**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS**

Leonardo Felipe dos Santos Scabini

**Análise de imagens multiespectrais através de redes
complexas**

São Carlos

2018

Leonardo Felipe dos Santos Scabini

**Análise de imagens multiespectrais através de redes
complexas**

Dissertação apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

Área de concentração: Física Aplicada
Opção: Física Computacional

Orientador: Prof. Dr. Odemir Martinez Bruno

Versão corrigida
(Versão original disponível na Unidade que aloja o Programa)

São Carlos
2018

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Scabini, Leonardo Felipe dos Santos

Análise de imagens multiespectrais através de redes complexas / Leonardo Felipe dos Santos Scabini; orientador Odemir Martinez Bruno - versão corrigida -- São Carlos, 2018.

117 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Física Aplicada Computacional) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2018.

1. Imagens multiespectrais. 2. Redes complexas. 3. Extração de características. 4. Classificação de imagens. I. Bruno, Odemir Martinez, orient. II. Título.

A todos aqueles com mente e imaginação indomáveis.

AGRADECIMENTOS

À vastidão do Cosmos e seus mistérios, pela existência e as infinitas possibilidades de descobrimento.

Aos meus pais, irmãos, avós e tios, pelo apoio e por sempre acreditarem nas minhas escolhas.

A todos Aesir e Vanir e principalmente Wotan, por se manterem incólumes, imutáveis em meio ao caos e permitirem que a antiga tradição continue.

À Karina Rocha Campos pelo amor, companhia e apoio em todos os momentos.

A Odemir Martinez Bruno pela orientação, amizade, discussões e ideias que contribuíram para o desenvolvimento deste trabalho.

Aos amigos e companheiros de laboratório, pelos momentos de descontração, discussões e trocas de conhecimento.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro que contribuiu para o desenvolvimento deste trabalho.

E a todos os funcionários, colegas e professores do Instituto de Física de São Carlos (IFSC-USP) e do Instituto de Ciências Matemáticas e de Computação (ICMC-USP), que contribuíram, direta ou indiretamente, para a minha formação acadêmica e o desenvolvimento deste trabalho.

*Døyr fe, døyr frender
Døyr sjølv det sama
men ordet om deg aldreg døyr
vinn du et gjetord gjevt
Døyr fe, døyr frender
Døyr sjølv det sama
Eg veit et som aldreg døyr
dom om daudan kvar
(Håvamål)*

RESUMO

SCABINI, L. F. S. **Análise de imagens multiespectrais através de redes complexas**. 2018. 117p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2018.

Imagens multiespectrais estão presentes na grande maioria de dispositivos de imageamento atuais, desde câmeras pessoais até microscópios, telescópios e satélites. No entanto, grande parte dos trabalhos em análise de texturas e afins propõem abordagens monocromáticas, que muitas vezes consideram apenas níveis de cinza. Nesse contexto e considerando o aumento da capacidade dos computadores atuais, o uso da informação espectral deve ser considerada na construção de modelos melhores. Ultimamente redes neurais convolucionais profundas pré-treinadas tem sido usadas em imagens coloridas de 3 canais, porém são limitadas a apenas esse formato e computam muitas convoluções, o que demanda por *hardware* específico (GPU). Esses fatos motivaram esse trabalho, que propõem técnicas para a modelagem e caracterização de imagens multiespectrais baseadas em redes complexas, que tem se mostrado uma ferramenta eficiente em trabalhos anteriores e possui complexidade computacional similar à métodos tradicionais. São introduzidas duas abordagens para aplicação em imagens coloridas de três canais, denominadas Rede Multicamada (RM) e Rede Multicamada Direcionada (RMD). Esses métodos modelam todos os canais da imagem de forma conjunta, onde as redes possuem conexões intra e entre canais, de forma parecida ao processamento oponente de cor do sistema visual humano. Experimentos em cinco bases de textura colorida mostram a proposta RMD supera vários métodos da literatura no geral, incluindo redes convolucionais e métodos tradicionais integrativos. Além disso, as propostas demonstraram alta robustez a diferentes espaços de cor (RGB, LAB, HSV e $I_1I_2I_3$), enquanto que outros métodos oscilam de base para base. Também é proposto um método para caracterizar imagens multiespectrais de muitos canais, denominado Rede Direcionada de Similaridade Angular (RDSA). Nessa proposta, cada pixel multiespectral é considerado como um vetor de dimensão equivalente à quantidade de canais da imagem e o peso das arestas representa sua similaridade do cosseno, apontando para o pixel de maior valor absoluto. Esse método é aplicado em um conjunto de imagens de microscopia por fluorescência de 32 canais, em um experimento para identificar variações na estrutura foliar do espécime Jacaranda Caroba submetidos à diferentes condições. O método RDSA obtém as maiores taxas de acerto de classificação nesse conjunto de dados, com 91,9% usando o esquema de validação cruzada *Leave-one-out* e $90,5(\pm 1,1)\%$ com 10-pastas, contra 81,8% e $84,7(\pm 2,2)$ da rede convolucional VGG16.

Palavras-chave: Imagens multiespectrais. Redes complexas. Extração de características. Classificação de imagens.

ABSTRACT

SCABINI, L. F. S. **Multispectral image analysis through complex networks.** 2018. 117p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2018.

Multispectral images are present in the vast majority of current imaging devices, from personal cameras to microscopes, telescopes and satellites. However, much of the work in texture analysis and the like proposes monochromatic approaches, which often consider only gray levels. In this context and considering the performance increase of current computers, the use of the spectral information must be considered in the construction of better models. Lately, pre-trained deep convolutional neural networks have been used in 3-channel color images, however they are limited to just this format and compute many convolutions, which demands specific hardware (GPU). These facts motivated this work, which propose techniques for the modeling and characterization of multispectral images based on complex networks, which has proved to be an efficient tool in previous works and has computational complexity similar to traditional methods. Two approaches are introduced for application in 3-channel color images, called Multilayer Network (RM) and Directed Multilayer Network (RMD). These methods model all channels of the image together, where the networks have intra- and inter-channel connections, similar to the opponent color processing of the human visual system. Experiments in five color texture datasets shows that the RMD proposal overcomes several methods of the literature in general, including convolutional networks and traditional integrative methods. In addition, the proposals have demonstrated high robustness to different color spaces (RGB, LAB, HSV and $I_1I_2I_3$), while other methods oscillate from dataset to dataset. Moreover it is proposed a new method to characterize multispectral images of many channels, called Directed Network of Angular Similarity (RDSA). In this proposal, each multispectral pixel is considered as a vector of dimensions equivalent to the number of channels of the image and the weight of the edges represents its cosine similarity, pointing to the pixel of greatest absolute value. This method is applied to a set of fluorescence microscopy images of 32 channels in an experiment to identify variations in the leaf structure of the Jacaranda Caroba specimen under different conditions. The RDSA method obtains the highest classification rates in this dataset, with 91.9% with the Leave-one-out cross-validation scheme and $90.5(\pm 1.1)\%$ with 10-folds, against 81.8% and $84.7(\pm 2.2)\%$ of the convolutional network VGG16.

Keywords: Multispectral image. Complex network. Feature extraction. Image classification.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplos variados de texturas naturais e de materiais encontrados no dia-dia (a) e de texturas presentes em imagens microscópicas (b).	24
Figura 2 – Diferentes comprimentos de onda do espectro eletromagnético (a), e o respectivo pico de resposta dos cones L, M e S dentro do espectro visível. Esquema de processamento oponente de cor (b) a partir da diferença entre as respostas dos cones.	32
Figura 3 – Esquemas da maioria dos métodos da literatura para caracterização de textura colorida. O modelo integrativo computa descritores tradicionais de textura para cada canal individual, e então os combina. O esquema paralelo computa descritores de cor pura não espaciais, e os combina com descritores obtidos da imagem em nível de cinza. Descritores de cor pura também são comumente utilizados individualmente.	34
Figura 4 – Topologia das redes complexas aleatórias de Erdős e Rényi. (a) Exemplo de uma rede aleatória construída com $m = 50$ e $p = 0, 1$. (b) distribuição do grau para 100 redes geradas com as mesmas configurações de (a).	37
Figura 5 – Topologia das redes complexas de pequeno mundo de Watts e Strogatz. (a) Exemplo de uma rede de pequeno mundo construída com $m = 50$, $i = 2$ e $p = 0, 1$. (b) distribuição do grau para 100 redes geradas com as mesmas configurações de (a).	38
Figura 6 – Topologia das redes complexas livres de escala de Barabási e Albert. (a) Exemplo de uma rede livre de escala construída com $m_0 = 2$, $m = 50$ e $x = 1$. (b) distribuição do grau para 100 redes geradas com as mesmas configurações de (a).	39
Figura 7 – Modelagem de formas com RCs, onde cada pixel do contorno do objeto é um vértice, e arestas são definidas conectando pixels próximos se sua distância Euclidiana é menor que um dado limiar t	41
Figura 8 – Modelagem de redes N^r a partir de uma imagem em nível de cinza de tamanho $ I = w * h$, onde $w = 10$ e $h = 8$. O parâmetro r limita a criação de conexões para uma região local de raio r , portanto quanto maior seu valor, mais conexões são criadas.	42
Figura 9 – Estrutura de redes $N^{r,t}$ modeladas a partir da textura de entrada (a). O padrão textural é refletido no padrão topológico da rede resultante. Quanto maior o raio r e o limiar t , mais densa será a rede $N^{r,t}$	44
Figura 10 – Diagrama geral das propostas deste trabalho para a caracterização de imagens multiespectrais, onde são introduzidas 3 novas abordagens.	48

Figura 11 – Primeira etapa da modelagem de imagens coloridas (no exemplo RGB) em redes multicamada. Um pixel RGB é representado por 3 vértices, sendo 1 para cada canal de cor, e combinados compõem o conjunto de camadas $V = \{V_R, V_G, V_B\}$	49
Figura 12 – Distribuição dos pesos das arestas de acordo com diferentes equações de peso, calculada para a textura de entrada com 3 diferentes valores de raio (r). A frequência é a função densidade de probabilidade calculada usando uma discretização de 100 <i>bins</i>	51
Figura 13 – Processo de definição de arestas em uma rede multicamada para uma imagem RGB. Os vértices entre todas as camadas são conectados dentro de um raio de distância r . Apenas vértices válidos são considerados para evitar o efeito de borda.	53
Figura 14 – Etapa de remoção de arestas com limiares $t_1 < t_2 < t_3$. Quanto maior t , menos arestas são mantidas e a rede se torna mais esparsa.	54
Figura 15 – Diferenças na distribuição das arestas (c, e, g) de duas imagens diferentes (a e b) usando 3 raios diferentes e discretização de 100 intervalos. Quando a rede $N^{r,t}$ é modelada usando $t = a$, o padrão do grau médio μ_k também varia de acordo com a imagem, com pouca variação entre raios (d, f, h).	56
Figura 16 – Frequência de arestas e grau médio em 4 bases de textura colorida. Em azul os limites $[t_1, t_m]$ estimados pela abordagem proposta.	58
Figura 17 – Exemplo de modelagem de redes $N^{r,t}$ (1 camada) com o método proposto a partir de uma textura em nível de cinza, utilizando 3 raios e 2 limiares, onde $t_1 < t_2$. Arestas de valor inferior ao limiar são removidas, portanto quanto maior t , mais esparsa será a rede $N^{r,t}$	59
Figura 18 – Processo de obtenção das redes W e B a partir da derivação do conjunto de arestas da rede N	60
Figura 19 – Representação visual (b, d) e distribuição da força (c, e) (grau ponderado) calculada para a rede de conexões intra-canal W e extra-canal B , usando diferentes limiares t e $r = 4$. A representação visual é gerada normalizando os valores da força nas camadas V_1, V_2 e V_3 para $[0,255]$, obtendo um pixel RGB. A distribuição das forças é calculada com uma discretização de $(\frac{(2r+1)^2-1}{2}z) + z$ bins (grau máximo possível + 1).	62
Figura 20 – Visão geral do funcionamento do Método proposto 1, desde uma textura de entrada até a obtenção do vetor final de características.	65
Figura 21 – Rede N direcionada modelada para uma textura com 3 canais de cor (RGB) (a). (b-d) Uma das camadas da rede direcionada N^r	68
Figura 22 – Exemplo de redes multicamada direcionadas W e B de conexões intra (a) e entre (b) canal, respectivamente.	69

Figura 23 – Representação visual (b, d) e (c, e) distribuição do grau de entrada, força de entrada e de saída calculados para a rede de conexões intra-canal W e extra-canal B , usando $r = 4$. A representação visual é gerada normalizando os valores da medida nas camadas V_1 , V_2 e V_3 para $[0,255]$, obtendo um pixel RGB. Quanto maior a medida nas três camadas, mais o pixel tende para o branco.	70
Figura 24 – Visão geral do funcionamento do método baseado em RC multicamada direcionada, desde uma textura de entrada até a obtenção do vetor final de características.	73
Figura 25 – Modelagem de imagens multiespectrais em redes direcionadas N^r de similaridade angular. Cada pixel da imagem é representado por um vértice, e as conexões são criadas apontando para o pixel de maior valor absoluto (soma de todos canais). O peso da conexão é definido pela similaridade do cosseno multiplicada pela distância Euclidiana dos pixels.	75
Figura 26 – Representação visual (b) da distribuição do grau de entrada, força de entrada e de saída (c-e) calculados para a rede direcionada de similaridade angular, usando $r = 4$. A imagem de entrada possui 32 canais espectrais e pixels de 16 bits, e a conversão para RGB (a) é feita calculando a média dos canais e normalizando os valores para 8 bits. A representação visual é gerada normalizando os valores de cada medida para $[0,255]$, onde o canal R é o grau de entrada, o canal G a força de entrada e o canal B a força de saída. Quanto maior o valor das 3 medidas, mais o pixel tende para o branco.	76
Figura 27 – Amostras de 20 classes de cada uma das 5 bases de textura colorida utilizada nos experimentos.	82
Figura 28 – Acurácia do Método proposto 1 em várias bases de textura utilizando diferentes valores de raio, individuais e combinados, e variando o número de limiares m . Os resultados são mostrados utilizando dois classificadores diferentes (LDA e KNN) e validação L1O.	83
Figura 29 – Performance média dos métodos propostos (RM e RMD) e de métodos da literatura m 5 bases de textura colorida, usando dois classificadores (KNN e LDA). A performance é calculada como a média entre os resultados de classificação obtidos nos esquemas L1O e 10-pastas.	95
Figura 30 – Desempenho médio dos métodos, definido pela média da acurácia obtida em cada um dos 4 espaços de cor considerados (RGB, LAB, HSV e $I_1I_2I_3$). São mostrados os 5 melhores métodos em cada base utilizando o classificador LDA e validação cruzada L1O.	101

Figura 31 – Amostras da estrutura foliar de espécimes de Jacaranda Caroba obtidas por microscopia por fluorescência. Cada espécime foi submetido à doses diárias de fluoreto de potássio, como indicado em cada imagem, por um período de 60 dias (o grupo de controle não recebeu dose alguma). As imagens, originalmente com 32 canais espectrais, são convertidas para o formato tradicional RGB.	103
Figura 32 – Taxa de acerto usando o método proposto de Redes complexas Direcionadas de Similaridade Angular (RDSA) na base de dados microscópicos do experimento com o espécime Jacaranda Caroba. Foram utilizados o classificador LDA e os esquemas de validação cruzada L1O (a) e 10-pastas (b).	104
Figura 33 – Acurácia dos métodos na caracterização de cada uma das 6 classes do problema de análise de plasticidade foliar do Jacaranda Caroba. Foram utilizados o classificador LDA e o esquema de validação cruzada L1O. .	106

LISTA DE TABELAS

- Tabela 1 – Equações de cada medida estatística usada para compor o vetor de características. Elas podem ser obtidas para cada uma das 3 redes $N^{r,t}$, $W^{r,t}$ ou $B^{r,t}$, e são calculadas para ambas as distribuições do grau $P(k)$ (μ_k , σ_k , λ_k , κ_k , e_k e ϵ_k) e da força $P(s)$ (μ_s , σ_s , λ_s , κ_s , e_s e ϵ_s). 63
- Tabela 2 – Acurácia do método de RM usando diferentes vetores de característica φ , concatenando raios e usando $m = 4$ limiares. Foram usados os classificadores LDA e o esquema de validação cruzada L1O. Em negrito são mostrados os melhores resultados para cada vetor φ 85
- Tabela 3 – Acurácia do método de RMD usando diferentes vetores de característica φ e concatenando raios de 2 à 6. Foram usados os classificadores LDA e o esquema de validação cruzada L1O. Em negrito são mostrados os melhores resultados para cada vetor φ 86
- Tabela 4 – Comparação de performance de diferentes métodos de análise de textura colorida nas bases Usptex e Vistex. São utilizados os classificadores KNN e LDA com validação cruzada L1O e 10-pastas, as imagens de cada bases estão representadas no formato RGB. Em negrito são mostrados os melhores resultados em cada esquema, assim como os dois métodos que atingiram os maiores resultados. 90
- Tabela 5 – Comparação de performance de diferentes métodos de análise de textura colorida nas bases Outex13, MBT e CURET. Em negrito são mostrados os melhores resultados em cada esquema, assim como os dois métodos que atingiram os maiores resultados no geral. 93
- Tabela 6 – Influência de diferentes espaços de cor na performance dos métodos nas bases de textura colorida Usptex e Vistex. Foram considerados o classificador LDA e o esquema de validação cruzada L1O. Células em negrito destacam os melhores resultados em cada espaço de cor e o melhor resultado médio, os dois métodos com maior desempenho médio também são destacados em negrito. 97
- Tabela 7 – Influência de diferentes espaços de cor na performance dos métodos nas bases de textura colorida Outex13, MBT e CURET. Foram considerados o classificador LDA e o esquema de validação cruzada L1O. Células em negrito destacam os melhores resultados em cada espaço de cor e o melhor resultado médio, os dois métodos com maior desempenho médio também são destacados em negrito. 99

Tabela 8 – Performance de diferentes métodos na classificação de folhas da espécie Jacaranda Caroba com diferentes níveis de poluição. É utilizado o classificador LDA e os esquemas de validação cruzada L1O e 10-pastas. 105

LISTA DE ABREVIATURAS E SIGLAS

VC	Visão Computacional
RC	Rede Complexa
RCs	Redes Complexas
RM	Rede Multicamada
RMD	Rede Multicamada Direcionada
RDSA	Rede Direcionada de Similaridade Angular
LBP	<i>Local Binary Patterns</i>
LPQ	<i>Local Phase Quantization</i>
CLBP	<i>Complete Local Binary Patterns</i>
nm	nanômetro
L	<i>Long</i>
M	<i>Medium</i>
S	<i>Short</i>
RGB	<i>Red Green Blue</i>
CIE	<i>Commission internationale de l'éclairage</i>
HSV	<i>Hue Saturation Value</i>
L1O	<i>Leave-one-out</i>
KNN	<i>k-nearest-neighbours</i>
LDA	<i>Linear discriminant analysis</i>
Vistex	<i>Vision Texture</i>
MBT	<i>Multi-Band Texture</i>
CUReT	<i>Columbia-Utrecht Reflectance and Texture</i>
CM	Caminho mínimo

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Contextualização e Motivação	23
1.2	Trabalhos Correlatos	25
1.3	Objetivos	27
1.4	Organização do Texto	27
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	Análise de Textura	29
2.1.1	Textura Colorida	30
2.1.1.1	Visão e Representação Colorida	30
2.1.1.2	Métodos	33
2.2	Redes Complexas	35
2.2.1	Medidas Topológicas	36
2.2.2	Modelos de Rede Complexa	37
2.3	Modelagem de Redes Complexas a partir de Imagens	39
2.3.1	Modelagem de Formas	40
2.3.2	Modelagem de Textura	41
2.3.3	Detecção de Pontos de Interesse	43
3	PROPOSTAS	47
3.1	Método Proposto 1: Modelagem de Imagens multiespectrais com Redes Multicamada (RM)	47
3.1.1	Análise de diferentes equações de peso	49
3.1.2	De uma rede regular para uma rede complexa	50
3.1.3	Seleção Automática de Limiares	54
3.1.4	Caracterização das Redes Complexas	58
3.1.5	Vetor de Características	63
3.1.6	Complexidade Computacional	64
3.2	Método Proposto 2: Rede Multicamada Direcionada (RMD)	66
3.2.1	Caracterização Estatística	67
3.2.2	Vetor de Características	71
3.2.3	Complexidade Computacional	72
3.3	Método Proposto 3: Análise de Imagens Multi-espectrais com Redes Direcionadas de Similaridade Angular (RDSA)	72
3.3.1	Caracterização Estatística das Redes e Vetor de Características	75
3.3.2	Complexidade Computacional	77

4	EXPERIMENTOS E RESULTADOS	79
4.1	Configuração Experimental	79
4.1.1	Bases de imagens	80
4.2	Análise dos Métodos Propostos	81
4.2.1	Análise do método Proposto 1: Rede Multicamada (RM)	81
4.2.1.1	Análise dos parâmetros R e T	81
4.2.1.2	Análise da caracterização estatística das redes	84
4.2.2	Análise do método proposto 2: Rede Multicamada Direcionada (RMD)	84
4.3	Comparação com Métodos da Literatura	86
4.3.1	Métodos da literatura	86
4.3.2	Comparação de Resultados	89
4.3.3	Influência de Diferentes Espaços de Cor	96
4.4	Aplicação no Reconhecimento de imagens Microscópicas	100
4.4.1	Plasticidade Foliar do Jacaranda Caroba	102
5	CONCLUSÃO	107
	REFERÊNCIAS	111

1 INTRODUÇÃO

1.1 Contextualização e Motivação

A visão é um dos sentidos mais importantes na vida dos mamíferos e outros animais, permitindo rápida identificação de ambientes, alimentos, presas e ameaças. Nesse contexto, os olhos tem sido um órgão definitivamente vantajoso em termos de seleção natural. Foram aproximadamente 500 milhões de anos de evolução desde pequenas cavidades capazes de detectar a direção da luz até os olhos modernos. O olho humano, por exemplo, é capaz de captar alguns milhões de diferentes cores resultantes da combinação dos comprimentos de onda do espectro visível. Além disso, conseguimos identificar finos detalhes, perceber profundidade e focalizar a visão em diferentes distâncias. Mas é muito importante ressaltar que toda essa informação, da forma como é captada pela retina, não é o suficiente para que de fato entendamos o que a imagem representa. O cérebro é o responsável por este processo, começando por terminações nervosas na própria retina que realizam o primeiro pré-processamento da luz captada. Em seguida, a informação viaja um longo caminho por neurônios e sinapses até o córtex visual no lobo occipital, na parte de trás da cabeça. Como um todo, nosso sistema de visão funciona como um *pipeline* com diversas etapas de detecção desde bordas, retas, formas até cores. No fim, entendemos os elementos presentes na imagem captada pelos olhos.

A ciência hoje compreende uma significativa parte dos processos da visão humana, o que tornou possível o estudo de formas de reproduzi-la artificialmente. Essa abordagem tem sido adotada na Visão Computacional (VC), área que busca desenvolver métodos de inteligência artificial para o reconhecimento de imagens. Porém ainda não existe um modelo capaz de realizar, paralelamente e em tempo real, todas as tarefas que inconscientemente realizamos quando olhamos para algo. Primeiro porque não se sabe exatamente o que acontece em etapas mais avançadas do processamento visual, e segundo pela limitação de *hardware* dos computadores atuais. Mesmo os computadores comerciais recentes alcançando taxas de processamento em torno de alguns *gigaflops* (10^9 operações de ponto flutuante por segundo), a quantidade de informação e os cálculos complexos necessários demandam por máquinas mais rápidas. Por esses e outros motivos, a VC procura focar na resolução de tarefas individuais da visão, como detecção de formas, objetos, textura e outros padrões. Os sistemas de VC mais comuns procuram explorar essas características em imagens para realizar tarefas clássicas de reconhecimento de padrões, como por exemplo classificação, recuperação, etc.

Dentre as diferentes abordagens exploradas na VC, cor e textura estão entre as características visuais mais expressivas.¹ O elemento textura é uma propriedade física abundante na natureza, e podemos discriminar materiais, animais e plantas apenas observando

(a) Texturas no contexto natural

(b) Texturas de imagens microscópicas

Figura 1 – Exemplos variados de texturas naturais e de materiais encontrados no dia-dia (a) e de texturas presentes em imagens microscópicas (b).

Fonte: Elaborada pelo autor.

seus aspectos texturais. Além disso, processos realizados pelo homem também produzem imagens com texturas de características visuais igualmente expressivas. Alguns exemplos de textura de materiais naturais e de imagens geradas por microscopia são mostrados na Figura 1. Apesar de sua relevância, não há uma definição formal para o termo textura que seja amplamente aceita pela comunidade científica. De uma forma simples podemos dizer que é uma combinação de aspectos físicos da superfície de objetos, como aspereza e cor. Em uma imagem digital estes elementos são evidenciados por padrões na mudança de intensidade dos pixels. O mundo físico impõe sobre essas mudanças de intensidade uma grande variedade de organizações espaciais, de forma bastante independente em diferentes escalas.² Isso implica que entender e medir estes padrões em imagens não é uma tarefa simples. Nesse contexto, a área de análise de textura se empenha há décadas para propor técnicas eficientes, variando entre modelos matemáticos ou bio inspirados.

Além da necessidade de entender e reproduzir artificialmente o processo de visão, as técnicas de VC possuem alta aplicabilidade em problemas reais. Como exemplo, algumas propriedades visuais podem ser subjetivas ou demandar muito esforço para serem analisadas a olho nu por uma pessoa, como distâncias geométricas, área ou volume. Entretanto, tais tarefas podem ser realizadas fácil e rapidamente por um sistema de VC. Em caracterização de textura, encontramos aplicações de VC em inspeção industrial,³ sensoriamento remoto,⁴ taxonomia vegetal,⁵ imageamento médico,⁶ etc.

Nos primórdios da análise de textura digital, por volta da década de 70, pesquisas pioneiras⁷ analisavam imagens em níveis de cinza, i.e. a luminância dos pixels. As primeiras técnicas consideravam estatísticas computadas sobre os níveis de cinza em um contexto

espacial, métodos que ainda são populares. Recentemente, modelos de análise de textura mais sofisticados tem surgido. Alguns exemplos são métodos baseados em geometria fractal⁸ e também Redes Complexas (RCs),⁹⁻¹² com resultados promissores. A maioria dos métodos de análise de textura encontrados na literatura não considera informações além dos níveis de cinza da imagem para caracterizá-la. Mas é válido ressaltar que a grande maioria das imagens capturadas por dispositivos de fotografia possuem 3 canais de cor, onde o formato mais utilizado é o RGB (do inglês vermelho, verde e azul). Também existem imagens multiespectrais capturadas em n canais, como por exemplo com a utilização de microscópios, satélites e telescópios. Porém, a maioria das técnicas de análise de textura normalmente descarta essa informação através de conversões para o nível de cinza, ou a processa separadamente com abordagens simples. Alguns trabalhos tem estudado os benefícios da adição da cor na análise de textura,¹³⁻¹⁶ porém pouca atenção se tem dado à relação espacial entre pixels coloridos.

Considerando a importância da cor na caracterização visual, esse trabalho estuda formas de incluir a informação espectral das imagens, de uma forma mais intrínseca, para a análise de textura. O foco principal é introduzir RCs como forma de modelar toda a informação espectral das imagens, captando sua relação em diferentes níveis espaciais e espectrais.

1.2 Trabalhos Correlatos

RCs tem atraído a atenção da comunidade de VC pelo fato de permitirem a modelagem de imagens. A partir desta proposta, varias possibilidades surgem, onde a solução de problemas de visão se tornam um problema de análise de redes. Por exemplo, alguns trabalhos recentes tem abordado segmentação em imagens como um problema de detecção de comunidades em redes.^{17,18} Em caracterização de texturas, a ideia geral é modelar a imagem como uma rede representando cada pixel como um vértice. Essa abordagem tem sido explorada em trabalhos anteriores.^{9-12,19,20} O que varia entre estes trabalhos são os parâmetros de modelagem e caracterização, como a função de definição das conexões, método de análise dos vértices, etc. Nestes trabalhos, as imagens consideradas contém apenas uma dimensão espectral, ou seja a luminosidade dos pixels (nível de cinza).

