

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola de Engenharia
Programa de Pós-Graduação em Engenharia Elétrica

Antônio Hamilton Magalhães

SÍNTESE DE IMAGENS HIPERESPECTRAIS A PARTIR DE CÂMERAS RGB
APLICADA À DETECÇÃO DE INCÊNDIOS FLORESTAIS

Belo Horizonte

2024

Antônio Hamilton Magalhães

**SÍNTESE DE IMAGENS HIPERESPECTRAIS A PARTIR DE CÂMERAS RGB
APLICADA À DETECÇÃO DE INCÊNDIOS FLORESTAIS**

Tese apresentada ao Programa de Pós-Graduação
em Engenharia Elétrica da Universidade Federal de
Minas Gerais, como requisito parcial à obtenção do
título de Doutor em Engenharia Elétrica.

Área de concentração: Sistemas de Computação e
Telecomunicações

Linha de pesquisa: Inteligência Computacional

Orientador: Prof. Dr. Hermes Aguiar Magalhães

Coorientador: Prof. Dr. Hani Camille Yehia

Belo Horizonte

2024

M188s	Magalhães, Antônio Hamilton. Síntese de imagens hiperespectrais a partir de câmeras RGB aplicada à detecção de incêndios florestais [recurso eletrônico] / Antônio Hamilton Magalhães. – 2024. 1 recurso online (181 f. : il., color.) : pdf. Orientador: Hermes Aguiar Magalhães. Coorientador: Hani Camille Yehia. Tese (doutorado) – Universidade Federal de Minas Gerais, Escola de Engenharia. Inclui bibliografia. 1. Engenharia elétrica – Teses. 2. Incêndios florestais – Prevenção e controle – Teses. 3. Câmeras – Teses. 4. Redes neurais (Computação) – Teses. 5. Aprendizado do computador – Teses. I. Magalhães, Hermes Aguiar. II. Yehia, Hani Camille. III. Universidade Federal de Minas Gerais. Escola de Engenharia. IV. Título. CDU: 621.3(043)
-------	--



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

FOLHA DE APROVAÇÃO

"SÍNTESE DE IMAGENS HIPERESPECTRAIS A PARTIR DE CÂMERAS RGB APLICADA À DETECÇÃO DE INCÊNDIOS FLORESTAIS"

ANTÔNIO HAMILTON MAGALHÃES

Tese de Doutorado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Doutor em Engenharia Elétrica. Aprovada em 16 de abril de 2024. Por:

Prof. Dr. Hermes Aguiar Magalhães
Orientador - DELT (UFMG)

Prof. Dr. Hani Camille Yehia
Coorientador - DELT (UFMG)

Prof. Dr. Adriano Vilela Barbosa
DELT (UFMG)

Prof. Dr^a. Luciana Pedrosa Salles
DEE (UFMG)

Prof. Dr. Adriano Chaves Lisboa
(Gaia Soluções em Engenharia Ltda e CEFET-MG)

Prof. Dr. Ana Paula Marques Ramos
(UNESP)

Prof. Dr. Wesley Nunes Gonçalves
(UFMS)



Documento assinado eletronicamente por **Hermes Aguiar Magalhaes, Professor do Magistério Superior**, em 16/04/2024, às 19:04, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Hani Camille Yehia, Professor do Magistério Superior**, em 17/04/2024, às 12:32, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Luciana Pedrosa Salles, Professora do Magistério Superior**, em 17/04/2024, às 15:53, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Adriano Vilela Barbosa, Professor do Magistério Superior**, em 17/04/2024, às 18:33, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ana Paula Marques Ramos, Usuário Externo**, em 18/04/2024, às 14:15, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Adriano Chaves Lisboa, Usuário Externo**, em 19/04/2024, às 15:51, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Wesley Nunes Gonçalves, Usuário Externo**, em 24/04/2024, às 08:11, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3182949** e o código CRC **6F9696E8**.

Dedico este trabalho à minha mãe, Maria, o ser mais sublime que conheço.

Para minha esposa Wagna.

Para meus filhos Bárbara, Jean e Júlia.

Para Vanusa, que estará comigo para sempre.

