmente, as descrições de textura com o método GBF são obtidas por intermédio da filtragem das imagens por um núcleo de *Gabor*. Esse processo é realizado para todo o banco em que se obtém uma resposta específica para cada um dos núcleos, por meio da escolha de um intervalo de frequência e um número de orientações (LAHMIRI; BOUKADOUM, 2013).

4.1.3.5 Descritor *Local Binary Pattern*

O descritor de características Padrão Binário Local (*Local Binary Pattern* – LBP) também é baseado em texturas. Esse método é definido por intermédio de uma função de variação espacial que é obtida a partir da intensidade dos *pixels* de uma imagem. O conceito apresentado nesse descritor foi adaptado por Ojala, Pietikäinen e Harwood (1996), com o objetivo de obter características relevantes em situações no qual as imagens sofrem mudanças de iluminação.

O objetivo do LBP é determinar a intensidade de um *pixel* por meio de um valor numérico no qual as suas coordenadas precisam ser definidas de acordo com uma vizinhança. Esse valor é definido por $W = (x, y)$ na qual as amostras P definem uma vizinhança circular de raio R , com o centro definido sobre o *pixel* W . A partir da intensidade g_c do pixel na posição W , obtém-se uma sequência de valores binários $T_p = (S(g_0 - G_c), S(g_1 - g_c), \dots, S(g_{p-1} - g_c))$. O valor de S determina uma função sinal que quando nula possui valor negativo, caso contrário seu valor é 1 (MERY, 2011).

Uma sequência de bits T_p é mapeada de acordo com as transições entre 0 e 1. Os valores iguais a 1 determinam a relevância na caracterização de textura. O descritor de $LBP_{P,R}$ é invariante a translação quando ocorre mais de duas transições entre 0 e 1. Esse processo pode ser descrito pela Equação 4.15,

$$LBP_{P,R} = \begin{cases} 1, & \sum_{p=0}^{p-1} (g_p - g_c) \\ 0, & p + 1 \text{ caso contrário} \end{cases}, \quad (4.15)$$

em que $LBP_{P,R}$ pode assumir $P + 2$ valores distintos que representam as texturas de um *pixel* em relação a uma vizinhança. Os valores de P e R podem ser alterados para representar texturas em outras escalas na qual essa operação pode ser realizada em cada canal de uma imagem, extraindo características distintas de textura (OJALA; PIETIKÄINEN; HARWOOD, 1996).

4.1.3.6 Descritor de *Haralick*

O descritor de *Haralick* é baseado em matrizes de co-ocorrência para extrair características de texturas entre os níveis de cinza de uma imagem. Esses dados servem para tabular as diferentes combinações de intensidade de *pixels* por intermédio das suas vizinhanças (HARALICK, 1979).

De modo geral, a co-ocorrência é definida como uma frequência relativa $P(i, j; d, \theta)$ em que i e j são as coordenadas dos vizinhos de um *pixel*, separados por uma distância d e uma orientação θ . O valor de θ pode assumir as direções 0° , 45° , 90° e 135° . A composição da matriz de co-ocorrência é feita mediante as variações dos níveis de cinza de uma imagem. Entretanto, é necessário calcular a probabilidade da frequência relativa em cada vizinhança de *pixels*, que é dada por meio da célula $M_{i,j}$ (HARALICK, 1979).

A matriz de co-ocorrência final é calculada a partir de todas as iterações de vizinhanças de uma imagem para cada *pixel*. A sua definição é dada por uma matriz quadrada $M_{i,j}$, com probabilidade de frequências relativas $P_{i,j}$. Essa operação é definida na Equação,

$$P_{i,j} = \frac{M_{i,j}}{\sum_{i,j=0}^{N-1}}, \quad (4.16)$$

no qual N representa o número de linhas e colunas, já que M é uma matriz quadrada (HARALICK, 1979).

Os conceitos do método de *Haralick* oferecem a possibilidade de extração de características em imagens multicanais. Nesse caso, cada canal é utilizado para extrair as matrizes de co-ocorrência. A partir dessas matrizes, é possível determinar dados de texturas em um único vetor de atributos por intermédio das medidas estatísticas dos itens seguintes em que C representa um canal específico e μ os valores médios das distribuições (MERY, 2011).

- segundo momento angular $\rightarrow \int_{sma} = \sum_{i,j=0}^C P_{i,j}^2$;
- entropia $\rightarrow \int_{ent} = - \sum_{i,j=0}^C P_{i,j} \log P_{i,j}$;
- contraste $\rightarrow \int_{con} = \sum_{i,j=0}^C (i - j)^2 P_{i,j}$;
- variância $\rightarrow \int_{var(i)} = \sum_{i,j=0}^C (i - \mu_i)^2 P_{i,j}$ e $\int_{var(j)} = - \sum_{i,j=0}^C (j - \mu_j)^2 P_{i,j}$;
- correlação $\rightarrow \int_{corr} = \frac{1}{\sigma_x \sigma_y} \sum_{i,j=0}^C (i - \mu_i)(j - \mu_j) P_{i,j}$, e;
- homogeneidade $\rightarrow \int_{hom} = \sum_{i,j=0}^C \frac{1}{1+(i-j)^2} P_{i,j}$

4.1.4 Classificação

A etapa de classificação é responsável por diagnosticar as lesões de pele mediante as características extraídas das imagens dermatoscópicas. Os métodos classificadores

empregados nessa etapa são responsáveis por predizer classes desconhecidas nos conjuntos de dados por meio de um reconhecimento sistemático. Essa etapa permite ainda a utilização de vários métodos com a mesma finalidade na qual é possível desenvolver um comparativo de resultados (BARATA et al., 2014a).

A eficácia de qualquer abordagem de reconhecimento de lesões de pele também depende dos classificadores empregados. Uma diversidade de métodos já foram estudados na literatura para esse problema. Entretanto, alguns deles são preferíveis, pois podem proporcionar melhores resultados de diagnóstico do melanoma. Nessa perspectiva, esse trabalho se baseou na promoção dos estudos de Korotkov e Garcia (2012) e Maglogiannis e Doukas (2009) para definir os métodos de *K-Nearest Neighbor*, *MultiLayer Perceptron* e *Support Vector Machine* que são considerados tradicionais para esse problema. As próximas subseções apresentam uma descrição desses métodos.

4.1.4.1 *K-Nearest Neighbor*

O *K-Nearest Neighbor* (KNN) é um dos classificadores mais simples já desenvolvidos. Esse método é baseado no aprendizado por analogia em que cada amostra de um conjunto de dados pertence a um espaço *n-dimensional*. O seu objetivo é identificar grupos de dados considerando as medidas de distância entre os *K* atributos que são chamados de *K*-vizinhos mais próximos (HASTIE; TIBSHIRANI, 1996).

O classificador KNN é representado por um conjunto de dados *D*, com um subconjunto de treinamento $\{x_i, y_i\}$ em que $i = \{1, 2, 3, \dots, D\}$. Seja *z* um novo dado a ser agrupado e x_i um dado do conjunto treino *D*, um cálculo de similaridade é aplicado para determinar a distância entre eles. Nesta situação, a distância euclidiana é considerada a forma mais indicada para determinar a proximidade entre *z* e x_i . Em seguida, se a distância entre eles estiver dentro dos *K* valores mais próximos, o dado *z* é agrupado à classe c_i mais próxima. A Figura 11 apresenta um exemplo de classificação com esse método, utilizando os valores de $K = \{1, 3 \text{ e } 5\}$. Por votação, *z* é agrupado à classe azul para $K = 5$ (PETERSON, 2009).

O processo de classificação com o KNN pode ser computacionalmente exaustivo, se considerado um conjunto com muitos dados. Contudo, em algumas aplicações, o processo é bem aceitável e produz bons resultados (HASTIE; TIBSHIRANI, 1996).

4.1.4.2 *MultiLayer Perceptron*

O Perceptron Multicamadas (*MultiLayer Perceptron* – MLP) é considerado uma das forma mais simples de Redes Neurais Artificial (RNAs) usada para a classificação de dados. O *Perceptron* (neurônio artificial) é a sua unidade fundamental e a sua composição

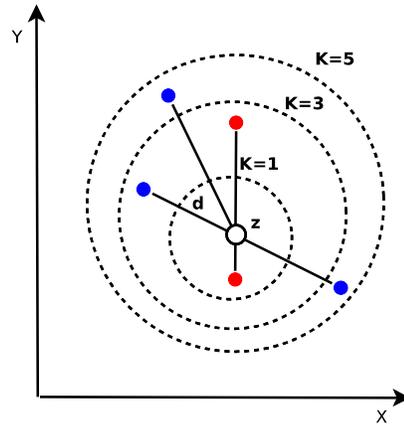


Figura 11 – Classificação de dados com o KNN.

é feita por sinais de entrada (x_i), pesos sinápticos (w_j), um integrador Σ , uma função de ativação $f(\cdot)$ e os sinais de saída (y_i). A Figura 12 representa esta estrutura (HAYKIN, 2001).

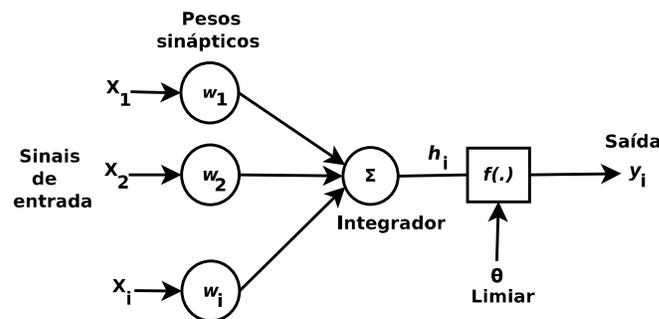


Figura 12 – Modelo de neurônio artificial. Adaptado de (HAYKIN, 2001)

A estrutura do MLP é composta por vários neurônios organizados em múltiplas camadas e cada um deles possui o seguinte comportamento que é descrito na Equação 4.17,

$$y_i = f\left(\sum_{i=1}^n w_i \cdot x_i - \theta\right), \quad (4.17)$$

em que f é uma função de ativação e θ um limiar conhecido como *bias*.

A principal característica do MLP na classificação de dados é a capacidade de separar dados não lineares. Para esse fim, é necessário que a função de ativação utilizada nas camadas intermediárias seja não linear. A função mais utilizada nessas redes é a sigmóide. Ademais, essa rede é progressiva (*feedforward*), pois as suas entradas propagam-se mediante a estrutura em um único sentido em que uma camada se conecta unicamente à outra seguinte (BRAGA; CARVALHO; LUDERMIR, 2000).

Uma rede MLP só é capaz de classificar dados de um problema por meio de um treinamento prévio. Para tal, tem sido aplicado com sucesso o treinamento supervisionado com o método retropropagação (*backpropagation*). As operações executadas por esse método consistem em dois passos, o primeiro é a propagação da entrada na rede pelas camadas posteriores e o segundo a retropropagação do erro que ocorre no sentido inverso. Esse processo serve para ajustar os pesos sinápticos e são representados pelas Equações 4.18 e 4.19,

$$\rho_i(n) = d_i(n) - y_i(n), \quad (4.18)$$

em que $\rho_i(n)$ representa a diferença entre o valor ótimo para uma determinada entrada e saída que é gerada pela rede,

$$\varepsilon(n) = \frac{1}{2} \sum_{i \in C} \rho_i^2(n), \quad (4.19)$$

em que C é o conjunto de todos os neurônios que pertencem a camada de saída da rede e $\varepsilon(n)$ o valor temporal do erro (JÚNIOR; MONTGOMERY, 2007).

Por fim, a aprendizagem do MLP consiste em encontrar um conjunto de pesos w que reduz o custo da função de erro. A eficiência disso é observada pela medição dos erros que é realizada com as entradas que ainda não foram apresentadas à rede (HAYKIN; HAYKIN, 2001).

4.1.4.3 Support Vector Machine

A Máquina de Vetor de Suporte (*Support Vector Machine* – SVM) é uma teoria de aprendizado estático que tem sido amplamente utilizada para resolver problemas de classificação e regressão. A ideia principal apresentada por esse método é a separação de dados por meio de hiperplanos que são funções de separação. Existem várias possibilidades de obter hiperplanos no SVM, entretanto, é necessário definir qual deles distingue melhor as classes (YUAN et al., 2006). A Figura 13 mostra um exemplo visual da criação destes hiperplanos em duas classes: C1 e C2.

A qualidade das funções de separação no SVM é diretamente proporcional à distância apresentada entre as classes analisadas. A margem entre elas pode ser compreendida como a soma entre as distâncias dos dados mais próximos entre essas funções. O conceito de proximidade é dado pelo vetor de suporte que estabelece valores e designa as distâncias entre as margens das classes. A Figura 14 descreve um hiperplano ótimo em que a linha sólida o representa por intermédio da função $Y = w * x + b = 0$. As margens $y = w * x + b > 1$ e $y = w * x + b < -1$ são as regiões limitadoras desse hiperplano, as amostras ζ que estão dispostas sobre hiperplanos ($Y > 1$ e $Y < -1$) são os vetores suporte, e ε uma variável de

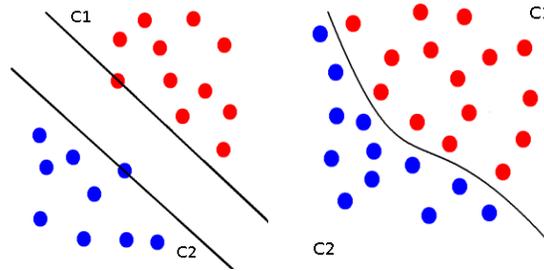


Figura 13 – Exemplo de classificação com SVM.

"folga". A variável "folga" permite uma suavização do classificador, permitindo que alguns dados permaneçam entre os hiperplanos (YUAN et al., 2006).

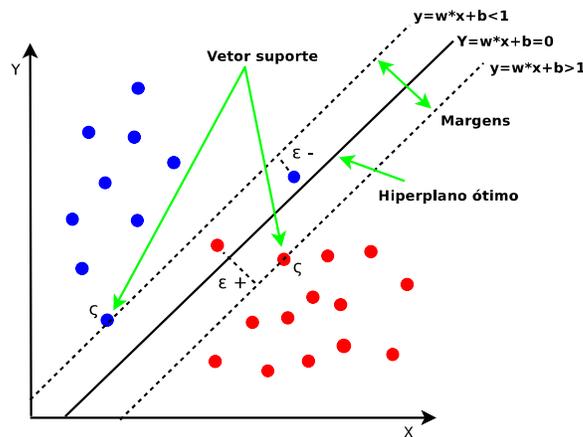


Figura 14 – Hiperplano ótimo.

Na prática, a aplicação do SVM dificilmente trata de problemas linearmente separáveis. Por isso, ele implementa formas que permitem que as amostras possam estar nas margens ou, até mesmo, na região de outra classe. Para isso, ele introduz a folga ε_n , que resulta no método de margens suaves (GUNN et al., 1998). A Equação 4.20 descreve essa proposta.

$$y_i(w * x + b) \geq 1 - \varepsilon_i, \quad \varepsilon \geq 0, \quad \forall i = 1, \dots, n. \quad (4.20)$$

O SVM pode ser descrito como um conjunto de dados D , no qual é determinado um conjunto de treinamento $\{x_i, y_i\}$, com $i = \{1, 2, 3, \dots, D\}$. Os valores de $x_i \in R^m$ são considerados representações vetoriais de um conjunto $y_i \in \{-1, 1\}$ de classes. Neste contexto, existe uma distribuição de probabilidade $P\{x, y\}$ que ainda não é conhecida e será retirada do conjunto de treino (YUAN et al., 2006).

A etapa de treinamento do SVM consiste em preparar o classificador de forma que ele aprenda uma relação de $x \rightarrow y$ por meio das classes de treinamento $\{x_i, y_i\}$. Desta

forma, ao fim desse treinamento, o método será capaz de classificar o conjunto teste $\{x, y\}$ que ainda não foi apresentado, seguindo a distribuição de probabilidade P (RAZMJOOY et al., 2013).