No trabalho de Gonçalves et. al.¹² é proposto o uso de RCs para modelar texturas dinâmicas (vídeos) conectando pixels de diferentes *frames* (quadros do vídeo). Neste caso, a rede é modelada sobre uma sequência de imagens em nível de cinza, i.e. 3 dimensões, duas espaciais e uma temporal. A caracterização é feita através de medidas estatísticas dos vértices considerando conexões em um mesmo *frame* ou entre *frames* subsequentes. Em um trabalho mais recente,²⁰ é explorado o conceito de difusão e caminhadas aleatórias em RCs, modeladas também a partir de imagens em nível de cinza.

Em um outro trabalho,²¹ é proposta uma técnica para modelar imagens coloridas,

com 3 canais de cor, como grafos. De uma forma parecida com o trabalho de Gonçalves et. al. em texturas dinâmicas, esta técnica modela um dado em 3 dimensões (imagem colorida), 2 espaciais e 1 espectral. Neste contexto, cada pixel é representado por 3 vértices diferentes, sendo um para cada canal de cor. Conexões espaço-espectral são formadas entre pixels em uma determinada região local, conectando vizinhos tanto de um mesmo canal de cor como de canais diferentes. Como só é considerada a distância entre os vértices para conectá-los, a rede resultante é chamada grafo por ser uma rede regular, onde todo vértice tem o mesmo número de conexões. Este grafo é analisado por medidas de caminho mínimo entre vértices, compondo um descritor de textura colorida da imagem. Apesar de ser similar as técnicas de RC, este método analisa uma rede regular, portanto métricas de RC não são aplicáveis. As formas mais comuns de análise de RCs consistem em medir padrões nas conexões dos vértices, o que não é possível quando a distribuição de conexões é constante.

RCs foram analisadas em aplicações envolvendo imagens multiespectrais nos trabalhos de Xia et. al. (2013)²² e Sui et. al. (2016).²³ Na primeira proposta,²² os autores propõem uma técnica de modelagem e caracterização de uma rede não ponderada com o intuito de selecionar bandas relevantes em imagens multiespectrais. Cada banda da imagem é convertida em um conjunto de dados de 1 dimensão (série ou sinal), onde pixels são mapeados em vértices e conectados se seus valores são maiores do que todos os valores de pixels entre eles. Essa técnica é conhecida como Grafo de Visibilidade Horizontal (HVG do inglês),²⁴ e é usualmente aplicada na modelagem de séries temporais. Para definir a relevância de uma banda, os autores propõem o uso do coeficiente de agrupamento (*clustering* do inglês). Este método, porém, descarta parte da informação de organização espacial dos pixels, considerando apenas sua organização horizontal. No método proposto por Sui et. al. (2016)²³ a proposta é similar ao trabalho anterior, onde cada banda da imagem é modelada como uma rede, porém a modelagem proposta é similar à modelagem encontrada em trabalhos de análise de textura. Isto é, a modelagem considera devidamente a informação espacial da imagem na construção da rede, conectando vértices/pixels se sua distância Euclidiana é menor que um limiar (raio), e o peso da conexão é a diferença absoluta de intensidade entre os pixels. Arestas com peso superior a um outro limiar de similaridade são removidas, mantendo apenas pixels semelhantes conectados. A densidade de arestas das redes de cada banda da imagem é calculada, que é a fração entre a quantidade de arestas da rede e a quantidade total de arestas que podem ocorrer. Quando a densidade é baixa, a rede é esparsa, portanto há uma tendência de que os vértices sejam mais heterogêneos. A partir desse critério, os autores propõem o uso da densidade para selecionar bandas relevantes. Uma outra proposta deste trabalho é combinar as redes de todas as bandas, ou apenas de bandas relevantes, para formar uma rede multibanda, onde o conjunto de vértices é o mesmo (pixels da imagem) e o conjunto de arestas de cada vértice é a concatenação de suas arestas em cada banda considerada. Nesta nova rede, o

grau dos vértices é usado para detectar regiões de interesse nas imagens multiespectrais.

1.3 Objetivos

O objetivo geral desta dissertação é explorar formas de modelar e caracterizar imagens multiespectrais. Para isso são exploradas técnicas baseadas em RC e no processamento oponente de cores, para realizar uma representação e caracterização mais rica da imagem. Foram estudadas técnicas de análise considerando a interação espacial e espectral dos pixels, ou seja os padrões locais entre pixels em um mesmo canal e em canais diferentes. Esse tipo de abordagem é pouco usado na literatura, onde a maioria dos descritores de imagens coloridas são integrativos, ou seja computa características locais de canais individuais, e então as concatena. A abordagem proposta visa considerar toda a informação espectral de forma mais intrínseca, diretamente conectadas da forma como realmente estão presentes na natureza. O uso de RC permite representações dessa informação nesse contexto, portanto propomos técnicas onde pixels de diferentes canais espectrais se conectam, formando uma única rede multi-camada. A estrutura dessa rede é então analisada topologicamente, de forma a obter medidas quantitativas sobre sua estrutura, que refletem as propriedades espaço-espectrais da imagem.

São propostos três novos métodos onde o primeiro, denominado Rede Multicamada (RM), foca na modelagem de imagens coloridas comuns de 3 canais como redes multicamada, e a caracterização é feita através de medidas topológicas extraídas de sua estrutura. O segundo método, Rede Multicamada Direcionada (RMD), segue a mesma ideia do primeiro, no entanto as arestas da rede possuem direção, apontando para a direção do gradiente entre pares de pixel. O terceiro método foca em imagens multiespectrais de muitos canais, denominado Rede Direcionada de Similaridade Angular (RDSA). Esse método modela uma rede de uma única camada, onde cada pixel é um vértice, e define arestas que apontam para o pixel de maior valor absoluto. O peso das arestas representa a similaridade do cosseno entre os pixels, que são considerados como vetores em um plano z -dimensional, onde z é a quantidade de canais da imagem. Para analisar os métodos propostos são realizados diversos experimentos de classificação de imagens em 5 bases de textura colorida da literatura e uma base multiespectral de microscopia por fluorescência, obtida no Instituto de Física de São Carlos - USP. Para validação das abordagens propostas, são comparados resultados com diversos métodos da literatura em diferentes cenários.

1.4 Organização do Texto

O texto dessa dissertação segue dividido em outros 4 capítulos. O Capítulo 2 apresenta a base literária da área de análise de textura, apontando conceitos e técnicas tradicionais assim como metodologias mais recentes. Também é introduzida uma apresentação e discussão a respeito de textura colorida e dos processos de representação de cor

tanto no sistema visual humano quanto computacionalmente, e métodos da literatura para diferentes formas de caracterização. Ainda nesse capítulo são introduzidos os conceitos de RC, suas definições matemáticas e os modelos estruturais de redes mais tradicionais na literatura. No Capítulo 3 são introduzidas as abordagens propostas nesse trabalho de mestrado, assim como uma análise sobre suas propriedades e funcionamento. Sua formulação matemática é devidamente apresentada e cada método é detalhadamente descrito. O Capítulo 4 apresenta, em detalhes, todos os resultados quantitativos de classificação obtidos nesse trabalho. São descritos os protocolos experimentais, as bases de dados utilizadas e também os métodos da literatura considerados para comparação de resultado. Por fim, o Capítulo 5 faz uma breve discussão do trabalho em geral e dos resultados obtidos.

2 FUNDAMENTAÇÃO TEÓRICA

Nesse capítulo são introduzidos os conceitos da literatura estudados durante o desenvolvimento desse trabalho de mestrado. Para abordar o problema de análise de imagens multiespectrais com redes, primeiramente é feita uma revisão da área de Análise de Textura, considerando também a questão da visão e representação das cores. Em seguida, são apresentados os conceitos da área de RCs, incluindo métricas e estratégias utilizadas no desenvolvimento das propostas desse trabalho.

2.1 Análise de Textura

Toda superfície natural possui uma propriedade física única que denominamos textura. Apesar disso, o termo textura é ambíguo, e não existe uma definição formal que seja amplamente aceita pela comunidade científica. Por exemplo, pode ser considerada a definição de Julesz,^{25,26} que assimila duas texturas se seu primeiro e segundo momento estatístico são semelhantes. De uma forma mais simples e abstrata, podemos definir textura como uma combinação de aspectos de cor e superfície que formam padrões espaciais, locais e globais. É possível discriminar entre diferentes materiais, animais e plantas apenas observando seus padrões texturais, o que torna a análise de textura uma habilidade essencial do sistema de visão de diversas formas de vida. Durante décadas, pesquisas tem explorado este aspecto visual para desenvolver métodos computacionais que executem a caracterização de texturas automaticamente. Diversas áreas constantemente necessitam de novos e sofisticados métodos para análise de textura, por exemplo inspeção industrial,³ sensoriamento remoto,⁴ biologia,^{5,27,28} medicina,⁶ nanotecnologia,²⁸ biometria,²⁹ etc.

Em imagens digitais, o padrão textural é evidenciado por variações no nível de intensidade dos pixels. Porém a natureza impõe sobre essas mudanças uma ampla variedade de organizações espaciais, em diferentes escalas.² Isso torna a tarefa de caracterização complexa, pois não existe uma métrica única que determine essas escalas, que dependem de diversos fatores. Em suma, o desafio das pesquisas de análise de textura é desenvolver formas de captar padrões texturais espaciais em diferentes escalas, levando em conta complexidade computacional e performance. Métodos encontrados na literatura usualmente exploram formas de combinar a análise de padrões locais, em diferentes escalas, e globais. Porém a forma como estas análises são feitas varia amplamente. Décadas de pesquisa resultaram em muitas diferentes abordagens propostas.³⁰ Os métodos mais difundidos são estatísticos, e exploram métricas baseadas na ocorrência/co-ocorrência de pixels com determinados níveis de cinza, em diferentes regiões locais. Como exemplos desse tipo de método, é possível citar as matrizes de co-ocorrência⁷ e os padrões locais binários (LBP do inglês).^{31,32} Esses métodos influenciaram técnicas futuras, baseadas nos mesmos princípios. Por exemplo, o

método de Quantização Local de Fase (LPQ, do inglês),³³ computa descritores locais da forma como o descritor LBP funciona, porém focando no problema do desfoque simétrico central na textura. Um outro método bastante utilizado, também inspirado no método LBP, é conhecido como Padrões Locais Binários Completos (CLBP, do inglês).³⁴ Esse método inclui informações não abrangidas pelo LBP original através da transformada de diferença de sinal de magnitude (LDSMT, do inglês).

Um outro tipo de método de análise de textura consiste em transformar a imagem para o espaço de frequência. Dois exemplos amplamente difundidos desse tipo de método são aqueles baseados em filtros de Gabor^{35–37} ou espectro de Fourier.³⁸ De uma forma similar, também existem métodos de transformada baseados na análise da complexidade da imagem usando medidas de dimensão fractal.^{8,39} Também considerando a complexidade da imagem, métodos baseados em RC tem sido estudados,^{9–12,19,20} onde uma imagem é modelada e analisada como uma rede de vértices. A topologia dessas redes pode ser quantificada com medidas de RC da literatura, compondo um descritor textural. Esta abordagem apresenta resultados promissores para caracterização de texturas em nível de cinza.

2.1.1 Textura Colorida

A maioria dos métodos de análise de textura considera apenas imagens em nível de cinza. A informação de cor é normalmente descartada ou processada separadamente com métodos não-espaciais de estatísticas de cor (histogramas, momentos estatísticos, etc). Porém, a grande maioria das imagens hoje deriva de dispositivos coloridos, como por exemplo da internet, câmeras pessoais, de vigilância, satélites, microscópios, e muito mais. O aumento do poder de processamento de computadores modernos também torna possível analisar maiores quantidades de dados, permitindo manter a informação de cor para análise de texturas. Autores então tem questionado¹⁴ se a inclusão da informação de cor é viável para análise de textura, argumentando que cor e textura devem ser tratadas separadamente. Porém deve ser levado em conta que poucas técnicas consideram toda a informação de cor de uma forma mais intrínseca, considerando aspectos espaço-espectrais. Existem diferentes conceitos sobre o processamento de cores, até mesmo biológicos, que devem ser levados em conta no processo de elaboração de um método de análise de textura colorida.

2.1.1.1 Visão e Representação Colorida

O conceito de visão colorida está relacionado à capacidade de um indivíduo de distinguir imagens de acordo com o comprimento de onda refletido ou emitido. Seres humanos estão limitados a uma fração do espectro, conhecido como espectro visível, como pode ser visto na Figura 2 (a). Um olho humano comum responde a comprimentos de onda em torno de aproximadamente 390 e 700 nanômetros (nm).⁴⁰ Por outro lado, existem

máquinas e até outros seres vivos capazes de captar outras regiões do espectro, como por exemplo o ultravioleta, abaixo de $380nm$. Em qualquer caso, o objetivo é o mesmo, captar e interpretar esses dados multispectrais. Se considerado o sistema visual humano, a captação da luz acontece a partir de células foto-receptoras na retina responsáveis por captar respostas em diferentes comprimentos de onda, dentro do espectro visível. As células conhecidas como cones são responsáveis por captar a informação da luz referente às cores, de forma tricromática, e estão concentradas na região da fóvea. Essas células são classificadas em três diferentes tipos, cada um responsável por uma faixa do espectro, sendo os cones do tipo L, M e S, do inglês *Long*, *Medium* e *Short*, respectivamente.⁴¹ Isso porque cones L respondem à luz de comprimento de onda longo, na faixa entre $564nm$ e $580nm$, referente à cor vermelha. Os cones M respondem na faixa de $534nm$ e $545nm$, que corresponde à cor verde, e os cones S estão relacionados à cor azul, na faixa $420nm$ e $440nm$. Todos os outros tipos de cores que percebemos são resultado da combinação das respostas dos cones à essas três cores primárias.

É importante ressaltar que o processo descrito até o momento apenas concerne ao sistema de captação da luz pela retina, que é a primeira etapa da visão humana. Os passos cruciais para que, de fato, compreendamos o que se vê estão concentrados no córtex visual. Existem mecanismos que tratam cores de forma complexa, combinando respostas das células da retina para extrair informações visuais. A teoria do processamento oponente de cor determina que o sistema visual humano combina as respostas de diferentes células foto-receptoras de uma maneira antagônica. Isso acontece pois o comprimento de onda de luz para o qual os três tipos de cones (L, M e S) respondem se sobrepõem. Portanto existe a hipótese de que o processo de interpretação funcione desta forma pois, para o sistema visual, é mais eficiente armazenar as diferenças entre as respostas dos cones ao invés da resposta da cada tipo individual de cone. A teoria sugere que existem dois canais oponentes de cor: vermelho versus verde e azul versus amarelo,⁴² da forma como mostra a Figura 2 (b). Em outras palavras, segundo essa teoria a evolução fez com que o sistema visual humano focasse em variações entre pares de cores ao invés de cores individuais, como forma de melhorar a visão colorida.

Em computação, a representação de imagens coloridas se dá por diferentes abordagens, denominadas espaço de cor. Os espaços de cor mais conhecidos podem ser classificados em quatro grupos principais,⁴³ denominados por:

- Espaços de cor primário, baseados no sistema tricromático da retina, como apresentado anteriormente. Esses métodos assumem que é possível representar qualquer cor através da mistura apropriada de quantidades das três cores primárias. O formato mais comum é denominado RGB, que vem do inglês *red*, vermelho, *green*, verde, e *blue*, azul. Esse formato é amplamente usado em várias situações devido a maioria dos dispositivos de aquisição de imagens funcionarem desta forma. Em outras palavras,

(a) Espectro eletromagnético e espectro visível.

(b) Processamento oponente de cor.

Figura 2 – Diferentes comprimentos de onda do espectro eletromagnético (a), e o respectivo pico de resposta dos cones L, M e S dentro do espectro visível. Esquema de processamento oponente de cor (b) a partir da diferença entre as respostas dos cones.

Fonte: Elaborada pelo autor.

uma imagem RGB é representada por uma matriz de três canais, sendo cada uma para R, G e B, respectivamente.

- Os espaços de cor de luminância-crominância, onde um canal representa a luminância e dois canais descrevem a crominância. Um formato desse tipo bastante utilizado é conhecido como LAB, ou $L^*a^*b^*$, ou ainda CIELAB, e é um espaço de cor definido pela Comissão Internacional de Iluminação (CIE), em 1976. Esse formato tem inspiração na teoria do processamento de cor oponente, pois o canal L representa a luminosidade, o canal A representa o componente vermelho versus verde, e o canal B representa o componente azul versus amarelo.
- Uma outra forma de representação de cores é conhecida como espaço de cor perceptual.

Um dos métodos mais conhecidos desse tipo é o HSV, do inglês *hue*, matiz, *saturation*, saturação, e *value*, valor, proposto em 1970. Nesse modelo as cores de cada matiz são organizadas em uma fatia radial, ou seja, diferente do espaço tridimensional dos modelos citados anteriormente, o formato HSV possui uma geometria cilíndrica. Nesse espaço a matiz representa a dimensão angular, variando do vermelho em 0° , passando pelo verde em 120° e então o azul em 240° . O eixo vertical representa V, que define o nível de luminosidade variando de 0 (preto), na parte inferior, até 1 (branco), na parte superior. A distância do centro até a borda do cilindro representa o valor de S, que define a saturação da cor, onde as cores puras se encontram no valor máximo de S, nas bordas do cilindro.

- Os espaços de cor baseados no conceito de eixos independentes, que são obtidos através de métodos estatísticos. Um exemplo desse tipo de método é o espaço de cor $I_1 I_2 I_3$, que consiste no cálculo de componentes com a menor correlação possível. Normalmente esse espaço de cor é obtido através de uma imagem RGB, convertida a partir de uma transformação linear. Nesse caso o canal I_1 representa a luminosidade ($I_1 = \frac{R+G+B}{3}$), o canal I_2 o oponente entre vermelho e azul ($I_2 = \frac{R-G}{2}$), e o canal I_3 o oponente entre verde, vermelho e azul ($I_3 = \frac{2G-R-B}{4}$).

Dada uma imagem representada por um determinado espaço de cor, um método de análise de textura colorida deve explorar formas de caracterizar a informação espectral presente. Existem diferentes métodos que abordam este problema, e alguns focam em problemas específicos. Em seguida, uma breve revisão sobre diferentes métodos de análise de textura colorida é apresentada.

2.1.1.2 Métodos

Existem diferentes formas de análise de textura colorida encontradas na literatura, e a maioria é integrativa ou paralela, que separam cor de textura. Métodos integrativos comumente computam descritores de nível de cinza tradicionais de cada canal de cor, separadamente. Nesse caso, qualquer método de análise de textura em nível de cinza pode ser aplicado, obtendo descritores de cada canal de cor que podem ser combinados. Um outro tipo de abordagem é denominada cor pura, ou *pure color* do inglês. Esse tipo de método considera apenas as cores da imagem, não levando em conta aspectos espaciais comumente analisados em imagens em nível de cinza. Nesse tipo de métodos, os mais difundidos são baseados em histogramas de cor.⁴⁴ Esta técnica computa uma sumarização compacta da distribuição de cores da imagem. Normalmente, isso é feito discretizando os valores de cores para um valor menor (normalmente de 255 para 32, 64, etc), e então computando a ocorrência de cores de diferentes formas. Como exemplo, histogramas conjuntos podem ser calculados combinando pares de canais de cor, ou também histogramas 3D que consideram os 3 canais de cor. Porém, esses métodos não consideram a interação espacial dos pixels,

Figura 3 – Esquemas da maioria dos métodos da literatura para caracterização de textura colorida. O modelo integrativo computa descritores tradicionais de textura para cada canal individual, e então os combina. O esquema paralelo computa descritores de cor pura não espaciais, e os combina com descritores obtidos da imagem em nível de cinza. Descritores de cor pura também são comumente utilizados individualmente.

Fonte: Adaptada de CERNADAS et. al.¹⁶

como é feito na análise de textura. Portanto, uma abordagem comum é combinar descritores de cor pura com descritores nível de cinza em métodos paralelos. O esquema geral das abordagens integrativa e paralela é mostrado na Figura 3. Nos trabalhos,^{14,16} o leitor pode consultar uma análise entre diferentes métodos integrativos, de cor pura e paralelos. Em alguns casos, métodos integrativos geram resultados similares a métodos tradicionais em nível de cinza, o que levou alguns autores a questionar sua aplicabilidade. Portanto acreditamos que esta não é a melhor forma de incorporar cor na análise de texturas. Cor e textura devem ser analisadas de uma forma mais intrínseca, considerando padrões espaço-espectrais, que é a relação espacial e de cor entre pixels próximos.

Considerando o sistema visual humano, a presença de cor em imagens acelera o reconhecimento e ajuda a memorizar coisas com mais facilidade.^{45,46} Pesquisas em psicofísica⁴⁷ mostraram que a percepção humana de cor depende da frequência com que a cor está espacialmente distribuída, ao invés de cores formando áreas uniformes. Também é possível considerar a teoria do processamento oponente de cor,⁴² que sugere que o sistema visual humano foca em variações entre pares de cores ao invés de cores individuais. Levando

isso em conta, e como a textura consiste em padrões de mudanças de intensidade, uma rica informação pode ser extraída analisando a interação espacial entre diferentes cores. Um dos primeiros trabalhos a considerar a interação intra e entre canais de cor para análise de textura foi proposto por Rosenfeld et. al.⁴⁸ Os autores exploraram a interação entre pares de canais de cor, e resultados indicam um aumento na informação textural obtida, que não está disponível na análise de canais de cores individuais. Em um trabalho de 1998⁴⁹ uma técnica baseada em bancos de filtros é introduzida para a caracterização de textura colorida. São obtidas características monocromáticas e de canais oponentes, calculadas a partir da saída de filtros de Gabor. Um outro trabalho a considerar este aspecto⁵⁰ propõem a segmentação de textura colorida abordando a interação de cores e sua distribuição espacial de forma inseparável, como elas são realmente percebidas durante a fase pré-atencional da visão humana. Canais oponentes também são abordados por Jacob et. al. (2014),⁵¹ onde é possível acessar características inter-relacionadas de padrões cromáticos de textura.

2.2 Redes Complexas

As pesquisas em RCs surgem da combinação entre a teoria dos grafos, física e estatística, com o objetivo de analisar grandes redes que derivam de processos naturais complexos. Inicialmente, trabalhos mostraram que existem padrões estruturais na maioria dessas redes, algo que não é esperado em uma rede aleatória.^{52,53} Isso levou a definição de modelos de RC que permitem entender propriedades estruturais de redes reais. Os modelos mais populares são os de redes livre de escala⁵⁴ e de pequeno mundo.⁵⁵ Portanto, uma nova porta foi aberta para o reconhecimento de padrões, onde RCs são adotadas como uma ferramenta de modelagem e caracterização de fenômenos naturais.

Os conceitos de RC são aplicadas em diversas áreas como física,^{56,57} biologia,^{58,59} nanotecnologia⁶⁰ neurociência,⁶¹ sociologia,⁶² entre outras.⁶³ Aplicar RCs em algum problema usualmente consiste em dois passos principais: i) a modelagem do problema como uma rede; e ii) a análise da estrutura da RC resultante. Quantificar a topologia da RC permite chegar a importantes conclusões relacionadas ao sistema que ela representa. Por exemplo, medidas locais de vértices podem evidenciar importantes regiões da rede, estimar sua vulnerabilidade, encontrar grupos de vértices similares, etc.

Uma rede pode ser definida matematicamente por $N = \{V, E\}$, onde $V = \{v_1, \dots, v_n\}$ é um conjunto de n vértices e $E = \{a(v_i, v_j)\}$ um conjunto de arestas (ou conexões). As arestas podem ser ponderadas, representando um valor que descreve o peso da conexão entre dois vértices, ou não ponderada, indicando apenas se a conexão existe. Uma rede também pode ser direcionada, implicando que suas arestas satisfazem $a(v_i, v_j) \neq a(v_j, v_i)$, ou não ter direção alguma, sendo $a(v_i, v_j) = a(v_j, v_i)$.

2.2.1 Medidas Topológicas

A topologia de uma RC é definida pelos padrões de suas conexões. Para quantificá-la, medidas podem ser extraídas tanto para vértices individuais, grupos de vértices ou globalmente para a rede toda. Uma das medidas mais comumente utilizadas é o grau dos vértices, que nada mais é do que a soma de conexões partindo de um determinado vértice. Considerando os conjuntos V e E , o grau k de cada vértice v_i pode ser calculado da seguinte forma:

$$k(v_i) = \sum_{\forall v_j \in V} \begin{cases} 1, & \text{se } a(v_i, v_j) \in E \\ 0, & \text{caso contrário} \end{cases} \quad (2.1)$$

Note que nesse caso o grau é calculado para redes não-ponderadas, pois o somatório considera apenas 1 caso haja a aresta, ou 0 caso ela não exista. Se a rede for ponderada, o grau pode ser calculado ponderado, métrica comumente conhecida por força. Portanto, basta apenas somar o peso de todas as arestas incidentes no vértice

$$s(v_i) = \sum_{\forall v_j \in V} \begin{cases} a, & \text{se } a(v_i, v_j) \in E \\ 0, & \text{caso contrário} \end{cases} \quad (2.2)$$

O grau e a força descrevem como um vértice interage com sua vizinhança. A distribuição do grau pode evidenciar importantes propriedades a respeito da estrutura da rede, como por exemplo a existência de *hubs* (vértices altamente conectados). Um outro tipo de medida que pode ser calculada localmente em uma rede é o coeficiente de agrupamento (*clustering coefficient* do inglês). Também é uma medida obtida para cada vértice, porém depende do comportamento de sua vizinhança. São consideradas as conexões do vértice em questão e as conexões entre seus vizinhos. Basicamente, essa medida conta o número de triângulos que ocorrem entre o vértice e seus vizinhos. Um triângulo em uma rede acontece sempre que existe uma tripla de vértice totalmente conectada. O coeficiente de agrupamento c de um vértice v_i pode ser calculado da seguinte forma:

$$c(v_i) = \frac{2 \nabla(v_i)}{k(v_i)(k(v_i) - 1)} \quad (2.3)$$

onde $\nabla(v_i)$ é o número de triângulos em que v_i participa, ou simplesmente o número de arestas entre os vizinhos de v_i . A equação é normalizada pelo número máximo de possíveis triângulos que v_i pode participar ($k(v_i)(k(v_i) - 1)$), ou seja $c(v_i) = 1$ quando todos os vizinhos de v_i estão conectados entre si. O coeficiente de agrupamento pode ser usado de várias formas, como para medir o nível de aglomeração da rede, evidenciando regiões de alta ou baixa densidade. As características de grau, força e agrupamento são conhecidas como medidas de centralidade, e foram utilizadas para descrever alguns modelos clássicos de redes. Em seguida, alguns desses modelos são introduzidos e detalhados.

Figura 4 – Topologia das redes complexas aleatórias de Erdős e Rényi. (a) Exemplo de uma rede aleatória construída com $m = 50$ e $p = 0,1$. (b) distribuição do grau para 100 redes geradas com as mesmas configurações de (a).

Fonte: Elaborada pelo autor.

2.2.2 Modelos de Rede Complexa

Dentro da literatura de RCs, existem 3 modelos topológicos clássicos para descrever tipos estruturais de rede. Nessa seção, esses modelos são explicados de forma mais detalhada. Um dos modelos mais básicos representa redes aleatórias, onde vértices são ligados aleatoriamente. Considere uma rede com m vértices e uma probabilidade de ligação de arestas p . Segundo o modelo de Erdős e Rényi,⁵² uma rede aleatória é construída partindo inicialmente de m vértices desconexos. Em seguida, cada vértice tem uma probabilidade p de se conectar a cada outro vértice da rede. Um exemplo de rede complexa aleatória com $m = 50$ e $p = 0,1$ pode ser visto na Figura 4 (a). Para entender a estrutura das redes aleatórias, o gráfico na Figura 4 (b) mostra o comportamento da distribuição do grau dos vértices, para 100 redes geradas com $m = 50$ e $p = 0,1$. Pode ser observado um comportamento em relação à distribuição do grau deste modelo de rede, que se encaixa em uma distribuição de *Poisson* do tipo $P(k) = \frac{(mp)^k e^{-mp}}{k!}$.