Agradecimentos

"Jesus curou dez leprosos, mas somente um deles voltou para agradecer."(Lc 17,11-19)

Agradeço a Deus por sua ação misteriosa em minha vida que me conduziram, por seus desígnios, a este momento. A ti minha reverência pelo intrincado e complexo arcabouço da criação, da vida, da natureza, do universo com sutilezas e emaranhamentos que escapam à compreensão humana.

Agradeço à minha mãe que sempre foi o esteio da minha vida, por seus ensinamentos, por seu exemplo de coragem e fé inabaláveis, pela serenidade em todos os momentos, mesmo naqueles mais difíceis, pelos inúmeros sacrifícios que de várias formas me conduziram até aqui. Deus a abençoe sempre!

Agradeço à minha esposa Wagna, pelo amor, companheirismo e incentivo, e por sempre acreditar em minha capacidade.

Minha gratidão aos meus filhos, Bárbara, Jean e Júlia, que trouxeram o sentido de pertencimento e o significado amplo do amor em minha vida, e ao Anthony pela vida renovada.

Agradeço ao José Geraldo pelo exemplo e por me mostrar o caminho nos primeiros passos na ciência; à Celi pelo olhar otimista; à Mary por sempre trazer graça às coisas; ao Júnior pelo exemplo de serenidade e retidão; à Márcia pelo espírito livre; à Vanusa pelos muitos momentos de alegria, mesmo partindo tão cedo. Você está fazendo muita falta; à Rita por ser um alicerce moral e espiritual, e pela disposição em servir.

À Meg por sempre estar do meu lado, em todos esses anos.

Agradeço aos muitos amigos que, de uma forma ou outra, me ajudaram a chegar nesse momento. Agradeço à turma do Escalar, da CEL, Audiolab, Batik, Lucent, FITec, Clamper e aos amigos e colegas professores e funcionários da PUC Minas.

Agradeço aos colegas do CEFALA, da GAIA e do projeto CEMIG D619 pelo convívio e oportunidade de aprendizado.

Meu muito obrigado ao pessoal do BEMAD (Batalhão de Emergências Ambientais e Resposta a Desastres) do Corpo de Bombeiros Militares de Minas Gerais, nas pessoas do Capitão Warley Barbosa, do Capitão Rafael Lima e do Sargento Douglas Paulino pela oportunidade de participar dos eventos de queima prescrita que muito contribuíram para a composição da base de dados de imagens utilizados nesse trabalho. Ao Sargento Douglas Paulino agradeço também por sempre me notificar dos eventos de treinamento e MIF. Meu obrigado ao André, do Parque Estadual Serra Verde, e ao Henrique, do Parque Estadual do Rola Moça, pela partilha de informações e por viabilizar minha participação nos eventos de MIF.

Gostaria de agradecer ao Gibran, por me apresentar o Hani que me aceitou no projeto D619 e no Doutorado.

Agradeço ao Samuel, ao Geraldo e ao Manuel pela amizade e por serem referências para mim no campo das engenharias.

Por fim, o meu mais profundo agradecimento aos meus orientadores Prof. Dr. Hermes Aguiar Magalhães e Prof. Dr. Hani Camille Yehia, pelos ensinamentos, conhecimentos partilhados, pelas orientações e discussões técnicas nestes anos, que, para mim, foram de muito aprendizado, em todos os sentidos. Foram muitas conversas, e diversos assuntos, para além das questões científicas e tecnológicas, com leveza e bom humor. Obrigado pelo espírito de acolhimento, compreensão e cooperação. Vocês foram fundamentais, não poderia ter tido melhores orientadores. A vocês minha sincera gratidão.

“... pois a sabedoria é mais preciosa do que rubis; nada do que se possa desejar compara-se a ela”

Provérbios 8:11

“Infinita diversidade, em infinitas combinações.”