Um erro no conjunto de treinamento é indicado pelo valor de $\varepsilon_i > 1$, no qual a soma de seus valores representam o limite de erros de treinamento que precisa ser minimizado pela Equação 4.21,

$$\text{Minimizar}_{(w,b,\varepsilon)} \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^n \varepsilon_i \right), \quad (4.21)$$

em que C representa um termo que regulariza os pesos à minimização dos erros de treinamento e $\sum_{i=1}^n \varepsilon_i$ uma redução dos erros marginais que serve para indicar valores entre as margens (GUNN et al., 1998).

O mapeamento do espaço de características com o SVM pode ter uma dimensão altíssima que faz com que a computação desse mapeamento seja custosa. Para a solução dessa problema, o cálculo dos produtos escalares de x_i e x_j de duas amostras do conjunto é obtida com uma função *kernel* (K), descrita na Equação 4.22. Ela é a mais utilizada para a construção de hiperplanos não lineares.

$$K(x_i, x_j) = \exp\left(\frac{\|x_i, x_j\|}{2\sigma^2}\right), \quad (4.22)$$

em que $\|x_i, x_j\|$ corresponde a distância euclidiana entre dois vetores de características e σ um parâmetro livre (YUAN et al., 2006).

4.2 Métodos de Aprendizado Profundo

O Aprendizado Profundo é uma abordagem de RNA que desenvolve arquiteturas com um grande número de camadas intermediárias. Os métodos que reproduzem essas redes são baseados em representações distribuídas, que são compostas por vários níveis de abstrações de um conjunto de dados, selecionados por intermédio das concepções adquiridas ao longo da arquitetura (LECUN; BENGIO; HINTON, 2015).

As motivações do Aprendizado Profundo exploram conceitos de aprendizado em alto nível, desenvolvidos por conceitos de baixo nível que são obtidos por meio da escolha dos melhores recursos. Este processo gera uma transformação dos dados e determina um melhor aprendizado da estrutura neural profunda. As operações realizadas com esse propósito podem ser representadas por funções não-lineares que extraem ou identificam características. Ainda, é possível utilizar nesse modelo de arquitetura profunda, ao mesmo tempo, as técnicas de aprendizado supervisionado e não-supervisionado (SCHMIDHUBER, 2015).

A maioria das Redes Neurais Profundas (*Deep Neural Network*) mantém uma estrutura com dois tipos de camadas. As camadas iniciais recebem os dados de entrada e realizam as suas modificações, juntamente com a remoção das redundâncias. Já as últimas camadas recebem os dados ajustados e realizam a sua classificação em grupos (BENGIO; GOODFELLOW; COURVILLE, 2015).

Atualmente, muitos problemas têm sido resolvidos com a abordagem de Aprendizado Profundo. Os métodos que desenvolvem os seus conceitos permitem representar qualquer função lógica por meio das camadas de uma RNA. Entretanto, a quantidade de camadas pode crescer de maneira polinomial, de acordo com a quantidade de entradas do problema. A vantagem disso é uma maior proporção de abstrações que podem ser utilizadas na identificação de melhores soluções (DENG; YU, 2014).

4.2.1 Redes Neurais Convolucionais

Alguns dos métodos de Aprendizado Profundo mais utilizados para identificar atributos relevantes são as Redes Neurais Convolucionais (RNCs). A sua utilização está obtendo grandes progressos em problemas que envolvem extração de características e classificação de imagens. Essas redes são baseadas no processo visual humano de interpretar cenários com características próximas e relacionáveis (GOODFELLOW et al., 2013).

O aprendizado nas RNCs é feito por intermédio de uma estrutura de neurônios localmente conectadas que servem para representar pequenas porções de um conjunto de dados por meio das suas camadas. Os seus conceitos são baseados em campos receptivos locais, convolução, compartilhamento de pesos e subamostragem. Geralmente, as arquiteturas convolucionais são formadas por três tipos de camadas: convolução, *pooling* e totalmente conectada (BENGIO; GOODFELLOW; COURVILLE, 2015).

4.2.2 Camada de convolução

A convolução é a principal operação realizada nas RNCs, pois ela é responsável por encontrar os campos receptivos locais de uma imagem por meio de seus *pixels*. Essa operação é definida como um operador linear entre duas funções e o seu resultado é uma terceira função. Nela é representada uma combinação s entre as matrizes f e g , em que o valor de f representa a matriz de entrada e g um filtro pré definido (GOODFELLOW et al., 2013). A convolução é representada na Figura 15.

As operações realizadas na camada de convolução podem ser representadas por matrizes $v = (v_1, v_2, \dots, v_m)$ em que os neurônios $y(v)$ possuem coordenadas v na vizinhança $x(v - u)$. Os valores de u representam a distância entre os neurônios na posição das coordenadas v . Estes valores são adquiridos por meio dos campos receptivos de $y(v)$. Já

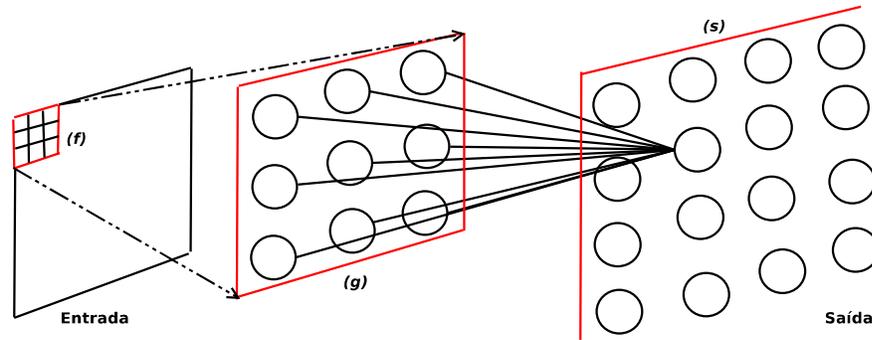


Figura 15 – Convolução.

os valores de $w(u)$ são os pesos dos neurônios $y(v)$ em relação aos seus vizinhos e z os resultados da convolução (SERMANET; LECUN, 2011).

A camada de convolução serve para captar características em um conjunto de dados. Essa camada é representada por conjuntos de neurônios organizados e subdivididas em séries de agrupamentos, conhecidos como canais. Em cada um dessas estruturas existe um conjunto reticular de neurônios localmente conectados que é descrito na Figura 16. Nessa situação, a convolução é dada por um neurônio $y_i(v)$ no canal i , com posição na coordenada (v) , em uma camada em que podemos definir um conexão com uma vizinhança na camada anterior $x_j(v - u)$, por um canal j e uma distância u em relação às coordenadas (v) . O neurônio $y_i(v)$ recebe conexões de todos os canais j da camada anterior e as conexões locais são representadas ao longo da estrutura reticular do canal, em que o domínio é dado por (u) e o campo receptivo de $y_i(v)$ são todos os neurônios da sua vizinhança (BENGIO; GOODFELLOW; COURVILLE, 2015).

A camada de convolução é definida em arquitetura profunda para realizar filtragens espaciais com o intuito de encontrar características relevantes na estrutura dos seus elementos. Essa fase é realizada por intermédio de máscaras que indicam bordas, arestas e cores. Essas estruturas são formas de representar filtros de convolução e a sua propagação é denotada como uma operação de vizinhança entre neurônios $x(v - u)$ e o seu resultado é uma combinação linear de valores $h(v)$, conhecidos como mapa de características (*features map*) (LEE; CHEN, 2015).

A cada operação de convolução, utilizam-se os mesmos parâmetros para que uma imagem seja completamente coberta por uma máscara. Essa ação só é possível por meio do compartilhamento de pesos entre os neurônios das camadas convolucionais. Essa operação é proporcionada pelos canais da rede, em que cada um dos seus neurônios pode responder igualmente para um conjunto de características relacionadas, de acordo com a movimentação das máscaras. A propagação desse sinal é descrita pelas Equações 4.23 e 4.24, em que f é a função de ativação e $h(v)$ a combinação linear do neurônio $y_i(v)$

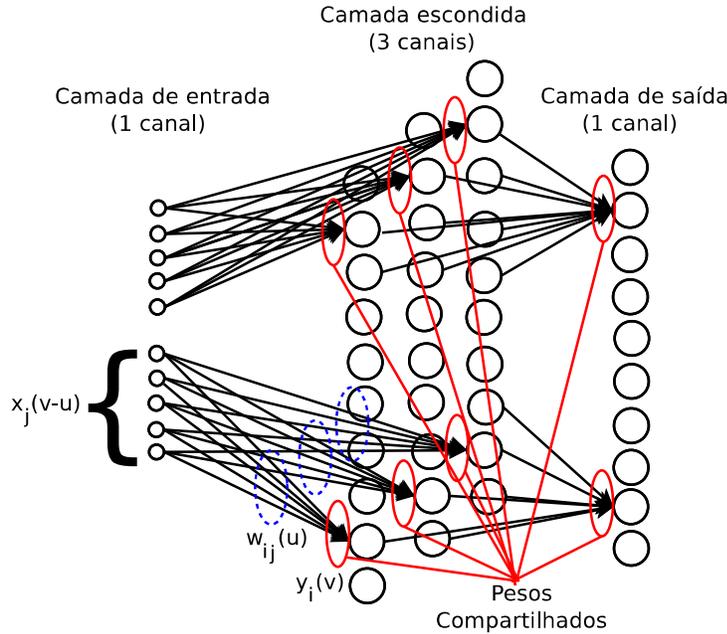


Figura 16 – Reticulado de neurônios convolucionais.

(GOODFELLOW et al., 2013).

$$h(v) = \sum_j \sum_u w_{ij}(u) x_j(v-u) \quad (4.23)$$

$$y_i(v) = f(h(v)) \quad (4.24)$$

A distribuição de pesos iguais permite definir valores para conexões que representam um campo receptivo local, permitindo a equivariância à translação. Se uma função translar um conjunto de entrada por meio de determinados valores, a saída da convolução também sofrerá essa operação. A igualdade lógica entre essas funções permite que filtros subsequentes tenham acesso a filtros anteriores por meio dessa distribuição de pesos. Essas operações acontecem sem as perdas de parâmetros, já que a tendência dos campos receptivos é aumentar em cada camada convolucional. A Figura 17 apresenta um exemplo da distribuição de pesos iguais entre conexões de neurônios por canal, em que um mesmo campo receptivo é definido em três camadas convolucionais (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

Outra característica do compartilhamento de pesos é a diminuição do consumo de memória por camada, pois a quantidade de parâmetros é preservada a cada iteração. Por fim, a camada de convolução é equivalente a estrutura RNA com pesos compartilhados, entretanto, a sua definição é mais precisa e sua notação mais concisa (LEE; GALLAGHER; TU, 2016).

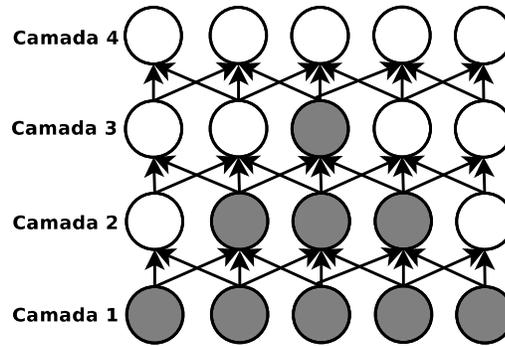


Figura 17 – Compartilhamento de pesos de campos receptivos em três camadas.

4.2.3 Camada de *Pooling*

Apesar do processo realizado pela camada de convolução reduzir a dimensionalidade dos dados, ainda é necessário realizar mais reduções para viabilizar as RNCs. Nesse problema, a camada de *pooling* serve para realizar uma subamostragem dos dados. A sua estrutura é responsável por tratar mapas de características, reduzido as suas linhas e colunas e enviando os seus valores às camadas seguintes (LECUN; BENGIO; HINTON, 2015). A Figura 18 ilustra uma redução de um mapa de características por meio de uma função f de *pooling*.

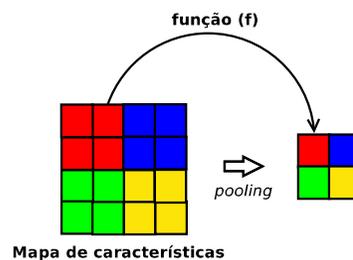


Figura 18 – *Pooling* em mapa de características.

O cálculo realizado pela camada de um *pooling* pode ser definido de diversas formas em que é possível utilizar funções de subamostragem para os elementos de uma mesma região. Geralmente, as RNCs utilizam as operações de *average pooling* e *max-pooling*. A primeira função realiza o cálculo da média de todos os valores por região, já a segunda determina apenas o maior valor. Essas funções são representadas pelas Equações 4.25 e 4.26,

$$f_{avg}(x) = \frac{1}{N} \sum_{i=1}^N x_i, \quad (4.25)$$

$$f_{max}(x) = \max(x_i), \quad (4.26)$$

em que o valor de x é um mapa com os valores de ativação de uma região local com N *pixels* (LEE; GALLAGHER; TU, 2016).

Existe ainda outras função especial de *pooling* que só são utilizadas na última camada convolucional. Essas funções são conhecidas como *global average pooling* e *global max-pooling* e têm o objetivo de reduzir o conjunto de dados para apenas uma única dimensão. As suas aplicações fazem com que sejam tratadas as propensões à má generalização nas próximas camadas das RNCs (LEE; CHEN, 2015).

4.2.4 Camada Totalmente Conectada

As próximas camadas das RNCs desempenham o papel de regressão e classificação. As suas arquiteturas são formadas por neurônios totalmente conectados (*Fully Connected* - FC). Essas camadas também estão presentes em redes neurais comuns. Elas servem para conectar as camadas anteriores sem a utilização dos pesos compartilhados. Ademais, essas camadas recebem as melhores ativações geradas no processo convolucional, produzindo uma saída de acordo com o número de características ou classes desejadas. Em qualquer rede desse tipo é necessário ao menos uma camada totalmente conectada após o processo convolucional e de subamostragem, pois ela realiza a logística de representação dos dados (BENGIO; GOODFELLOW; COURVILLE, 2015).

Apesar das camadas de convolução e de *pooling* possuírem um bom desempenho no aprendizado das RNCs, as camadas totalmente conectadas exercem um papel significativo na combinação de características e na distinção de elementos, pois a saída dessa última camada é responsável por normalizar os dados de modo que o somatório seja igual a 1. Para isso, ela aplica uma função conhecida como *softmax* que é descrita na Equação 4.27,

$$\text{softmax}(x)_j = \frac{e^{x_j}}{\sum_{i=1}^N e^{x_i}}, \quad (4.27)$$

em que e^{x_j} denota um produto interno de x que permite a equivalência de uma operação linear entre um vetor multidimensional, transformando-o para um vetor com N dimensões (PARKHI; VEDALDI; ZISSERMAN, 2015).

Por exemplo, a submissão de uma imagem em arquitetura de rede convolucional treinada pode produzir o seguinte resultado probabilístico em duas classes: $x = 0,8$ e $y = 0,2$. Nessa situação, a imagem é classificada como pertencente a classe x , pois a probabilidade de ocorrência dos seus *pixels* é maior para essa classe e a sua estrutura é semelhante as imagens que a compõe.

4.3 Treinamento de Redes Neurais Convolucionais

O processo de treinamento das RNCs é bem semelhante a forma realizada em RNA. A sua execução é baseada na estimação de parâmetros que é designada por um procedimento de otimização não linear. Esse processo tem o objetivo de minimizar uma função de custo que serve para medir a precisão de um modelo a partir da sua capacidade de distinção em classes (BENGIO; GOODFELLOW; COURVILLE, 2015).

Em redes convolucionais, a função de custo não é medida pelo erro de classificação no treinamento e, sim pela distância entre as classes, ou seja, até em que uma amostra aparenta pertencer a uma classe. Em cada modelo é possível obter essa medida de acordo com as classes impostas à rede. Nessa situação, a função de custo $J(\theta)$ considera o erro médio quadrático e entropia cruzada. A sua definição é dada pela Equação 4.28,

$$J(\theta) = \frac{1}{T} \left(\sum_{t=1}^T (J(x^{(t)}; \theta), y^{(t)}) \right), \quad (4.28)$$

em que $x^{(t)}$ é uma amostra do conjunto de dados, $(J(x^{(t)}; \theta), y^{(t)})$ os erros cometidos no treino e θ a dependência do custo J aos parâmetros θ da rede (SIMONYAN; ZISSERMAN, 2014).