As redes de pequeno mundo, de Watts e Strogatz,⁵⁵ é um tipo de estrutura onde vértices podem ser alcançados através de poucas arestas. Em termos de estrutura, isso significa que o caminho mínimo médio da rede, no geral, é pequeno se comparado à outros tipos de rede. Outra propriedade encontrada em redes de pequeno mundo é o alto coeficiente de aglomeração (Equação 2.3), ou seja, a quantidade de triângulos que ocorrem entre os vértices. Para construir uma rede de pequeno mundo segundo o modelo de Watts e Strogatz, primeiro considere um conjunto de m vértices. Cada vértice será conectado aos seus i vizinhos mais próximos em cada direção (direita e esquerda), ou seja, cada vértice

(a) Rede de pequeno mundo, $m = 50$, $i = 2$ e $p = 0, 1$. (b) Distribuição do grau de 100 redes, em escala logarítmica.

Figura 5 – Topologia das redes complexas de pequeno mundo de Watts e Strogatz. (a) Exemplo de uma rede de pequeno mundo construída com $m = 50$, $i = 2$ e $p = 0, 1$. (b) distribuição do grau para 100 redes geradas com as mesmas configurações de (a).

Fonte: Elaborada pelo autor.

terá, inicialmente, $2i$ arestas. Para complementar a topologia da rede, uma probabilidade p é aplicada para que cada aresta seja religada. Portanto, uma aresta tem probabilidade p de ter seus extremos (vértices que originam a aresta) modificados. Para ilustrar o modelo da rede de pequeno mundo, a Figura 5 (a) mostra a estrutura resultante construída com $m = 50$, $i = 2$ e $p = 0, 1$. Novamente, para analisar a distribuição do grau dos vértices, 100 redes foram geradas e a ocorrência de cada grau foi medida, como pode ser visto na Figura 5 (b). A distribuição do grau nesse modelo de rede é parecida com a das redes aleatórias, porém com pico em $2i$. Apesar da similaridade na distribuição do grau, é possível notar as estruturas modulares que se formam na rede, isso é, os triângulos entre vértices próximos. Essa característica evidencia o fenômeno de pequeno mundo, onde a informação tende a fluir através da rede passando por poucos vértices.

Por último, o modelo de redes livre de escala, proposto por Barabási e Albert,⁵⁴ possui um padrão diferente em relação ao grau dos vértices. Neste modelo existem alguns vértices com o grau muito alto, enquanto a maioria dos vértices tem grau baixo. Essa característica evidencia a existência de *hubs* (ou centros), que são vértices conectados à uma grande parte do restante da rede. Esse modelo se baseia em duas regras: crescimento e ligação preferencial. Para isso, a construção da rede é feita de forma dinâmica, partindo inicialmente de um conjunto de m_0 vértices desconectados e seguindo uma sequência de iterações. A cada iteração um novo vértice é acrescentado à sua estrutura, e x novas arestas são adicionadas entre o novo vértice e os previamente existentes. Este processo é repetido até atingir um conjunto final de m vértices. Os vértices que receberão as novas arestas, a

Figura 6 – Topologia das redes complexas livres de escala de Barabási e Albert. (a) Exemplo de uma rede livre de escala construída com $m_0 = 2$, $m = 50$ e $x = 1$. (b) distribuição do grau para 100 redes geradas com as mesmas configurações de (a).

Fonte: Elaborada pelo autor.

cada iteração, são escolhidos de acordo com uma regra de preferência linear relacionada ao grau do respectivo vértice. Portanto, a probabilidade p_i de um vértice i receber uma conexão do novo vértice adicionado a rede é proporcional à $k(i)$

$$p_i = \frac{k(i)}{\sum_{\forall j \in V} k(j)} \quad (2.4)$$

ou seja, quanto maior o grau de j , mais chance ele terá de receber novas arestas. Um exemplo de rede livre de escala é mostrado na Figura 6 (a) juntamente com o comportamento da distribuição do grau para 100 redes geradas usando $m_0 = 2$, $m = 50$ e $x = 1$ (b). É possível observar que como 1 aresta é adicionada a cada iteração, muitos vértices possuem grau 1. Os vértices que tendem a se tornar *hubs* são os iniciais (m_0), pois sua probabilidade de receber arestas cresce a partir das primeiras arestas recebidas. Esse tipo de rede tende a ter uma distribuição do tipo lei de potência, ou seja $P(k) \approx k^{-\gamma}$, onde o expoente γ varia de acordo com sua estrutura. Diversos trabalhos tem demonstrado que é comum redes modeladas a partir de fenômenos reais terem expoente $\gamma \approx 3$.

2.3 Modelagem de Redes Complexas a partir de Imagens

Como RCs são estruturas flexíveis que permitem a análise de diversos fenômenos do mundo real, é possível empregá-las na modelagem de imagens. Nesse caso, torna-se possível transformar um problema de visão computacional em um problema de RCs, que

pode ser tratado de diferentes formas. A primeira etapa desta abordagem é a modelagem da rede a partir da imagem, ou seja a definição do que é vértice e o que é aresta. Essa definição é variável e normalmente depende do problema que se pretende tratar. Nesta seção serão revisadas técnicas de RCs que tratam de problemas como análise de formas, detecção de pontos de interesse e segmentação de textura.

2.3.1 Modelagem de Formas

No trabalho de Backes et. al.,⁶⁴ RCs são aplicadas para a modelagem de formas em imagens. Nesse caso, cada vértice é um pixel do contorno do objeto. Considere uma imagem I , contendo uma determinada forma com n pixels de contorno. Uma rede $N = \{V, E\}$ é obtida construindo o conjunto de vertices $V = \{v_1, \dots, v_n\}$. A cada vértice são associadas duas coordenadas espaciais $x(v)$ e $y(v)$ referentes à posição do pixel que este representa. O conjunto de arestas $E = \{a(v_i, v_j)\}$ é definido conectando todos os vértices i e j , e o peso das conexões corresponde à sua distância Euclidiana:

$$d(v_i, v_j) = \sqrt{(x(v_i) - x(v_j))^2 + (y(v_i) - y(v_j))^2} \quad (2.5)$$

que é então normalizada por seu maior valor ($\max(d)$).

A princípio, o conjunto de arestas $E = \{a(v_i, v_j)\} \forall i, j \in V$, ou seja todos os pixels de borda se conectam entre si. Isso resulta em uma rede regular, pois todos os vértices possuem o mesmo grau. Os autores então propõem um corte nas conexões de acordo com um dado limiar t , ou seja, conexões com peso maior que t são descartadas:

$$a(v_i, v_j) = \begin{cases} d(v_i, v_j), & \text{se } d(v_i, v_j) \leq t \\ \emptyset, & \text{caso contrário} \end{cases} \quad (2.6)$$

Assim, apenas pixels espacialmente próximos permanecem conectados. O funcionamento deste método é ilustrado na Figura 7. Experimentos⁶⁴ demonstraram que esse processo produz uma rede de pequeno mundo, que possui alto coeficiente de agrupamento e baixo caminho mínimo médio entre os vértices. A topologia dessa rede reflete as propriedades geométricas da borda que está sendo modelada, como pode ser observado na Figura 7.

Como o valor adotado para t afeta diretamente a rede resultante, diferentes valores são combinados, o que permite também analisar a dinâmica da rede. Neste contexto, conforme t aumenta, cresce a quantidade de conexões, e a rede evolui até que todos os vértices estejam totalmente conectados ($t = 1$). Para caracterizar a forma em questão, os autores propõem um conjunto de descritores composto pelo grau médio e grau máximo de redes modeladas com diferentes valores de t . Esta abordagem foi futuramente aprimorada,⁶⁵ onde medidas angulares da topologia da rede são usadas ao invés do grau dos vértices.

Figura 7 – Modelagem de formas com RCs, onde cada pixel do contorno do objeto é um vértice, e arestas são definidas conectando pixels próximos se sua distância Euclidiana é menor que um dado limiar t .

Fonte: SCABINI et al.⁶⁵

Esse método consiste em calcular a ocorrência de ângulos entre 0° e 180° formados por combinações de arestas adjacentes a um vértice. Resultados indicam⁶⁵ que o uso das medidas angulares é mais eficiente do que o uso do grau no caso de análise de formas.

2.3.2 Modelagem de Textura

Para a modelagem de textura com RCs, a técnica usualmente empregada consiste em considerar cada pixel da imagem como um vértice da rede. Considere uma imagem I em nível de cinza, com $w * h$ pixels, com níveis de intensidade $p(i)$ entre $[0, L]$. A constante L é o maior valor de intensidade possível nas imagens, e no caso de arquivos comuns de 8 bits, esse valor equivale a $(2^8) - 1 = 255$ (são 256 níveis possíveis, variando entre $[0, 255]$). Uma rede $N = \{V, E\}$ é obtida construindo o conjunto $V = \{v_1, \dots, v_{w*h}\}$, ou seja, cada pixel é mapeado como um vértice. Os primeiros trabalhos^{9,10} consideram a diferença absoluta de intensidade entre os pixels para definir o peso de sua conexão. É importante destacar que a diferença de intensidade não é afetada por mudanças na iluminação média da imagem.⁶⁶ Dado um raio r que define uma janela de limite espacial, uma nova rede N^r é obtida, em que cada vértice está conectado aos seus vizinhos dentro de um raio r

$$a(v_i, v_j) = \begin{cases} \frac{|p(v_i) - p(v_j)|}{L}, & \text{se } d(v_i, v_j) \leq r \\ \emptyset, & \text{caso contrário} \end{cases} \quad (2.7)$$

onde $\frac{|p(v_i) - p(v_j)|}{L}$ (normalização opcional) é a diferença absoluta normalizada entre os níveis de intensidade dos pixels representados pelos vértices v_i e v_j . Essa mesma abordagem de modelagem é também utilizada pelos métodos propostos em.^{12,20} A rede resultante contém

Figura 8 – Modelagem de redes N^r a partir de uma imagem em nível de cinza de tamanho $|I| = w * h$, onde $w = 10$ e $h = 8$. O parâmetro r limita a criação de conexões para uma região local de raio r , portanto quanto maior seu valor, mais conexões são criadas.

Fonte: Elaborada pelo autor.

arestas que respeitam uma vizinhança local de raio r , este processo é ilustrado na Figura 8.

O peso da conexão representa a similaridade entre dois pixels, de uma forma inversa. Isso é, quanto menor o valor da aresta, mais similares serão os pixels. Para incluir informação espacial ao peso das conexões, alguns trabalhos incluem d à equação.^{11,19} De qualquer forma, não existe um estudo aprofundado que constate que tipo de equação é mais eficiente. De acordo com os resultados encontrados na literatura e juntamente com os testes realizados durante este trabalho, supomos que incluir d à equação adiciona mais informação à modelagem, o que pode trazer vantagens. Desta forma, a variação da Equação 2.7 com um fator espacial adicional, para análise de textura, pode ser definida por

$$a(v_i, v_j) = \frac{d(v_i, v_j) + r \left(\frac{|p(v_i) - p(v_j)|}{L} \right)}{2r} \quad (2.8)$$

segundo,¹¹ ou ainda

$$a(v_i, v_j) = \left(\frac{|p(v_i) - p(v_j)|}{L} \right) \left(\frac{d(v_i, v_j)}{r} \right) \quad (2.9)$$

segundo.¹⁹

Com esta abordagem, o peso da conexão entre vértices respeita a distância entre os pixels. Por exemplo, imagine um pixel central de intensidade 20 conectado a outro pixel de intensidade 100, a uma distância 2. Agora imagine um outro pixel também de intensidade 100, porém a uma distância 3 do pixel central. De acordo com a Equação 2.7, ambas as conexões teriam o mesmo peso. Porém, utilizando as Equações 2.8 ou 2.9 o peso da conexão com o segundo pixel, mais distante, seria maior, dando relevância à sua organização espacial.

Os passos de modelagem descritos até então produzem uma rede N^r com escala proporcional a r , como pode ser observado na Figura 8. Essa rede não pode ser considerada uma RC devido à sua topologia regular, ou seja seus vértices possuem a mesma quantidade de conexões (desconsiderando bordas). Uma nova transformação é necessária para que padrões estruturais surjam na rede. É importante ressaltar que estes padrões devem evidenciar propriedades texturais, que é o foco da aplicação. Uma abordagem simples e utilizada em praticamente todos trabalhos consiste na aplicação de um limiar de corte de arestas. Arestas de valor relativamente alto são removidas da rede, mantendo apenas conexões entre pixels similares. Dado um limiar t , uma nova rede $N^{r,t}$ é obtida onde arestas com peso maior que t são removidas:

$$a(v_i, v_j) = \begin{cases} a(v_i, v_j), & \text{se } a(v_i, v_j) \leq t \\ \emptyset, & \text{caso contrário} \end{cases} \quad (2.10)$$

Dessa forma, a estrutura da rede reflete a informação de textura da imagem, isso é, os padrões de variação de intensidade dos pixels. A etapa de corte de arestas é ilustrada na Figura 9 de acordo com a textura de entrada (a), usando diferentes parâmetros r e t (b-g).

Como pode ser observado nas Figuras 8 e 9, os valores adotados para r e t impactam diretamente na estrutura da rede resultante $N^{r,t}$. Porém, esse fator permite a análise da evolução de uma rede esparsa para uma rede densa, conforme r e t crescem. Nesse contexto, trabalhos tem aplicado um conjunto de raios e um conjunto de limiares, combinados, para caracterizar a dinâmica das RCs em termos da textura que esta representa.^{11,12,19} Essa caracterização é feita a partir de medidas topológicas extraídas de redes $N^{r,t}$, que combinadas compõem um vetor de características que descreve a textura em questão.

2.3.3 Detecção de Pontos de Interesse

Conceitos de RCs podem ser considerados para a detecção de pontos de interesse em imagens, como bordas e regiões de variações de gradiente. Em⁶⁷ os autores propõem a modelagem de uma rede que conecta pixels dentro de uma região local definida por uma segmentação do tipo *watershed* *. Considere uma imagem $|I| = w * h$, onde w é a largura e h a altura e onde cada pixel possui intensidade $p(i) \in [0, L]$. A segmentação *watershed* encontra regiões $R(I) = \{r_1, \dots, r_l\}$, e uma rede $N = \{V, E\}$ é construída onde $V = \{v_1, \dots, v_l\}$ é o conjunto de vértices de tamanho $|V| \ll w * h$. Isso significa que, diferente da abordagem em textura, poucos pixels da imagem serão mapeados em vértices na rede, e essa definição depende das regiões encontradas pela segmentação *watershed*. O conjunto de arestas $E = \{a(v_i, v_j)\}$ é definido por conexões entre pixels de regiões adjacentes

* *Watershed* é uma transformação que considera uma imagem em nível de cinza como um mapa topográfico, com a intensidade de cada pixel representando a altura, e encontra as linhas que correm ao longo dos picos.

Figura 9 – Estrutura de redes $N^{r,t}$ modeladas a partir da textura de entrada (a). O padrão textural é refletido no padrão topológico da rede resultante. Quanto maior o raio r e o limiar t , mais densa será a rede $N^{r,t}$.

Fonte: Elaborada pelo autor.

$$a(v_i, v_j) = \begin{cases} |p(v_i) - p(v_j)|, & \text{se } r_i \text{ e } r_j \text{ são adjacentes} \\ \emptyset, & \text{caso contrário} \end{cases} \quad (2.11)$$

onde $|p(v_i) - p(v_j)|$ é a diferença absoluta entre o nível de intensidade dos pixels representados pelos vértices v_i e v_j . Para detectar os pontos de interesse, a força (grau ponderado, veja a Equação 2.2) dos vértices da rede resultante é calculada. Dessa forma, quando a força é grande, existe um alto gradiente em relação ao vértice e seus vizinhos. Essa medida então é usada para ranquear vértices e determinar os pontos de interesse da imagem.

Em⁶⁸ os conceitos de RCs são aplicados de forma semelhante. Nesse trabalho, o conjunto de vértices $V = \{v_1, \dots, v_{w \times h}\}$, ou seja, cada pixel é um vértice na rede $N = \{V, E\}$. Para definir o conjunto de conexões de cada vértice v_i , uma janela de raio r é centralizada no vértice para definir um limite espacial, e vizinhos com distância Euclidiana $d \leq r$ são conectados. O peso das conexões considerado é, novamente, a diferença absoluta de intensidade entre os pixels considerados. Para a detecção dos pontos de interesse, os autores propõem a análise do grau e das comunidades[†] da rede resultante. Em um trabalho de 2015, Wu et. al.⁶⁹ aplicam os conceitos de RCs para a detecção de bordas em imagens, de forma similar aos métodos de detecção de pontos de interesse. Nesse trabalho, da mesma forma que em,⁶⁸ as conexões são definidas dentro de uma janela de raio r . Porém, para a definição do peso das conexões, é usada uma combinação da diferença de nível de intensidade e da distância Euclidiana entre os pixels

$$a(v_i, v_j) = \begin{cases} \sqrt{(\alpha|p(v_i) - p(v_j)|)^2 + (\beta d)^2}, & \text{se } d(v_i, v_j) \leq r \\ \emptyset, & \text{caso contrário} \end{cases} \quad (2.12)$$

onde α e β representam um parâmetro de escalamento para a diferença de nível de cinza e para a distância, respectivamente. Em outras palavras, esses parâmetros são fatores que vão ponderar a importância da informação espectral e espacial no peso da conexão. Para a detecção de bordas, a dimensão local dos vértices (ou dimensão fractal) é considerada (o leitor pode consultar⁶⁹ para detalhes aprofundados).

[†] Comunidades são regiões de vértices que são altamente conectados entre si, e por outro lado possuem poucas conexões com os demais vértices da rede.

3 PROPOSTAS

A proposta principal desse trabalho de mestrado é abordar RCs para a modelagem e caracterização de imagens multiespectrais. É considerado também o mecanismo de processamento de cores oponentes do sistema visual humano, que da prioridade a regiões de variação de cores. A Figura 10 ilustra a ideia geral do trabalho onde três novas metodologias são propostas, sendo duas para a análise de imagens tradicionais coloridas e uma com o foco em imagens multiespectrais, de muitos canais. De forma geral o primeiro método proposto, denominado Rede Multicamada (RM), modela uma imagem colorida de 3 canais em uma rede multicamada, também de três canais, onde cada pixel é representado por 3 vértices e arestas conectam pixels no mesmo canal e entre canais diferentes. O segundo método, denominado Rede Multicamada Direcionada (RMD) segue o mesmo princípio do primeiro, porém as arestas da rede tem direções que apontam na direção do gradiente dos pixels. O terceiro método foca em imagens multiespectrais de muitos canais, e consiste na modelagem de uma Rede Direcionada de Similaridade Angular (RDSA), onde cada pixel multiespectral é considerado como um vetor, e o peso das conexões é a similaridade do cosseno entre pixels. Nas seções seguintes cada método proposto é descrito em detalhes, e suas propriedades são discutidas.

3.1 Método Proposto 1: Modelagem de Imagens multiespectrais com Redes Multicamada (RM)

Uma das propostas desse trabalho é estender a abordagem anterior de modelagem de imagens com RCs para incluir informação espectral, i.e. os canais de cor da imagem. Neste contexto, é introduzida uma abordagem multidimensional com redes multicamada. É possível construir uma RC considerando cada canal de cor como uma camada e cada pixel, em cada canal de cor, como um vértice. Considere uma imagem multiespectral I , com largura w , altura h e contendo z canais espectrais, portanto $w * h * z = |I|$ valores arbitrários. Uma rede multicamada $N = \{V, E\}$ é então obtida, onde $V = \{V_1, \dots, V_z\}$ representa o conjunto total de vértices, composto pela combinação de z camadas, e $V_i = \{v_1, \dots, v_{w*h}\}$ são os vértices pertencentes à camada i . Em outras palavras, cada pixel será representado por z vértices na rede N , e o número total de vértices é dado por $|V| = w * h * z$. Considerando imagens comuns que usualmente consistem de 3 canais de cor (como as RGB), um pixel será mapeado como 3 vértices na rede N , este processo é ilustrado na Figura 11.

O passo principal na modelagem de uma RC é a definição de suas arestas. Isso acontece pois uma RC é analisada em termos de sua estrutura topológica, que depende de como suas arestas estão organizadas. Primeiramente, pares de vértices que estejam a

Figura 10 – Diagrama geral das propostas deste trabalho para a caracterização de imagens multiespectrais, onde são introduzidas 3 novas abordagens.

Fonte: Elaborada pelo autor.

uma distância Euclidiana máxima r (raio, limiar de distância local) são conectados. A distância espacial entre pixels, no contexto multiespectral, é calculada considerando apenas suas coordenadas espaciais, x e y . Ou seja, a posição espacial de ambos os z vértices que representam um mesmo pixel é a mesma. Ao definir uma conexão entre um par de vértices v_i e v_j , considerando que $d(v_i, v_j) \leq r$, um valor é calculado para representar o peso desta aresta na rede. Como apresentado no Capítulo 2, Seção 2.3.2, trabalhos anteriores tem considerado diferentes abordagens na definição do peso das arestas. Neste contexto, discutiremos em seguida a influência de diferentes equações de peso na estrutura da rede multicamada.

Figura 11 – Primeira etapa da modelagem de imagens coloridas (no exemplo RGB) em redes multicamada. Um pixel RGB é representado por 3 vértices, sendo 1 para cada canal de cor, e combinados compõem o conjunto de camadas $V = \{V_R, V_G, V_B\}$.

Fonte: Elaborada pelo autor.

3.1.1 Análise de diferentes equações de peso

Teoricamente, a equação de peso deve representar devidamente a relação textural e espacial entre dois pixels. Diferentes equações propostas na literatura adotam a diferença absoluta de intensidade entre os pixels e/ou sua distância Euclidiana. Em^{9,10,12} (Equação 2.7), apenas a diferença absoluta de intensidade entre pixels é usada na definição do peso das conexões. Em⁶⁴ (Equação 2.8), para análise de formas, apenas a distância Euclidiana entre os pixels é usada como peso das conexões, porém essa técnica não é aplicável para análise de texturas uma vez que apenas a posição dos pixels não fornece informação de seus padrões texturais. Em^{11,19} é incluída a informação espacial combinada à informação de diferença de intensidade.

Primeiramente, a Equação 2.7 é analisada considerando a distribuição de arestas produzida dada uma determinada textura de entrada e um raio r , os resultados são mostrados na Figura 12b para $r = 1, 2, 3$. É possível observar que, independente dos valores de raio testados, a maioria das arestas tem valores menores que 0.2. Isso é esperado pois os locais onde há variações bruscas de intensidade são espacialmente pequenos se comparados ao tamanho total da imagem, enquanto que regiões de intensidade homogênea são mais frequentes. Além disso, a região local limitada por r é relativamente pequena em relação ao tamanho total da imagem, e pixels a pequenas distâncias tendem a ter intensidades similares. Isso também é evidenciado pela pequena variação na curva do gráfico conforme o raio r é aumentado, aumentando a frequência de arestas com valor em torno de 0.4. Apesar de representar devidamente a informação de variação de intensidade

dos pixels, essa equação não inclui a informação sobre sua organização espacial.

A equação 2.8 proposta em¹¹ inclui a informação espacial dos pixels durante o cálculo do peso de sua conexão. A proposta é adicionar a distância Euclidiana juntamente com a diferença absoluta de intensidade dos pixels, normalizadas. Porém há alguns problemas na equação, que não pondera devidamente os valores considerados. A diferença de intensidade é normalizada para o intervalo $[0, r]$, e quando o raio r utilizado é pequeno, há poucas distâncias distintas possíveis entre os pixels. Nesse caso, a distribuição de peso das arestas tende a formar picos de acordo com a quantidade de valores possíveis para d , como pode ser observado na Figura 12c. Conforme se aumenta o raio r , a quantidade de picos aumenta, portanto a equação depende de um raio relativamente grande para produzir uma distribuição mais uniforme. Em outras palavras, a equação não é robusta a mudanças de raio. A Equação 2.9, proposta em¹⁹ não apresenta este problema pois os fatores normalizados de distância e diferença de intensidade são multiplicados diretamente. O problema com essa abordagem acontece quando algum dos lados da multiplicação resultam em 0, o que cancela o outro lado. Por exemplo, pixels com o mesmo nível de intensidade vão ter conexão com peso sempre 0, independente de sua distância. Da mesma forma, pixels com distância 0 causam o mesmo problema, que correspondem a um mesmo pixel conectado com seu equivalente em outro canal espectral da imagem.

Considerando os problemas citados anteriormente, uma nova equação de peso é proposta. A ideia é incluir ambas as informações espaciais e espectrais dos pixels, ambas com o mesmo peso e sem ser fortemente influenciada pelo parâmetro r . A abordagem proposta ajusta a Equação 2.9 para gerar valores entre $[0, 1]$, sem que algum dos lados da Equação anule o valor do outro. Dessa forma, a definição da conexão entre dois pixels representados pelos vértices v_i e v_j é definida por

$$a(v_i, v_j) = \begin{cases} \frac{(|p(v_i) - p(v_j)| + 1)(d + 1) - 1}{(L + 1)(r + 1) - 1}, & \text{se } d(v_i, v_j) \leq r \\ \emptyset, & \text{caso contrário} \end{cases} \quad (3.1)$$

onde $p(v_i)$ retorna o valor de intensidade do pixel representado pelo vértice v_i em seu canal correspondente, de acordo com a camada que v_i pertence. A Figura 12d mostra a distribuição do peso das arestas produzida usando a equação proposta. É possível observar que a curva segue o mesmo padrão da curva produzida pela Equação 2.7, que respeita a organização das intensidades dos pixels, porém agora contém também a informação sobre sua organização espacial.

3.1.2 De uma rede regular para uma rede complexa

O processo de modelagem descrito até o momento mapeia cada pixel em z vértices, isso é, um vértice para cada canal espectral. A definição das conexões gera o conjunto de arestas E , que é composto por conexões entre vértices pertencentes a qualquer camada.

(a) Textura de entrada

(b) Distribuição de pesos usando a equação segundo ^{9,10,12}

(c) Distribuição de pesos usando a equação segundo ¹¹

(d) Distribuição de pesos usando a equação proposta (ajuste da equação de ¹⁹)

Figura 12 – Distribuição dos pesos das arestas de acordo com diferentes equações de peso, calculada para a textura de entrada com 3 diferentes valores de raio (r). A frequência é a função densidade de probabilidade calculada usando uma discretização de 100 *bins*.

Fonte: Elaborada pelo autor.

Isso ocorre pois a janela de raio r cobre os vizinhos do vértice em todas as camadas, como se estivessem sobrepostas, uma vez que consideramos apenas as coordenadas espaciais x e y para calcular sua distância Euclidiana. Esse processo resulta em uma rede N^r , escalada em função de r , que é ilustrada na Figura 13a. É possível observar que acontece um efeito de borda, onde vértices próximos dos limites da imagem tendem a ter menos conexões dando uma forma "quadrada" para a rede. Isso prejudica sua estrutura, o que pode afetar o desempenho na caracterização da textura. Para contornar esse efeito, somente são considerados vértices que estejam a uma distância de pelo menos r da borda, que chamamos de vértices válidos. Eles podem se conectar a vértices não válidos, porém as informações dos vértices não válidos não são usadas na análise posterior da rede. Nas Figuras 13b e 13d é possível observar a estrutura da rede quando apenas os vértices válidos são considerados.