Surak

Resumo

A detecção precoce de incêndios florestais é crucial para salvar vidas, minimizar impactos ecológicos e ambientais e reduzir danos causados pelo fogo. Sistemas de detecção de incêndios baseados em aprendizado de máquina e no uso de imagens, fotográficas e de vídeo, têm sido desenvolvidos nos últimos anos com desempenho notável, mas ainda são susceptíveis a detecções incorretas em cenários específicos. O imageamento hiperespectral tem o potencial de prover dados adicionais que possibilitem aprimorar a detecção de incêndios e minimizar tais erros. Entretanto, câmeras hiperespectrais são dispendiosas, de manuseio complexo e frágeis. O objetivo desse trabalho foi desenvolver uma metodologia baseada em aprendizado de máquina para sintetizar imagens com alta resolução espectral a partir de imagens de câmeras RGB convencionais, de baixo custo, acessíveis, e robustas, cujos dados são espectralmente esparsos. O problema de sintetizar imagens hiperespectrais a partir de imagens RGB é do tipo malposto (*ill-posed*), com infinitas soluções. Entretanto, é possível utilizar restrições impostas pelas características de iluminação da cena e pelos elementos que a compõem para encontrar a solução viável. Para isso, supõe-se que dados de imagens hiperespectrais utilizados nesse trabalho residam em uma variedade (*manifold*) de baixa dimensionalidade intrínseca, contido em um espaço de alta dimensionalidade. Essa condição pode ser confirmada caso haja baixa frequência de metâmeros nas distribuições espectrais dos píxeis da imagem. Foram implementados e testados algoritmos para síntese de dados hiperespectrais utilizando técnicas de regressão linear, composição de agrupamentos e redes neurais do tipo *perceptron* multicamada. Em uma primeira etapa, foi utilizada uma base de imagens hiperespectrais disponibilizada pelo ICVL (*Interdisciplinary Computational Vision Laboratory*). Posteriormente, foram capturadas e processadas imagens de cenas urbanas, rurais, de chamas, de queimas prescritas e de incêndios florestais com uso de câmeras hiperespectrais e RGB. Essas imagens foram utilizadas para a construção de bases de dados de imagens utilizadas nessa pesquisa. A metodologia empregada propiciou a síntese de imagens hiperespectrais, tanto na faixa de luz visível, quanto na faixa do infravermelho próximo (NIR), a partir de imagens RGB obtidas por projeção de imagens hiperespectrais no espaço de um sensor hipotético, assim como as obtidas diretamente de sensores RGB convencionais. Um importante resultado foi a confirmação da identificação de um sinal na faixa NIR oriundo do decaimento de elétrons em átomos de potássio, que pode ser usado como um marcador para rastreamento de incêndios florestais.

Palavras-chave: aprendizado de máquina; rede neural; *clustering*; regressão linear, imagem hiperespectral; detecção de incêndio florestal.

Abstract

Early detection of wildfires is crucial for saving lives, minimizing ecological and environmental impacts, and reducing damage caused by wildfire. While wildfire detection systems utilizing machine learning and RGB images have shown notable performance in recent years, they still exhibit susceptibility to incorrect detections in specific scenarios. Hyperspectral imaging has the potential to provide additional data to improve fire detection and minimize such errors. However, hyperspectral cameras are expensive, complex to handle, and fragile. This work aimed at developing a methodology based on machine learning to synthesize high-spectral-resolution images from low-cost, accessible, and robust conventional RGB cameras whose data is spectrally sparse. The problem of synthesizing hyperspectral images from RGB images is ill-posed, with infinite solutions. However, it is possible to use constraints imposed by the scene's lighting characteristics and its elements to find the most realistic solution. It is assumed that hyperspectral image data resides in a low-dimensional manifold contained in a high-dimensional space. This condition can be confirmed if there is a low frequency of metamers in the pixel spectral distribution of the image. Algorithms for hyperspectral data synthesis were implemented and tested using linear regression techniques, clustering, and multilayer perceptron neural networks. In the initial stage, a hyperspectral image dataset provided by ICVL (Interdisciplinary Computational Vision Laboratory) was used. Subsequently, images of urban and rural scenes, flames, prescribed burns, and wildfires were captured and processed using hyperspectral and RGB cameras to compose the image datasets utilized in this research. The employed methodology enabled the synthesis of hyperspectral images in both the visible light and near-infrared (NIR) ranges, from RGB images projected onto the image sensor space. It also provided methods for hyperspectral images synthesis based on RGB images generated by the Specim IQ camera, as well as those produced by a conventional low-cost RGB camera. An important result of this research was to confirm the identification of a signal in the NIR range, originating from the decay of electrons in potassium atoms, which can be used as a marker for Wildfire tracking.

Keywords: machine learning; neural network; clustering; linear regression; hyperspectral image; wildfire detection.