As RNCs também geram restrições no compartilhamento de pesos por possuir conexões locais. Essa característica faz com que algumas conexões possuam o mesmo valor, sendo que a estrutura da rede não precisa ser alterada. No entanto, o método de treinamento utilizado para realizar a captação do conhecimento não pode alterar algumas funções, pois a arquitetura dessas redes depende tanto das conexões locais, como das totalmente conectadas (LEE; GALLAGHER; TU, 2016).

O método mais utilizado para treinar RNCs é o *backpropagation*. Entretanto, ele necessita de algumas modificações para tratar esse compartilhamento de pesos. Esse método deve realizar a atualização dos pesos de maneira igual, por intermédio de uma média de fatores que atribui mesmos valores para algumas conexões. Essas alterações fazem com que todos os pesos compartilhados sofram as mesmas variações ao longo do treinamento (BENGIO; GOODFELLOW; COURVILLE, 2015).

A adaptação do *backpropagation* pode ser definida pela inicialização dos pesos (w_i) em que é feita a seleção de um padrão (e ou s). A partir disso, esse padrão escolhido é aplicado à camada de entrada (a_i^0) em relação a uma vizinhança v . A Equação 4.29 demonstra essa operação (LECUN; BENGIO; HINTON, 2015).

$$a_i^0(v) = e_i(v) \quad (4.29)$$

A propagação do sinal por intermédio da alteração anterior faz com que a camada

l até a camada de saída (a^L) sofrem cálculos de atualizações dos pesos. Após esse procedimento, o cálculo de delta (δ^L) é realizado para cada uma das saídas por meio da comparação do resultado obtido com o padrão desejado. Logo em seguida, todos os deltas são calculados para as camadas intermediária pela retropropagação do erro. A modificação dos valores dos pesos deve ser calculada de acordo com uma taxa de aprendizado η que é definida na Equação 4.30,

$$wnovo_{i,j}^l = w_{i,j}^l + \Delta w_{i,j}^l, \quad (4.30)$$

em que $wnovo_{i,j}^l$ é o peso da próxima iteração que substituirá o peso da iteração atual (w) por meio do aprendizado de $\Delta w_{i,j}^l$ que é um produto do aprendizado acumulado em cada camada. Esse aprendizado pode ser definido pela Equação 4.31 (BENGIO; GOODFELLOW; COURVILLE, 2015).

$$\Delta w_{i,j}^l = \eta \frac{1}{n} (a_j^{l-1} * \delta_i^l) \quad (4.31)$$

4.3.1 Ativação

O poder de qualquer rede neural está na sua capacidade de se ajustar a dados não lineares. As funções de ativação são operações indispensáveis nessa tarefa, pois elas são responsáveis por aplicar às saídas de cada camada um cálculo matemático que permite uma representação dos dados. Normalmente, as funções não lineares são priorizadas nas RNCs, dado que os seus conceitos permitem aplicações sucessivas de distorções, fazendo com que os dados se tornem separáveis (CUN et al., 1990).

Nos pré-requisitos de treino de uma rede convolucional é vital a utilização de uma função de ativação, pois o processo de convolução realizado por suas camadas é linear e necessita da adição de propriedades não lineares para aumentar a capacidade de representação. Após a computação da combinação linear do neurônio $y_i(v)$ (ver Equação 4.23), essa função produz um estímulo a esse neurônio (ver Equação 4.24) que é apresentado a um conjunto de propriedades para realizar a tarefa de ativação (PRICE; ISO; SHINODA, 2016).

Alguns exemplos de funções de ativação utilizadas nas redes convolucionais são: logística ($\sigma(z)$), tangente hiperbólica ($\tanh(z)$) e *Rectified Linear Unit* ($ReLU(z)$). As funções logística e tangente hiperbólicas são *sigmóides* e exibem um comportamento linear e não linear que é dado pela razão de um fator de um ganho. Já a *Rectified Linear Unit* é uma função *rampa* que é restrita a regiões lineares. O comportamento dessas funções é apresentado na Figura 19 (HAYKIN, 2001).

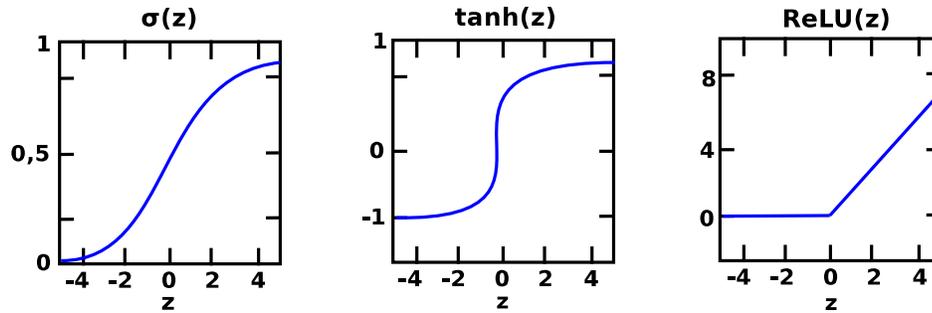


Figura 19 – Comportamento das Funções de Ativação em RNCs.

A utilização de funções *sigmóides* é incomum nas redes convolucionais, pois seus aspectos fazem com que haja um rápido decréscimo do gradiente. Já a ReLU é, por padrão, a função mais utilizada nessas redes. Ela possui propriedades positivas que são dadas por intermédio das derivadas constantes unitárias em valores de combinações entre neurônios. Essa característica faz com que a adição de camadas proporcione um cálculo do gradiente pela multiplicação dessas derivadas. Isso proporciona excelentes resultados em modelos voltados para descrição e classificação de dados (BENGIO; GOODFELLOW; COURVILLE, 2015).

4.3.2 Dropout

Um dos principais problemas encontrados na execução do treino de uma RNC é a quantidade de parâmetros livres utilizados na determinação do aprendizado. Esses parâmetros servem para criar fronteiras entre as classes no processo de decisão. Um modelo treinado com muitos parâmetros livres pode se adaptar a ruídos e detalhes desnecessários, formando um classificador com capacidade reduzida. Esse evento é conhecido como *overfitting* (sobreajuste) e deve ser evitado em qualquer situação (BENGIO; GOODFELLOW; COURVILLE, 2015).

O treinamento de redes convolucionais deve direcionar seus esforços para criação de modelos simples com regularidade de dados para atenuar os efeitos do *overfitting*. Os métodos que tratam esse problema são chamados de regularização. Eles são utilizados para reduzir a diferença entre os erros de estimação do treino e de validação (KAWAHARA; BENTAIEB; HAMARNEH, 2016).

As redes que operam convoluções geralmente utilizam a técnica de *dropout* para minimizar a relação dos valores dos erros de treino e concordância. Esse método é definido como um descarte aleatório de unidades entre as camadas convolucionais. A aplicação desse método serve para dificultar a adaptação da rede durante a etapa de treinamento. Ademais, esse método ainda dificulta a coordenação entre parâmetros, diminuindo a capacidade de inferência do modelo. Entretanto, essa característica melhora a capacidade final de

generalização. O *dropout* pode ser representado pela Figura 20 (BENGIO; GOODFELLOW; COURVILLE, 2015).

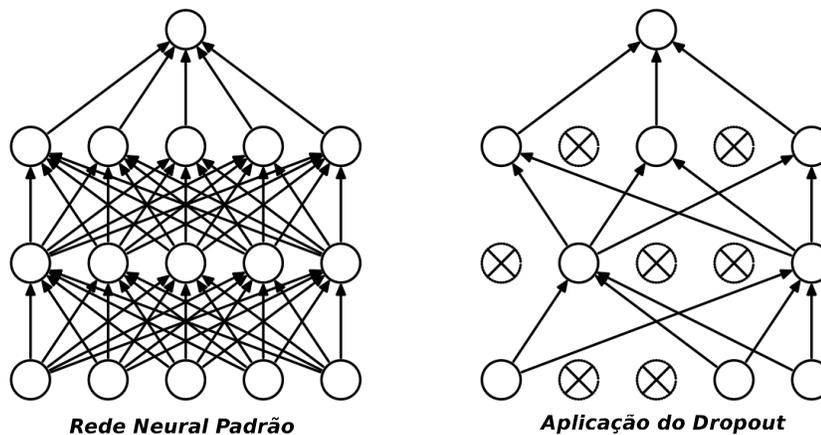


Figura 20 – Comportamento do *Dropout*

A aplicação do *dropout* causa uma queda considerável no número de classificação errada de amostras de dados pelas redes convolucionais. Isso mostra que a sua utilização é de grande importância no processo de aprendizado dessas redes. Na descrição e classificação de imagens, ela pode ser utilizada juntamente com método *Data Augmentation*, pois ambas possuem resultados semelhantes (LEE; CHEN, 2015).

4.3.2.1 *Data Augmentation*

Além da técnica de *dropout*, as redes convolucionais que tratam imagem geralmente utilizam o método de *data augmentation* (aumento de dados) no seu treinamento. Esse método proporciona um crescimento artificial de um conjunto de dados, evitando o estado de *overfitting* e produzindo modelos de predição robustos e sensíveis a detalhes. Essa técnica é amplamente utilizada em situações em que é possível perceber uma semelhança entre classes (CHATFIELD et al., 2014).

Em bases de imagens, o *data augmentation* consiste em aplicar operações básicas como rotação, translação, brilho, contraste, cor e ruído para criar novas amostras a partir das existentes. Essas operações podem modificar a estrutura de uma imagem, desenvolvendo variações de padrões que são consideradas amostra sintética que não são, necessariamente, imagens iguais às originais (PARKHI; VEDALDI; ZISSERMAN, 2015). Na Figura 21 é apresentado um exemplo de aplicação dessa técnica para imagens dermatoscópicas.

O aumento de dados também pode ser aplicado visando melhorar o aprendizado das redes convolucionais. Nesse caso, as variações de amostras produzidas podem servir para o aprendizado de características invariantes. A quantidade de transformações realizadas melhora significativamente a estrutura de aprendizado das camadas convolucionais, por

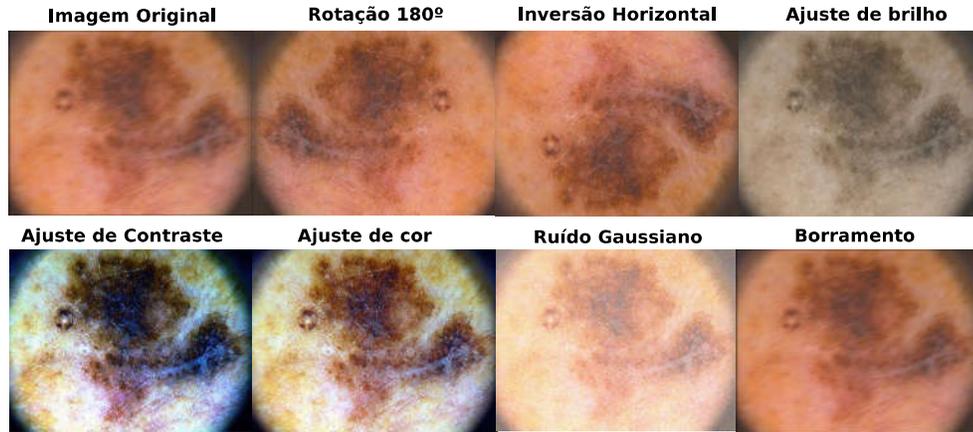


Figura 21 – Exemplos de *Data Augmentation* para imagens dermatoscópicas.

computar apenas características relevantes de geometria, bordas, cores e texturas (FAWZI et al., 2016; KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

4.3.3 *Batch Normalization*

Um problema típico do treinamento das redes de convolução é a alteração das propriedades estatísticas entre as suas camadas. A função de ativação é a principal responsável por essa operação. O treino realizado nesses moldes pode retardar os resultados de aprendizado e, em alguns casos, aumentar o tempo de convergência da rede. Para evitar esse problema, é necessário uma transformação no conjunto de treinamento. Nesse contexto, um dos métodos mais utilizados é o *Batch Normalization* (BN) (normalização em lote). Esse método é baseado na distribuição normal padrão que é uma probabilidade absolutamente contínua e parametrizada, pois considera o desvio padrão para evitar problemas de comparação causados pelas diferentes escalas dos dados.

O método de *Batch Normalization* propõe a adição de uma camada entre as camadas convolucionais. A sua aplicação é desenvolvida por meio de um mini lote (*small lot*) de dados x em que é subtraído um valor médio μ_i de cada atributo x_{k_i} das n amostras de x_k pelos desvios padrões σ_i . A aplicação dessa operação multiplica o atributo x'_{k_i} por um parâmetro γ e, por fim, adiciona um parâmetro β . Essas operações são descritas nas Equações 4.32 e 4.33,

$$x'_{k_i} = \frac{x_{k_i} - \mu_i}{\sigma_i} \quad (4.32)$$

$$x''_{k_i} = \gamma x'_{k_i} + \beta, \quad (4.33)$$

em que x''_{k_i} é considerado o valor de saída do atributo i de uma amostra k para um mini lote.

Em alguns casos de treinamento, o *batch normalization* elimina a necessidade de aplicar o *dropout* nas redes convolucionais, proporcionando uma aceleração significativa no tempo de treinamento e na quantidade de passos executados (LEE; GALLAGHER; TU, 2016).

4.4 Considerações Finais

Este capítulo apresentou os principais conceitos sobre a estrutura convencional de detecção do melanoma, além de dos princípios do aprendizado profundo voltados para as Redes Neurais Convolucionais. Esses materiais e métodos foram utilizados para o desenvolvimento de duas formas distintas de reconhecimento do melanoma para imagens dermatoscópicas. O próximo capítulo apresenta todas as estruturas desenvolvidas na proposta deste trabalho, além das respectivas ações realizadas para alcançar o diagnóstico de lesões de pele.

5 Classificação com Aprendizado Profundo

As RNCs foram desenvolvidas especificamente para o reconhecimento de imagens e sons. Essas redes têm mostrado um grande poder na captação de atributos relevantes, proporcionando uma melhor identificação de elementos em classes. A partir dessa característica, este capítulo apresenta duas formas de classificação do melanoma utilizando os princípios de aprendizado profundo com essas redes. A primeira forma é baseada em estruturas convencionais de diagnóstico. Já a segunda forma é embasada em novas maneiras de classificação que são desenvolvidas por meio de processos convolucionais.

5.1 Reconhecimento do Melanoma pela Classificação com RNCs

Como dito anteriormente, as abordagens tradicionais de detecção do melanoma em imagens dermatoscópicas são baseadas em quatro etapas básicas: pré-processamento, segmentação, extração de características e classificação. Essa arquitetura é uma das mais desenvolvidas em uma boa parte das pesquisas realizadas na literatura ([MAGLOGIANNIS; DOUKAS, 2009](#)). Nesse contexto, surge a necessidade de agregar o aprendizado profundo com RNCs nesse modelo de diagnóstico de lesões de pele.

A proposta de reconhecimento do melanoma pela classificação com RNCs sugerida neste trabalho desenvolve a estrutura convencional de detecção do melanoma, entretanto, implementa os processos convolucionais na etapa de classificação. O desenvolvimento dessa proposta visa tratar as imagens dermatoscópicas, corrigindo as suas imperfeições e, em seguida, separando as lesões do restante da imagem. Essas imagens segmentadas são utilizadas no processo de extração de características que determina os dados relevantes para um diagnóstico. O diferencial dessa proposta em relação às demais é a utilização de um processo convolucional na classificação das características de assimetria, bordas irregulares, variações de cor e textura. A Figura 22 apresenta um esboço de todas as etapas desenvolvidas na referida sugestão.