A rede N^r obtida até o momento apresenta uma topologia regular, mesmo removendo o efeito de borda. Isso porque todos os vértices possuem o mesmo grau, ou seja a mesma quantidade de arestas. Uma topologia regular não pode ser analisada em termos de redes complexas, uma vez que busca-se padrões estruturais. A rede então é transformada removendo um conjunto de arestas para que sua nova topologia evidencie padrões relacionados à textura da imagem que ela representa. Trabalhos anteriores^{9,11,19,20} adotam um limiar de corte t , e arestas com valor superior a t são removidas. Porém, propomos uma forma diferente de corte, onde arestas de valor inferior ao limiar são removidas. Considerando a Equação 3.1, valores baixos significam alta similaridade entre pixels, tanto espectral como espacial. Nesse contexto, como a textura consiste nos padrões de variações da imagem, manter vértices distintos conectados foca especificamente nesses padrões. Portanto, dado um limiar t , arestas de peso inferior a t são descartadas, mantendo apenas arestas de valor relevante

$$a(v_i, v_j) = \begin{cases} a(v_i, v_j), & \text{se } a(v_i, v_j) > t \\ \emptyset, & \text{caso contrário} \end{cases} \quad (3.2)$$

As conexões da rede resultante da combinação de r e t ($N^{r,t}$) representam regiões onde há variações de intensidade relevante, dependendo do valor adotado para t . Quanto menor t , mais arestas são consideradas, e a rede resultante será mais densa, enquanto altos valores de t implicam em uma rede mais esparsa. Da mesma forma, o raio de distância r limita a região local em que os padrões de variação de intensidade são analisados. A Figura 14 ilustra a etapa de remoção de arestas com diferentes limiares t .

Os valores adotados para r e t influenciam na topologia da rede resultante, conforme discutido na Seção 2.3.2. A influência de cada parâmetro permite realizar combinações distintas para obter redes correlatas, onde apenas o conjunto de arestas muda conforme r e t crescem. Desta forma, é possível analisar a dinâmica da rede através de medições com-

(a) Rede N^r com efeito de borda

(b) Rede N^r com apenas vértices válidos

(c) Efeito de borda em uma única camada de N^r

(d) Vértices válidos de uma camada de N^r

Figura 13 – Processo de definição de arestas em uma rede multicamada para uma imagem RGB. Os vértices entre todas as camadas são conectados dentro de um raio de distância r . Apenas vértices válidos são considerados para evitar o efeito de borda.

Fonte: Elaborada pelo autor.

Figura 14 – Etapa de remoção de arestas com limiares $t_1 < t_2 < t_3$. Quanto maior t , menos arestas são mantidas e a rede se torna mais esparsa.

Fonte: Elaborada pelo autor.

binando um conjunto de raios $R = \{r_1, \dots, r_n\}$ e um conjunto de limiares $T = \{t_1, \dots, t_m\}$. O conjunto de raios é composto por valores que cobrem uma região aproximadamente circular em torno do pixel central, isso é $R = \{1, 2, 3, 4, \dots\}$. Por outro lado, a definição do conjunto de limiares não é tão intuitiva quanto a dos raios. Os valores de t variam no intervalo $[0, 1]$, porém os casos $t = 0$ ou $t = 1$ produzem redes regulares, que não são úteis na caracterização. Além disso, como é possível observar na Figura 12, a grande maioria dos valores de arestas aparecem abaixo de 0,4, o que também pode variar de textura para textura. Nesse contexto, a próxima seção apresenta uma forma de determinar o intervalo de valores de t em diferentes casos.

3.1.3 Seleção Automática de Limiares

Para definir o conjunto de limiares para modelagem de RC a partir de imagens, trabalhos anteriores propõem testes extensivos de diversos valores. Nesse caso, 3 parâmetros são testados: um limiar inicial t_1 , um fator incremental e um limiar final t_m . Porém a

quantidade de combinações possíveis entre os três parâmetros é relativamente grande, e testar mesmo que uma fração dessas combinações é computacionalmente custoso. Além disso, um determinado conjunto de limiares pode se encaixar melhor para um determinado conjunto de imagens. Isso acontece pois a distribuição do peso das arestas varia para imagens diferentes, como é mostrado na Figura 15 (c), (e) e (g). Para contornar este problema, uma nova técnica é proposta para a definição automática do conjunto de limiares. Para isso, consideramos a distribuição das arestas e o grau médio das redes. Dada uma rede $N^{r,t}$, o grau médio (não-direcionado) é calculado por

$$\mu_k(N^{r,t}) = \frac{2|E|}{|V|} \quad (3.3)$$

onde E é o conjunto de arestas e V o conjunto de vértices de N .

No contexto de redes, o limiar inicial t_1 ideal deve descartar os casos extremos de redes densas, como é o caso de $t \approx 0$ que permite que todos os vértices se conectem com todos seus vizinhos dentro do raio r . Isso pode ser observado na Figura 15 (d), (f) e (h), que mostram o grau médio das redes $N^{r,t}$ modeladas para as imagens (a) e (b), usando o limiar indicado no eixo x . Estas curvas indicam que o raio r tem pouca influência no padrão geral do grau médio da rede. Quando t é muito pequeno, o grau médio de ambas as redes se aproxima do máximo, e suas topologias se tornam regulares e similares. É possível observar que o pico da distribuição se encontra entre $[0, 0.1]$, região que produz redes com o grau médio alto. Portanto, como durante a modelagem a ideia é descartar arestas de peso baixo para destacar regiões de variações de intensidade, podemos considerar essa região para definir o primeiro limiar t_1 a ser utilizado. Para isso, considere $P(a)$ como a função da distribuição das arestas (frequência). O limiar inicial t_1 é então definido por

$$t_1 = \arg \max_{a \in [0,1]} P(a) \quad (3.4)$$

onde a função $\arg \max$ retorna um índice entre $[0, 1]$ que produz o máximo da função $P(a)$, isso é, seu pico. Desta forma, o primeiro limiar irá produzir redes que contém apenas arestas com peso acima do valor de pico da distribuição, contornando casos de redes densas.

Da mesma forma que valores de limiar próximos de 0 produzem redes muito densas que devem ser evitadas, valores próximos de 1 produzem redes muito esparsas. Quando $t = 1$, a rede produzida será totalmente desconectada. Seguindo o mesmo conceito usado para definir t_1 , a definição do limiar final t_m consiste em contornar os casos extremos próximos à cauda da distribuição $P(a)$, removendo redes que possuam grau médio inferior a 1. Considerando o grau médio k_μ em função do limiar t , t_m é definido por

$$t_m = \arg \max_{t \in [0,1]} \delta(\mu_k(t), 1) \quad (3.5)$$

Figura 15 – Diferenças na distribuição das arestas (c, e, g) de duas imagens diferentes (a e b) usando 3 raios diferentes e discretização de 100 intervalos. Quando a rede $N^{r,t}$ é modelada usando $t = a$, o padrão do grau médio μ_k também varia de acordo com a imagem, com pouca variação entre raios (d, f, h).

Fonte: Elaborada pelo autor.

onde δ é o delta de Kronecker e retorna 1 quando $\mu_k(t) = 1$, e 0 caso contrário. Neste caso, a função $\arg \max_{t=0}^1$ retorna o valor de t onde a função do grau médio atinge 1, definindo o limite final de corte t_m .

Os limites t_1 e t_m se adaptam às distribuições de arestas de diferentes imagens. Por exemplo, é possível observar pela Figura 15 (c-d) que, para cada imagem, são diferentes ambos os picos de frequência de arestas e o momento em que o grau médio da rede resultante atinge 1. Usando as definições apresentadas, para uma determinada imagem é possível construir o conjunto de limiares adaptativo $T = \{t_1, \dots, t_m\}$ dado um valor m . Porém, em problemas comuns de reconhecimento de padrões visuais, um conjunto de imagens previamente conhecido é usado para treinar um classificador. Neste contexto, propomos o uso de todas as imagens do conjunto de treinamento para a definição de $[t_1, t_m]$, generalizando para futuras imagens similares que venham a ser analisadas. Portanto, a função $P(a)$ é construída considerando a ocorrência de arestas em todas as imagens do conjunto de treinamento, e t_1 é obtido pela Equação 3.4. Da mesma forma, a função do grau médio pode ser obtida para o conjunto de imagens de treinamento considerando a média de grau de todas as imagens para um determinado limiar t . Nessa função, o limiar final é então definido da mesma forma pela Equação 3.5. Essa abordagem é mostrada na Figura 16 utilizando 4 bases de imagens diferentes (detalhes de cada base no Capítulo 4, Seção 4.1.1). É possível observar que em cada base a frequência das arestas $P(a)$ varia, refletindo suas propriedades individuais de textura. Por exemplo, em imagens com muitas variações de intensidade em todos os canais de cor, como é o caso da base MBT, a função tende a se estender, apresentando mais ocorrência de arestas de peso além de 0.5, por exemplo. Da mesma forma, a função do grau médio também varia, seguindo o padrão da função $P(a)$. Portanto, o momento em que a função do grau médio atinge 1 irá variar dependendo dos padrões do conjunto de imagens em análise. De acordo com esses resultados, é notável que usar um intervalo de limiares fixo não é a melhor abordagem para se obter melhores resultados em diferentes casos.

Uma vez que o intervalo $[t_1, t_m]$ é obtido, o conjunto de limiares $T = \{t_1, \dots, t_m\}$ é construído dividindo o intervalo em m valores equidistantes. Portanto, os únicos parâmetros necessários para modelar RCs com o método proposto são o conjunto R e a quantidade de limiares desejada m para compor T . Combinando diferentes valores de raio e limiar, redes distintas podem ser obtidas para uma mesma imagem. Porém, diferente do que é mostrado na Figura 9 que ilustra o funcionamento de métodos anteriores, o método proposto foca nas variações de intensidade da imagem, removendo arestas de peso baixo. Desta forma, os padrões topológicos das redes resultantes são diferentes, enaltecendo regiões de variação de intensidade. Para ilustrar essa diferença a Figura 17 mostra a estrutura da rede $N^{r,t}$ modelada para uma textura em nível de cinza, incrementando-se r e t . Combinando diversas redes $N^{r,t}$ é possível realizar a análise dinâmica de sua topologia conforme esta evolui com o incremento de r e t , como apresenta o Capítulo 2, Seção 2.3.2.

Figura 16 – Frequência de arestas e grau médio em 4 bases de textura colorida. Em azul os limites $[t_1, t_m]$ estimados pela abordagem proposta.

Fonte: Elaborada pelo autor.

Nesse contexto, a próxima seção explora esse conceito para introduzir formas de analisar a estrutura das redes quantitativamente.

3.1.4 Caracterização das Redes Complexas

A rede $N^{r,t}$ obtida até então contém informações importantes a respeito da textura representada. De acordo com sua estrutura multi-camada, há vários meios de se analisar sua topologia. Considere a rede N e seu conjunto de vértices $V^N = \{V_1, \dots, V_z\}$, representando vértices de cada uma de suas z camadas, e seu conjunto de arestas E^N . Propomos então uma derivação da estrutura de N em 2 sub-redes W e B , onde $E^N = E^W \cup E^B$ e $V^N = V^W = V^B = V$. Em outras palavras, W e B possuem o mesmo conjunto de vértices de N , porém seu conjunto de arestas é uma derivação do conjunto E^N . Primeiramente, para a construção da rede W são consideradas apenas arestas que conectam vértices pertencentes à mesma camada. De forma contrária, somente arestas entre camadas são consideradas na construção da rede B . Portanto, dada uma rede $N^{r,t} = \{V, E^N\}$, o conjunto de arestas das novas redes $W^{r,t} = \{V, E^W\}$ e $B^{r,t} = \{V, E^B\}$ são obtidos selecionando arestas de N

$$E^W = \{a(v_i, v_j) \in E^N \Rightarrow V(v_i) = V(v_j)\} \quad (3.6)$$

Figura 17 – Exemplo de modelagem de redes $N^{r,t}$ (1 camada) com o método proposto a partir de uma textura em nível de cinza, utilizando 3 raios e 2 limiares, onde $t_1 < t_2$. Arestas de valor inferior ao limiar são removidas, portanto quanto maior t , mais esparsa será a rede $N^{r,t}$.

Fonte: Elaborada pelo autor.

(a) Rede original $N^{r,t}$ (b) Rede de conexões intra-canal $W^{r,t}$ (c) Rede de conexões extra-canal $B^{r,t}$

Figura 18 – Processo de obtenção das redes W e B a partir da derivação do conjunto de arestas da rede N .

Fonte: Elaborada pelo autor.

$$E^B = \{a(v_i, v_j) \in E^N \Rightarrow V(v_i) \neq V(v_j)\} \quad (3.7)$$

onde $V(v_i)$ retorna a camada a qual pertence o vértice v_i . Portanto as novas redes W e B apenas contém, respectivamente, arestas intra e extra canal/camada. As redes W e B são ilustradas na Figura 18, onde N é modelada a partir de uma imagem RGB (3 canais de cor). A rede W contém um padrão topológico relacionado ao comportamento local das intensidades de um pixel em um determinado canal. Por exemplo, há casos em que a cor geral da imagem muda devido a variações em algum canal, porém a intensidade em outros canais pode permanecer constante. Já a rede B capta variações locais que acontecem entre um canal e outro, ou seja a relação de variância espectral entre canais. De certa forma este processo lembra a teoria da cor oponente, do sistema visual humano,⁴² que alega que células específicas no cérebro processam variações entre pares de cor.

Dada uma rede N , W ou B , sua estrutura pode ser analisada de diversas formas. A literatura de redes complexas é abrangente, e diferentes tipos de medidas topológicas podem ser calculadas. Em,⁷⁰ os autores apresentam uma revisão detalhada de diferentes métricas e como podem ser aplicadas. Um problema comum com a maioria das medidas

encontradas na literatura é seu custo computacional. Medidas comuns como por exemplo proximidade e intermediação (do inglês *closeness* e *betweenness*) possuem complexidade $O(|I| * |E|)$, pois analisam a centralidade de um vértice em relação a todos os outros vértices da rede. No contexto de análise de textura multiespectral, essas medidas são custosas pois $|I| = |V| = w * h * z$, e normalmente $|E| \gg |V|$, portanto a complexidade tende a ser maior que $O(|I|^2)$. Portanto, propomos o uso de métricas de baixo custo computacional. Duas medidas que podem ser calculadas paralelamente à modelagem das redes, portanto sem custo adicional, são o grau e a força dos vértices (Equações 2.1 e 2.2), que medem a centralidade de um vértice em relação aos seus vizinhos adjacentes. A combinação das duas medidas proporciona uma melhor caracterização de um vértice do que o uso de cada uma individualmente. Por exemplo, vértices de mesmo grau podem ter forças diferentes, da mesma forma que quando as forças são iguais, o grau também pode ser diferente. Em termos de textura esse tipo de análise é útil para medir o padrão textural local, e a combinação de diferentes raios r permite uma análise multi-escala. A força calculada para as redes W e B é mostrada de diferentes formas na Figura 19, usando $r = 4$ e diferentes valores de t . Uma representação visual das forças é construída convertendo seus valores nas camadas V_1, V_2 e V_3 para um pixel RGB, onde altas medidas significam regiões mais claras. É possível observar como a força de cada rede destaca padrões diferentes da imagem de entrada, e como isso varia conforme t é aumentado. Por exemplo, a rede W realça bordas em diferentes regiões, porém conforme t cresce, apenas bordas entre as folhas e o fundo permanecem. Já na rede B é possível observar o realçamento das regiões internas das folhas, relacionadas à estruturas foliares de cor mais predominante. Conforme se refinam as arestas de B aumentando t , se destacam as estruturas centrais das folhas, de cor branca, e a cor verde representa a variação do branco para o verde.

Calculando o grau e a força de todos os vértices de uma rede (N , W ou B) é possível obter suas distribuições ($P(k)$ e $P(s)$) calculando a ocorrência de cada valor. Primeiramente, é necessária uma discretização dos valores para que a ocorrência seja contabilizada em uma quantidade inteira de *bins*. Para isso, consideramos o grau máximo que um vértice pode atingir de acordo com o raio r usado mais 1 (considerando o grau 0), dado por $k_{max}^r = (\frac{(2r+1)^2-1}{2}z) + z$, onde z é a quantidade de canais da imagem. É possível observar que o valor máximo da força também será k_{max}^r , pois como o valor máximo de uma aresta é 1, seu somatório será k_{max}^r . Analisando $P(k)$ e $P(s)$ é possível obter informações importantes sobre a estrutura topológica da rede. As diferenças na distribuição da força de W e B podem ser observadas na Figura 19 (c, e). Ambas redes tem maior ocorrência de vértices poucos conectados, porém B possui mais vértices próximos ao valor máximo (1) do que W . Ambas demonstram um comportamento do tipo lei de potência, como redes livre de escala.⁵⁴

(a) Entrada

(b) Imagens geradas a partir das forças $s(v_i)$ de $W^{4,t}$, com $t_1 < t_2 < t_3 < t_4$. (c) Distribuição da força s da rede W correspondente acima. (d) Imagens geradas a partir das forças $s(v_i)$ de $B^{4,t}$, com $t_1 < t_2 < t_3 < t_4$. (e) Distribuição da força s da rede B correspondente acima.

Figura 19 – Representação visual (b, d) e distribuição da força (c, e) (grau ponderado) calculada para a rede de conexões intra-canal W e extra-canal B , usando diferentes limiares t e $r = 4$. A representação visual é gerada normalizando os valores da força nas camadas V_1 , V_2 e V_3 para $[0,255]$, obtendo um pixel RGB. A distribuição das forças é calculada com uma discretização de $(\frac{(2r+1)^2-1}{2}z) + z$ bins (grau máximo possível + 1).

Fonte: Elaborada pelo autor.

Tabela 1 – Equações de cada medida estatística usada para compor o vetor de características. Elas podem ser obtidas para cada uma das 3 redes $N^{r,t}$, $W^{r,t}$ ou $B^{r,t}$, e são calculadas para ambas as distribuições do grau $P(k)$ (μ_k , σ_k , λ_k , κ_k , e_k e ϵ_k) e da força $P(s)$ (μ_s , σ_s , λ_s , κ_s , e_s e ϵ_s).

<p>Média</p> $\mu_k = \frac{1}{ V } \sum_{i=1}^{ V } k(i)$	<p>Desvio Padrão</p> $\sigma_k = \sqrt{\sum_{i=1}^{ V } (k(i) - \mu_k)^2}$	<p>Assimetria</p> $\lambda_k = \frac{\frac{1}{k_{max}} \sum_k (P(k) - \mu_k)^3}{\left(\sqrt{\frac{1}{k_{max}} \sum_k (P(k) - \mu_k)^2} \right)^3}$
<p>Curtose</p> $\kappa_k = \frac{\frac{1}{k_{max}} \sum_k (P(k) - \mu_k)^4}{\left(\sqrt{\frac{1}{k_{max}} \sum_k (P(k) - \mu_k)^2} \right)^2}$	<p>Energia</p> $e_k = \sum_k P(k)^2$	<p>Entropia</p> $\epsilon_k = - \sum_k P(k) \log(P(k))$

Fonte: Elaborada pelo autor.

3.1.5 Vetor de Características

Através das distribuições do grau e da força é possível obter medidas quantitativas para caracterizar as redes correspondentes, portanto caracterizando a textura que estas representam. De acordo com a análise dinâmica das redes, descrita na Seção 2.3.2, a caracterização completa é feita combinando os conjuntos $R = \{r_1, \dots, r_n\}$ e $T = \{t_1, \dots, t_m\}$. Para r_i e t_j onde $i \in [1, n]$ e $j \in [1, m]$, são obtidas as redes N^{r_i, t_j} , W^{r_i, t_j} e B^{r_i, t_j} , e a distribuição do grau e força são calculadas. Porém, utilizar as distribuições $P(k)$ e $P(s)$ inteiras pode gerar vetores de característica muito grandes. O comprimento das distribuições, $k_{max}^r = \left(\frac{(2r+1)^2 - 1}{2} z \right) + z$, para os raios $r = 1, 2, 3, 4, 5, 6$ é, respectivamente, 15, 29, 75, 123, 183, 255. Portanto o tamanho do vetor de características seria $2mk_{max}$, e usando 2 como a menor quantidade de limiares m indicada, os tamanhos dos vetores de característica para cada raio seriam 60, 116, 300, 492, 732, 1020. Realizando a concatenação dos raios, o vetor final de características teria tamanho 2720, isso considerando apenas uma das redes N , W ou B . Nesse contexto, propomos uma sumarização das distribuições $P(k)$ e $P(s)$ usando medidas estatísticas, para reduzir e padronizar a quantidade de características para cada raio. São usados os 4 primeiros momentos estatísticos da distribuição, que são respectivamente a Média, Desvio Padrão (raiz quadrada da Variância), Assimetria e Curtose (do inglês *Skewness* e *Kurtosis*). Para complementar, são usadas também a Energia e a Entropia, totalizando 6 medidas estatísticas. A equação de cada uma das medidas é mostrada na Tabela 1. Desta forma, combinando as medidas para o grau e a força, o tamanho do vetor de características de uma rede será sempre 12, independente do raio usado.

Cada medida estatística descrita anteriormente pode ser calculada para uma das redes $N^{r,t}$, $W^{r,t}$ ou $B^{r,t}$. Dada uma rede, um vetor de característica ϖ é obtido para descrever sua estrutura topológica combinando as 6 medidas para grau e força

$$\varpi(N^{r,t})_k = [\mu_k, \sigma_k, \lambda_k, \kappa_k, e_k, \epsilon_k] \quad (3.8)$$

$$\varpi(N^{r,t})_s = [\mu_s, \sigma_s, \lambda_s, \kappa_s, e_s, \epsilon_s] \quad (3.9)$$

$$\varpi(N^{r,t}) = [\varpi(N^{r,t})_k, \varpi(N^{r,t})_s] \quad (3.10)$$

Para realizar a análise completa da evolução dinâmica de uma rede, diferentes vetores ϖ são obtidos ao concatenar os valores de raio R e limiar T . Dada uma rede e um conjunto $R = \{r_1, \dots, r_n\}$ e $T = \{t_1, \dots, t_m\}$, um vetor de características φ de tamanho $12nm$ é obtido realizando combinações de valores de raio r_i e limiar t_j . Portanto, o vetor de característica que descreve a estrutura de cada uma das redes é obtido por

$$\varphi_N = [\varpi(N^{r_1,t_1}), \dots, \varpi(N^{r_n,t_m})] \quad (3.11)$$

$$\varphi_W = [\varpi(W^{r_1,t_1}), \dots, \varpi(W^{r_n,t_m})] \quad (3.12)$$

$$\varphi_B = [\varpi(B^{r_1,t_1}), \dots, \varpi(B^{r_n,t_m})] \quad (3.13)$$

A critério do usuário, cada vetor φ pode ser utilizado como descritor final para a textura modelada, e também sua combinação

$$\varphi = [\varphi_N, \varphi_W, \varphi_B] \quad (3.14)$$

O vetor de características relacionado a cada diferente rede possui certas particularidades, que são discutidas no Capítulo 4. O uso padrão do método, portanto, é a concatenação de todos os vetores, ou seja φ , pois inclui diferentes tipos de informação e apresenta boa performance de forma geral. Cada passo do Método proposto 1 é ilustrado na Figura 20, da textura de entrada ao vetor de características.

3.1.6 Complexidade Computacional

A complexidade computacional é uma função matemática que define o consumo de tempo demandando por um algoritmo. Nessa seção, a complexidade do método de modelagem e caracterização de imagens multiespectrais com RCs multi-camada (Método proposto 1) é calculada. Inicialmente, na etapa de modelagem, é necessário visitar cada pixel da imagem, em cada canal de cor, e verificar sua vizinhança dentro do raio r para criar as conexões da rede. Portanto, considere uma imagem I de tamanho $w * h$ e z

Figura 20 – Visão geral do funcionamento do Método proposto 1, desde uma textura de entrada até a obtenção do vetor final de características.

Fonte: Elaborada pelo autor.

canais espectrais ($|I| = w * h * z$), e uma região de vizinhança de tamanho k_{max}^r (definida anteriormente). A complexidade de modelar a imagem I como uma rede multi-camada $N^{r, t} = \{V, E\}$ é dada por $O(|I|k_{max}^r)$. As redes W e B são obtidas paralelamente ao processo de modelagem através da checagem do canal correspondente aos vértices sendo conectados, portanto não adicionam custo adicional à modelagem. Considerando a análise dinâmica das redes, o processo de modelagem é repetido para a concatenação do conjunto de parâmetros de raio $R = \{r_1, \dots, r_n\}$ e limiar $T = \{t_1, \dots, t_m\}$. Nesse caso, são nm iterações de modelagem, e como k_{max}^r depende do raio usado, as iterações de cada raio tem complexidade diferente. Considerando o pior caso, ou seja o maior raio usado, a complexidade total é dada por $O(m|I|k_{max}^{r_n})$. É importante destacar que conforme se use raios que cubram regiões muito grandes, isso é valores próximos do tamanho total da imagem, a complexidade então tende a $O(m|I|^2)$. Entretanto, seria necessário um raio

muito grande para se aproximar desse valor, por exemplo em uma imagem de tamanho $w = 200, h = 200$ e $z = 3$, onde $|I| = 120000$, seria necessário aproximadamente $r \approx 140$. Além disso, não sugerimos o uso de raios muito grandes, pois testes empíricos indicam que a performance estabiliza em torno de $r = [4, 6]$ para imagens desses tamanhos.

Para calcular a complexidade total do método proposto, é necessário calcular também a complexidade de caracterização das redes. Entretanto, as medidas topológicas consideradas, isso é o grau k e a força s , foram escolhidas pelo fato de poderem ser calculadas paralelamente à modelagem da rede. Isso é feito contabilizando as arestas de cada vértice durante a modelagem, ou seja, para o grau conta-se a quantidade de arestas que satisfaçam $a > t$, e para a força soma-se o peso dessas arestas. Dessa forma, a rede completa em si não precisa ser armazenada na memória, pois estamos extraindo apenas as medidas localmente. Portanto, a complexidade total do Método proposto 1 é a complexidade de modelagem calculada anteriormente, ou seja

$$O(m|I|k_{max}^{r_n}) \quad (3.15)$$

3.2 Método Proposto 2: Rede Multicamada Direcionada (RMD)

Alguns fatores de métodos anteriores baseados em RCs,^{9,11,19,20} e também do Método proposto 1, limitam sua performance e eficácia. Por exemplo, o uso de limiares t torna o processamento mais lento, pois a rede precisa ser remodelada m vezes. Além disso, é difícil determinar valores ideais de limiar, mesmo com as técnicas de aproximação. Nesse contexto, é introduzida uma nova abordagem que exclui a necessidade de utilização de limiares de corte. Durante a modelagem das RC, inicialmente todos os vértices possuem a mesma quantidade de conexão, k_{max}^r . Os limiares de corte são usados para que a rede, inicialmente regular, se torne complexa e possa ser caracterizada. Porém, se considerarmos a rede como direcionada, as distribuições de grau e força de entrada e saída dos vértices contém informação topológica suficiente para sua caracterização. De forma semelhante ao método de RC multicamada não-direcionado, dada uma imagem multiespectral I de tamanho $w * h * z$, uma rede multicamada $N = \{V, E\}$ é obtida mapeando cada pixel, em cada canal, em um vértice. O conjunto de vértices $V = \{V_1, \dots, V_z\}$ é composto pela combinação de z camadas, e cada uma é descrita por $V_i = \{v_1, \dots, v_{w*h}\}$, onde v_j representa o pixel j , na camada i . Para construir a rede de forma direcionada, um critério de direção é adotado durante a definição de seu conjunto de arestas $E = \{a(v_i, v_j)\}$, ou seja $a(v_i, v_j) \neq a(v_j, v_i)$. Portanto, dados dois vértices v_i e v_j , uma conexão é criada se sua distância Euclidiana é menor que um dado raio r e se o gradiente da sua intensidade é positivo

$$\forall(i, j), a(v_i, v_j) = \begin{cases} \frac{(|p(v_i) - p(v_j)| + 1)(d + 1) - 1}{(L + 1)(r + 1) - 1}, & \text{se } d(v_i, v_j) \leq r \text{ e } p(v_i) \leq p(v_j) \\ \emptyset, & \text{caso contrário} \end{cases} \quad (3.16)$$

onde d é a distância Euclidiana e $p(v_i)$ retorna o nível de intensidade do pixel representado pelo vértice v_i . Em outras palavras, a conexão aponta para o pixel de maior intensidade, e se as intensidades forem iguais, a aresta é bidirecional ($a(v_i, v_j) = a(v_j, v_i)$). Note que o peso da aresta é calculado pela mesma equação usada pelo Método 1 (Equação 3.1). É importante ressaltar que vértices de borda são ignorados, da mesma forma descrita para o Método proposto 1. A Figura 21 (a) ilustra o funcionamento do método de modelagem direcionada para uma imagem com 3 canais de cor (RGB), e as figuras (b-d) mostram como uma camada de N muda conforme se aumenta o raio. Essa abordagem elimina a necessidade do uso de limiares de corte pois a informação direcional das arestas de um vértice é suficiente para caracterizar sua interação local, pois indica o padrão de variação de intensidade entre os pixels.