Os métodos definidos nesta proposta de arquitetura tradicional têm o intuito de proporcionar melhores resultados na identificação do melanoma. Esses métodos foram selecionados de acordo com um levantamento bibliográfico que serviu para direcionar a modelagem, por intermédio do diagnóstico eficaz dessa doença em outros trabalhos. Ademais, esse estudo ainda proporcionou algumas técnicas que foram úteis no desenvolvimento das operações que precedem o diagnóstico das lesões. As próximas seções descrevem todas as etapas desenvolvidas por essa estrutura tradicional com as RNCs, além dos seus respectivos objetivos e operações.

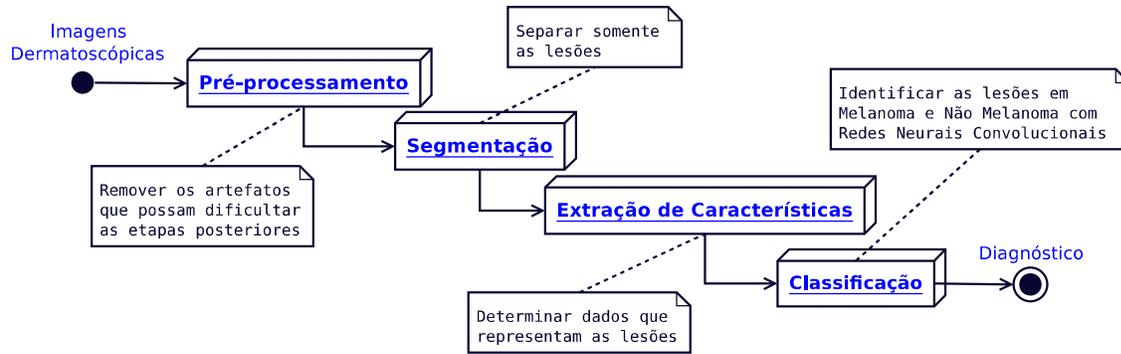


Figura 22 – Proposta de reconhecimento do melanoma com a classificação por RNCs.

5.1.1 Pré-processamento

O excesso de pelos encontrados nas imagens dermatoscópicas é um problema que pode afetar significativamente o diagnóstico das lesões. A solução proposta nesta etapa para corrigir ou aliviar seus efeitos segue os princípios desenvolvidos por Lee et al. (1997) que foram descritos na Seção 4.1.1. As operações desenvolvidas nesse método são amplamente utilizadas para remoção dos pelos em imagens dermatoscópicas (MAGLOGIANNIS; DOUKAS, 2009).

A aplicação das operações sugeridas nesse método identifica a estrutura de *pixels* que compõem os pelos. Em seguida, faz algumas alterações para que eles sejam amenizados, sem produzir perdas significativas na estrutura e na qualidade das imagens. Para tal, são aplicadas operações de fechamento morfológico em cada canal de cor com as direções 0° , 45° e 90° . A união desses componentes produz uma máscara otimizada dos pelos. A Figura 23a apresenta uma imagem dermatoscópica com excesso de pelos e a Figura 23b uma máscara obtida com o referido processo de identificação.

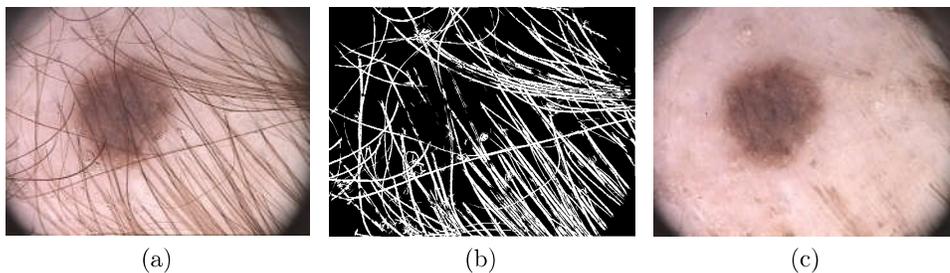


Figura 23 – Remoção de pelos pelo método de Lee et al. (1997). (a) Imagem de entrada. (b) Máscara dos pêlos (c) Imagem corrigida. Adaptado de (MENDONÇA et al., 2013)

Após o processo de identificação dos pelos, a máscara passa por uma operação de limiar que define a localização dos *pixels* que se encontram em uma estrutura curvilínea. Essa fase é utilizada para guiar a remoção dos *pixels* que represam essa estrutura. Já

a interpolação bilinear se encarrega de reconstruir os locais afetados a partir da sua vizinhança. Por fim, uma leve suavização com um elemento estruturante de 5×5 melhora o resultados da correção. A Figura 23c representa o resultado final dessas operações.

5.1.2 Segmentação

O processo de segmentação das imagens dermatoscópicas proposta nesta forma é baseada nos conceitos apresentados na Seção 4.1.2. A sugestão do método *Fuzzy K-means* determina o agrupamento de *pixels* que melhor representa as lesões. Esse agrupamento é alcançado por meio dos princípios de similaridade entre regiões que serve para definir a melhor estrutura de *pixels* em classes.

O desenvolvimento da sugestão de segmentação visa combinar diferentes valores de K em busca do melhor contorno das lesões. Esse valor é alcançado por intermédio do agrupamento de *pixels* que possuem as maiores taxas de similaridades entre as suas respectivas regiões (lesão e fundo). Nessa situação, os valores $K = \{1, 2, \dots, 16\}$ são variados onde os seus respectivos agrupamentos são armazenados e, em seguida, comparados. A melhor estrutura sempre é selecionada para representar as respectivas regiões das lesões. A Figura 24a apresenta uma imagem dermatoscópica pré-processada e a Figura 24b o resultado da aplicação do *Fuzzy K-means* em uma imagem dermatoscópica.

Em alguns casos, as imagens dermatoscópicas podem conter bordas laterais que são facilmente removidas nessa fase por meio da transformada de *Hough*. Essa operação identifica as áreas circulares que ficam próximas das lesões (BALLARD, 1981), possibilitando a sua remoção. Após esse processo, os ruídos encontrados são removidos por intermédio das menores áreas conectadas. Finalmente, as imagens de saída são binarizadas, gerando uma máscara da lesão. Essa máscara é utilizada para separar o contorno da região de interesse do restante da imagem. O resultado dessas operações são apresentados nas Figuras 24c e 24d.

5.1.3 Extração de Características

As operações de extração de características realizadas nesta etapa sugerem o uso dos descritores apresentados na Seção 4.1.3. Esses métodos determinam valores numéricos que representam as variações de assimetria, bordas, cores e texturas nas lesões segmentadas, seguindo os princípios apresentados pela regra ABCD e produzindo os dados necessários para o diagnóstico das lesões.

O descritor de geometria capta as características de irregularidade de formas. Esses dados são definidos por intermédio das operações de área, perímetro, momentos invariantes e geométricos de Hu (1962), envoltório convexo de Berg et al. (2000), delimitação de

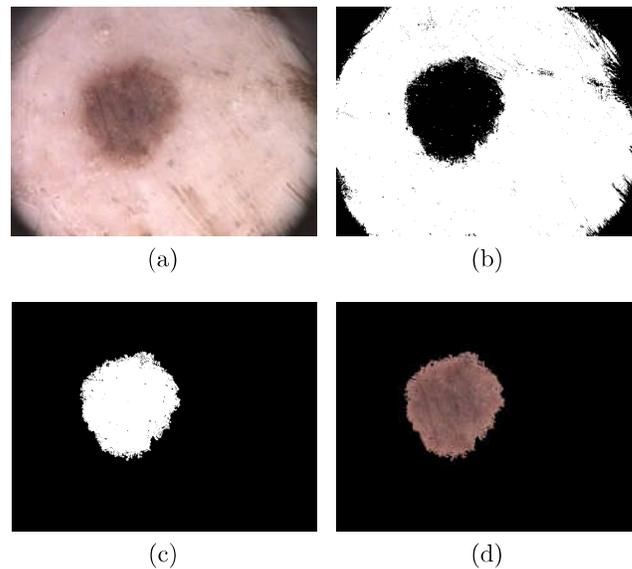


Figura 24 – Segmentação da lesão com método *Fuzzy K-means*. (a) Imagem de entrada (b) Segmentação. (c) Máscara. (d) Imagem de saída. Adaptado de (MENDONÇA et al., 2013).

regiões, centroide e orientações. Esse descritor é aplicado sobre a máscara binária das lesões, produzindo 47 características distintas que representam a assimetria.

As descrições realizadas por HOG representam as irregularidades de bordas das lesões. As operações realizadas por esse método determinam 9 *bins* de gradientes horizontal e vertical por meio de uma máscara 2 x 2, que é aplicada sobre a imagem entre as orientações 0° e 180°. As grades de *pixels* resultantes são divididas em células e cada uma delas gera um histograma. Esses histogramas são normalizados e concatenados, produzindo 27 características referentes às variações de estrutura de contorno das lesões. Essas operações são realizadas em cada canal de cor das imagens, produzindo um vetor final com 81 (27 por canal) características de bordas.

Já as descrições realizadas pelo método ECC produz dados que determinam irregularidades de cores nas lesões. Esse descritor extrai características baseadas em variações estatísticas por agrupamento de *pixels*. O diferencial desse método é a sua possibilidade de tratar os diferentes canais de cores de uma imagem por meio das operações de cor mínima, média, máxima, mediana e moda. Ademais, esse método ainda desenvolve as operações de variância, covariância, entropia, desvio padrão, *kurtosis* e correlação, em valores mínimos, médios e máximos. A aplicação desse descritor é sequencial por canal e, ao fim da sua execução, é produzido um vetor com 54 características de cor das lesões. Nessa situação, quanto maior a variação de valores entre os canais, maior é a possibilidade de identificação do melanoma.

Os dados extraídos pelo descritor de GBF são determinados pela Transformada

Discreta de *Fourier*. Esses valores levam em consideração a função de *Gabor* em duas dimensões, com filtros de passagem de banda de *Gauss*. Esse padrão de extração de dados que representam variações de textura foi desenvolvido por Kumar e Pang (2002), onde as operações realizadas aplicam 16 variações de frequências em cada canal de uma imagem, com 67 recursos de *Fourier*. Esse processo produz um total de 201 (67 por canal) características de textura para cada imagem.

As descrições de LBP determinam uma sequência de valores que descreve as variações de níveis de cinza das lesões. Esses valores também servem para representar dados de textura em imagens. Nessa situação, os raios $R = (1, 2, \dots, 5)$ são aplicados entre as vizinhanças de 4 a 8 *pixels*, resultando em uma estrutura com dados sobre as variações de brilho que são adquiridas por intermédio da contagem de transições entre 0 e 1 ao longo da imagem. O vetor resultante dessa operação produz 59 características, sendo possível a sua aplicação em cada canal de cor da imagem. O processo final desse descritor produz um total de 177 (59 por canal) dados de textura para cada imagem.

Finalmente, o descritor de *Haralick* extrai característica a partir das operações de segundo momento angular, entropia, contraste, variância, correlação e homogeneidade em matrizes de co-ocorrência. Esses dados são calculados para as variações mínimas, médias e máximas. O resultado parcial dessas operações é um vetor com 46 características que também representam dados sobre variações de textura das lesões. Ao fim da execução desse descritor, são produzidas 138 (46 por canal) características para uma imagem com três canais de cores.

5.1.4 Classificação

A etapa de classificação nesta etapa prevê a utilização de uma arquitetura de RNC para realizar as predições das características extraídas das lesões. O processo de diagnóstico do melanoma com esse método segue os princípios que foram descritos pelos conteúdos do Capítulo ??, selecionando dados na estrutura de entrada, por meio das operações de convolução e subamostragem. Por fim, as camadas totalmente conectadas definem uma relação que diferencia os dados das lesões melanômicas no diagnóstico.

O reconhecimento do melanoma pela classificação com RNCs também sugere a identificação em classes dos dados com as formas tradicionais, que foram descritas na Seção 4.1.4. Esses métodos são comumente empregados em muitos sistemas de detecção do melanoma. A proposta de classificação com esses métodos tem o objetivo de avaliar a eficácia dos processos convolucionais no diagnóstico de lesões de pele. Ademais, a análise desses resultados ainda serve para determinar um comparativo com outras pesquisas relacionadas.

5.2 Reconhecimento do Melanoma pela Descrição com RNCs

A forma de reconhecimento pela descrição com RNCs proposta neste trabalho tem uma estrutura diferente das abordagens tradicionais que tratam o problema de reconhecimento do melanoma. A referida proposta é baseada em novos métodos de identificação de elementos em imagens criados por Parkhi, Vedaldi e Zisserman (2015). Nesse arquétipo não é necessário desenvolver as etapas de pré-processamento e segmentação, pois as imagens originais são utilizadas diretamente para treinar as arquiteturas convolucionais e, em seguida, elas são utilizadas como descritores de características.

O processo de desenvolvimento dessa proposta que utiliza modelos treinados RNCs é baseada em quatro etapas. A primeira etapa utiliza as operações de aumento de dados nas bases de imagens dermatoscópicas. Já na segunda etapa, as arquiteturas convolucionais são treinadas com o ajuste dessas bases. Em seguida, a terceira etapa utiliza os modelos treinados como descritores de características. Por fim, os dados extraídos com as arquiteturas convolucionais são classificados para determinar o diagnóstico das lesões. A Figura 25 apresenta um esboço da sugestão dessa segunda forma de detecção do melanoma.

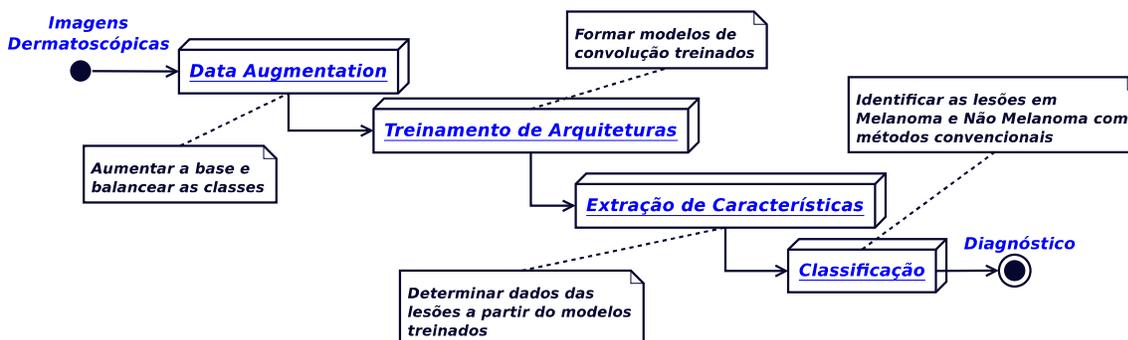


Figura 25 – Proposta de reconhecimento do melanoma pela descrição com RNCs.

O aprendizado profundo desenvolvido nesta referida proposta sofre variações a partir de outras formas tradicionais de classificação de lesões de pele, pois os princípios da regra ABCD não são observados e o processo de aprendizado fica a critério das arquiteturas convolucionais. Essas estruturas são utilizadas para captar características relevantes por intermédio da estrutura espacial das lesões. Nesse contexto, é sugerido o uso das arquiteturas convolucionais AlexNet e VGG-F que são bem conhecidas na literatura. Os dados extraídos das últimas camadas dessas redes armazenam as melhores características que descrevem as lesões de pele. Esses valores são utilizados na etapa de classificação para proporcionar um diagnóstico eficaz. As próximas seções apresentam a sugestão dessas arquiteturas e todo o desenvolvimento da segunda proposta recomendada neste trabalho.

5.2.1 Arquitetura Convolucional AlexNet

A AlexNet é uma das primeiras arquiteturas de rede convolucional bem sucedida na *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) (RUSSAKOVSKY et al., 2015). A ILSVRC é uma competição que promove resultados de modelos de rede com aprendizado profundo em uma diversidade de problemas que envolvem classificação de imagens. O objetivo dessa competição é enaltecer arquiteturas que obtêm grandes progressos em seus resultados.

O modelo convolucional AlexNet foi desenvolvido por Krizhevsky, Sutskever e Hinton (2012) para reconhecimento de objetos diversos. A sua arquitetura é composta por oito camadas, onde as cinco primeiras representam as operações de convolução, possuindo estruturas localmente conectadas. Já as demais camadas são totalmente conectadas. A última camada dessa arquitetura é a saída da rede, sendo alimentada por uma operação de *softmax* com 1000 saídas, cada uma com rótulo de classe. A Figura 26 apresenta uma visão dessa arquitetura convolucional.

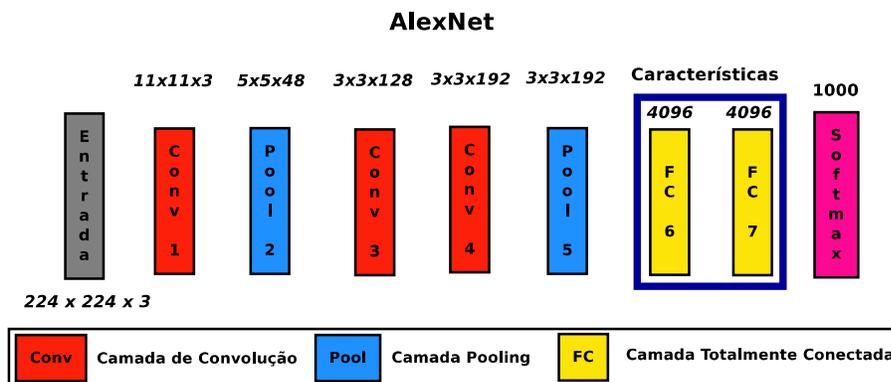


Figura 26 – Arquitetura da RNC AlexNet.

A primeira camada de convolução da AlexNet filtra a imagem de entrada com tamanho $224 \times 224 \times 3$, utilizando uma máscara de tamanho 11×11 em 3 filtros por camada ($11 \times 11 \times 3$), com *stride*¹ de 4 *pixels*. Já a segunda camada recebe as matrizes de convolução (mapas de características) e aplica máscaras com a função de *max-pooling* de $5 \times 5 \times 48$, com *stride* de 1. Na terceira e na quarta camada são realizadas outras operações de convolução, com máscaras de $3 \times 3 \times 128$ e $3 \times 3 \times 192$ e um *stride* 1. Ao fim dessas convoluções, as matrizes de saída da camada anterior recebem a última operação de *global max-pooling* com uma máscara $3 \times 3 \times 192$ e *stride* 1, que reduz o vetor de características a uma única dimensionalidade.

O produto do processo convolucional é atribuído as camadas totalmente conectadas seguintes. Elas rechem a matriz dos pesos e, em seguida, associam aos 4096 neurônios. Após

¹ O *stride* representa a quantidade de *pixels* que deverão ser saltados a cada iteração da máscara sobre a imagem.

esse processo, a operação de *softmax* é aplicada na última camada rede e o seu resultado é enviado para os 1000 neurônios de saída. Esse processo é responsável por normalizar o resultado por meio da ocorrência de probabilidades entre as classes, onde o somatório de todas as saídas sempre será igual a 1 e o maior valor determina a classificação da entrada.

5.2.2 Arquitetura Convolucional VGG - F

Após o sucesso da AlexNet, novas redes convolucionais surgiram a fim de superar seu resultado na ILSVRC. Algumas delas foram desenvolvidas pelo *Visual Geometry Group* (VGG). Uma das suas principais versões de rede convolucional é a VGG - *Fast* (VGG - F). Essa arquitetura é bem semelhante a AlexNet, entretanto, aplica mais filtros nas camadas convolucionais. A Figura 27 apresenta a estrutura dessa rede.

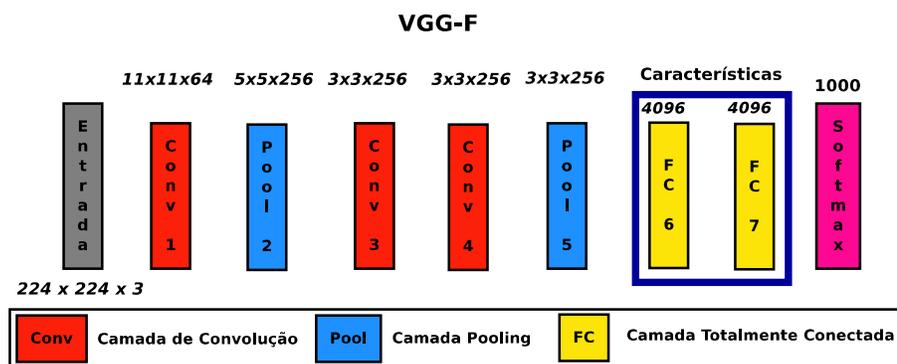


Figura 27 – Arquitetura da RNC VGG-F.

A arquitetura VGG-F reduz a imagem de entrada para um tamanho de 224 x 224 x 3 onde a primeira camada aplica um filtro 11 x 11 x 64, com *stride* de 4 *pixels*. A segunda camada recebe os mapas de características e aplica outro filtro de 5 x 5 x 256 e *stride* de 1 pixel, com a função de *max-pooling*. Já a terceira e a quarta camada aplicam um mesmo filtro de 3 x 3 x 256 e *stride* de 1 *pixel*. Finalmente, a última camada convolucional aplica a última operação de *global max-pooling*, com um filtro de 3 x 3 x 256 e *stride* de 1 *pixel*.

O resultado da convolução é repassado para as próximas camadas totalmente conectadas da arquitetura. Essa estrutura de camadas é igual a AlexNet e produz o mesmo efeito de classificação por meio de duas camadas de 4096 neurônios e uma última camadas com a operação de *softmax* para 1000 neurônios de saída.

5.2.3 Definição do Aumento de Dados

Os critérios utilizados nesta etapa para desenvolver o aumento de quantidade no conjunto de treino levou em consideração os estudos de [Chatfield et al. \(2014\)](#) e [Parkhi, Vedaldi e Zisserman \(2015\)](#) onde são aplicadas operações que proporcionam mudanças

nesse conjunto de treino. Essas mudanças são necessárias para evitar um possível estado de *overfitting* no processo de treinamento.

As imagens dermatoscópicas mantém características próximas de *pixels*, dificultando um treinamento eficaz das redes convolucionais por conjuntos com poucas imagens. Nesse sentido, essa a estrutura de reconhecimento do melanoma pelas descrições de RNCs também utiliza o aumento de dados para possibilitar a criação de modelos mais robustos e precisos. Nesse contexto, foram selecionadas as seguintes operações:

- rotações nos ângulos de 45°, 90°, 135° e 180°;
- translações nos ângulos de 90° e 180°;
- borramentos por heurística em 4 classes;
- ruídos de densidade, estatístico, multiplicativo e impulsivo, e;
- ajustes de brilho e cor por intensidade.

No fim da execução das operações do método, o aumento das bases de imagens produz uma multiplicação do conjunto de treino em 16 vezes, ou seja, para cada imagem são produzidos 16 exemplos sintéticos que pertencem a mesma classe. Entretanto, as suas estruturas são diferenciadas pelas mudanças de aspecto proporcionadas pelas operações anteriores.

5.2.4 Extração de Características

Na sugestão dessa etapa de extração de características, o processo de identificação dos dados utilizando os modelos de RNCs sugeridos é uma variação da aplicação do aprendizado profundo que já foi desenvolvida em outros trabalhos relacionados. Essas redes, quando treinadas, partilham pesos entre seus neurônios nas camadas convolucionais. Esses valores servem para produzir ativações de acordo com os picos de características que são identificados em estruturas espaciais nas imagens, ou seja, quando uma classe é identificada pelas camadas convolucionais, alguns neurônios em camadas subsequentes produzem mesmas respostas, ativando em circunstâncias que identificam estruturas de *pixels* semelhantes.

O resultado do processo convolucional com as redes AlexNet e VGG - F produz para cada imagem um total de 4096 dados que corresponde as características das lesões de pele. Esses valores representam informações espaciais invariantes que foram subtraídas de acordo com a estrutura dos *pixels* que melhor descrevem as classes. A utilização desses dados permite obter uma probabilidade de frequência que serve para definir um limiar de diagnóstico. Entretanto, o reconhecimento das lesões não utiliza essa última camada com a operação *softmax*, pois o objetivo é apenas descrever as imagens.

Na referida forma de reconhecimento, os atributos classificáveis são os dados das camadas FC_6 e FC_7 , descritas anteriormente nas Figuras 26 e 27 das respectivas redes convolucionais sugeridas. Essa forma de extração de dados já foi trabalhada nos experimentos de Parkhi, Vedaldi e Zisserman (2015) onde é possível perceber excelentes resultados. As últimas camadas das redes convolucionais treinadas possuem a capacidade de responder a aspectos espaciais semelhantes, ou seja, produzir ativações em circunstâncias onde uma possível estrutura visual é identificada. Seguindo esse princípio, as imagens que possuem *pixels* que correspondem a uma probabilidade de frequência para uma classe específica são descritas por meio dos dados armazenados por essas camadas.

5.2.5 Classificação

A fim de avaliar o desempenho das descrições de lesões com modelos de redes convolucionais, as imagens dermatoscópicas são submetidas às arquiteturas AlexNet e VGG - F já treinadas, onde cada imagem que é submetida obtém 4096 características espaciais por intermédio das camadas FC_6 e FC_7 que, em seguida, são armazenadas em vetores.

Os métodos utilizados nesta etapa para determinar o diagnóstico das lesões de pele por meio das arquiteturas sugeridas foram selecionados por meio dos princípios apresentados na forma convencional de detecção, com os classificadores KNN, SVM e MLP. Nesse contexto, eles podem proporcionar melhores resultados de diagnóstico com a imposição dessas configurações.

5.3 Considerações Finais

Neste capítulo foram apresentadas duas formas de reconhecer o melanoma utilizando o aprendizado profundo por intermédio das RNCs na descrição e classificação de imagens dermatoscópicas. A primeira desenvolve uma arquitetura já conhecida na bibliografia, entretanto, implementa uma nova forma de classificação com processos convolucionais. Já a segunda forma desenvolve o aprendizado profundo na extração de características, utilizando arquiteturas convolucionais treinadas com imagens específicas que, em seguida, são empregadas como descritores de características. O próximo capítulo apresenta os resultados obtidos com essas configurações no diagnóstico de lesões de pele.

6 Experimentos e Discussões

Após o desenvolvimento das abordagens sugeridas no capítulo anterior, alguns testes foram ajustados para avaliação de desempenhos na classificação de lesões em imagens dermatoscópicas. Para esse cenário, o presente capítulo apresenta todos os aparatos e configurações utilizados, além de todos os resultados alcançados no diagnóstico do melanoma. Esse conteúdo é apresentado por meio de uma análise quantitativa entre as arquiteturas de detecção, com alguns aspectos observados ao longo do desenvolvimento.

6.1 Bases de Imagens

Inicialmente, as abordagens com aprendizado profundo necessitam de um conjunto de dados para realizar os procedimentos experimentais. Para tal, é necessário definir as bases de imagens que precisam estar previamente diagnosticadas por um dermatologista. Nesse contexto, os experimentos realizados nessa pesquisa foram desenvolvidos com o uso das bases PH² (MENDONÇA et al., 2013) e *International Skin Imaging Collaboration* (ISIC) (GUTMAN et al., 2016).

A base PH² pertence ao Hospital Pedro Hispano de Matosinhos, Portugal. Ela é composta por 200 imagens dermatoscópicas, em que 80 são consideradas lesões de comum, 80 são definidas como lesões anormais e 40 diagnosticadas como melanoma. A resolução de todas as imagens é de 768×560 pixels no sistema de cor RGB.

Já a ISIC é uma base projetada em parceria com academias e juntas médicas que facilitam o desenvolvimento de ferramentas que constroem o diagnóstico do melanoma. Ela é composta por 900 imagens dermatoscópicas diagnosticadas, das quais, 887 são lesões não melanômicas e 172 lesões melanômicas. As imagens dessa base possuem uma resolução de 1022×767 pixels também no sistema de cor RGB.

6.2 Avaliação de Resultados

As formas de detecção do melanoma precisam ser testadas e avaliadas para certificar a sua eficácia na classificação de lesões de pele. Para isso, uma análise é necessária para definir quais são os melhores classificadores utilizados. Esses dados possibilitam delimitar modelos capazes de generalizar o conhecimento, produzindo informações que auxiliam na definição de testes mais precisos (ARGENZIANO et al., 2003).

A execução de cada estrutura de reconhecimento por meio de um método de

classificação produz uma matriz de confusão. Essa matriz serve para exibir dados numéricos sobre as instâncias que apresenta o diagnóstico do melanoma como correto ou incorreto. Uma instância de dados faz referência a uma imagem dermatoscópica que pode ser classificada em quatro formas: Verdadeiro Positivo (VP) e Verdadeiro Negativo (VN) para presença do melanoma e Falso Positivo (FP) e Falso Negativo (FN) para sua ausência. As métricas que utilizam esses dados determinam a eficácia do diagnóstico (LANDIS; KOCH, 1977).

As principais métricas utilizadas na literatura para avaliar o diagnóstico em imagens médicas são: acurácia (A), precisão (P), *recall* (R), *f-measure* (FM), índice *kappa* (K) e a Área sob a curva ROC (*Receiver Operating Characteristic*) (*Area Under Curve* - AUC). Essas medidas possibilitam várias análises quantitativas que avaliam as configurações de cada experimento de detecção do melanoma (KOROTKOV; GARCIA, 2012).

A medida de acurácia apresenta a porcentagem de amostras classificadas corretamente no conjunto de teste por um classificador. Essa operação é realizada por intermédio do número de predições corretas (VP e VN) dividido pelo número de predições feitas (n) que é apresentada pela Equação 6.1. A utilização da acurácia requer cautela em qualquer análise de classificação, pois ela pode induzir interpretações erradas de uma estrutura de detecção, caso as classes de dados impostas estejam desbalanceadas (FLEISS; LEVIN; PAIK, 2013).

$$A = \frac{VP + VN}{n} \quad (6.1)$$

O cálculo da medida de precisão indica as amostras classificadas corretamente como VP sobre o total de amostras por um classificador. Nessa situação, essa medida determina o número de amostras que são positivas sobre a soma de todas as amostras que foram determinadas como positivas. Já o cálculo do *recall* indica a proporção de amostras classificadas corretamente como positivas (VP) sobre todas as amostras que são realmente positivas. Essas duas operações são definidas pelas Equações 6.2 e 6.3 (LANDIS; KOCH, 1977).

$$P = \frac{VP}{VP + FP} \quad (6.2)$$

$$R = \frac{VP}{VP + FN} \quad (6.3)$$

O valor de *f-measure* determina a capacidade de predição de uma estrutura de reconhecimento por intermédio de um classificador. Essa operação é considerada uma média entre os cálculos de precisão e *recall* em que são refletidas as porções de VP em relação às

predições positivas (corretas e incorretas), somadas às amostras de classe efetivamente positivas. O *f-measure* é representado pela Equação 6.4 (LANDIS; KOCH, 1977).

$$FM = 2 * \frac{P * R}{P + R} \quad (6.4)$$

Já o índice *kappa* mostra a concordância entre os testes dos classificadores em cada forma de detecção, de acordo com os resultados obtidos. Essa medida fornece uma ideia de quanto um classificador se afasta ou aproxima do diagnóstico esperado, indicando a sua legitimidade. A estatística do índice *kappa* é determinada pela taxa de aceitação relativa (p_0) e pela taxa hipotética de aceitação (p_e), que é definida pela Equação 6.5 (FLEISS; LEVIN; PAIK, 2013).

$$K = \frac{p_o - p_e}{1 - p_e} \quad (6.5)$$

O índice *kappa* pode ser interpretado de acordo a Tabela 3 em que os valores menores que 0,2 são considerados pobres e os maiores que 0,8 são considerados excelentes (FLEISS; LEVIN; PAIK, 2013). Nesse contexto, quanto mais próximo de 1 é o valor de *kappa* melhor é a concordância do classificador com respectivos resultados (LANDIS; KOCH, 1977).

Tabela 3 – Nível de exatidão de classificação de dados com o índice kappa.