3.2.1 Caracterização Estatística

Para caracterizar a rede, a mesma abordagem de divisão proposta para o método não-direcionado é usada (veja 3.1.4 e as Equações 3.6 e 3.7). Portanto, são obtidas as redes W^r , de conexão intra-canal, e B^r , de conexão extra-canal, a partir da seleção das arestas da rede N^r . A Figura 22 ilustra a estrutura das redes W^r e B^r obtidas usando $r = 1$. Analisando cada rede individualmente é possível realizar uma caracterização mais completa e efetiva.

Para a caracterização quantitativa das redes, o conceito de direção deve ser levado em conta, pois todos seus vértices possuem k_{max}^r conexões. Nesse caso é possível calcular o grau de entrada, que conta quantas arestas apontam para o vértice, e o grau de saída, que conta a quantidade de vértices em que v_i aponta

$$k(v_i)_{in} = \sum_{\forall v_j \in V} \begin{cases} 1, & \text{se } a(v_i, v_j) \in E \\ 0, & \text{caso contrário} \end{cases} \quad (3.17)$$

$$k(v_i)_{out} = \sum_{\forall v_j \in V} \begin{cases} 1, & \text{se } a(v_j, v_i) \in E \\ 0, & \text{caso contrário} \end{cases} \quad (3.18)$$

Note que seguindo a Equação do grau em redes não direcionadas (2.1), $k(v_i) = k(v_i)_{in}$, porém em uma rede direcionada $a(v_i, v_j) \neq a(v_j, v_i)$. É intuitivo notar que $k(v_i)_{in} + k(v_i)_{out} = k_{max}^r$, para qualquer vértice. Ou seja, usar ambos os graus de entrada e de saída para caracterizar a rede é redundante, portanto definimos como padrão o uso apenas do grau de entrada. A força de um vértice também pode ser calculada para entrada e saída

(a) Rede N completa. (b) Uma camada de $N^r=1$. (c) Uma camada de $N^r=2$. (d) Uma camada de $N^r=3$.

Figura 21 – Rede N direcionada modelada para uma textura com 3 canais de cor (RGB)
 (a). (b-d) Uma das camadas da rede direcionada N^r .

Fonte: Elaborada pelo autor.

$$s(v_i)_{in} = \sum_{\forall v_j \in V} \begin{cases} a(v_i, v_j), & \text{se } a(v_i, v_j) \in E \\ 0, & \text{caso contrário} \end{cases} \quad (3.19)$$

$$s(v_i)_{out} = \sum_{\forall v_j \in V} \begin{cases} a(v_j, v_i), & \text{se } a(v_j, v_i) \in E \\ 0, & \text{caso contrário} \end{cases} \quad (3.20)$$

Ao contrário do grau, as forças de entrada e saída de um vértice são diferentes pois dependem da intensidade e distância dos pixels. Dessa forma, a caracterização da rede multicamada direcionada é feita combinando as medidas de grau de entrada $k(v_i)_{in}$, força de entrada $s(v_i)_{in}$ e de saída $s(v_i)_{out}$. Cada medida capta uma característica diferente da imagem modelada, como é possível observar na Figura 23 (b,d), que mostra uma representação visual das medidas calculada para cada vértice, em cada canal, e então convertida para um pixel RGB. O grau de entrada da rede W funciona como um detector

(a) Rede direcionada W de conexões intra-canal. (b) Rede direcionada B de conexões extra-canal.

Figura 22 – Exemplo de redes multicamada direcionadas W e B de conexões intra (a) e entre (b) canal, respectivamente.

Fonte: Elaborada pelo autor.

de bordas, destacando as muitas variações de intensidade presentes na imagem, de uma forma geral. Isso acontece pois regiões de borda tendem a receber ou enviar muitas arestas em uma mesma direção, em canais simultâneos, destacando os pixels ali presentes. As forças de W marcam um padrão similar ao grau, porém com respostas balanceadas entre canais diferentes, uma vez que a medida pondera as conexões. Na rede B , a resposta das medidas está relacionada à informação espectral da imagem, pois as conexões consideram a variação entre os três canais de cor. Por exemplo, nas regiões centrais das folhas o grau e a força de entrada no canal verde parecem receber mais conexões dos outros canais. É possível notar que a força de entrada faz uma melhor distinção das estrutura internas da folha, enquanto que o grau demonstra uma resposta mais geral. A força de saída contrasta com a força de entrada, tendo uma resposta maior no canal vermelho e destacando as estrutura das folhas, porém de forma diferente.

Para usar as medidas de grau e força de entrada e saída, sua distribuição é calculada da mesma forma como descrito para o método não-direcionado. Portanto, são calculadas as distribuições $P(k_{in})$, $P(s_{in})$ e $P(s_{out})$, que representam a função densidade de probabilidade da ocorrência de cada medida em uma rede N , W ou B . A discretização usada no cálculo da ocorrência das medidas é a mesma utilizada para o método anterior, utilizando k_{max} bins. No exemplo da Figura 23, a distribuição de cada medida das redes consideradas é mostrada ((c) e (e)). De uma forma geral, é possível notar que o padrão da distribuição do grau de entrada é diferente das forças, onde o grau na rede W apresenta um pico próximo ao centro da distribuição. As distribuições das forças em ambas as redes seguem um comportamento do tipo lei de potência, similar às redes mostradas para o método não

(a) Entrada

(b) Grau de entrada k_{in} , força de entrada s_{in} e saída s_{out} de W^4 . (c) Distribuição do grau de entrada k_{in} , força de entrada s_{in} e saída s_{out} de W^4 . (d) Grau de entrada k_{in} , força de entrada s_{in} e saída s_{out} de B^4 . (e) Distribuição do grau de entrada k_{in} , força de entrada s_{in} e saída s_{out} de B^4 .

Figura 23 – Representação visual (b, d) e (c, e) distribuição do grau de entrada, força de entrada e de saída calculados para a rede de conexões intra-canal W e extra-canal B , usando $r = 4$. A representação visual é gerada normalizando os valores da medida nas camadas V_1, V_2 e V_3 para $[0,255]$, obtendo um pixel RGB. Quanto maior a medida nas três camadas, mais o pixel tende para o branco.

direcionado (veja a Figura 19).

3.2.2 Vetor de Características

Considerando que o uso de limiares não é necessário na modelagem de redes direcionadas, é possível incluir mais informações na caracterização da rede para compensar a redução no vetor de características. Portanto, propomos uma nova segmentação da estrutura das redes já obtidas. Isso é feito analisando cada camada da rede individualmente, através das medidas topológicas de grau e força de entrada e saída. Dessa forma, o método inclui diferentes formas de análise da imagem, por exemplo analisar cada camada da rede W é similar a abordagem utilizada por métodos de textura integrativos. Esses métodos realizam análise espaciais tradicionais em textura monocromática, aplicados individualmente em cada canal. Nesse contexto, o método proposto estará incluindo tanto esse tipo de análise como também a análise espaço-espectral realizada pela rede B . Para a caracterização estatística das redes, as mesmas 6 medidas adotadas para o método não-direcionado são usadas (veja a Tabela 1), compondo o vetor de estatísticas ϖ . Então dada uma rede N , W ou B , seu conjunto de vértices $V = \{V_i, \dots, V_z\}$ é analisado considerando cada camada V_i individualmente. Nesse caso, o vetor de estatísticas ϖ é calculado para cada medida, em cada camada, e então concatenado compondo o vetor de estatísticas de uma rede

$$\varpi(N^r) = [\varpi(V_1)_{kin}, \varpi(V_1)_{sin}, \varpi(V_1)_{sout}, \dots, \varpi(V_z)_{kin}, \varpi(V_z)_{sin}, \varpi(V_z)_{sout}] \quad (3.21)$$

Considerando a análise dinâmica da evolução das redes, discutida na seção anterior, diferentes valores de raio $R = \{r_1, \dots, r_n\}$ são combinados para uma caracterização mais completa da textura. Nesse caso, o vetor de características de uma rede é a combinação das estatísticas calculadas para cada valor de raio

$$\varphi_N = [\varpi(N^{r_1}), \dots, \varpi(N^{r_n})] \quad (3.22)$$

$$\varphi_W = [\varpi(W^{r_1}), \dots, \varpi(W^{r_n})] \quad (3.23)$$

$$\varphi_B = [\varpi(B^{r_1}), \dots, \varpi(B^{r_n})] \quad (3.24)$$

Por fim, o descritor final do Método proposto 2 é a combinação dos vetores de característica de cada rede

$$\varphi = [\varphi_N, \varphi_W, \varphi_B] \quad (3.25)$$

Como discutido na seção do método não-direcionado, o vetor de característica pode ser utilizado de diferentes formas, considerando cada vetor individual ou sua combinação (3.25). Nós indicamos o uso da combinação de todos os vetores, que inclui diferentes tipos de informação e apresenta boa performance em diferentes cenários. O vetor de características é analisado em mais detalhes no Capítulo 4, Seção 4.2.2. A Figura 24 ilustra a estrutura geral de cada passo do método de RC multicamada direcionada.

3.2.3 Complexidade Computacional

O processo de modelagem do método de RC multicamada direcionada segue a mesma implementação do método não-direcionado, pois cada um dos k_{max}^r vizinhos de cada pixel são visitados, o que muda é apenas a definição da direção das arestas. Portanto, a modelagem das redes N , W e B tem complexidade $O(|I|k_{max}^r)$ para uma imagem I de tamanho $w * h$ e z canais espectrais ($|I| = w * h * z$). Da mesma forma, a caracterização utilizando o grau e a força dos vértices não inclui complexidade adicional ao método. Considerando a análise dinâmica das redes, diferentes parâmetros são combinados na modelagem. Como o método não usa limiares de corte, remove-se o termo m da equação de complexidade, restando apenas os raios como parâmetro. Dessa forma, a complexidade total para se obter os descritores segundo o Método proposto 2 é

$$O(|I|k_{max}^{r_n}) \quad (3.26)$$

onde $k_{max}^{r_n}$ é o grau máximo do maior raio usado, r_n .

3.3 Método Proposto 3: Análise de Imagens Multi-espectrais com Redes Direcionadas de Similaridade Angular (RDSA)

Os métodos propostos descritos nas seções anteriores são focados em imagens com poucos canais espectrais, mais especificamente os formatos mais usados de 3 canais (RGB, HSV, etc). É possível estender seu uso para imagens multiespectrais com mais canais, como por exemplo imagens científicas com $z \gg 3$, no entanto não recomendamos isso. Para esse novo problema, essa seção introduz um novo método baseado em RC para lidar com imagens com muitos canais espectrais, que podem representar respostas fora do espectro visível. O grande desafio de analisar esse tipo de imagem é encontrar formas eficientes de lidar com o grande conjunto de dados disponível. Por exemplo, nesse tipo de imagem, além da elevada quantidade de canais, seus pixels também podem armazenar valores maiores que 2^8 , que é o padrão em imagens comuns de 8 *bits*. Como exemplo, existem imagens registradas por microscopia confocal de fluorescência que contém 32 canais, captando a resposta de comprimentos de onda entre $400nm$ e $700nm$, e pixels de 16 *bits*, ou seja 2^{16} valores. Se comparado um pixel RGB de 8 *bits* a um pixel multiespectral desse tipo, temos uma diferença de informação na ordem de $1,25 \cdot 10^{145}\%$. Portanto fica evidente a

Figura 24 – Visão geral do funcionamento do método baseado em RC multicamada direcionada, desde uma textura de entrada até a obtenção do vetor final de características.

Fonte: Elaborada pelo autor.

quantidade de informação presente em uma imagem desse tipo, o que pode tornar sua análise custosa computacionalmente.

Para modelar uma imagem multiespectral com $z \gg 3$, propomos o seguinte método. Considere uma imagem I com largura w , altura h e z canais espectrais. Construímos uma rede $N = \{V, E\}$ onde o conjunto de vértices é definido por $V = \{v_1, \dots, v_{w*h}\}$. Portanto, diferente das abordagens introduzidas anteriormente, a rede possui apenas uma camada, onde cada pixel é mapeado por um vértice. Dessa forma, a modelagem da rede não depende da quantidade de canais da imagem, reduzindo o custo computacional. Para a definição do conjunto de arestas, é usada uma abordagem similar à do método de RM direcionada, onde as arestas possuem direção. Isso exclui a necessidade do uso de limiares, portanto reduzindo a quantidade de parâmetros do método. Para criar uma aresta, dois vértices são conectados se sua distância Euclidiana é menor que um dado raio r , da mesma forma introduzida anteriormente. Para definir o peso da aresta e sua direção, introduzimos um novo conceito de similaridade entre os pixels. Como um vértice representa um pixel da imagem, e um pixel possui z valores correspondentes à sua informação espectral, podemos considerar cada vértice como um vetor de tamanho z no espaço Euclidiano. Nesse caso, é possível medir a similaridade entre dois vértices, considerados como vetores, através de medidas geométricas. Uma medida bastante conhecida e utilizada é a similaridade do cosseno, que mede o ângulo entre dois vetores. Dados dois vértices v_i e v_j , que representam os pixels i e j , contendo z valores correspondentes a sua resposta em cada canal espectral, sua similaridade do cosseno é calculada por

$$\cos\theta(v_i, v_j) = \frac{\sum_{l=1}^z p(v_i, l)p(v_j, l)}{\sqrt{\sum_{l=1}^z p(v_i, l)^2} \sqrt{\sum_{l=1}^z p(v_j, l)^2}} \quad (3.27)$$

onde $p(v_i, l)$ retorna o valor do pixel i no canal l . Como os valores dos pixels variam em um intervalo positivo $([0, L])$, a similaridade do cosseno retorna valores entre $[0, 1]$. A similaridade máxima, ou seja 1, acontece quando os dois vetores formam um ângulo de 0° graus. A mínima similaridade acontece quando os vetores são ortogonais, formando um ângulo de 90° . A partir da similaridade do cosseno é possível estabelecer uma relação entre os vértices baseada na informação espectral dos pixels. Para incluir a informação espacial da distribuição dos pixels na imagem, consideramos sua distância Euclidiana $d(v_i, v_j)$ entre os pixels i e j . Para definir a direção das arestas, é usada uma abordagem similar a do método de RM direcionada, considerando o pixel de maior valor. Nesse caso, são somados os valores de cada canal dos pixels, e a aresta aponta na direção do pixel de maior valor absoluto

Figura 25 – Modelagem de imagens multiespectrais em redes direcionadas N^r de similaridade angular. Cada pixel da imagem é representado por um vértice, e as conexões são criadas apontando para o pixel de maior valor absoluto (soma de todos canais). O peso da conexão é definido pela similaridade do cosseno multiplicada pela distância Euclidiana dos pixels.

Fonte: Elaborada pelo autor.

$$\forall(i, j), a(v_i, v_j) = \begin{cases} \frac{\cos\theta(v_i, v_j)d}{r}, & \text{se } d(v_i, v_j) \leq r \text{ e } \sum_{l=1}^z p(v_i, l) \leq \sum_{l=1}^z p(v_j, l) \\ \emptyset, & \text{caso contrário} \end{cases} \quad (3.28)$$

onde r normaliza os valores da equação para $[0, 1]$. A Figura 25 ilustra o processo de modelagem de uma imagem multiespectral para uma rede direcionada.

3.3.1 Caracterização Estatística das Redes e Vetor de Características

Para caracterizar a estrutura da rede direcionada, são usadas as medidas de grau de entrada k_{in} , força de entrada s_{in} e de saída s_{out} , como descrito na Seção 3.2. Uma representação visual das medidas é mostrada na Figura 26, calculada para uma imagem de microscopia por fluorescência obtida de um corte foliar, com 32 canais espectrais e

(a) Imagem de entrada de 32 canais convertida para RGB

(b) Medidas topológicas de N convertidas em pixels RGB. Grau de entrada (R), força de entrada (G) e saída (B).

(c) Distribuição de cada medida topológica

Figura 26 – Representação visual (b) da distribuição do grau de entrada, força de entrada e de saída (c-e) calculados para a rede direcionada de similaridade angular, usando $r = 4$. A imagem de entrada possui 32 canais espectrais e pixels de 16 bits, e a conversão para RGB (a) é feita calculando a média dos canais e normalizando os valores para 8 bits. A representação visual é gerada normalizando os valores de cada medida para $[0,255]$, onde o canal R é o grau de entrada, o canal G a força de entrada e o canal B a força de saída. Quanto maior o valor das 3 medidas, mais o pixel tende para o branco.

Fonte: Elaborada pelo autor.

pixels de 16 bits. A imagem mostrada na Figura (a) corresponde à entrada convertida em uma imagem RGB através da média entre os canais e a normalização dos valores para 8 bits. A rede direcionada N de similaridade angular é construída usando $r = 4$, e as medidas obtidas de cada vértice são concatenadas e normalizadas para formar um pixel RGB (Figura (b)). É possível notar as regiões com maior resposta das medidas, destacando as estruturas celulares da planta.

Obtendo as distribuições P da ocorrência de cada medida é possível caracterizar a

topologia da rede N . No exemplo da Figura 26 é possível notar o comportamento de cada distribuição. De forma similar as RM direcionadas, a distribuição do grau de entrada tem um pico próximo da média, se assemelhando a uma distribuição de Poisson. Também é possível notar a semelhança nas medidas de força, que apresentam um comportamento do tipo lei de potência. A combinação da informação de cada distribuição permite quantificar as redes e obter medidas relevantes sobre a imagem multiespectral modelada. Para isso, cada distribuição é quantificada usando 6 medidas estatísticas (veja a Tabela 1), compondo um vetor de estatísticas para uma dada rede N^r

$$\varpi(N^r) = [\varpi(N^r)_{kin}, \varpi(N^r)_{sin}, \varpi(N^r)_{sout}] \quad (3.29)$$

Combinando vários valores de raio $R = \{r_1, \dots, r_n\}$, a dinâmica da rede é quantificada através da concatenação do vetor de características final

$$\varphi = [\varpi(N^{r_1}), \dots, \varpi(N^{r_n})] \quad (3.30)$$

Dessa forma, o método caracteriza uma imagem multiespectral com diversos canais espectrais em um vetor de características de tamanho único, dependendo apenas da quantidade de raios n utilizada.

3.3.2 Complexidade Computacional

Dada uma imagem I de tamanhos $|I| = whz$, para construir uma rede direcionada N^r de similaridade angular, é necessário percorrer, para cada pixel da imagem ($[1, w * h]$), uma vizinhança de tamanho k_{max} para conectar os vértices. Durante essa etapa é necessário percorrer o vetor de intensidades dos pixels, de tamanho z , para determinar seu valor absoluto e calcular similaridade do cosseno. No entanto isso só precisa ser feito uma vez para cada pixel da imagem, independente do raio usado, ou seja são $|I|$ computações. Isso implica que o processo de modelagem tem custo computacional $O(whr_{max} + |I|)$. Considerando que o cálculo das medidas topológicas é feito paralelamente à modelagem da rede e generalizando o custo para o maior raio utilizado, a complexidade total do método é

$$O(whk_{max}^{r_n} + |I|) \quad (3.31)$$

O custo do método de RC de similaridade angular direcionada é menor que o custo dos métodos descritos anteriormente, como por exemplo do método de RM direcionada, que possui complexidade $O(|I|r_{max}^{r_n})$.

4 EXPERIMENTOS E RESULTADOS

Neste capítulo são descritos os experimentos realizados no decorrer deste trabalho, com o intuito de avaliar o desempenho dos métodos propostos em relação ao estado da arte.

4.1 Configuração Experimental

Os experimentos realizados para analisar a performance dos métodos propostos e de métodos da literatura consistem na classificação de imagens. No contexto de aprendizagem de máquina, considera-se uma base de imagens que é dividida em dois conjuntos, um para treino e outro para teste. Um determinado método extrai descritores das imagens do conjunto de treinamento, que são usados para alimentar um classificador de forma supervisionada. Isso é, o classificador é treinando ajustando sua função de classificação sabendo qual a classe das imagens do conjunto de treinamento. Em seguida, o classificador tenta prever a classe das imagens do conjunto de teste, a partir de seus descritores. A performance do descritor é então medida pela porcentagem de imagens corretamente classificadas pelo classificador. Existem diferentes formas de dividir bases de imagens em conjuntos de treino e teste, de forma a avaliar a performance de um determinado método da forma mais ampla possível. Por esse motivo, escolhemos 2 esquemas diferentes de forma a obter resultados imparciais. O primeiro esquema é conhecido como *Leave-One-Out* (L1O), do inglês, que significa "deixar um de fora". Nesse caso, uma das amostras da base de imagens é escolhida para teste, enquanto que o restante das imagens é usada para o treinamento. Esse processo é repetido para todas as amostras da base, ou seja, em cada iteração uma imagem diferente é escolhida para teste. Essa abordagem é conhecida como validação cruzada (*cross-validation*, do inglês), pois repete o processo de forma a realizar a análise a partir de diferentes conjuntos de treino e teste. Outro esquema de validação cruzada que foi considerado é conhecido como *k*-pastas (*k-fold*, do inglês). Nesse caso, a base de imagens é dividida em *k* pastas, ou conjuntos, de tamanho igual, de forma aleatória. A divisão das pastas é feita de forma estratificada, ou seja, mesmo sendo um processo aleatório é garantido que cada pasta tenha uma quantidade igual de amostras de cada classe. Em seguida, uma das pastas é escolhida para teste, enquanto que o treinamento é realizado com as outras $k - 1$ pastas. Da mesma forma que no esquema L1O, o processo é repetido até que cada uma das pastas é usada para teste. Como a divisão das pastas é aleatória, repetimos o processo 10 vezes, onde são consideradas a média e o desvio padrão da porcentagem de acerto do classificador.

Da mesma forma que são usados dois esquemas de validação cruzada, também são considerados dois classificadores diferentes. O primeiro classificador é considerado

um dos mais simples da literatura, denominado k -vizinhos mais próximos (do inglês *k-nearest neighbours*, KNN). Este é um classificador supervisionado que classifica amostras de acordo com a distância Euclidiana de seus descritores em relação ao conjunto de treinamento. O parâmetro k determina quantos vizinhos mais próximos são considerados para a classificação, no caso foi adotado o caso mais simples, com $k = 1$, i.e. a classe é determinada pela amostra mais próxima. Outro classificador considerado nos experimentos é conhecido como *Linear discriminant analysis* (LDA),⁷¹ que significa análise discriminante linear. Esse é um classificador estatístico, supervisionado, que consiste em encontrar uma combinação linear das características onde a variância entre classes é maior do que a variância intra-classe. Em comparação com o classificador KNN esse é um método mais complexo, e normalmente produz melhores resultados.

4.1.1 Bases de imagens

Os primeiros experimentos analisam o desempenho dos métodos propostos RM e RMD em relação a métodos da literatura em bases coloridas, de três canais. Cada base é descrita em detalhes a seguir:

- **Usptex:** Essa base,⁷² construída pela Universidade de São Paulo, contém 191 classes de texturas coloridas naturais, encontradas no dia-dia. As imagens originais tem tamanho 512x384, e são divididas em 12 amostras de tamanho 128x128 sem sobreposição, totalizando 2292 imagens no total. Amostras de 20 classes dessa base são mostradas na Figura 27 (a).
- **Vistex:** A base *Vision Texture*,⁷³ do MIT, possui 54 imagens de textura colorida natural, de tamanho 512x512. Cada imagem é dividida em 16 amostras de tamanho 128x128, sem sobreposição, compondo uma base de 864 imagens. 20 classes da base Vistex são mostradas na Figura 27 (b).
- **Outex13:** Essa base esta presente no *framework* Outex,⁷⁴ proposto para a avaliação empírica de métodos de análise de textura. Esse *framework* é composto por vários conjuntos de imagens diferentes, e a base Outex13 (conjunto Outex_TC_00013 no site da base *) foca na análise de textura considerando cor como uma propriedade discriminatória. A base contém 1360 imagens divididas em 68 classes, ou seja, 20 amostras por classe, de tamanho 200x200. Algumas classes da base Outex13 são mostradas na figura Figura 27 (c).
- **MBT:** A base *Multi-Band Texture*⁷⁵ é composta por 154 imagens coloridas de classes formadas pelo efeito combinado de variações espaciais intra e entre canais. Esse tipo de padrão aparece em imagens com alta resolução espacial, que são comuns em áreas

* www.outex.oulu.fi

como astronomia e sensoriamento remoto. Cada uma das 154 imagens originais, de tamanhos 640x640, são divididas em 16 amostras sem sobreposição, de tamanho 160x160, compondo um conjunto de 2464 imagens. Como é possível observar na Figura 27 (d), que mostra 20 das classes da base MBT, suas características de textura são relativamente diferentes se comparadas às outras bases.

- **CUReT**: Proposta pelas Universidades de Columbia e Utrecht,⁷⁶ a base *Columbia-Utrecht Reflectance and Texture* é composta por imagens coloridas de materiais. A base contém 61 classes com 92 amostras cada, onde há uma ampla variedade de propriedades geométricas e fotométricas como variações intra-classe na rotação, iluminação e ângulo de visão. As classes representam texturas das superfícies de diversos materiais como folha de alumínio, grama artificial, gesso, concreto, couro, isopor, tecidos, entre outras. A Figura 27 (e) mostra amostras de 20 dessas classes.

4.2 Análise dos Métodos Propostos

Essa seção é composta por uma série de experimentos com o intuito de analisar a performance dos métodos propostos em função de seus parâmetros. Os experimentos são realizados seguindo a metodologia descrita na Seção 4.1.

4.2.1 Análise do método Proposto 1: Rede Multicamada (RM)

A primeira técnica proposta consiste em modelar uma rede multicamada não-direcionada dada uma imagem colorida, um conjunto de raios R e um conjunto de limiares T , que então é caracterizada com medidas topológicas. Na próxima seção, diferentes valores são testados para cada conjunto de parâmetro.

4.2.1.1 Análise dos parâmetros R e T

Primeiramente, é definido um conjunto geral de valores de raio r para compor o conjunto R . Por convenção, adotamos valores que correspondem às distâncias aproximadamente circulares na malha bidimensional da imagem, isto é $R = \{1, 2, 3, 4, \dots, r_n\}$. Para a definição do conjunto de limiares $T = \{t_1, \dots, t_m\}$, onde $0 \leq t_i \leq 1$, a abordagem apresentada no Capítulo 3, Seção 3.1.3, encontra os limites t_1 e t_m de acordo com o conjunto de treinamento. Em seguida o conjunto T é construído dividindo o intervalo $[t_1, t_m]$ em m valores equidistantes. Nesse contexto, são analisados o uso de raios r_i individuais, a combinação de $1, \dots, n$ raios e a quantidade de limiares m usada. Para quantificar a influência dos parâmetros, a performance do método é medida em termos da acurácia de classificação nas bases de textura colorida Usptex, Vistex, Outex13 e MBT, usando o vetor de características φ_N , referente à rede N . Os resultados são mostrados na Figura 28, utilizando os classificadores LDA e KNN e o esquema de validação cruzada L10.

(a) Usptex

(b) Vistex

(c) Outex13

(d) MBT

(e) CURET

Figura 27 – Amostras de 20 classes de cada uma das 5 bases de textura colorida utilizada nos experimentos.

Fonte: Imagens das bases: Usptex⁷²; Vistex⁷³; Outex13⁷⁴; MBT⁷⁵; CURET.⁷⁶

Figura 28 – Acurácia do Método proposto 1 em várias bases de textura utilizando diferentes valores de raio, individuais e combinados, e variando o número de limiares m . Os resultados são mostrados utilizando dois classificadores diferentes (LDA e KNN) e validação L1O.