Índice Kappa	Qualidade
$K \leq 0,2$	<i>Ruim</i>
$0,2 < K \leq 0,4$	Razoável
$0,4 < K \leq 0,6$	<i>Bom</i>
$0,6 < K \leq 0,8$	<i>Muito Bom</i>
$K \geq 0,8$	<i>Excelente</i>

Por fim, a última medida utilizada para avaliar um classificador é a área sob a curva ROC (AUC). Essa operação exhibe o conflito entre a taxa de VP e FP, além do respectivo desempenho. Essa medida é representada por gráficos, em que é possível analisar a probabilidade de detecção dos métodos classificadores utilizados nesse trabalho. Quanto maior a área sob a curva ROC no gráfico, mais preciso é o classificador. Essa área é delimitada entre 0 e 1, em que a interpretação desses valores são definidos na Tabela 4 (FLEISS; LEVIN; PAIK, 2013).

6.2.1 Definição dos Conjuntos de Treino, Validação e Teste

Além das medidas de desempenho dos métodos classificadores, a avaliação das formas de reconhecimento do melanoma necessita de um método para organizar as amostras

Tabela 4 – Nível de precisão por meio da AUC.

AUC	Precisão
Entre 0,5 e 0,6	<i>Não aceitável</i>
Entre 0,6 e 0,7	<i>Ruim</i>
Entre 0,7 e 0,8	<i>Regular</i>
Acima de 0,9	<i>Excelente</i>

de treino, validação e teste. Muitas vezes, a má distribuição desses conjuntos podem ocasionar falhas gravíssimas na modelagem da estrutura, além de invalidar os seus resultados. Para isso, é necessário definir um método que faça essa operação (KOROTKOV; GARCIA, 2012).

Os principais métodos utilizados para determinar a divisão do conjunto de amostras em abordagens de detecção do melanoma são baseadas no modelo de validação cruzada (*cross validation*). Esse modelo possibilita técnicas para divisão de um conjunto de amostras que será utilizada em modelos de treinamento e predição. As técnicas mais utilizadas de validação cruzada são *k-folds* (k subconjuntos) e *holdout* (FLEISS; LEVIN; PAIK, 2013).

A técnica *k-fold* consiste em dividir um mesmo conjunto de amostras em k subconjuntos, em que $k - 1$ subconjuntos serão utilizados para treinamento do classificador e o último conjunto que restou é utilizado para teste. Essa operação é repetida no método até que todos os subconjuntos k sejam utilizados como conjunto de teste (CELEBI et al., 2007).

Na avaliação de um conjunto de amostras por um classificador, o número de *k-folds* geralmente é definido como sendo 10. Esse valor proporciona mais confiabilidade com um maior número de estimativas, sendo que essa validação é repetida $k = 10$ vezes (KOROTKOV; GARCIA, 2012).

Já a técnica de *holdout* consiste em dividir um conjunto de amostras em dois subconjuntos (treinamento e teste), em que essa divisão pode ser ou não proporcional. Geralmente a literatura define 80% das amostras para o treino e 20% para testes. O conjunto de validação é definido como 10% do conjunto de treino. Esses valores definem um ajuste que pode avaliar o desempenho do classificador, já que maior parte das amostras é destinada para o conjunto de treino, possibilitando modelos de predição mais complexos (FLEISS; LEVIN; PAIK, 2013).

6.3 Configuração dos Experimentos

As propostas de abordagens sugeridas nesse trabalho foram desenvolvidas na ferramenta MATLAB (VEDALDI; LENC, 2015). Os testes realizados utilizaram um computador com processador Intel Core i5 5200-U CPU @ 2.2GHz com 8GB de memória

RAM e uma placa de vídeo NVIDIA GeForce GT 710M 2GB 384 Núcleos CUDA 1,8 GHz.

Em ambas as abordagens sugeridas, o processo de diagnóstico das lesões é tratado em duas classes (melanoma e não melanoma). As operações realizadas nessas formas consideram o balanceamento entre classes que determina um conjunto de imagens selecionadas aleatoriamente das bases imagens. Nesse sentido, foram utilizadas apenas 426 imagens dermatoscópicas (213 melanoma e 213 não melanoma).

A forma de reconhecimento do melanoma pela classificação com RNCs realiza todo o processo de pré-processamento e segmentação de maneira convencional com os métodos sugeridos. Entretanto, na etapa de extração de características se definiu o espaço de cor HSV para todos os descritores na identificação de dados. Esse espaço foi selecionado com base nos estudos apresentados por [Maglogiannis e Doukas \(2009\)](#) e [Celebi et al. \(2007\)](#) em que os melhores resultados são atingidos por meio desse espaço de cor. Já na etapa de classificação são utilizados os métodos KNN, SVM e MLP. Os resultados são comparados com a predição de um modelo de rede convolucional, selecionado de acordo com a quantidade de dados extraído das imagens. Essa arquitetura foi definida com apenas um processo convolucional, que contém uma camada de convolução $5 \times 5 \times 100$, uma camada de *pooling* com função *max-pooling* de $2 \times 2 \times 100$, uma camada totalmente conectada com 10 neurônios e, por fim, 2 saídas com a função *softmax*. Todos os classificadores utilizados nesta forma estão implementados na ferramenta WEKA ([HALL et al., 2009](#)) e o modo de execução segue o *k-folds cross validation* ($k = 10$).

O desenvolvimento do melanoma pela descrição com RNCs utiliza a biblioteca Mat-ConvNet ([VEDALDI; LENC, 2015](#)). As arquiteturas AlexNet e VGG-F são disponibilizadas, por padrão, nessa ferramenta. O treinamento desses modelos com imagens dermatoscópicas seguem duas situações diferentes. A primeira utiliza *Batch Normalization* (BN) para organizar os conjuntos de treino. A GPU é utilizada nessa fase apenas para calcular a média total de imagens, a média total de cores por canal e covariância. Essas operações servem para ajustar o conjunto de treino à arquitetura das redes. Já a segunda forma de treinamento utiliza o *Dropout*. Essa operação já é definida por padrão na ferramenta utilizada. Os resultados serão apresentados em duas formas: com o *Batch Normalization* (sim) e sem o *Batch Normalization* (não) (esse último com o *Dropout*).

A operação de aumento de dados nesta segunda forma de reconhecimento considera os princípios apresentados na Seção 5.2.3. Nesse contexto, o conjunto de treino obteve um total de 5.780 imagens (2.890 de melanoma e 2.890 de não melanoma) para treino. No entanto, as etapas de extração de características e de classificação levaram em consideração apenas o total de 426 imagens do conjunto para realizarem seus objetivos. Os modelos convolucionais treinados foram utilizados em partes, ou seja, apenas as camadas totalmente conectadas (FC₆ e FC₇). Os dados extraídos dessas camadas serviram para a discriminação das lesões. A classificação desses dados também utilizou os métodos KNN, SVM e MLP.

Os conjuntos utilizados nos treinos foram definidos por meio da técnica de *holdout*. Já a classificação utilizou o *k-folds cross validation* ($k = 10$).

6.4 Resultados do Reconhecimento pela Classificação com RNCs

A partir das configurações definidas para o reconhecimento do melanoma pela classificação com RNCs, a Tabela 5 apresenta os resultados atingidos com o classificador SVM. Nela se observa que os dados extraídos pelo descritor de *Haralick* proporcionaram uma acurácia de 92,4% na identificação correta das lesões, com índice kappa satisfatório de 0,849 e AUC de 0,925. Já o descritor de GBF obteve dados que atingiram uma acurácia de 90,6%, com um índice kappa também satisfatório de 0,812 e AUC de 0,906. Já os dados do descritor de Geometria proporcionaram o pior resultado neste teste, com uma acurácia de 81,2%, índice kappa de 0,624 e AUC de 0,812.

Tabela 5 – Resultados da classificação com o SVM.

Descritores	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
Geometria	81,2	0,624	0,812	0,188	0,814	0,812	0,812	0,812
HOG (hsv)	79,5	0,591	0,796	0,204	0,801	0,796	0,795	0,796
GBF (hsv)	90,6	0,812	0,906	0,094	0,906	0,906	0,906	0,906
Haralick (hsv)	92,4	0,849	0,925	0,075	0,925	0,925	0,925	0,925
LBP (hsv)	83,5	0,671	0,836	0,164	0,841	0,836	0,835	0,836
ECC (hsv)	83,5	0,671	0,836	0,164	0,836	0,836	0,836	0,836

A Tabela 6 apresenta os resultados obtidos com o método KNN. Os dados mostram que o melhor resultado é alcançado pelas características do descritor GBF. Esses dados determinam uma acurácia de 93,1%, com índice kappa satisfatório de 0,863 e AUC de 0,980. Ademais, se percebe ainda que os dados do descritor de *Haralick* proporcionaram uma acurácia de 91,7%, com índice kappa de 0,835 e AUC de 0,976. Já o pior resultado atingido com este teste também foi proporcionado pelo dados do descritor de Geometria, com uma acurácia de 76,7%, índice kappa de 0,535 e AUC de 0,765.

Tabela 6 – Resultados da classificação com o KNN.

Descritores	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
Geometria	76,7	0,535	0,768	0,232	0,771	0,768	0,765	0,815
HOG (hsv)	79,5	0,591	0,796	0,204	0,796	0,796	0,796	0,872
GBF (hsv)	93,1	0,863	0,932	0,068	0,934	0,932	0,932	0,980
Haralick (hsv)	91,7	0,835	0,918	0,082	0,919	0,918	0,918	0,976
LBP (hsv)	83,5	0,671	0,836	0,164	0,841	0,836	0,835	0,836
ECC (hsv)	80,5	0,610	0,805	0,195	0,809	0,805	0,805	0,850

A Tabela 6 apresenta os resultados obtidos com o método MLP. Os dados mostram que o melhor resultado é proporcionado pelas características do descritor GBF. Esses dados determinam uma acurácia de 93,4%, com índice kappa satisfatório de 0,868 e AUC de 0,967. Esses valores foram bem próximos dos alcançados com o método KNN. Ademais, se percebe que os dados do descritor de *Haralick* também proporcionaram os mesmos valores que foram alcançados com o método anterior.

Tabela 7 – Resultados da classificação com o MLP.

Descritores	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
Geometria	76,5	0,530	0,765	0,235	0,765	0,765	0,765	0,853
HOG (hsv)	78,1	0,563	0,782	0,218	0,786	0,782	0,781	0,866
GBF (hsv)	93,4	0,868	0,934	0,066	0,936	0,934	0,934	0,967
Haralick (hsv)	91,7	0,835	0,918	0,082	0,919	0,918	0,918	0,977
LBP (hsv)	82,1	0,643	0,822	0,178	0,823	0,822	0,821	0,878
ECC (hsv)	81,2	0,624	0,812	0,188	0,812	0,812	0,812	0,892

Finalmente, a Tabela 8 apresenta os resultados obtidos com a arquitetura de RNC sugerida. Os dados mostram que o melhor resultado é alcançado pelas características do descritor de *Haralick*. Esses dados determinam uma acurácia de 93,1%, índice kappa satisfatório de 0,836 e AUC de 0,975. Esses valores foram bem próximos do alcançado com o método KNN. Além disso, se percebe que os dados do descritor de GBF proporcionaram uma acurácia de 88%, com um bom índice kappa de 0,760 e AUC de 0,959. O pior resultado apresentado neste teste foi proporcionado pelos dados do descritor de HOG, com uma acurácia de 76,2%, índice kappa de 0,525, índice kappa de 0,525 e AUC de 0,762.

Tabela 8 – Resultados da classificação com a RNC.

Descritores	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
Geometria	80,7	0,615	0,808	0,192	0,808	0,808	0,807	0,874
HOG (hsv)	76,2	0,525	0,763	0,237	0,765	0,763	0,762	0,851
GBF (hsv)	88	0,760	0,880	0,120	0,881	0,880	0,880	0,959
Haralick (hsv)	93,1	0,836	0,932	0,068	0,932	0,932	0,932	0,975
LBP (hsv)	78,1	0,563	0,782	0,218	0,782	0,782	0,782	0,852
ECC (hsv)	83,3	0,666	0,833	0,167	0,834	0,833	0,833	0,916

Após analisar esses dados, a proposta de reconhecimento pela classificação com RNCs obteve resultados próximos ou similares com uma arquitetura de convolução, em relação aos métodos SVM, KNN e MLP. Entretanto, o classificador MLP manteve o melhor resultado por intermédio das configurações impostas.

6.5 Resultados do Reconhecimento pela Descrição de RNCs

A partir das configurações definidas para o reconhecimento do melanoma pela descrição com RNCs, a Tabela 9 apresenta os resultados para o classificador SVM, de acordo com a variação dos modelos e seus respectivos dados extraídos das camadas FC₆ e FC₇ em todas as imagens do conjunto. Nessa amostra é observada uma acurácia de mais de 89% de instâncias classificadas corretamente, com um índice *kappa* de 0,798, *recall* e precisão maior que 0,8. Esse resultado foi proporcionado pela descrição da camada FC₇ com o treinamento utilizando a operação de *Batch Normalization* na arquitetura VGG-F. Os testes realizados nesse classificador com a arquitetura AlexNet não obtiveram bons resultados.

Tabela 9 – Resultados da classificação mediante a descrição de RNCs no SVM.

RNC	BN	FC	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
AlexNet	sim	6	71,3	0,427	0,714	0,286	0,716	0,714	0,713	0,653
		7	67,7	0,352	0,676	0,324	0,682	0,676	0,673	0,676
	não	6	81,9	0,638	0,819	0,181	0,821	0,819	0,819	0,819
		7	75,11	0,502	0,751	0,249	0,761	0,751	0,749	0,751
VGG-F	sim	6	82,6	0,652	0,826	0,174	0,841	0,826	0,824	0,826
		7	89,9	0,798	0,899	0,101	0,899	0,899	0,899	0,876
	não	6	70,18	0,403	0,702	0,298	0,712	0,702	0,698	0,702
		7	83,1	0,662	0,831	0,169	0,836	0,831	0,830	0,831

Já a Tabela 10 apresenta os resultados obtidos com o classificador KNN. Nessa tabela, a camada citada anteriormente obteve uma acurácia de mais de 91% na identificação de instâncias corretas, um índice *kappa* maior que 0,8 e *recall* e precisão maior que 0,9. Nesse mesmo cenário, a Alexnet obteve seu primeiro resultado satisfatório com a camada FC₆, implementando a operação de *Batch Normalization*. Esse teste atingiu uma acurácia de 87,5%, com um índice *kappa* maior que 0,7 e *recall* e precisão maior que 0,8.

Tabela 10 – Resultados da classificação mediante a descrição de RNCs no KNN.

RCN	BN	FC	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
AlexNet	sim	6	87,5	0,751	0,876	0,124	0,876	0,876	0,876	0,891
		7	85,4	0,708	0,854	0,146	0,856	0,854	0,854	0,875
	não	6	74,41	0,488	0,744	0,256	0,757	0,744	0,741	0,750
		7	71,59	0,431	0,716	0,284	0,728	0,716	0,712	0,713
VGG-F	sim	6	86,8	0,737	0,869	0,131	0,869	0,869	0,869	0,860
		7	91,5	0,831	0,915	0,085	0,916	0,915	0,915	0,943
	não	6	74,41	0,744	0,256	0,758	0,758	0,744	0,741	0,739
		7	73,0	0,460	0,730	0,270	0,743	0,730	0,726	0,728

Por fim, a Tabela 11 apresenta os resultados com o classificador MLP. Nessa amostra se observa que a AlexNet obteve o melhor resultado sem a operação de *Batch Normalization*, utilizando a camada FC₆. Esse teste atingiu uma acurácia de 89,6% e índice

$kappa$ maior que 0,7. Já a VGG-F com a *Batch Normalization* utilizando a camada FC₇ obteve uma acurácia de 85,3% e índice $kappa$ maior que 0,7, *recall* igual a 0,9 e precisão maior que 0,8.

Tabela 11 – Resultados da classificação mediante a descrição de RNCs no MLP.