Fonte: Elaborada pelo autor.

Analisando as curvas da acurácia mostradas na Figura 28 é possível observar que, na grande maioria dos casos, o uso de diferentes raios combinados é mais efetivo do que o uso de raios individuais. O único caso onde isso não acontece é quando consideramos a base Vistex e o classificador LDA, combinando 5 ou 6 raios e $m > 8$, onde há uma queda de performance. Isso provavelmente ocorre devido ao aumento na quantidade de características utilizadas, que pode causar redundância. Em relação ao uso dos raios, combinar de 4 a 6 raios gera os maiores resultados alcançados em cada base de textura. Em relação ao parâmetro m , é possível observar que independente do classificador utilizado o padrão da curva da acurácia é similar, e $m = 2$ gera os menores resultados no geral. Quando $m \geq 4$ a performance do método tende a estabilizar ou até cair, em alguns casos. Portanto, estes resultados indicam que o uso de $4 \leq m \leq 8$ é a melhor abordagem, em diferentes cenários. Nos experimentos seguintes, $m = 4$ é adotado como valor padrão, e é o valor indicado para uso no geral, mantendo uma boa relação entre performance, complexidade e tamanho do vetor de características.

4.2.1.2 Análise da caracterização estatística das redes

Conforme introduzido na descrição do método de RM, a estrutura da rede N pode ser dividida em outras 2 sub-redes W e B . Portanto, são analisadas as performances de classificação considerando cada rede individualmente ou suas combinações. A Tabela 2 mostra a acurácia de classificação de cada abordagem considerando combinações de raios e $m = 4$, em 5 bases de textura colorida. A respeito da quantidade de raios utilizadas, novamente é possível notar que os melhores resultados aparecem com o uso de entre 5 e 6 raios, exceto na base Vistex, onde o melhor resultado aparece usando 2 raios. Utilizar características extraídas de cada rede, combinadas (vetor φ), provou ser a abordagem com os melhores resultados no geral. Apenas na base MBT é que o melhor resultado é alcançado usando o vetor φ_{WB} , porém os resultados obtidos usando φ tem uma diferença de apenas 0,5%. Considerando a ampla variedade de todas as 5 bases de textura colorida, consideramos que o uso da combinação total das características φ é a melhor abordagem para se obter bons resultados em diferentes cenários.

Os melhores resultados do método de RM se mantiveram acima de 95% em todas bases, onde a bases Outex13 e MBT se mostraram as mais desafiadoras, com acurácia de 95.7% e 95.8%, respectivamente. Na base Vistex, o maior resultado alcançado foi 99.9%, mostrando que o método foi capaz de caracterizar corretamente praticamente todas as amostras.

4.2.2 Análise do método proposto 2: Rede Multicamada Direcionada (RMD)

O segundo método proposto, que consiste na modelagem de redes multicamada direcionadas, necessita apenas do conjunto R de raios para a modelagem. Conforme a seção anterior descreve, a combinação de raios se mostrou mais efetiva que o uso de raios individuais, em qualquer base analisada. Portanto, são analisadas as combinações de raios em relação às diferentes formas de caracterização, isso é, o uso das redes N , W e B , individuais ou combinadas. Os resultados de classificação são mostrados na Tabela 3 para 5 bases de textura colorida e usando o classificador LDA com esquema de validação cruzada L1O. De forma semelhante ao método de RM, os melhores resultados no geral aparecem combinando todas as redes, ou seja usando o vetor φ . O única exceção acontece para a base MBT, onde o melhor resultado foi atingido usando apenas a rede W . Considerando o resultado obtido pelo método RM, isso evidencia a natureza complexa dessa base, que apresenta texturas de tipos relativamente diferentes aos comumente analisados. De qualquer forma, os resultados na base MBT usando o vetor φ são próximos aos resultados usando φ_W , com uma diferença de 0,3% de acurácia. Novamente, podemos concluir que o uso das combinações de redes aparenta ser mais efetivo na maioria dos casos.

Os resultados do método RM, em contraste com a segunda proposta RMD, demonstram aumento de performance obtido pela segunda abordagem. Por exemplo, na

Tabela 2 – Acurácia do método de RM usando diferentes vetores de característica φ , concatenando raios e usando $m = 4$ limiares. Foram usados os classificadores LDA e o esquema de validação cruzada L1O. Em negrito são mostrados os melhores resultados para cada vetor φ .

Base	R	φ_W	φ_B	φ_{WB}	φ_N	φ
Usptex	$R = \{1, 2\}$	82.3	88.6	94.3	89.1	95.2
	$R = \{1, 2, 3\}$	88.6	92.6	96.1	93.0	96.9
	$R = \{1, 2, 3, 4\}$	90.4	94.0	97.1	93.5	97.6
	$R = \{1, 2, 3, 4, 5\}$	91.4	94.8	97.6	94.4	97.9
	$R = \{1, 2, 3, 4, 5, 6\}$	91.4	95.2	97.5	94.5	97.7
Vistex	$R = \{1, 2\}$	96.6	96.9	99.3	97.5	99.9
	$R = \{1, 2, 3\}$	97.7	98.5	99.4	98.3	99.7
	$R = \{1, 2, 3, 4\}$	97.8	99.0	99.3	98.7	99.3
	$R = \{1, 2, 3, 4, 5\}$	98.5	98.7	99.2	99.0	98.5
	$R = \{1, 2, 3, 4, 5, 6\}$	98.6	98.7	99.0	98.8	96.8
Outex13	$R = \{1, 2\}$	79.8	87.3	91.5	87.9	93.2
	$R = \{1, 2, 3\}$	83.9	90.7	93.6	90.6	94.4
	$R = \{1, 2, 3, 4\}$	86.4	92.5	94.1	92.3	95.5
	$R = \{1, 2, 3, 4, 5\}$	86.3	92.9	94.9	92.5	95.6
	$R = \{1, 2, 3, 4, 5, 6\}$	86.5	93.4	95.0	92.8	95.7
MBT	$R = \{1, 2\}$	87.9	69.1	90.4	76.5	90.8
	$R = \{1, 2, 3\}$	92.9	79.3	94.3	84.4	94.0
	$R = \{1, 2, 3, 4\}$	94.7	82.2	95.0	86.8	94.6
	$R = \{1, 2, 3, 4, 5\}$	95.6	83.8	95.7	87.4	95.3
	$R = \{1, 2, 3, 4, 5, 6\}$	95.7	84.8	95.8	87.7	95.3
CURET	$R = \{1, 2\}$	74.6	73.0	86.0	74.3	88.1
	$R = \{1, 2, 3\}$	82.7	81.4	90.9	81.6	93.0
	$R = \{1, 2, 3, 4\}$	85.9	85.2	93.2	86.0	94.7
	$R = \{1, 2, 3, 4, 5\}$	88.8	88.2	94.6	88.3	96.0
	$R = \{1, 2, 3, 4, 5, 6\}$	90.3	89.6	95.5	89.9	96.6

Fonte: Elaborada pelo autor.

base Usptex, o melhor resultado subiu de 97.9% para 99.7%, um aumento de 1.8%. Na base Vistex, o resultado praticamente se manteve, com uma pequena variação de 99.9% para 99.8%. Na base Outex13, o aumento não é tão significativo, subindo de 95.7% para 96%. Porém, na base MBT foi observado o maior aumento de performance, de 95.8% para 98.5%, uma diferença de 2.7%. Na base CURET também temos um aumento expressivo, de 96.6% para 98.8%, um ganho de 2.2% de acurácia. Nessa base, que tem uma maior quantidade de amostras, esse ganho representa um adicional de 124 imagens corretamente classificadas.

Tabela 3 – Acurácia do método de RMD usando diferentes vetores de característica φ e concatenando raios de 2 à 6. Foram usados os classificadores LDA e o esquema de validação cruzada L1O. Em negrito são mostrados os melhores resultados para cada vetor φ .

Base	R	φ_W	φ_B	φ_{WB}	φ_N	φ
Usptex	$R = \{1, 2\}$	97.8	96.9	99.1	97.5	99.2
	$R = \{1, 2, 3\}$	98.4	98.1	99.4	98.1	99.4
	$R = \{1, 2, 3, 4\}$	98.7	98.2	99.4	98.1	99.5
	$R = \{1, 2, 3, 4, 5\}$	98.6	98.5	99.5	98.3	99.7
	$R = \{1, 2, 3, 4, 5, 6\}$	98.6	98.6	99.4	98.4	99.5
Vistex	$R = \{1, 2\}$	98.8	99.2	99.8	99.5	99.8
	$R = \{1, 2, 3\}$	99.4	99.7	99.8	99.7	99.8
	$R = \{1, 2, 3, 4\}$	99.7	99.8	99.8	99.5	99.5
	$R = \{1, 2, 3, 4, 5\}$	99.7	99.7	99.8	99.5	95.4
	$R = \{1, 2, 3, 4, 5, 6\}$	99.5	99.7	99.8	99.4	98.7
Outex13	$R = \{1, 2\}$	89.8	91.5	94.9	92.7	94.9
	$R = \{1, 2, 3\}$	91.0	93.1	95.4	94.1	96.0
	$R = \{1, 2, 3, 4\}$	91.5	93.8	95.7	94.7	95.1
	$R = \{1, 2, 3, 4, 5\}$	91.1	94.2	95.0	94.6	94.9
	$R = \{1, 2, 3, 4, 5, 6\}$	92.3	94.0	95.4	94.7	94.3
MBT	$R = \{1, 2\}$	97.0	82.1	97.6	90.2	98.1
	$R = \{1, 2, 3\}$	98.3	87.3	97.9	94.3	98.1
	$R = \{1, 2, 3, 4\}$	98.5	88.4	98.0	95.0	98.1
	$R = \{1, 2, 3, 4, 5\}$	98.4	89.5	98.1	95.1	98.2
	$R = \{1, 2, 3, 4, 5, 6\}$	98.5	89.4	98.2	95.5	98.2
CURET	$R = \{1, 2\}$	86.2	88.9	94.9	89.4	96.4
	$R = \{1, 2, 3\}$	91.8	92.8	97.3	93.8	98.0
	$R = \{1, 2, 3, 4\}$	93.3	94.9	98.3	95.7	98.5
	$R = \{1, 2, 3, 4, 5\}$	94.4	96.2	98.6	96.5	98.8
	$R = \{1, 2, 3, 4, 5, 6\}$	95.0	96.8	98.6	97.2	98.8

Fonte: Elaborada pelo autor.

4.3 Comparação com Métodos da Literatura

Nessa seção, os resultados obtidos pelos métodos RM e RMD são comparados ao resultado obtido por vários métodos da literatura, nas mesmas bases de imagem.

4.3.1 Métodos da literatura

Os seguintes métodos da literatura são considerados para comparação:

- **Hist₃₂**: Esse método computa histogramas de cor, uma das técnicas mais simples de caracterização, também denominado estatísticas de primeira ordem, pois não considera a distribuição espacial dos pixels. Os descritores são obtidos calculando o

histograma de intensidades, para cada canal da imagem, usando discretização em 32 intervalos, que então é normalizado para uma função densidade de probabilidade ($h(i) = \frac{h(i)}{\sum_{j=1}^{32} h(j)}$). Concatenando os histogramas dos três canais da imagem, o vetor de característica tem tamanho 96.

- **LPQ:** *Local Phase Quantization*³³ é uma técnica que usa informação de fase da transformada de Fourier obtida de regiões da imagem. Dado um pixel, coeficientes locais são obtidos para quatro pontos diferentes $\{u_1, u_2, u_3, u_4\}$. O descritor final é o histograma calculado a partir da análise de frequência de quantização para os quatro pontos, usando uma binagem de 256 intervalos. Esse é um método bastante utilizado em análise de textura monocromática, i.e. níveis de cinza. Nos experimentos deste trabalho, esse método é utilizado de forma integrativa, ou seja, descritores são obtidos para cada canal individualmente, que concatenados compõem um vetor de 768 características.
- **CLBP:** *Complete Local Binary Patterns*³⁴ é uma versão melhorada que adiciona informação não considerada pelo LBP tradicional através da transformação de diferença local de magnitude de sinal (LDSMT). Medidas locais são calculadas para uma vizinhança de tamanho $P = 8$ e um raio de distância $r = 1$, que resultam em um descritor de tamanho 118 para um canal da imagem. Da mesma forma que é feito com o método LPQ, o método CLBP é usado de forma integrativa, combinando descritores dos três canais da imagem, ou seja são 354 características.
- **Gabor-oponente:** *Gabor opponent color features*⁴⁹ é uma técnica baseada em bancos de filtros introduzida para a caracterização de textura colorida. As características são calculadas a partir da diferença normalizada de transformadas de Gabor, com diferentes escalas e orientações, aplicadas em canais de cor individuais e em pares de canais oponentes. Os descritores são obtidos usando os mesmos bancos de filtros e transformadas como em,¹⁵ resultando em 96 características monocromáticas, 96 de canais oponentes e 72 características oponentes de escala variada, que concatenadas compõem um vetor de características de tamanho 264.
- **CM em grafos:** Esse método usa caminho mínimo em grafos para análise de textura colorida.²¹ A imagem é modelada em um grafo regular de 3 camadas, sendo 1 camada para cada canal de cor, e o caminho mínimo nessa estrutura é usado para caracterização. São citados os resultados apresentado no artigo utilizando o classificador LDA e validação L10.
- **RC integrativo:** O método¹¹ se baseia em RCs para a análise de texturas monocromáticas através da modelagem da imagem como uma rede, e a caracterização é feita a partir de medidas de grau. São usados os parâmetros indicados pelos autores, ou seja

um raio de conexão $r = 3$ e um vetor de limiares de corte $T = \{0.005, 0.02, \dots, 0.53\}$, compondo um vetor de 108 descritores. Para a análise de textura colorida, o método é então aplicado de forma integrativa, concatenando descritores para os 3 canais da imagem, ou seja são 324 características.

- **AlexNet:** Essa⁷⁷ é uma rede neural convolucional profunda de 8 camadas, onde 5 são camadas convolucionais e as últimas 3 são camadas totalmente conectadas. Essa rede aceita como entrada imagens de tamanho 227x227x3.
- **VGG16:** *Very Deep Convolutional Network*⁷⁸ é uma arquitetura convolucional de profundidade relativamente maior se comparada à AlexNet. O modelo VGG16 possui 16 camadas, sendo 13 convolucionais e 3 camadas totalmente conectadas, e sua camada de entrada aceita imagens de tamanho 224x224x3.. De acordo com o artigo desse modelo, a rede VGG16 é descrita como a configuração D.
- **VGG19:** Essa é uma variação da rede convolucional proposta em,⁷⁸ onde apenas sua profundidade é aumentada, compondo 19 camadas no total. Essa é a configuração E apresentada no artigo, com 16 camadas convolucionais e 3 camadas totalmente conectadas.
- **InceptionV3:** Essa rede⁷⁹ é uma versão aprimorada da arquitetura *GoogLeNet*,⁸⁰ com uma estrutura diferente se comparada a suas predecessoras. As 5 primeiras camadas convolucionais funcionam de forma similar, no esquema *feed-forward* (a informação flui em uma única direção), porém o restante da rede é composto por módulos do tipo *inception*, que paralelizam convoluções locais, e então as recombina. A rede InceptionV3 possui 42 camadas no total, e sua camada de entrada aceita imagens de tamanho 299x299x3.
- **ResNet50:** *Deep Residual Network*⁸¹ é uma rede convolucional profunda com arquitetura de aprendizado residual de 50 camadas, que processa imagens de tamanho 224x224x3. Esse modelo tem como objetivo facilitar o treinamento de redes profundas através de conexões de atalho, que pulam uma ou mais camadas convolucionais.
- **ResNet101:** Essa rede tem a mesma arquitetura residual da rede ResNet50,⁸¹ porém estendida para 101 camadas. Nesse trabalho, essa é a rede convolucional com maior profundidade considerada.

Para usar as redes convolucionais profundas como extrator de características, é considerada uma arquitetura pré-treinada na base de imagens *ImageNet*, da competição de mesmo nome,⁸² disponível na biblioteca *Neural Network Toolbox* do software Matlabtm. Para obter os descritores, a imagem é redimensionada para o tamanho suportado de cada rede convolucional, e então é processada até as últimas camadas, removendo a camada de

classificação e a última camada totalmente conectada. São considerados então os mapas de ativação produzidos nas últimas camadas convolucionais, abordagem que tem sido utilizada recentemente.⁸³⁻⁸⁶ Em seguida, de acordo com uma abordagem proposta em,⁸³ é aplicada uma poda global média (*global average pooling*, ou GAP) nos mapas de ativação. Isso gera um vetor de características de tamanho correspondente à quantidade de canais da ativação da camada considerada. Caso a camada produza um mapa de ativação com mais de 800 canais, camadas anteriores são consideradas, até que a condição de tamanho seja satisfeita. Isso é feito pois normalmente as últimas camadas dessas redes geram mapas de ativação de tamanhos 2048 ou até 4096. A redução é feita para evitar o mal da dimensionalidade (do inglês *curse of dimensionality*), causado pelo aumento exponencial no volume associado à adições de dimensões extras ao espaço Euclidiano.^{87,88} Além disso, um efeito também conhecido no classificador LDA é o problema da singularidade, tamanho de amostras pequeno (da sigla em inglês SSS), que acontece quando existem vetores de alta dimensionalidade, ou um número pequeno de imagens de treinamento.^{89,90} Dessa forma a quantidade de características extraídas das redes convolucionais se mantem em um valor razoável em comparação aos outros métodos. Além disso, resultados do laboratório ainda não publicados indicam que camadas convolucionais anteriores são melhores para a caracterização de textura.

4.3.2 Comparação de Resultados

Considerando os método descritos para comparação, são realizadas classificações utilizando os classificadores KNN e LDA com validação cruzada L1O e 10-pastas. Os resultados nas bases Usptex e Vistex são mostrados na Tabela 4, e o resultado obtido pelo método de CM em grafos disponível em²¹ é indicado. É possível observar que estatísticas simples de cor não são suficientes para uma caracterização eficiente nessa base, de acordo com os resultados obtidos com o método de histogramas de cor *Hist*₃₂. Em relação aos outros métodos, se considerados ambos os classificadores os menores resultados são alcançados pelos métodos integrativos, incluindo o método baseado em RC, onde o método LPQ apresenta os menores resultados no geral. O método de filtros de Gabor em canais oponentes se mostrou mais efetivo que os métodos integrativos em ambos os classificadores. Com o classificador KNN as redes neurais convolucionais profundas apresentam os melhores resultados, com taxas de acerto acima de 98%, onde o melhor resultado é atingido pela rede AlexNet em ambos os esquemas de validação cruzada, com 99.2% e 99.2(±0.1)%. Considerando o classificador LDA, as redes convolucionais mantem o desempenho entre os melhores resultados, onde a rede ResNet50 atinge 99.7% em ambos os esquemas de validação. O método proposto de RMD também atinge esse resultado com o esquema L1O, e 99.6(±0.1)% com o esquema de 10-pastas, o que é estatisticamente equivalente ao resultado da rede ResNet50. O método proposto de RM apresenta resultados similares aos métodos integrativos, na casa dos 97%, assim como o método de CM em grafos. Em

Tabela 4 – Comparação de performance de diferentes métodos de análise de textura colorida nas bases Usptex e Vistex. São utilizados os classificadores KNN e LDA com validação cruzada L1O e 10-pastas, as imagens de cada bases estão representadas no formato RGB. Em negrito são mostrados os melhores resultados em cada esquema, assim como os dois métodos que atingiram os maiores resultados.

Base	Método	KNN		LDA	
		L-1-O	10-pastas	L-1-O	10-pastas
Usptex	<i>Hist₃₂</i>	54.7	55.0(± 0.3)	45.0	44.6(± 0.4)
	LPQ	86.6	86.7(± 0.2)	90.4	90.9(± 0.3)
	CLBP	89.4	89.5(± 0.1)	97.4	97.2(± 0.3)
	Gabor-oponente	96.1	96.1(± 0.1)	99.1	99.0(± 0.1)
	CM em grafos ²¹			96.9	
	RC integrativo	88.1	88.1(± 0.2)	97.9	98.1(± 0.1)
	AlexNet	99.2	99.2 (± 0.1)	99.6	99.6(± 0.1)
	VGG16	98.3	98.3(± 0.1)	99.5	99.5(± 0)
	VGG19	98.6	98.6(± 0.1)	99.5	99.5(± 0.1)
	InceptionV3	98.5	98.5(± 0.1)	99.5	99.5(± 0.1)
	ResNet50	99.1	99.1(± 0)	99.7	99.7 (± 0)
	ResNet101	98.5	98.6(± 0.1)	99.5	99.5(± 0)
	RM	86.3	86.5(± 0.2)	97.9	97.8(± 0.1)
	RMD	95.0	95.0(± 0.1)	99.7	99.6(± 0.1)
Vistex	<i>Hist₃₂</i>	87.0	86.8(± 0.3)	83.7	83.4(± 0.4)
	LPQ	98.6	98.6(± 0.1)	74.8	75.8(± 1.9)
	CLBP	99.7	99.6(± 0.2)	99.5	99.4(± 0.1)
	Gabor-oponente	99.0	98.9(± 0.1)	99.9	99.9(± 0)
	CM em grafos ²¹			99.1	
	RC integrativo	90.5	94.6(± 0.2)	99.7	99.6(± 0.1)
	AlexNet	99.8	99.8 (± 0.1)	100	100 (± 0)
	VGG16	99.1	98.9(± 0.1)	99.9	99.8(± 0.1)
	VGG19	99.0	98.8(± 0.1)	99.7	99.6(± 0.2)
	InceptionV3	99.3	99.3(± 0)	99.7	99.7(± 0.1)
	ResNet50	99.5	99.5(± 0.1)	99.5	99.4(± 0.1)
	ResNet101	99.0	98.9(± 0.1)	99.5	99.5(± 0.1)
	RM	92.0	91.9(± 0.4)	99.9	99.7(± 0.1)
	RMD	98.6	98.6(± 0.2)	99.8	99.8(± 0)

Fonte: Elaborada pelo autor.

relação a ambos os métodos propostos, é possível notar uma variação na performance em relação ao classificador, onde os resultados com LDA se mostraram superiores. Em suma, o método de RMD alcança resultados equivalentes ao estado da arte utilizando LDA, além de melhorar significativamente a performance em relação ao método original de RC. Por outro lado, o método proposto de RM apresenta performance relativamente semelhante aos métodos integrativos e ao método de RC original.

Na base Vistex os melhores resultados são similares aos obtidos na base Usptex, relativamente próximos de 100%. Por outro lado, se observada a performance do método de histogramas de cor é possível notar que grande parte das imagens dessa base podem ser classificadas apenas com estatísticas básicas de cor. Com o classificador KNN, os menores resultados são obtidos pelo método proposto de RM e pelo método original de RC integrativo, na faixa de 90% e 94%. Os demais métodos apresentam performance acima de 98%, onde os melhores resultados são obtidos pela rede AlexNet, com 99.8% e 99.8(± 0.1)%, respectivamente. De forma geral, as redes convolucionais com os menores resultados são ResNet, InceptionV3 e VGG19, indicando que sua profundidade pode ser um fator prejudicial em bases mais simples. O resultado obtido pela rede AlexNet corrobora com esse argumento, pois essa rede atingiu os maiores resultados e possui a menor profundidade entre as redes comparadas. Com o classificador LDA, vários métodos apresentam taxas de acerto muito próximas do máximo, variando entre 99.8% e 100%, como ambos os métodos propostos, as redes convolucionais AlexNet e VGG16 e o método de filtros de Gabor oponentes. Nesse caso, fica difícil medir de forma eficaz a performance desses métodos, devido as altas taxas de acerto obtidas nessa base. Além disso, na base Usptex os resultados obtidos por vários métodos estão acima de 99.5%, indicando que o desafio de caracterização de textura colorida em ambas essas bases já está praticamente superado.

Considerando as altas taxas de acerto obtidas nas bases Usptex e Vistex, consideramos bases mais desafiadoras para medir o desempenho dos métodos de caracterização de textura colorida. Na Tabela 5 são mostrados os resultados em três bases, Outex13, MBT e CURET. É possível observar que os resultados são relativamente menores se comparados às bases Usptex e Vistex. Na base Outex13, os resultados obtidos pelo método de histograma de cores, na faixa dos 50%, indicam que simples estatísticas não são suficientes para uma caracterização eficiente. Com o classificador KNN, os menores resultados são obtidos pelo método LPQ, porém os outros métodos integrativos e o método Gabor-oponente apresentam resultados similares às redes convolucionais, variando entre aproximadamente 85% e 87%. A melhor rede convolucional, com esse classificador, foi a rede VGG16, com resultados de 87.6% e 87.5(± 0.2)%, respectivamente, faixa de performance onde se encontram também os métodos de RC integrativo, a rede ResNet101 e o método proposto de RM. Por outro lado, as redes com menor resultado foram InceptionV3 e AlexNet, respectivamente. Isso evidencia a queda de desempenho da rede AlexNet, que obteve os melhores resultados nas bases Usptex e Vistex. A maior performance com o classificador KNN é atingida pelo método proposto de RMD, com 89.7% e 89.8(± 0.2)% nos esquemas L10 e 10-pastas, respectivamente, uma diferença de performance de mais de 2% em relação aos outros métodos. Com o classificador LDA os melhores resultados são obtidos por ambos os métodos propostos, onde o máximo é atingido pelo método de RMD, com 96% e 95.6(± 0.2)% de taxa de acerto, uma diferença, novamente, de mais de 2% em relação

a todos os outros métodos. No geral, todos os métodos da literatura atingem resultados abaixo de 92%, com exceção do método Gabor-oponente, com 93.5% e 93.4(± 0.3)%. De acordo com esse e os resultados obtidos pelos métodos propostos, há indícios de que a análise de características oponentes, juntamente com o classificador LDA, proporciona uma melhor caracterização do fenômeno de textura colorida presente nas classes da base Outex13.

Os resultados obtidos na base MBT, mostrados na Tabela 5, evidenciam características diferentes de cada método analisado. As taxas de acerto baixas do método de histograma de cores mostra que o desafio dessa base está na análise espacial. Isso acontece pois essa base foi feita de forma que estatísticas simples de primeira ordem não fossem capazes de caracterizar suas classes, formadas por variações espaço-espectrais entre canais. A natureza diferente dessa base permite analisar a capacidade dos métodos de lidar com esse tipo de informação. Com o classificador KNN, o método CLBP atinge os maiores resultados, 96.2% e 96.4(± 0.1)% em cada esquema de validação cruzada, respectivamente. O restante dos métodos, no geral, atingem resultados em torno de 90%, e os menores resultados são obtidos pelo método RM com 90.9% e 90.9(± 0.3)%. Em linhas gerais, os maiores resultados alcançados nessa base estão na faixa de 98% de taxa de acerto, obtidos com o classificador LDA e os métodos CLBP, RC integrativo e o método proposto de RMD. As redes convolucionais AlexNet e VGG16 obtêm resultados em torno de 97%, performance também similar ao método Gabor-oponente. As redes convolucionais mais profundas (InceptionV3, ResNet50 e 101) apresentaram os menores resultados, na faixa de 94%, indicando novamente um fator prejudicial na performance relacionada à quantidade de camadas dessas redes. O método proposto de RM supera essas redes, obtendo 95.8% e 95.6(± 0.2)% de acurácia, respectivamente, uma performance similar também ao método LPQ. É possível notar que praticamente todos os métodos tiveram a performance significativamente melhorada em relação ao uso do classificador KNN para o LDA. Isso indica certa complexidade na classificação dessa base, onde técnicas simples de distância, como feito pelo KNN, não são tão eficientes na separação das classes.