CNN	BN	FC	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
AlexNet	sim	6	84,7	0,694	0,847	0,153	0,847	0,847	0,847	0,862
		7	85,6	0,713	0,857	0,143	0,859	0,857	0,857	0,932
	não	6	89,6	0,793	0,897	0,103	0,900	0,897	0,896	0,898
		7	80,0	0,600	0,800	0,200	0,801	0,800	0,800	0,901
VGG-F	sim	6	84,5	0,690	0,845	0,155	0,845	0,845	0,845	0,875
		7	85,3	0,713	0,857	0,143	0,857	0,857	0,857	0,881
	não	6	70,18	0,403	0,702	0,298	0,712	0,702	0,698	0,702
		7	73,0	0,460	0,730	0,270	0,743	0,730	0,726	0,728

Os valores de precisão e *recall* obtidos nos melhores testes de classificação foram maiores que 0,8, como visto nas Tabelas 9, 10 e 11, indicando a eficiência na definição de verdadeiros positivos no conjunto de amostras. No entanto, o classificador KNN obteve o melhor resultado entre todos, com os dados da camada FC₇ na arquitetura VGG-F utilizando a normalização. Esses valores de *recall* e precisão foram superiores a 0,9.

6.6 Discussões e Comparações

Com base nos testes apresentados pelas sugestões de abordagens utilizando RNCs, observou-se bons resultados na etapa de classificação na maioria dos testes realizados. A forma de reconhecimento pela classificação com RNCs proporcionou bons diagnósticos das lesões de pele. A partir das configurações submetidas na estrutura sugerida, o melhor resultado foi obtido por meio dos dados extraídos com o descritor de GBF, com uma acurácia de 93,4%, dada pelo método MLP. Em seguida, o seu segundo melhor resultado é apresentado pelas descrições de *Haralick*, com uma acurácia de 93,1%, dada pela arquitetura de convolução sugerida na predição. A tabela 12 mostra um resumo desses principais resultados.

Tabela 12 – Melhores resultados para a classificação com RNCs.

Classificadores	Descritores	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
SVM	<i>Haralick (hsv)</i>	92,4	0,849	0,925	0,075	0,925	0,925	0,925	0,925
KNN	GBF (hsv)	93,1	0,863	0,932	0,068	0,934	0,932	0,932	0,980
MLP	GBF (hsv)	93,4	0,868	0,934	0,066	0,936	0,934	0,934	0,967
RNC	GBF (hsv)	93,1	0,836	0,932	0,068	0,932	0,932	0,932	0,975

Os testes apresentados no reconhecimento do melanoma pela descrição com RNCs também obtiveram bons resultados. A melhor configuração de teste é mostrada pelas descrições da arquitetura VGG-F. A classificação dos seus dados definiu uma acurácia de

91,5% das lesões de pele com a camada FC₇, utilizando um treinamento com o método de *Batch Normalization*, dada pelo método KNN. No segundo melhor teste, a mesma configuração também atingiu uma acurácia de 89,9% com o classificador SVM. A Tabela mostra um resumo dos principais resultados obtidos.

Tabela 13 – Melhores resultados para classificação mediante a descrição de RNCs.

Classificadores	Modelos	FC	BN	A	K	Taxa VP	Taxa FP	P	R	FM	AUC
SVM	VGG-F	7	sim	89,9	0,798	0,899	0,101	0,899	0,899	0,899	0,876
KNN	VGG-F	7	sim	91,5	0,831	0,915	0,085	0,916	0,915	0,915	0,943
MLP	AlexNet	6	não	89,6	0,793	0,897	0,103	0,900	0,897	0,896	0,898

A partir da análise dos melhores resultados, as abordagens sugeridas neste trabalho foram comparadas com alguns dos principais estudos semelhantes. No caso do reconhecimento do melanoma pela classificação com uma arquitetura de RNC, se percebe que, mesmo não atingindo o melhor resultado nos experimentos executados, os dados obtidos foram superiores aos trabalhos apresentados na Tabela 14, levando em consideração a arquitetura tradicional como pré-processamento (1^a), segmentação (2^a), extração de características (3^a) e classificação (4^a), além dos dados de acurácia e precisão.

Tabela 14 – Comparações com a Abordagem de classificação utilizando RNC.

Trabalhos	Etapas				Resultados	
	1 ^a	2 ^a	3 ^a	4 ^a	A (%)	P (%)
Ercal et al. (1994)			x	x	*	80%
Rahman, Bhattacharya e Desai (2008)		x	x	x	83,75%	*
Cavalcanti e Scharcanski (2011)	x	x	x	x	91,7%	*
Classificação com RNCs	x	x	x	x	93,1%	93,2%

Já no caso da abordagens com modelos treinados de RNCs, as comparações realizadas levaram em consideração as etapas de *data augmentation* (1^a), treinamento de RNCs (2^a), extração de características (3^a) e classificação (4^a), além dos dados de acurácia e precisão. A Tabela 15 apresenta um resumo obtido com a comparação de trabalhos semelhantes que desenvolveram os mesmos princípios, em que se percebe um melhor resultado para a forma em questão.

Tabela 15 – Comparações com a Abordagem de Descrição Utilizando Modelos de RNCs.

Trabalhos	Etapas				Resultados	
	1 ^a	2 ^a	3 ^a	4 ^a	A (%)	P (%)
Codella et al. (2015)			x	x	72,3%	*
Majtner, Yildirim-Yayilgan e Hardeberg (2016)			x	x	*	80,5%
Kawahara, BenTaieb e Hamarneh (2016)				x	81,8%	*
Descrição com Modelos de RNCs	x	x	x	x	91,5%	91,6%

6.7 Considerações Finais

Esse capítulo apresentou todos os aparatos e configurações que foram realizados para determinar o diagnóstico das lesões de pele com as formas sugeridas neste trabalho,

além dos principais resultados obtidos ao longo dos experimentos, associando as Redes Neurais Convolucionais em abordagens de detecção do melanoma nas etapas de extração de características e de classificação para imagens dermatoscópicas. O próximo capítulo descreve as principais conclusões obtidas no desenvolvimento deste trabalho.

7 Conclusões e Trabalhos Futuros

A partir dos conceitos da abordagem tradicional de detecção do melanoma e do Aprendizado Profundo, foram apresentadas duas formas distintas de reconhecimento de lesões de pele para imagens dermatoscópicas. Essas formas conduziram as Redes Neurais Convolucionais ao diagnóstico do melanoma, em que foi possível obter resultados significativos em situações que envolvem descrição e classificação nesse tipo de imagem. Para tal, a primeira estrutura apresentada é baseada nos sistemas convencionais, porém ela implementa as operações de convolução na etapa de classificação de dados. Já a segunda estrutura sugerida utiliza esses processos convolucionais na etapa de extração de características para determinar dados espaciais relevantes no diagnóstico em classes.

O diagnóstico das lesões de pele com o reconhecimento do melanoma pela classificação com RNC mostra seu melhor resultado por meio dos dados extraídos com o descritor de GBF, com uma acurácia de 93,4%, dada pelo método MLP. Entretanto, o resultado apresentado pelas descrições de *Haralick* se aproximou desse resultado, com uma acurácia de 93,1%, dada pela arquitetura de RNC sugerida. Isso mostra que as operações de convolução conseguem atingir resultados semelhantes aos métodos convencionais. O aumento na quantidade de dados pode influenciar diretamente os resultados nessa estrutura. Portanto, um maior conjunto de dados pode reverter o quadro e proporcionar resultados mais precisos com outras arquiteturas de RNC. Além disso, as descrições de texturas se mostraram indispensáveis em qualquer configuração que trata do reconhecimento do melanoma nesses moldes. A sua utilização determina uma descrição que proporciona um melhor diagnóstico de lesões cancerígenas.

Os testes apresentados pelo reconhecimento do melanoma pela descrição com RNCs também obtiveram bons resultados. O melhor resultado é apresentado pelas descrições da arquitetura VGG - F, em que definiu-se uma acurácia de 91,5% das lesões de pele com a camada FC_7 , utilizando um treinamento com o método de *Bach Normalization*, dada pelo método KNN. Entretanto, a mesma configuração também atingiu uma acurácia de 89,9% com o classificador SVM. A arquitetura AlexNet obteve resultados inferiores nessa estrutura por possuir uma menor quantidade de filtros por camadas convolucionais. O treinamento das arquiteturas sugeridas neste trabalho com a técnica de *Dropout* não conseguiu produzir modelos capazes de captar dados relevantes das lesões e seu tempo de execução é superior a normalização com lotes.

As propostas sugeridas neste trabalho atingiram bons resultados na classificação do melanoma. Entretanto, a forma de reconhecimento pela classificação com RNCs ainda é mais vantajosa em relação à estrutura de reconhecimento com modelos treinados de RNCs,

pois se percebe melhores resultados para essa configuração. A diferença de resultados entre as abordagens é explicada por meio da quantidade de imagens que foram submetidas nos experimentos. A operação de *data augmentation* aplicada ao conjunto de treino não foi suficiente para mediar esse problema, sendo necessário uma maior quantidade de imagens e classes para produzir um melhor aprendizado em arquiteturas convolucionais. E outra situação, as perspectivas com a correção desse problema podem demonstrar um melhor resultado para a abordagem com modelos treinados de RNCs, pois seria possível produzir um descritor mais robusto.

A etapa de seleção de características também foi implementada na descrição com modelos treinados de RNCs, já que a quantidade de dados obtidos das camadas FC₆ e FC₇ é muito grande (4096 características). Entretanto, os resultados de diagnóstico do melanoma com essa etapa adicional foram inferiores a proposta apresentada neste trabalho, pois a utilização de seletores de características precedendo a etapa de classificação os valores significativos das descrições dos modelos AlexNet e VGG-F. As camadas FC₆ e FC₇ extraem dados que estão entre os intervalos 0 e 1 e quando um algum método seletor determina uma relevância de atributos é possível que eles tenham mais valores nulos. Nessa situação de teste, o conjunto selecionado possui perda de dados, prejudicando a eficácia do diagnóstico pelos métodos classificadores.

Em trabalhos futuros, pretende-se aprimorar o reconhecimento do melanoma pela classificação com RNCs por meio da combinação de novos espaços de cores, juntamente com novos descritores. Essa estrutura de detecção ainda deverá testar outras arquiteturas convolucionais para realizar a classificação dos dados obtidos. Já a forma de reconhecimento do melanoma pela descrição com RNCs sugere como melhoria o acréscimo de mais bases de imagens, além de novas arquiteturas convolucionais para treinamento, com intuito de verificar a efetividade das descrições proporcionadas pelo processo convolucional.

Referências

- ABBAS, Q.; CELEBI, M. E.; GARCÍA, I. F. Hair removal methods: a comparative study for dermoscopy images. *Biomedical Signal Processing and Control*, Elsevier, v. 6, n. 4, p. 395–404, 2011. Citado na página 26.
- ALBERTI, K. G. M. M.; ZIMMET, P. f. Definition, diagnosis and classification of diabetes mellitus and its complications. part 1: diagnosis and classification of diabetes mellitus. provisional report of a who consultation. *Diabetic medicine*, Wiley Online Library, v. 15, n. 7, p. 539–553, 1998. Citado na página 28.
- ANTHIMOPOULOS, M. et al. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE transactions on medical imaging*, IEEE, v. 35, n. 5, p. 1207–1216, 2016. Citado na página 2.
- ARGENZIANO, G. et al. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Archives of dermatology*, American Medical Association, v. 134, n. 12, p. 1563–1570, 1998. Citado na página 15.
- ARGENZIANO, G. et al. Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. *Journal of the American Academy of Dermatology*, Elsevier, v. 48, n. 5, p. 679–693, 2003. Citado 3 vezes nas páginas 3, 15 e 61.
- BALLARD, D. H. Generalizing the hough transform to detect arbitrary shapes. *Pattern recognition*, Elsevier, v. 13, n. 2, p. 111–122, 1981. Citado na página 53.
- BARATA, C. et al. Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Systems Journal*, IEEE, v. 8, n. 3, p. 965–979, 2014. Citado 7 vezes nas páginas 1, 3, 21, 22, 26, 29 e 35.
- BARATA, C. et al. A bag-of-features approach for the classification of melanomas in dermoscopy images: The role of color and texture descriptors. In: *Computer Vision Techniques for the Diagnosis of Skin Cancer*. [S.l.]: Springer, 2014. p. 49–69. Citado na página 25.
- BENGIO, Y.; GOODFELLOW, I. J.; COURVILLE, A. Deep learning. *An MIT Press book in preparation. Draft chapters available at <http://www.iro.umontreal.ca/bengioy/dlbook>*, Citeseer, 2015. Citado 7 vezes nas páginas 40, 41, 44, 45, 46, 47 e 48.
- BERG, M. D. et al. Computational geometry. In: *Computational geometry*. [S.l.]: Springer, 2000. p. 1–17. Citado na página 53.
- BEZDEK, J. C.; EHRLICH, R.; FULL, W. Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, Elsevier, v. 10, n. 2-3, p. 191–203, 1984. Citado na página 4.
- BRAGA, A. d. P.; CARVALHO, A.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: Livros Técnicos e Científicos, 2000. Citado na página 36.

- CARNEIRO, G.; NASCIMENTO, J.; BRADLEY, A. P. Unregistered multiview mammogram analysis with pre-trained deep learning models. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.], 2015. p. 652–660. Citado 2 vezes nas páginas 23 e 24.
- CAVALCANTI, P. G.; SCHARCANSKI, J. Automated prescreening of pigmented skin lesions using standard cameras. *Computerized Medical Imaging and Graphics*, Elsevier, v. 35, n. 6, p. 481–491, 2011. Citado 3 vezes nas páginas 20, 22 e 70.
- CELEBI, M. E. et al. Lesion border detection in dermoscopy images. *Computerized medical imaging and graphics*, Elsevier, v. 33, n. 2, p. 148–153, 2009. Citado na página 30.
- CELEBI, M. E. et al. A methodological approach to the classification of dermoscopy images. *Computerized Medical Imaging and Graphics*, Elsevier, v. 31, n. 6, p. 362–373, 2007. Citado 5 vezes nas páginas 20, 22, 26, 64 e 65.
- CELEBI, M. E.; MENDONCA, T.; MARQUES, J. S. *Dermoscopy Image Analysis*. [S.l.]: CRC Press, 2015. v. 10. Citado 3 vezes nas páginas 1, 3 e 14.
- CHATFIELD, K. et al. Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*, 2014. Citado 2 vezes nas páginas 48 e 58.
- CODELLA, N. et al. Deep learning, sparse coding, and svm for melanoma recognition in dermoscopy images. In: *Machine Learning in Medical Imaging*. [S.l.]: Springer, 2015. p. 118–126. Citado 4 vezes nas páginas 2, 23, 24 e 70.
- CRUZ-ROA, A. A. et al. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.], 2013. p. 403–410. Citado 2 vezes nas páginas 22 e 24.
- CUN, B. B. L. et al. Handwritten digit recognition with a back-propagation network. In: CITESEER. *Advances in neural information processing systems*. [S.l.], 1990. Citado 2 vezes nas páginas 22 e 46.
- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: IEEE. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. [S.l.], 2005. v. 1, p. 886–893. Citado 3 vezes nas páginas 4, 30 e 31.
- DÂNGELO, J. G.; FATTINI, C. Sistema tegumentar. *Anatomia humana sistêmica e segmentar, para o estudante de medicina*. Edit Ateneu, São Paulo, p. 173–175, 1998. Citado 3 vezes nas páginas 9, 10 e 12.
- DENG, L.; YU, D. Deep learning: Methods and applications. *Foundations and Trends® in Signal Processing*, v. 7, n. 3–4, p. 197–387, 2014. ISSN 1932-8346. Disponível em: <<http://dx.doi.org/10.1561/20000000039>>. Citado na página 40.
- DIO, L. J. D. Tratado de anatomia aplicada. In: *Tratado de anatomia aplicada*. [S.l.]: Póluss, 1998. Citado 3 vezes nas páginas 10, 11 e 12.
- ELDER, D. E.; ELENITSAS, R.; JR, B. L. J. *Lever: Histopatologia Da Pele*. [S.l.]: Grupo Gen-Guanabara Koogan, 2000. Citado 3 vezes nas páginas 9, 10 e 14.