A última base analisada, CURET, contém o maior número de imagens entre todas as bases consideradas. Os resultados de classificação nessa base são mostrados na Tabela 5. Nessa base, de forma geral, houve pouca variação entre as taxas de acerto dos classificadores KNN e LDA. Os casos de maior variação acontecem com o método de histogramas de cor, com maior acurácia com KNN, e o método proposto de RM, com LDA. O método de RMD também foi levemente melhorado do KNN para o LDA, reforçando as evidências de que os métodos propostos funcionam melhor com esse classificador. Com o classificador KNN, os melhores resultados são obtidos pelas redes convolucionais ResNet101 e VGG19, com 98.9% de acurácia, seguidas pelas redes ResNet50, VGG16 e AlexNet, com resultados entre 98.4% e 98.8%. Com o classificador LDA, os maiores resultados são obtidos pelo método proposto de RMD e pela rede convolucional ResNet101, empatando exatamente com 98.8% e 98.7(± 0.1)%, em cada esquema de validação cruzada. Outros métodos que

Tabela 5 – Comparação de performance de diferentes métodos de análise de textura colorida nas bases Outex13, MBT e CURET. Em negrito são mostrados os melhores resultados em cada esquema, assim como os dois métodos que atingiram os maiores resultados no geral.

Base	Método	KNN		LDA	
		L-1-O	10-pastas	L-1-O	10-pastas
Outex13	<i>Hist</i> ₃₂	50.8	50.7(± 0.4)	46.0	45.9(± 0.7)
	LPQ	81.4	82.0(± 0.3)	80.1	79.3(± 0.4)
	CLBP	84.9	84.9(± 0.4)	89.6	89.5(± 0.3)
	Gabor-oponente	86.7	86.9(± 0.2)	93.5	93.4(± 0.3)
	CM em grafos ²¹			91.5	
	RC integrativo	87.1	87.3(± 0.3)	92.3	92.8(± 0.5)
	AlexNet	84.9	85.3(± 0.2)	91.4	91.6(± 0.3)
	VGG16	87.6	87.5(± 0.2)	91.1	91.8(± 0.3)
	VGG19	86.3	86.6(± 0.3)	90.7	90.8(± 0.3)
	InceptionV3	84.3	84.5(± 0.2)	89.5	89.7(± 0.3)
	ResNet50	86.3	86.4(± 0.2)	91.5	91.2(± 0.2)
	ResNet101	87.1	87.4(± 0.3)	91.3	91.0(± 0.4)
	RM	87.4	87.4(± 0.3)	95.7	95.1(± 0.4)
	RMD	89.7	89.8(± 0.2)	96.0	95.6(± 0.2)
MBT	<i>Hist</i> ₃₂	34.0	34.1(± 0.3)	17.7	17.8(± 0.2)
	LPQ	92.7	92.9(± 0.3)	95.7	95.8(± 0.3)
	CLBP	96.2	96.4(± 0.1)	98.2	98.2(± 0.1)
	Gabor-oponente	89.7	89.9(± 0.2)	97.6	97.5(± 0.2)
	RC integrativo	91.2	91.3(± 0.2)	98.5	98.6(± 0.1)
	AlexNet	90.2	90.3(± 0.2)	97.8	97.7(± 0.2)
	VGG16	92.3	92.4(± 0.1)	97.2	97.2(± 0.2)
	VGG19	90.3	90.4(± 0.2)	96.3	96.3(± 0.3)
	InceptionV3	88.4	88.6(± 0.2)	94.4	94.6(± 0.1)
	ResNet50	89.9	90.0(± 0.3)	94.9	94.8(± 0.2)
	ResNet101	87.8	88.0(± 0.3)	94.6	94.4(± 0.2)
	RM	62.9	62.9(± 0.5)	95.8	95.6(± 0.2)
	RMD	90.9	90.9(± 0.3)	98.5	98.3(± 0.1)
CURET	<i>Hist</i> ₃₂	72.4	71.9(± 0.3)	46.7	46.5(± 0.2)
	LPQ	94.8	94.6(± 0.1)	91.7	91.5(± 0.1)
	CLBP	95.5	95.1(± 0.1)	91.8	91.7(± 0.2)
	Gabor-oponente	97.0	96.8(± 0.1)	95.8	95.7(± 0.1)
	RC integrativo	89.6	88.8(± 0.2)	91.9	91.9(± 0.1)
	AlexNet	98.5	98.4(± 0)	98.2	98.2(± 0.1)
	VGG16	98.7	98.5(± 0.1)	98.5	98.5(± 0)
	VGG19	98.9	98.7(± 0)	98.6	98.7(± 0)
	InceptionV3	97.1	96.8(± 0.1)	97.3	97.3(± 0.1)
	ResNet50	98.8	98.7(± 0.1)	98.7	98.7(± 0.1)
	ResNet101	98.9	98.9(± 0.1)	98.8	98.7(± 0.1)
	RM	89.4	88.7(± 0.2)	96.6	96.7(± 0.1)
	RMD	97.4	97.2(± 0.1)	98.8	98.7(± 0.1)

Fonte: Elaborada pelo autor.

obtiveram resultados próximos foram as redes ResNet50 e VGG19, entre 98.6% e 98.7%. Esses resultados indicam que o aumento de profundidade das redes convolucionais se torna benéfico no caso de bases de imagens maiores. Em contrapartida, é possível observar que os métodos integrativos obtêm as menores performances na base CURET, próximos de 92%, e o método de filtros de Gabor oponentes tem desempenho intermediário, em torno de 95.7%.

Para ilustrar o desempenho de todos os métodos de forma geral, a Figura 29 mostra um gráfico de barras que representa a média de acurácia obtida entre os esquemas L1O e 10-pastas, em cada classificador. Em linhas gerais, o classificador LDA melhora o resultado em relação ao classificador KNN para todos os métodos, salvo algumas poucas exceções. Os métodos propostos também funcionam significativamente melhor com esse classificador, em todas as bases. Na base Usptex e Vistex, os métodos propostos apresentam resultados equivalentes ao estado da arte, e métodos integrativos funcionam de maneira similar às redes convolucionais profundas. Na base Outex13, esse comportamento é diferente, onde métodos integrativos superam todas as redes consideradas. Porém, os melhores resultados nessa base são alcançados pelos métodos propostos, com um aumento significativo em relação aos concorrentes. Na base MBT, novamente é possível notar que a performance dos métodos integrativos supera as redes convolucionais, que tiveram performance equivalente ao método proposto RM. O método RMD atinge resultados similares aos melhores. Na base CURET, os métodos integrativos tem uma considerável queda de performance, enquanto que os maiores resultados são obtidos pelo método proposto de RMD e as redes convolucionais de maior profundidade, como a ResNet101.

Considerando os resultados apresentados, é possível constatar que o método proposto RMD tem a melhor performance média no geral, pois mantém resultados entre os melhores métodos e na base Outex13 atinge um aumento significativo. Por outro lado, na maioria dos casos a performance de métodos integrativos e de redes convolucionais oscila de base para base. Os resultados do método de filtros de Gabor indicam que a abordagem de processamento oponente de cor permite melhor caracterização em certos cenários, como por exemplo na base Outex13. No entanto, esse método tem uma performance relativamente inferior em outros casos, como na base CURET. Os métodos propostos se beneficiam dessa abordagem, obtendo os melhores resultados na base Outex13, entretanto eles não são prejudicados nos cenários em que o método de filtros de Gabor são. Isso demonstra a robustez da abordagem proposta, em especial do método RMD, que se manteve como o melhor método, ou entre os melhores, em todos os diferentes cenários, o que não foi observado em nenhum outro método comparado. Em relação aos métodos similares anteriores, de RC integrativo e de CM em grafos, os métodos propostos também apresentaram uma melhora de performance significativa.

Figura 29 – Performance média dos métodos propostos (RM e RMD) e de métodos da literatura m 5 bases de textura colorida, usando dois classificadores (KNN e LDA). A performance é calculada como a média entre os resultados de classificação obtidos nos esquemas L1O e 10-pastas.

Fonte: Elaborada pelo autor.

4.3.3 Influência de Diferentes Espaços de Cor

Em análise de textura colorida, diferentes espaços de cor em que as imagens estão representadas podem influenciar o desempenho dos métodos. Nesse contexto, esta seção apresenta experimentos em cada uma das bases de textura colorida consideradas, onde as imagens são transformadas em 4 diferentes espaços de cor. São escolhidos 1 espaço de cor para cada diferente tipo de abordagem, sendo esses o tradicional RGB, já analisado anteriormente, LAB, HSV e $I_1I_2I_3$. Cada uma dessas abordagens de espaço de cor são descritas no Capítulo 2, Seção 2.1.1.1. Por convenção, os valores de cada canal da imagem, em cada espaço de cor, são normalizados para o intervalo $[0, 1]$. Foram considerados o classificador LDA, que apresentou os melhores resultados anteriormente, e o esquema de validação cruzada L1O. Os resultados obtidos por cada método nas bases Usptex e Vistex são mostrados na Tabela 6. É possível observar como diferentes espaços de cor influenciam cada método, de diferentes formas. Por exemplo, os métodos integrativos, assim como o método $Hist_{32}$, tiveram a performance aumentada no espaço HSV, enquanto que o oposto acontece com as redes convolucionais. Na verdade, esse efeito nas redes convolucionais acontece também em todos os outros canais, ainda que com uma intensidade menor. Isso é de certa forma esperado, pois essas redes são pré-treinadas em um conjunto de imagens RGB. Ainda na base Usptex, o método proposto de RM teve sua performance levemente aumentada com os outros canais de cor. Nessa base, de forma geral, os dois métodos que apresentaram maior robustez aos diferentes tipos de espaços de cor foram, respectivamente, o método proposto RMD e o método de filtros de Gabor oponentes. Isso indica que a abordagem de processamento de cor oponente permite maior tolerância às diferentes formas de se representar as cores das imagens. O método proposto de RMD incorpora essa características e, ao mesmo tempo, aumenta a invariância aos espaços de cor, atingindo $99.4(\pm 0.2)\%$ de performance média contra $98.8(\pm 0.6)\%$ do método Gabor-oponente. Além disso, o método RMD também obteve os maiores resultados em cada espaço de cor individualmente.

Na base Vistex, novamente é possível observar uma queda de performance das redes convolucionais frente aos diferentes espaços de cor. Os métodos integrativos, de forma geral, tiveram novamente performance aumentada no espaço de cor HSV. O método CLBP obteve performance ótima nos espaços de cor LAB, HSV e $I_1I_2I_3$, ou seja 100% de acurácia. O método proposto obteve resultados similares nos espaços de cor LAB e HSV, com 99.9%, e 99.5% no espaço $I_1I_2I_3$. Por outro lado, o método proposto de RM obtém o segundo melhor resultado no espaço $I_1I_2I_3$, com 99.8%. Em relação à performance média, os dois métodos mais robustos foram o CLBP, com $99.9(\pm 0.2)\%$ e o método RMD, com $99.8(\pm 0.2)\%$. Em seguida, os dois métodos mais próximos foram os métodos Gabor-oponente, com $99.7(\pm 0.1)\%$ e o método proposto de RM, com $99.7(\pm 0.2)\%$, reforçando a ideia de que a abordagem oponente é robusta à diferentes espaços de cor.

Tabela 6 – Influência de diferentes espaços de cor na performance dos métodos nas bases de textura colorida Usptex e Vistex. Foram considerados o classificador LDA e o esquema de validação cruzada L10. Células em negrito destacam os melhores resultados em cada espaço de cor e o melhor resultado médio, os dois métodos com maior desempenho médio também são destacados em negrito.

Base	Método	RGB	LAB	HSV	$I_1 I_2 I_3$	Média
Usptex	<i>Hist</i> ₃₂	45.0	39.3	68.8	36.9	47.5(±14.6)
	LPQ	90.4	95.0	96.6	94.6	94.1(±2.7)
	CLBP	97.4	98.6	98.6	98.5	98.3(±0.6)
	Gabor-oponente	99.1	99.0	97.9	99.3	98.8(±0.6)
	RC integrativo	97.9	98.3	99.1	98.2	98.4(±0.5)
	AlexNet	99.6	99.0	94.7	99.3	98.2(±2.3)
	VGG16	99.5	98.6	94.4	98.7	97.8(±2.3)
	VGG19	99.5	98.2	92.1	98.6	97.1(±3.4)
	InceptionV3	99.5	97.7	94.7	98.1	97.5(±2.0)
	ResNet50	99.7	98.2	92.5	98.5	97.2(±3.2)
	ResNet101	99.5	98.0	92.8	98.5	97.2(±3.0)
	RM	97.9	98.1	98.2	98.2	98.1(±0.2)
	RMD	99.7	99.4	99.1	99.3	99.4(±0.2)
Vistex	<i>Hist</i> ₃₂	83.7	69.1	84.8	64.9	75.6(±10.1)
	LPQ	74.8	87.5	89.5	88.8	85.1(±7.0)
	CLBP	99.5	100	100	100	99.9(±0.2)
	Gabor-oponente	99.9	99.8	99.8	99.5	99.7(±0.1)
	RC integrativo	99.7	99.3	100	98.4	99.3(±0.7)
	AlexNet	100	99.7	99.2	99.5	99.6(±0.3)
	VGG16	99.9	99.2	98.8	99.5	99.4(±0.4)
	VGG19	99.7	99.0	98.6	99.2	99.1(±0.4)
	InceptionV3	99.7	99.1	98.6	99.0	99.1(±0.4)
	ResNet50	99.5	99.2	99.1	99.3	99.3(±0.2)
	ResNet101	99.5	99.0	98.3	98.8	98.9(±0.5)
	RM	99.9	99.4	99.8	99.8	99.7(±0.2)
	RMD	99.8	99.9	99.9	99.5	99.8(±0.2)

Fonte: Elaborada pelo autor.

Como comentado anteriormente, as altas taxas de acerto nas bases Vistex e Usptex tornam difícil uma análise precisa do desempenho dos métodos. Nesse contexto, as bases Outex13, MBT e CURET complementam os experimentos realizados com diferentes espaços de cor. Os resultados para essas três bases são mostrados na Tabela 7, para cada espaço de cor e utilizando o classificador LDA com validação cruzada L10. Na base Outex13, os métodos com o menor desempenho foram, respectivamente, LPQ e a rede convolucional InceptionV3. As outras redes, exceto pela AlexNet, obtêm uma média na faixa dos 89%. O método CLBP tem sua performance reduzida nos canais LAB e $I_1I_2I_3$, o que resulta em uma média abaixo das redes convolucionais, de $87.8(\pm 1.4)\%$. O método de RC supera os outros métodos integrativos, obtendo média de 91.9%, resultado também obtido aproximadamente pelo método Gabor-oponente e pela rede convolucional AlexNet. Isso mostra que os métodos integrativos tradicionais apresentam limitações nessa base, quando diferentes espaços de cor são considerados. Por exemplo, o método CLBP, nas bases anteriores, é beneficiado pela mudança de espaço de cor, o que não acontece na Outex13. Os métodos com maior robustez à diferentes espaços de cor nessa base são os métodos propostos RMD e RM, com desempenho médio de $94.9(\pm 1.2)\%$ e $92.5(\pm 2.7)\%$, respectivamente. Nesse contexto e considerando os resultados obtidos pelos métodos Gabor-oponente e RC integrativo, é possível notar que RC e técnicas oponentes apresentam os melhores desempenhos nessa base, onde o método RMD combina essas propriedades e aumenta o resultado ainda mais. Na base MBT, novamente essas técnicas atingem os maiores resultados, com $96.6(\pm 2.3)\%$ do método RC integrativo, $96.4(\pm 1.8)\%$ do RMD e 96.3% obtido por ambos os métodos Gabor-oponente e CLBP. Nessa base o desempenho médio das redes neurais mais profundas (InceptionV3 e ambas ResNet) é o menor observado, próximos de 92%, enquanto que as redes menores AlexNet e VGG16 obtêm, respectivamente, $95.6(\pm 4.8)\%$ e $94.7(\pm 4.2)\%$. O alto desvio padrão das redes convolucionais se deve à queda acentuada de desempenho com o espaço de cor HSV, com baixas de quase 10%. Diferentemente do que é observado nas outras bases, isso também acontece com todos os outros métodos, porém com uma menor variação em relação às redes convolucionais.

Os resultados obtidos na base CURET, com diferentes espaços de cor, corroboram com a robustez dos métodos propostos. Os maiores resultados são obtidos pelos métodos RMD e RM, com $98.8(\pm 0.3)\%$ e $97.5(\pm 0.7)\%$ de desempenho médio, respectivamente. Como discutido anteriormente, as redes convolucionais ResNet50 e 101, de maior profundidade, obtêm os melhores resultados no espaço RGB, resultados similares ao método proposto RMD. Além disso, nessa base as redes convolucionais superam os métodos integrativos no desempenho médio e nos espaços de cor individuais LAB e $I_1I_2I_3$. No espaço HSV, os maiores resultados são obtidos pelos métodos propostos e pelo método Gabor-oponente, seguidos pelos métodos integrativos e então as redes convolucionais. Em linhas gerais, nessa base tanto os métodos integrativos como as redes convolucionais apresentam os maiores desvio padrão, indicando que sua performance oscila mais entre diferentes espaços

Tabela 7 – Influência de diferentes espaços de cor na performance dos métodos nas bases de textura colorida Outex13, MBT e CURET. Foram considerados o classificador LDA e o esquema de validação cruzada L1O. Células em negrito destacam os melhores resultados em cada espaço de cor e o melhor resultado médio, os dois métodos com maior desempenho médio também são destacados em negrito.

Base	Método	RGB	LAB	HSV	$I_1 I_2 I_3$	Média
Outex13	<i>Hist</i> ₃₂	46.0	37.1	55.1	38.9	44.2(±8.2)
	LPQ	80.1	74.8	78.2	76.0	77.3(±2.4)
	CLBP	89.6	86.8	88.2	86.6	87.8(±1.4)
	Gabor-oponente	93.5	91.3	91.7	91.3	91.9(±1.1)
	RC integrativo	92.3	90.6	94.0	90.6	91.9(±1.7)
	AlexNet	91.4	91.6	91.0	91.1	91.3(±0.3)
	VGG16	91.1	89.2	87.6	89.8	89.4(±1.4)
	VGG19	90.7	89.1	85.9	90.8	89.1(±2.3)
	InceptionV3	89.5	85.0	84.2	85.3	86.0(±2.4)
	ResNet50	91.5	88.7	86.8	89.0	89.0(±1.9)
	ResNet101	91.3	90.1	86.4	88.8	89.2(±2.1)
	RM	95.7	90.0	94.0	90.4	92.5(±2.7)
	RMD	96.0	93.7	95.8	94.0	94.9(±1.2)
MBT	<i>Hist</i> ₃₂	17.7	10.6	14.3	8.0	12.6(±4.2)
	LPQ	95.7	94.2	90.6	93.9	93.6(±2.1)
	CLBP	98.2	95.9	94.6	96.6	96.3(±1.5)
	Gabor-oponente	97.6	96.9	93.9	96.7	96.3(±1.6)
	RC integrativo	98.5	96.7	93.3	97.8	96.6(±2.3)
	AlexNet	97.8	98.1	88.4	98.2	95.6(±4.8)
	VGG16	97.2	97.0	88.5	96.3	94.7(±4.2)
	VGG19	96.3	95.8	87.9	95.8	94.0(±4.0)
	InceptionV3	94.4	95.0	85.8	93.1	92.1(±4.3)
	ResNet50	94.9	94.2	85.1	93.6	91.9(±4.6)
	ResNet101	94.6	94.1	85.3	92.7	91.6(±4.3)
	RM	95.8	94.0	92.0	94.9	94.2(±1.6)
	RMD	98.5	96.2	94.1	96.7	96.4(±1.8)
CURET	<i>Hist</i> ₃₂	46.7	30.5	68.5	29.6	43.8(±18.2)
	LPQ	91.7	94.8	96.3	93.0	93.9(±2.0)
	CLBP	91.8	94.2	96.2	93.9	94.0(±1.8)
	Gabor-oponente	95.8	95.6	97.4	95.3	96.0(±0.9)
	RC integrativo	91.9	92.7	95.4	91.2	92.8(±1.8)
	AlexNet	98.2	96.6	92.9	96.8	96.1(±2.2)
	VGG16	98.5	97.6	93.7	97.4	96.8(±2.1)
	VGG19	98.6	97.9	92.3	97.5	96.6(±2.9)
	InceptionV3	97.3	95.5	89.4	95.0	94.3(±3.4)
	ResNet50	98.7	97.3	93.9	97.8	96.9(±2.1)
	ResNet101	98.8	98.0	93.4	97.9	97.0(±2.5)
	RM	96.6	97.6	98.2	97.7	97.5(±0.7)
	RMD	98.8	98.9	99.1	98.4	98.8(±0.3)

Fonte: Elaborada pelo autor.

de cor. Os métodos com menor oscilação de desempenho são os métodos propostos RMD e RM e o método Gabor-oponente, respectivamente, reforçando a robustez do processamento oponente de cor e das abordagens propostas.

Para analisar de forma geral a robustez aos diferentes espaços de cor analisados, a Figura 30 mostra um gráfico com o desempenho médio dos 5 melhores métodos em cada base de textura colorida. É possível observar que o método proposto de RMD apresenta o maior desempenho médio, aparecendo como o melhor método nas bases Usptex, Outex13 e CURET, casos em que a diferença de performance para o segundo colocado é significativa. Nos demais casos, ou seja nas bases Vistex e MBT, esse método aparece com performance praticamente similar aos melhores métodos, com pouca variação. Na base Vistex, o método integrativo CLBP aparece com o maior desempenho médio, seguido pelo método RMD, com diferença de apenas 0.1%. Além disso, os métodos que aparecem no top-5 dessa base tem performance muito similares, variando entre 99.7% e 99.9%. Nesse caso, o desafio dessa base está praticamente superado, mesmo com a variação de espaços de cor, tornando difícil avaliar com precisão a performance dos melhores métodos. Na base MBT, o método que aparece com o maior desempenho médio é o RC integrativo, seguido pelo método proposto RMD. No entanto, novamente a diferença entre as duas abordagens é pequena, de 0.2%, incluindo também os métodos CLBP e Gabor-oponente na mesma faixa de performance. A diferença maior nessa base é observada do quarto para o quinto melhor método, AlexNet, com 0.7% de diferença. De forma geral, é possível constatar que as redes convolucionais apresentam pouca robustez aos diferentes tipos de espaço de cor analisados. No entanto, na base CURET, essas redes superam os demais métodos da literatura, ficando atrás apenas dos métodos propostos RMD e RM. Por fim, os resultados apontam que o método RMD é o mais robusto dentre as abordagens comparadas, com desempenho consistente em todas as bases. Métodos que tiveram bom desempenho em alguns casos, como o CLBP e RC integrativo, não mantêm a performance nos outros casos. O mesmo acontece com as redes convolucionais, onde a rede AlexNet supera as demais nas quatro primeiras bases, mas é superada pelas redes mais profundas na base CURET.

4.4 Aplicação no Reconhecimento de imagens Microscópicas

Imagens multiespectrais não se restringem aos formatos mais usados de 3 canais tricromáticos. Existem imagens de áreas como astrofísica, sensoriamento remoto ou microscopia que captam respostas mais precisas dentro e fora do espectro visível, resultando em uma quantidade de canais muito maior. Nesse contexto o terceiro método proposto neste trabalho (RDSA), descrito no Capítulo 3, Seção 3.3, foca na caracterização desse tipo de dado. Nessa seção, são realizados experimentos com dados multiespectrais de 32 canais para avaliar o desempenho do método RDSA. Esses dados foram obtidos através de microscopia confocal por varredura a laser no Instituto de Física de São Carlos - USP.

Figura 30 – Desempenho médio dos métodos, definido pela média da acurácia obtida em cada um dos 4 espaços de cor considerados (RGB, LAB, HSV e $I_1I_2I_3$). São mostrados os 5 melhores métodos em cada base utilizando o classificador LDA e validação cruzada L10.

Fonte: Elaborada pelo autor.

4.4.1 Plasticidade Foliar do Jacaranda Caroba

Este experimento tem como objetivo analisar a possibilidade de usar a espécie Jacaranda Caroba como um bio-sensor de poluentes através da observação de sua estrutura foliar. A ideia é identificar na plasticidade da planta sinais da qualidade do ar em que o espécime habita. O Jacaranda Caroba é escolhido por ser comumente utilizado na arborização de cidades e devido à sua presença em diversos estados do Brasil como Amazonas, Tocantins, Goiás, Mato Grosso, Minas Gerais, São Paulo, Rio de Janeiro e Espírito Santo. Os espécimes são divididos em 6 grupos, semeados e cultivados em ambiente controlado, onde 5 grupos são expostos ao borrifamento de fluoreto de potássio (KF) por um período de 60 dias. Cada grupo de espécimes recebeu uma concentração diferente da substância: $5mg/L$, $15mg/L$, $30mg/L$, $60mg/L$ e $1162mg/L$. O sexto grupo não foi exposto à substância, e é usado como grupo de controle. Após o período de exposição, as análises são feitas utilizando entre 4 e 5 folhas de 6 espécimes de cada grupo. Imagens multiespectrais são obtidas de cada folha pelo microscópio confocal à laser Zeiss LSM 780, do Instituto de Física de São Carlos - USP. No total são 148 imagens, sendo 25 para os grupos de controle e o exposto a KF $5mg/L$, 26 para o grupo de $15mg/L$ e 24 para o grupo de $30mg/L$, $60mg/L$ e $1162mg/L$.⁹¹

O microscópio considerado no experimento utiliza a técnica de microscopia por fluorescência, ou seja os espectros captados correspondem à emissão de luz correspondente à energia restante dos elétrons do material após conversão interna, uma vez excitados por um feixe de laser de varredura. Como diferentes tipos de materiais possuem uma configuração de nuvem de elétrons distinta, em função de sua composição química e arranjo molecular, tem-se então a emissão de comprimentos de ondas distintos. Esse microscópio registra imagens de tamanho $2048 \times 2048 \times 32$ que captam comprimentos de onda entre $400nm$ e $700nm$ divididos entre os 32 canais, onde cada pixel armazena valores de 16 *bits*, ou seja entre $[0, 65535]$. De acordo com os dados obtidos, são então aplicadas técnicas de inteligência artificial para construir um modelo que consiga distinguir entre os diferentes grupos do experimento. Isso é feito para determinar se é possível medir o nível de alteração na estrutura foliar do Jacaranda Caroba conforme a qualidade do ar de seu *habitat* muda. Normalmente esses dados são convertidos em uma imagem colorida comum de 3 canais com pixels (RGB) de 8 *bits* através de médias dos espectros, técnica que reduz a informação de cada pixel multiespectral na ordem de $1,25 \cdot 10^{145}\%$. Isso é feito para que os pesquisadores possam observar a imagem em um monitor comum e então selecionar regiões de interesse, analisar os espectros das regiões, etc. Na Figura 31 são mostradas imagens convertidas para RGB de amostras de espécimes de Jacaranda Caroba de cada grupo do experimento.

Considerando a significativa perda de informação em relação à conversão das imagens multiespectrais para RGB, o método proposto de RDSA é desenvolvido para analisar os dados originais em sua totalidade. O método gera um conjunto de descritores para cada

Figura 31 – Amostras da estrutura foliar de espécimes de Jacaranda Caroba obtidas por microscopia por fluorescência. Cada espécime foi submetido à doses diárias de fluoreto de potássio, como indicado em cada imagem, por um período de 60 dias (o grupo de controle não recebeu dose alguma). As imagens, originalmente com 32 canais espectrais, são convertidas para o formato tradicional RGB.

Fonte: Elaborada pelo autor.

imagem, que então são usados na classificação dos diferentes grupos do experimento. O único parâmetro do método é o conjunto de raios utilizado na modelagem das redes. A Figura 32 mostra seu desempenho utilizando os raios individuais $r = 1, 2, 3, 4, 5, 6$, ou sua combinação $R = \{1, 2, 3, 4, 5, 6\}$, onde n corresponde ao raio ou quantidade de raios usada. É possível observar que os maiores resultados são obtidos com a combinação dos raios, em ambos os casos, comportamento também observado na performance dos métodos propostos para textura colorida (RM e RMD).

Os resultados obtidos pelo método de RDSA são comparados com os outros métodos da literatura descritos anteriormente. Para realizar a caracterização dos dados microscópicos usando esses métodos, alguns pontos devem ser consideradas. Os métodos integrativos foram desenvolvidos para a aplicação em imagens coloridas comuns de 3 canais, portanto sua aplicação nos 32 canais se torna inviável devido à quantidade de características geradas, como por exemplo seriam 8192 características para o método LPQ.

Figura 32 – Taxa de acerto usando o método proposto de Redes complexas Direcionadas de Similaridade Angular (RDSA) na base de dados microscópicas do experimento com o espécime Jacaranda Caroba. Foram utilizados o classificador LDA e os esquemas de validação cruzada L1O (a) e 10-pastas (b).

Fonte: Elaborada pelo autor.

Nesse contexto e considerando o foco desses métodos, seus descritores são obtidos a partir das imagens convertidas para RGB. Da mesma forma, as redes convolucionais utilizadas tem um padrão de entrada fixo de tamanho, e aceitam apenas imagens de 3 canais. Nesse caso, as imagens convertidas para RGB, de tamanho 2048x2048, são redimensionadas para o tamanho definido de cada rede convolucional. Os resultados de cada método são mostrados na Tabela 8, utilizando o classificador LDA e os esquemas de validação cruzada L1O e 10-pastas. Como é possível observar, o método proposto supera os demais, obtendo 91.9% e 90.5(± 1.1)% em cada esquema de validação, respectivamente. O segundo método com o melhor desempenho é a rede convolucional VGG16, com 81.8% e 87.2(± 1.1)% de acurácia. Nesse problema, os métodos LPQ e Gabor-oponente obtém performance superior aos demais métodos integrativos e redes convolucionais, e os menores resultados são obtidos pela rede InceptionV3 e pelo método RC integrativo.

Para entender melhor a performance de cada método, a Figura 33 mostra a acurácia obtida em cada classe do problema. Como é possível observar, o método proposto RDSA obtém as melhores taxas de acerto nas classes 5mg/L e 15mg/L, e empata com a rede VGG16 na classe 1162mg/L. Na classe controle, todos os métodos funcionam de forma similar, e as classes 30mg/L e 60mg/L se mostraram as mais difíceis no geral. Nessas classes, os métodos LPQ e Gabor-oponente superam os demais em cada caso, respectivamente, porém sua performance oscila em relação às outras classes, o que faz sua acurácia média cair. De forma geral, a maioria dos métodos demonstra instabilidade de classe para classe, enquanto que o método proposto se mostrou o mais estável, com a maior acurácia média.

Tabela 8 – Performance de diferentes métodos na classificação de folhas da espécie Jacaranda Caroba com diferentes níveis de poluição. É utilizado o classificador LDA e os esquemas de validação cruzada L1O e 10-pastas.

Método	LDA	
	L1O	10-pastas
<i>Hist</i> ₃₂	37.2	40.7(±2.2)
LPQ	82.4	84.9(±1.4)
CLBP	75.7	79.2(±1.9)
Gabor-oponente	81.1	84.7(±2.3)
RC integrativo	66.9	68.1(±3.0)
AlexNet	73.0	75.1(±3.0)
VGG16	81.8	87.2(±2.2)
VGG19	74.3	75.8(±2.2)
InceptionV3	63.5	67.1(±2.4)
ResNet50	71.6	72.2(±2.2)
ResNet101	70.9	74.6(±2.2)
RDSA	91.9	90.5(±1.1)

Fonte: Elaborada pelo autor.

Esses resultados também apontam que, no geral, é possível diferenciar diferentes espécimes de Jacaranda Caroba, expostos a diferentes condições, analisando apenas aspectos visuais de sua estrutura microscópica foliar. Em conjunto com o método proposto, isso demonstra a eficácia do uso dessa planta como bio-sensor de poluentes do ar e a capacidade do método RDSA de captar essas características, presentes na informação multiespectral da amostra, com cerca de 91% de precisão. Em comparação aos outros métodos que operam nas imagens RGB, é notável que o método proposto capta informações importantes para a caracterização que estão presentes na informação multiespectral, melhorando o desempenho.

Figura 33 – Acurácia dos métodos na caracterização de cada uma das 6 classes do problema de análise de plasticidade foliar do Jacaranda Caroba. Foram utilizados o classificador LDA e o esquema de validação cruzada L1O.

Fonte: Elaborada pelo autor.

5 CONCLUSÃO

Nesse trabalho foram estudadas formas de analisar imagens multiespectrais a partir de RCs e técnicas baseadas na teoria do processamento oponente de cor. São introduzidas duas novas abordagens para a caracterização de imagens coloridas de três canais: Rede Multicamada (RM) e Rede Multicamada Direcionada (RMD). O primeiro método foi desenvolvido como uma extensão de trabalhos anteriores de redes em imagens monocromáticas, onde é proposta uma rede multicamada que conecta todos os canais da imagem. Esse método utiliza um conjunto de raios e um conjunto de limiares de corte como parâmetro, e os descritores são obtidos analisando a estrutura da rede em termos de conexões intra e entre canais. Nos trabalhos anteriores, a definição dos limiares é feita de forma empírica e extensiva, portanto é proposta uma nova técnica para estimar valores de limiar de forma automática e adaptativa, que se mostrou eficiente. O segundo método, RMD, é desenvolvido com o intuito de aprimorar o método anterior e reduzir a quantidade de parâmetros. Portanto é abordado o conceito de redes direcionadas, que eliminam a necessidade do uso de limiares de corte, ou seja, esse método utiliza apenas um conjunto de raios. Para a caracterização da rede, a direção das arestas é usada para calcular o grau de entrada e a força de entrada e saída de vértices, analisando conexões intra e entre canais. Ambos os métodos podem ser aplicados em qualquer tipo de espaço de cor de até três canais, desde que os valores de cada canal sejam normalizados para um mesmo intervalo.

Foram realizados diversos experimentos de classificação para validar os métodos propostos em bases conhecidas da literatura de textura colorida. Especificamente, foram usadas as bases RGB Usptex, Vistex, Outex13, MBT e CURET, que combinadas oferecem imagens com diversas variações geométricas e espectrais, permitindo avaliar diferentes propriedades de cada método. Também foram considerados 12 abordagens da literatura para comparação de resultados, incluindo métodos tradicionais de análise de textura aplicados de forma integrativa, métodos baseados em grafos e RC e redes neurais convolucionais profundas. Na base Usptex, os métodos que aparecem com os maiores resultados são, respectivamente, a rede convolucional ResNet50 e o método proposto RMD, com 99.7% de acurácia. Na base Vistex, vários métodos obtêm um desempenho acima de 99.8%, como as redes convolucionais AlexNet e VGG16, o método Gabor-opponente, CLBP, e os métodos propostos RM e RMD. Na base Outex13, ambos os métodos propostos apresentam os melhores resultados, com 96.0% obtido pelo método RMD, uma margem significativa de diferença em relação aos outros métodos. Na base MBT, os métodos com os maiores resultados são RC integrativo e o método proposto RMD, com resultados em torno de 98.6%. Na base CURET, aparecem o método proposto RMD e a rede convolucional ResNet101 com as melhores taxas de acerto, de 98.8%. Em linhas gerais, o método

proposto RMD mostrou os melhores resultados em diferentes cenários, enquanto que as redes convolucionais e os métodos integrativos oscilam. Por exemplo, nas bases Usptex e CURET, as redes convolucionais superam os métodos integrativos, e tem performance semelhante ao método RMD. Na base MBT o oposto acontece, onde os métodos integrativos tem o melhor desempenho, juntamente com o método proposto RMD. Na base Outex13, o desempenho das redes convolucionais é superado pelo método RC integrativo, e o método RMD supera ambos. Além disso, é possível observar que a profundidade das redes convolucionais impacta no desempenho em alguns casos. Na maioria das bases as menores redes consideradas, AlexNet e VGG16, tem desempenho relativamente maior que a rede mais profunda ResNet101, porém o cenário se inverte na base CURET, que tem a maior quantidade de imagens.

Para analisar o desempenho dos métodos em diferentes espaços de cor, foram realizados experimentos de classificação considerando os modelos LAB, HSV e $I_1I_2I_3$. No geral, as redes convolucionais demonstraram alta sensibilidade aos diferentes espaços de cor. Esse comportamento é esperado, pois essas redes foram treinadas e ajustadas para imagens RGB. Em relação ao desempenho médio dos métodos para os diferentes espaços de cor, as abordagens que aparecem com a maior robustez são RMD, CLBP, Gabor-oponente, RC integrativo e RM. O método proposto RMD apresenta o maior desempenho médio nas bases Usptex, Outex13 e CURET, casos em que a diferença para os outros métodos é considerável. Na base Vistex, o método CLBP aparece com a maior média, no entanto os top-5 métodos tem performance relativamente próximas, acima de 99.6%. Na base MBT, o método RC integrativo obtém o maior desempenho médio, com 99.6%, e o método proposto RMD aparece com 99.4%. Esses resultados corroboram com a alta robustez do método RMD a diferentes espaços de cor, em diferentes cenários.

Uma nova abordagem também é proposta para a análise de imagens multiespectrais, com muitos canais, denominada Rede Direcionada de Similaridade Angular (RDSA). Essa técnica analisa aspectos locais espaciais considerando todos os canais espectrais da imagem, enquanto mantém uma complexidade computacional menor que os métodos RM e RMD. Nessa rede, os pixels multiespectrais são considerados como vetores, e as arestas descrevem a similaridade do cosseno entre pares de pixel. A direção da aresta aponta para o pixel de maior valor absoluto, isso é o somatório das intensidades de cada canal. Esse método é aplicado em uma base de imagens de 32 canais espectrais obtidas por microscopia de fluorescência no Instituto de Física de São Carlos - USP. A base representa um experimento de plasticidade foliar para determinar a usabilidade da espécie Jacaranda Caroba como bio-sensor da qualidade do ar. Diferentes espécimes submetidos a diferentes condições são separados em 6 classes distintas, e as imagens de microscopia captam os padrões estruturais de suas folhas. O método proposto RDSA obtém 91.9% e 90.5(± 1.1)% de taxa de acerto nessa base utilizando os esquemas de validação cruzada L10 e 10-pastas. Esse resultado representa uma diferença significativa em relação ao método com a segunda

melhor taxa de acerto, a rede VGG16 com 81.8% e $87.2(\pm 2.2)\%$.

Os resultados obtidos durante o desenvolvimento desse trabalho de mestrado corroboram que o uso de RC na análise de imagens multiespectrais proporciona resultados significantes se comparados ao estado da arte. Além de superar as redes convolucionais em vários casos, o método proposto não necessita de *hardware* específico (GPU) para calcular os descritores propostos em tempo aceitável, da forma como essas redes necessitam para calcular convoluções com eficácia. A baixa complexidade computacional dos métodos propostos permitem que isso seja feito em tempo similar à métodos comuns da literatura de análise de textura. Nesse contexto, acreditamos que essa abordagem deve ser mais investigada nesta área de aplicação. Como sugestões para aprimoramento das técnicas propostas nesse trabalho, novas métricas topológicas de redes devem ser investigadas, podendo melhorar o desempenho. Além disso, essas técnicas podem ser estendidas para outras áreas como segmentação, detecção e descrição de pontos de interesse, etc.

REFERÊNCIAS

- 1 MANJUNATH, B. S. et al. Color and texture descriptors. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 11, n. 6, p. 703–715, 2001.
- 2 MARR, D. **Vision**: a computational investigation into the human representation and processing of visual information. San Francisco: WH Freeman and Company, 1982. v. 1.
- 3 MALAMAS, E. N. et al. A survey on industrial vision systems, applications and tools. **Image and Vision Computing**, v. 21, n. 2, p. 171–188, 2003.
- 4 ZHU, C.; YANG, X. Study of remote sensing image texture analysis and classification using wavelet. **International Journal of Remote Sensing**, v. 19, n. 16, p. 3197–3203, 1998.
- 5 ROSSATTO, D. R. et al. Fractal analysis of leaf-texture properties as a tool for taxonomic and identification purposes: a case study with species from neotropical melastomataceae (miconieae tribe). **Plant Systematics and Evolution**, v. 291, n. 1-2, p. 103–116, 2011.
- 6 CHICKLORE, S. et al. Quantifying tumour heterogeneity in 18f-fdg pet/ct imaging by texture analysis. **European Journal of Nuclear Medicine and Molecular Imaging**, v. 40, n. 1, p. 133–140, 2013.
- 7 HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. **IEEE Transactions on Systems, Man and Cybernetics**, n. 6, p. 610–621, 1973. doi.org/10.1109/tsmc.1973.4309314.
- 8 VARMA, M.; GARG, R. **Locally invariant fractal features for statistical texture classification**. 2007. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4408876>>. Acesso em: 23 jan. 2017.
- 9 CHALUMEAU, T. et al. **Optimized texture classification by using hierarchical complex network measurements**. 2006. Disponível em: <<https://doi.org/10.1117/12.655592>>. Acesso em: 12 fev. 2017.
- 10 _____. Complex networks: application for texture characterization and classification. **Electronic Letters on Computer Vision and Image Analysis**, v. 7, n. 3, p. 93–100, 2009.
- 11 BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Texture analysis and classification: a complex network-based approach. **Information Sciences**, v. 219, p. 168–180, 2013. doi.org/10.1016/j.ins.2012.07.003.
- 12 GONÇALVES, W. N.; MACHADO, B. B.; BRUNO, O. M. A complex network approach for dynamic texture recognition. **Neurocomputing**, v. 153, p. 211–220, 2015. doi.org/10.1016/j.neucom.2014.11.034.
- 13 DRIMBAREAN, A.; WHELAN, P. Experiments in colour texture analysis. **Pattern Recognition Letters**, v. 22, n. 10, p. 1161 – 1167, 2001.

- 14 MÄENPÄÄ, T.; PIETIKÄINEN, M. Classification with color and texture: jointly or separately? **Pattern Recognition**, v. 37, n. 8, p. 1629–1640, 2004.
- 15 BIANCONI, F. et al. Theoretical and experimental comparison of different approaches for color texture classification. **Journal of Electronic Imaging**, v. 20, n. 4, p. 043006–043006, 2011.
- 16 CERNADAS, E. et al. Influence of normalization and color space to color texture classification. **Pattern Recognition**, v. 61, p. 120–138, 2017. doi.org/10.1016/j.patcog.2016.07.002.
- 17 BROWET, A.; ABSIL, P.-A.; DOOREN, P. V. **Community detection for hierarchical image segmentation**. 2011. Disponível em: <https://www.researchgate.net/profile/Arnaud_Browet/publication/221016951_Community_Detection_for_Hierarchical_Image_Segmentation>. Acesso em: 3 fev. 2017.
- 18 HU, H. et al. **Multislice modularity optimization in community detection and image segmentation**. 2012. Disponível em: <<https://arxiv.org/pdf/1211.7180.pdf>>. Acesso em: 20 fev. 2017.
- 19 SCABINI, L. F.; GONÇALVES, W. N.; JR, A. A. C. **Texture analysis by bag-of-visual-words of complex networks**. 2015. Disponível em: <https://www.researchgate.net/publication/283015227_Texture_Analysis_by_Bag-Of-Visual-Words_of_Complex_Networks>. Acesso em: 20 jan. 2017.
- 20 GONÇALVES, W. N. et al. Texture recognition based on diffusion in networks. **Information Sciences**, v. 364, p. 51–71, 2016. doi.org/10.1016/j.ins.2016.04.052.
- 21 JUNIOR, J. J. d. M. S.; CORTEZ, P. C.; BACKES, A. R. Color texture classification using shortest paths in graphs. **IEEE Transactions on Image Processing**, v. 23, n. 9, p. 3751–3761, 2014.
- 22 XIA, W.; WANG, B.; ZHANG, L. Band selection for hyperspectral imagery: a new approach based on complex networks. **IEEE Geoscience and Remote Sensing Letters**, v. 10, n. 5, p. 1229–1233, 2013.
- 23 SUI, Y. et al. Complex network modeling of spectral remotely sensed imagery: a case study of massive green algae blooms detection based on modis data. **Physica A: statistical mechanics and its applications**, v. 464, p. 138–148, 2016.
- 24 LUQUE, B. et al. Horizontal visibility graphs: exact results for random time series. **Physical Review E**, v. 80, n. 4, p. 046103, 2009.
- 25 JULESZ, B. Visual pattern discrimination. **IRE Transactions on Information Theory**, v. 8, n. 2, p. 84–92, 1962.
- 26 JULESZ, B. et al. Inability of humans to discriminate between visual textures that agree in second-order statistics—revisited. **Perception**, v. 2, n. 4, p. 391–405, 1973.
- 27 CASANOVA, D. **Identificação de espécies vegetais por meio da análise de textura foliar**. 2008. 113 p. Dissertação (Mestrado) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2008.

- 28 GONÇALVES, W. N. **Análise de texturas estáticas e dinâmicas e suas aplicações em biologia e nanotecnologia**. 2013. 189 p. Tese (Doutorado em Ciências) — Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.
- 29 GONÇALVES, W. N.; SILVA, J. de A.; BRUNO, O. M. **A rotation invariant face recognition method based on complex network**. 2010. Disponível em: <https://www.researchgate.net/publication/220843037_A_Rotation_Invariant_Face_Recognition_Method_Based_on_Complex_Network>. Acesso em: 10 fev. 2017.
- 30 ZHANG, J.; TAN, T. Brief review of invariant texture analysis methods. **Pattern Recognition**, v. 35, n. 3, p. 735–747, 2002.
- 31 OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, v. 29, n. 1, p. 51–59, 1996.
- 32 OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 7, p. 971–987, 2002.
- 33 OJANSIVU, V.; HEIKKILÄ, J. **Blur insensitive texture classification using local phase quantization**. 2008. Disponível em: <<http://www.ee.oulu.fi/research/imag/mvg/files/pdf/ICISP08.pdf>>. Acesso em: 18 jan. 2017.
- 34 GUO, Z.; ZHANG, L.; ZHANG, D. A completed modeling of local binary pattern operator for texture classification. **IEEE Transactions on Image Processing**, v. 19, n. 6, p. 1657–1663, 2010.
- 35 JAIN, A. K.; FARROKHNI, F. Unsupervised texture segmentation using gabor filters. **Pattern Recognition**, v. 24, n. 12, p. 1167–1186, 1991.
- 36 HOANG, M. A.; GEUSEBROEK, J.-M. et al. **Measurement of color texture**. 2002. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.5644&rep=rep1&type=pdf>>. Acesso em: 15 nov. 2016.
- 37 HOANG, M. A.; GEUSEBROEK, J.-M.; SMEULDERS, A. W. Color texture measurement and segmentation. **Signal Processing**, v. 85, n. 2, p. 265–275, 2005.
- 38 AZENCOTT, R.; WANG, J.-P.; YOUNES, L. Texture classification using windowed fourier filters. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 19, n. 2, p. 148–153, 1997.
- 39 BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Color texture analysis based on fractal descriptors. **Pattern Recognition**, v. 45, n. 5, p. 1984–1992, 2012.
- 40 STARR, C.; EVERS, C.; STARR, L. **Biology: concepts and applications without physiology**. Stanford: Cengage Learning, 2010.
- 41 WYSZECKI, G.; STILES, W. S. **Color science**. New York: Wiley, 1982. v. 8.
- 42 FOSTER, M. **A Text-book of Physiology**. Philadelphia: Lea Brothers & Company, 1895.

- 43 VANDENBROUCKE, N.; MACAIRE, L.; POSTAIRE, J.-G. Color image segmentation by pixel classification in an adapted hybrid color space. application to soccer image analysis. **Computer Vision and Image Understanding**, v. 90, n. 2, p. 190–216, 2003.
- 44 HAFNER, J. et al. Efficient color histogram indexing for quadratic form distance functions. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 17, n. 7, p. 729–736, 1995.
- 45 GEGENFURTNER, K. R.; RIEGER, J. Sensory and cognitive contributions of color to the recognition of natural scenes. **Current Biology**, v. 10, n. 13, p. 805 – 808, 2000.
- 46 WICHMANN, F. A.; SHARPE, L. T.; GEGENFURTNER, K. R. The contributions of color to recognition memory for natural scenes. **Journal of Experimental Psychology: learning, memory, and cognition**, v. 28, n. 3, p. 509, 2002.
- 47 ZHANG, X.; WANDELL, B. A. A spatial extension of cielab for digital color-image reproduction. **Journal of the Society for Information Display**, v. 5, n. 1, p. 61–63, 1997.
- 48 ROSENFELD, A.; CHENG-YE, W.; WU, A. Y. Multispectral texture. **IEEE Transactions on Systems, Man and Cybernetics**, v. 12, p. 79–84, 1980.
- 49 JAIN, A.; HEALEY, G. A multiscale representation including opponent color features for texture recognition. **IEEE Transactions on Image Processing**, v. 7, n. 1, p. 124–128, 1998.
- 50 MIRMEHDI, M.; PETROU, M. Segmentation of color textures. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 2, p. 142–159, 2000.
- 51 JACOB, I. J.; SRINIVASAGAN, K.; JAYAPRIYA, K. Local oppugnant color texture pattern for image retrieval system. **Pattern Recognition Letters**, v. 42, p. 72–78, 2014.
- 52 ERDOS, P.; RÉNYI, A. On random graphs i. **Publicationes Mathematicae**, v. 6, p. 290–297, 1959.
- 53 _____. On the evolution of random graphs. **Publication of the Mathematical Institute of the Hungarian Academy of Sciences**, v. 5, p. 17–61, 1960.
- 54 BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, n. 5439, p. 509–512, 1999.
- 55 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. **Nature**, v. 393, n. 6684, p. 440–442, 1998.
- 56 BIANCONI, G.; BARABÁSI, A.-L. Bose-Einstein condensation in complex networks. **Physical Review Letters**, v. 86, n. 24, p. 5632, 2001.
- 57 KITSACK, M. et al. Identification of influential spreaders in complex networks. **Nature Physics**, v. 6, n. 11, p. 888–893, 2010.
- 58 BARABASI, Z. O. A.-L. Network biology: understanding the cell’s functional organization. **Nature Reviews Genetics**, v. 5, n. 2, p. 101–113, 2004.

- 59 GUIMERA, R.; AMARAL, L. A. N. Functional cartography of complex metabolic networks. **Nature**, v. 433, n. 7028, p. 895–900, 2005.
- 60 MACHADO, B. B. et al. A complex network approach for nanoparticle agglomeration analysis in nanoscale images. **Journal of Nanoparticle Research**, v. 19, n. 2, p. 65, 2017.
- 61 RUBINOV, M.; SPORNS, O. Complex network measures of brain connectivity: uses and interpretations. **Neuroimage**, v. 52, n. 3, p. 1059–1069, 2010.
- 62 NEWMAN, M. E. J.; PARK, J. Why social networks are different from other types of networks. **Physical Review E**, v. 68, p. 036122, 2003.
- 63 COSTA, L. d. F. et al. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, v. 60, n. 3, p. 329–412, 2011.
- 64 BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. A complex network-based approach for boundary shape analysis. **Pattern Recognition**, v. 42, n. 1, p. 54–67, 2009.
- 65 SCABINI, L. F. et al. Angular descriptors of complex networks: a novel approach for boundary shape analysis. **Expert Systems with Applications**, v. 89, p. 362 – 373, 2017. doi.org/10.1016/j.eswa.2017.08.009.
- 66 GONÇALVES, W. N. et al. Texture descriptor based on partially self-avoiding deterministic walker on networks. **Expert Systems with Applications**, v. 39, n. 15, p. 11818–11829, 2012.
- 67 CRIADO, R.; ROMANCE, M.; SANCHEZ, A. Interest point detection in images using complex network analysis. **Journal of Computational and Applied Mathematics**, v. 236, n. 12, p. 2975–2980, 2012.
- 68 ZHANG, Y. et al. **A complex network-based approach for interest point detection in images**. 2012. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6264291>>. Acesso em: 30 abr. 2017.
- 69 WU, Z.; LU, X.; DENG, Y. Image edge detection based on local dimension: a complex networks approach. **Physica A: statistical mechanics and its applications**, v. 440, p. 9–18, 2015. doi.org/10.1016/j.physa.2015.07.020.
- 70 COSTA, L. d. F. et al. Characterization of complex networks: a survey of measurements. **Advances in Physics**, v. 56, n. 1, p. 167–242, 2007.
- 71 RIPLEY, B. D. **Pattern recognition and neural networks**. Cambridge: Cambridge University Press, 1996.
- 72 BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Color texture analysis based on fractal descriptors. **Pattern Recognition**, v. 45, n. 5, p. 1984–1992, 2012.
- 73 PICKARD, R. et al. Vistex vision texture database. Disponível em: <<http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>>. Acesso em: 2 nov. 2016.
- 74 OJALA, T. et al. **Outex-new framework for empirical evaluation of texture analysis algorithms**. 2002. Disponível em: <<https://pdfs.semanticscholar.org/4bf0/ca857419da52bc82bcd6197675c42ed7a638.pdf>>. Acesso em: 16 fev. 2017.

- 75 ABDELMOUNAIME, S.; DONG-CHEN, H. New brodatz-based image databases for grayscale color and multiband texture analysis. **International Scholarly Research Notices Machine Vision**, 2013. doi.org/10.1155/2013/876386.
- 76 DANA, K. J. et al. Reflectance and texture of real-world surfaces. **ACM Transactions on Graphics**, v. 18, n. 1, p. 1–34, 1999.
- 77 KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in Neural Information Processing Systems**, v. 25, n. 2, p. 1097–1105, 2012.
- 78 SIMONYAN, K.; ZISSERMAN, A. **Very deep convolutional networks for large-scale image recognition**. 2014. Disponível em: <<https://arxiv.org/pdf/1409.1556/>>. Acesso em: 10 jan. 2017.
- 79 SZEGEDY, C. et al. **Rethinking the inception architecture for computer vision**. 2016. Disponível em: <https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.pdf>. Acesso em: 20 jan. 2017.
- 80 _____. **Going deeper with convolutions**. 2015. Disponível em: <http://openaccess.thecvf.com/content_cvpr_2015/papers/Szegedy_Going_Deeper_With_2015_CVPR_paper.pdf>. Acesso em: 22 jan. 2017.
- 81 HE, K. et al. **Deep residual learning for image recognition**. 2016. Disponível em: <http://openaccess.thecvf.com/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf>. Acesso em: 23 jan. 2017.
- 82 RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. **International Journal of Computer Vision**, v. 115, n. 3, p. 211–252, 2015.
- 83 LIN, M.; CHEN, Q.; YAN, S. **Network in network**. 2013. Disponível em: <<https://arxiv.org/pdf/1312.4400.pdf>>.
- 84 CIMPOI, M. et al. **Describing textures in the wild**. 2014. Disponível em: <http://openaccess.thecvf.com/content_cvpr_2014/papers/Cimpoi_Describing_Textures_in_2014_CVPR_paper.pdf>. Acesso em: 25 jan. 2017.
- 85 CIMPOI, M.; MAJI, S.; VEDALDI, A. **Deep filter banks for texture recognition and segmentation**. 2015. Disponível em: <https://www.cv-foundation.org/openaccess/content_cvpr_2015/app/2B_053.pdf>. Acesso em: 26 jan. 2017.
- 86 CIMPOI, M. et al. Deep filter banks for texture recognition, description, and segmentation. **International Journal of Computer Vision**, v. 118, n. 1, p. 65–94, 2016.
- 87 BELLMAN, R. **Dynamic Programming**. Princeton: Princeton University Press, 1957.
- 88 KEOGH, E.; MUEEN, A. Curse of dimensionality. In: **Encyclopedia of machine learning and data mining**. Berlin: Springer, 2017. p. 314–315.

- 89 HUANG, R. et al. **Solving the small sample size problem of lda**. 2002. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1047787>>. Acesso em: 30 abr. 2017.
- 90 THARWAT, A. et al. Linear discriminant analysis: a detailed tutorial. **AI Communications**, v. 30, n. 2, p. 169–190, 2017.
- 91 FALVO, M. **Método de mapeamento espaço-espectral em imagens multi-espectrais e sua aplicação em tecidos vegetais**. 2016. 155 p. Tese (Doutorado em Ciências) — Instituto de Física de São carlos, Universidade de São Paulo, São Carlos, 2016.