- ERCAL, F. et al. Neural network diagnosis of malignant melanoma from color images. *IEEE Transactions on biomedical engineering*, IEEE, v. 41, n. 9, p. 837–845, 1994. Citado 3 vezes nas páginas 19, 22 e 70.
- FAWZI, A. et al. Adaptive data augmentation for image classification. In: IEEE. *Image Processing (ICIP), 2016 IEEE International Conference on*. [S.l.], 2016. p. 3688–3692. Citado na página 49.
- FLEISS, J. L.; LEVIN, B.; PAIK, M. C. *Statistical methods for rates and proportions*. [S.l.]: John Wiley & Sons, 2013. Citado 3 vezes nas páginas 62, 63 e 64.
- FOGEL, I.; SAGI, D. Gabor filters as texture discriminator. *Biological cybernetics*, Springer, v. 61, n. 2, p. 103–113, 1989. Citado 3 vezes nas páginas 4, 32 e 33.
- FRANGE, V. M. N.; ARRUDA, L. H. F.; DALDON, P. É. C. Dermatoscopia: importância para a prática clínica. *Revista de Ciências Médicas*, v. 18, n. 4, 2012. Citado 3 vezes nas páginas 1, 15 e 25.
- GANSTER, H. et al. Automated melanoma recognition. *IEEE transactions on medical imaging*, IEEE, v. 20, n. 3, p. 233–239, 2001. Citado 5 vezes nas páginas 4, 19, 20, 22 e 29.
- GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. [S.l.]: Prentice hall Upper Saddle River, 2002. Citado 3 vezes nas páginas 26, 27 e 29.
- GOODFELLOW, I. J. et al. Multi-digit number recognition from street view imagery using deep convolutional neural networks. *arXiv preprint arXiv:1312.6082*, 2013. Citado 2 vezes nas páginas 40 e 42.
- GUNN, S. R. et al. Support vector machines for classification and regression. *ISIS technical report*, v. 14, 1998. Citado 2 vezes nas páginas 38 e 39.
- GUTMAN, D. et al. Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1605.01397*, 2016. Citado na página 61.
- HALL, M. et al. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, ACM, v. 11, n. 1, p. 10–18, 2009. Citado na página 65.
- HARALICK, R. M. Statistical and structural approaches to texture. *Proceedings of the IEEE*, IEEE, v. 67, n. 5, p. 786–804, 1979. Citado 2 vezes nas páginas 4 e 34.
- HARGIS, A. Sistema tegumentar. *THOMSON, RG Patologia Veterinária especial*. São Paulo: Manole, 1990. Citado na página 11.
- HASTIE, T.; TIBSHIRANI, R. Discriminant adaptive nearest neighbor classification. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 18, n. 6, p. 607–616, 1996. Citado na página 35.
- HAYKIN, S. S. *Redes neurais*. [S.l.]: Bookman, 2001. Citado 3 vezes nas páginas 9, 36 e 46.
- HAYKIN, S. S.; HAYKIN, S. S.; HAYKIN, S. S. *Kalman filtering and neural networks*. [S.l.]: Wiley Online Library, 2001. Citado na página 37.

- HONG, S. et al. A novel coupled template for face recognition based on a convolutional neural network. In: IEEE. *2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA)*. [S.l.], 2015. p. 52–56. Citado na página 2.
- HU, M.-K. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, IEEE, v. 8, n. 2, p. 179–187, 1962. Citado na página 53.
- HUYNH, X.-P.; TRAN, T.-D.; KIM, Y.-G. Convolutional neural network models for facial expression recognition using bu-3dfe database. In: *Information Science and Applications (ICISA) 2016*. [S.l.]: Springer, 2016. p. 441–450. Citado na página 2.
- INCA, I. N. de C. *Estimativas do câncer para o ano de 2016*. 2016. Url: <http://www.inca.gov.br/estimativa/2016/>. Acessado em 01-05-2016. Citado 2 vezes nas páginas 1 e 14.
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. *ACM computing surveys (CSUR)*, Acm, v. 31, n. 3, p. 264–323, 1999. Citado na página 28.
- JONES, J. P.; PALMER, L. A. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of neurophysiology*, Am Physiological Soc, v. 58, n. 6, p. 1233–1258, 1987. Citado na página 32.
- JÚNIOR, O. L.; MONTGOMERY, E. *Redes neurais-fundamentos e aplicações com programas em c*. Rio de Janeiro: Editora Ciencia Moderna Ltda, 2007. Citado na página 37.
- KAWAHARA, J.; BENTAIEB, A.; HAMARNEH, G. Deep features to classify skin lesions. In: IEEE. *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. [S.l.], 2016. p. 1397–1400. Citado 5 vezes nas páginas 22, 23, 24, 47 e 70.
- KLIGERMAN, J. Estimativas sobre a incidência e mortalidade por câncer no brasil-2000 estimates of cancer incidence and mortality in brazil-2000. *Revista Brasileira de Cancerologia*, v. 46, n. 1, 2000. Citado 2 vezes nas páginas 11 e 13.
- KOROTKOV, K.; GARCIA, R. Computerized analysis of pigmented skin lesions: a review. *Artificial intelligence in medicine*, Elsevier, v. 56, n. 2, p. 69–90, 2012. Citado 10 vezes nas páginas 1, 3, 21, 22, 25, 26, 29, 35, 62 e 64.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105. Citado 4 vezes nas páginas 23, 42, 49 e 57.
- KUMAR, A.; PANG, G. K. Defect detection in textured materials using gabor filters. *IEEE Transactions on industry applications*, IEEE, v. 38, n. 2, p. 425–440, 2002. Citado na página 55.
- LAHMIRI, S.; BOUKADOUM, M. Hybrid discrete wavelet transform and gabor filter banks processing for features extraction from biomedical images. *Journal of medical engineering*, Hindawi Publishing Corporation, v. 2013, 2013. Citado na página 33.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *biometrics*, JSTOR, p. 159–174, 1977. Citado 2 vezes nas páginas 62 e 63.

- LAWRENCE, S. et al. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, IEEE, v. 8, n. 1, p. 98–113, 1997. Citado na página 22.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015. Citado 3 vezes nas páginas 39, 43 e 45.
- LEE, C.-Y.; GALLAGHER, P. W.; TU, Z. Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree. In: *International conference on artificial intelligence and statistics*. [S.l.: s.n.], 2016. Citado 4 vezes nas páginas 42, 44, 45 e 50.
- LEE, H.; CHEN, Y.-P. P. Image based computer aided diagnosis system for cancer detection. *Expert Systems with Applications*, Elsevier, v. 42, n. 12, p. 5356–5365, 2015. Citado 4 vezes nas páginas 5, 41, 44 e 48.
- LEE, T. et al. Dullrazor®: A software approach to hair removal from images. *Computers in biology and medicine*, Elsevier, v. 27, n. 6, p. 533–543, 1997. Citado 3 vezes nas páginas 9, 26 e 52.
- LOZANO-DIEZ, A. et al. An end-to-end approach to language identification in short utterances using convolutional neural networks. In: *Proc. of interspeech*. [S.l.: s.n.], 2015. p. 403–407. Citado na página 2.
- MAGLOGIANNIS, I.; DOUKAS, C. N. Overview of advanced computer vision systems for skin lesions characterization. *Information Technology in Biomedicine, IEEE Transactions on*, IEEE, v. 13, n. 5, p. 721–733, 2009. Citado 13 vezes nas páginas 1, 3, 20, 22, 25, 26, 27, 29, 31, 35, 51, 52 e 65.
- MAJTNER, T.; YILDIRIM-YAYILGAN, S.; HARDEBERG, J. Y. Combining deep learning and hand-crafted features for skin lesion classification. In: IEEE. *Image Processing Theory Tools and Applications (IPTA), 2016 6th International Conference on*. [S.l.], 2016. p. 1–6. Citado 3 vezes nas páginas 23, 24 e 70.
- MARIEB, E. N.; HOEHN, K. *Anatomia e fisiologia*. [S.l.]: Artmed Editora, 2009. Citado 4 vezes nas páginas 9, 10, 11 e 12.
- MENDONÇA, T. et al. Ph 2-a dermoscopic image database for research and benchmarking. In: IEEE. *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. [S.l.], 2013. p. 5437–5440. Citado 4 vezes nas páginas 9, 52, 54 e 61.
- MENZIES, S. W.; CROTTY, K. A.; MCCARTHY, W. H. The morphologic criteria of the pseudopod in surface microscopy. *Archives of dermatology*, American Medical Association, v. 131, n. 4, p. 436–440, 1995. Citado na página 15.
- MERY, D. *BALU: a toolbox Matlab for computer vision, pattern recognition and image processing*. 2011. Citado 4 vezes nas páginas 30, 31, 33 e 34.
- MU, Y. et al. Discriminative local binary patterns for human detection in personal album. In: IEEE. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. [S.l.], 2008. p. 1–8. Citado na página 4.
- NACHBAR, F. et al. The abcd rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, Elsevier, v. 30, n. 4, p. 551–559, 1994. Citado na página 15.

- NAZ, S.; MAJEED, H.; IRSHAD, H. Image segmentation using fuzzy clustering: A survey. In: IEEE. *Emerging Technologies (ICET), 2010 6th International Conference on*. [S.l.], 2010. p. 181–186. Citado na página 28.
- NGUYEN, N. H.; LEE, T. K.; ATKINS, M. S. Segmentation of light and dark hair in dermoscopic images: a hybrid approach using a universal kernel. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *SPIE Medical Imaging*. [S.l.], 2010. p. 76234N–76234N. Citado na página 26.
- OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, Elsevier, v. 29, n. 1, p. 51–59, 1996. Citado na página 33.
- PARKHI, O. M.; VEDALDI, A.; ZISSERMAN, A. Deep face recognition. In: *BMVC*. [S.l.: s.n.], 2015. v. 1, n. 3, p. 6. Citado 5 vezes nas páginas 44, 48, 56, 58 e 60.
- PETERSON, L. E. K-nearest neighbor. *Scholarpedia*, v. 4, n. 2, p. 1883, 2009. Citado na página 35.
- PHUNG, S. L.; BOUZERDOUM, A.; CHAI, D. Skin segmentation using color pixel classification: analysis and comparison. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 27, n. 1, p. 148–154, 2005. Citado na página 31.
- PRICE, R.; ISO, K.-i.; SHINODA, K. Wise teachers train better dnn acoustic models. *EURASIP Journal on Audio, Speech, and Music Processing*, Springer, v. 2016, n. 1, p. 1–19, 2016. Citado 2 vezes nas páginas 2 e 46.
- RAHMAN, M. M.; BHATTACHARYA, P.; DESAI, B. C. A multiple expert-based melanoma recognition system for dermoscopic images of pigmented skin lesions. In: IEEE. *BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on*. [S.l.], 2008. p. 1–6. Citado 3 vezes nas páginas 20, 22 e 70.
- RANGAYYAN, R. M.; AYRES, F. J.; DESAUTELS, J. L. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute*, Elsevier, v. 344, n. 3, p. 312–348, 2007. Citado na página 30.
- RAZMJOOY, N. et al. A computer-aided diagnosis system for malignant melanomas. *Neural Computing and Applications*, Springer, v. 23, n. 7-8, p. 2059–2071, 2013. Citado na página 39.
- RUIZ, D. et al. A decision support system for the diagnosis of melanoma: A comparative approach. *Expert Systems with Applications*, Elsevier, v. 38, n. 12, p. 15217–15223, 2011. Citado 2 vezes nas páginas 21 e 22.
- RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, v. 115, n. 3, p. 211–252, 2015. Citado na página 57.
- SAMPAIO, S. A.; RIVITTI, E. A. *Dermatologia*. [S.l.]: Artes Médicas, 2001. Citado na página 13.
- SBD, S. B. de D. *Tipos de câncer de pele*. 2016. Url: <http://www.sbd.org.br/informacoes/sobre-o-cancer-da-pele/tipos-de-cancer-da-pele/>. Acessado em 01-05-2016. Citado 5 vezes nas páginas 9, 1, 13, 14 e 16.

SCHMID, P. Segmentation of digitized dermatoscopic images by two-dimensional color clustering. *IEEE Transactions on Medical Imaging*, IEEE, v. 18, n. 2, p. 164–171, 1999. Citado na página 27.

SCHMID-SAUGEONA, P.; GUILLODB, J.; THIRANA, J.-P. Towards a computer-aided diagnosis system for pigmented skin lesions. *Computerized Medical Imaging and Graphics*, Elsevier, v. 27, n. 1, p. 65–78, 2003. Citado na página 4.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural Networks*, Elsevier, v. 61, p. 85–117, 2015. Citado na página 39.

SEELEY, R. R. Anatomia & fisiologia. Lusodidacta, 2001. Citado 2 vezes nas páginas 9 e 12.

SERMANET, P.; CHINTALA, S.; LECUN, Y. Convolutional neural networks applied to house numbers digit classification. In: IEEE. *Pattern Recognition (ICPR), 2012 21st International Conference on*. [S.l.], 2012. p. 3288–3291. Citado na página 2.

SERMANET, P.; LECUN, Y. Traffic sign recognition with multi-scale convolutional networks. In: IEEE. *Neural Networks (IJCNN), The 2011 International Joint Conference on*. [S.l.], 2011. p. 2809–2813. Citado 2 vezes nas páginas 22 e 41.

SHEN, W. et al. Multi-scale convolutional neural networks for lung nodule classification. In: SPRINGER. *International Conference on Information Processing in Medical Imaging*. [S.l.], 2015. p. 588–599. Citado na página 2.

SHI, Z.; HE, L. Current status and future potential of neural networks used for medical image processing. *Journal of multimedia*, v. 6, n. 3, p. 244–251, 2011. Citado 4 vezes nas páginas 2, 3, 5 e 22.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Citado na página 45.

SIOHAN, O.; RYBACH, D. Multitask learning and system combination for automatic speech recognition. In: IEEE. *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*. [S.l.], 2015. p. 589–595. Citado na página 2.

SOARES, H. B. Análise e classificação de imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte. Universidade Federal do Rio Grande do Norte, 2008. Citado na página 1.

SOCIETY, A. C. *Skin Cancer Lesion*. 2016. Url: <http://www.cancer.org>. Acessado em 01-06-2016. Citado 4 vezes nas páginas 9, 14, 15 e 16.

SOYER, H. P. et al. Is dermoscopy useful for the diagnosis of melanoma? *Archives of dermatology*, American Medical Association, v. 137, n. 10, p. 1361–1363, 2001. Citado 2 vezes nas páginas 15 e 16.

UZAN, L.; WOLF, L. I know that voice: Identifying the voice actor behind the voice. In: IEEE. *2015 International Conference on Biometrics (ICB)*. [S.l.], 2015. p. 46–51. Citado na página 2.

- VEDALDI, A.; LENC, K. Matconvnet: Convolutional neural networks for matlab. In: ACM. *Proceedings of the 23rd ACM international conference on Multimedia*. [S.l.], 2015. p. 689–692. Citado 2 vezes nas páginas 64 e 65.
- VEZHNEVETS, V.; SAZONOV, V.; ANDREEVA, A. A survey on pixel-based skin color detection techniques. In: MOSCOW, RUSSIA. *Proc. Graphicon*. [S.l.], 2003. v. 3, p. 85–92. Citado na página 31.
- XIE, F.; BOVIK, A. C. Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm. *Pattern Recognition*, Elsevier, v. 46, n. 3, p. 1012–1019, 2013. Citado na página 27.
- YUAN, X. et al. Svm-based texture classification and application to early melanoma detection. In: IEEE. *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*. [S.l.], 2006. p. 4775–4778. Citado 3 vezes nas páginas 37, 38 e 39.
- ZHAO, W.; DU, S. Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 54, n. 8, p. 4544–4554, 2016. Citado na página 2.
- ZHOU, H. et al. Anisotropic mean shift based fuzzy c-means segmentation of dermoscopy images. *IEEE Journal of Selected Topics in Signal Processing*, IEEE, v. 3, n. 1, p. 26–34, 2009. Citado na página 27.