Lista de Figuras

2.1	Arquitetura de um arranjo de sensores de imagem CMOS	45
2.2	Estrutura de um píxel ativo do tipo fotodiodo	45
2.3	Energia do elétron em função do momento linear na banda de valência	46
2.4	Curvas de coeficiente de absorção em função do comprimento de onda	47
2.5	Curva de profundidade de penetração em função do comprimento	48
2.6	Responsividade de diversos materiais semicondutores	49
2.7	Distorção radial	54
2.8	Distorção tangencial	54
2.9	Bandas espectrais consideradas na detecção de incêndios florestais.	58
2.10	Técnicas de sensoriamento hiperespectral com esquema de varredura de linha	60
2.11	Técnicas de sensoriamento hiperespectral com uso de filtro sintonizável	60
2.12	Técnicas de sensoriamento hiperespectral com uso de arranjo de difratores	61
2.13	Espectro de corpo negro a diferentes temperaturas	62
2.14	Assinatura espectral de fogo, grama e solo aquecido	63
2.15	Imagens de incêndio em diferentes comprimentos de onda (A) e (B)	64
2.16	Imagens de incêndio em diferentes comprimentos de onda (C) e (D)	64
2.17	Arquitetura de uma rede do tipo multilayer perceptron	69
2.18	Amostra de imagens renderizadas da base ICVL	71
3.1	Infográfico com as etapas da metodologia de pesquisa	77

3.2	Método de validação cruzada com $k_f = 5$ (<i>5-fold cross-validation</i>).	84
3.3	Preprocessamento com projeção de imagens hiperespectrais no espaço de um sensor (RGB)	88
3.4	Câmera Hiperespectral FX-10	89
3.5	Imagens monocromáticas de fogo em diversos comprimentos de onda (VIS e NIR) - FX-10	90
3.6	Operação de queima prescrita Parque Estadual Serra Verde	93
3.7	Estrutura mecânica para fixação das câmeras Specim IQ e Panasonic DMC-FZ5	94
3.8	Estrutura mecânica para fixação das câmeras Specim e Panasonic montada em um tripé	95
3.9	Calibração da câmera Specim IQ	97
3.10	Calibração da câmera Panasonic DMC-FZ5	98
3.11	Etapas de pré-processamento das imagens das câmeras RGB (a) e Specim IQ (b).	99
3.12	Amostra de imagem da base obtida com a montagem das câmeras Panasonic DMC-FZ5 e Specim	99
4.1	Autovalores em função dos fatores principais	102
4.2	Variância acumulada em função dos fatores associados a variáveis latentes	103
4.3	Diagrama de processo para recuperação do hiperespectro a partir de imagens RGB - Ridge	106
4.4	Métrica R^2 para cada combinação do conjunto de validação e fator de regularização	107
4.5	Diagrama de processo para recuperação do hiperespectro a partir de imagens RGB	109
4.6	Mapeamento dos centroides do espaço RGB para o espaço hiperespectral	110
4.7	Desempenho do modelo em função do número de agrupamentos - R^2 Score	111
4.8	Valor médio da métrica MARE, calculado para cada partição do conjunto de treinamento e de validação.	111
4.9	Valor médio da métrica MARE, calculado em função do número de agrupamentos	112

4.10	Diagrama de processo contendo uma rede neural	113
4.11	Arquitetura de redes MLP com duas e três camadas	113
4.12	Arquitetura de redes MLP com quatro, cinco e sete camadas	114
4.13	Arquitetura de redes MLP com nove e onze camadas	114
4.14	Curvas da função de custo para várias arquiteturas de redes neurais	115
4.15	Diagrama de processo contendo as etapas de processamento das imagens RGB e hiperespectrais, o algoritmo <i>minibatch K-Means</i>	116
4.16	Amostra de uma imagem hiperespectral do conjunto de teste - renderizada	117
4.17	Amostra de imagem hiperespectral, em tons de cinza, para 31 bandas espectrais.	118
4.18	Comparação das assinaturas espectrais de píxeis da imagem original e das ima- gens estimadas utilizando os quatro processos.	119
4.19	Comparação da distribuição da métrica MARE obtida utilizando os quatro pro- cessos de síntese hiperespectral.	121
4.20	Comparação da distribuição da métrica RMSE obtida utilizando os quatro pro- cessos de síntese hiperespectral.	122
4.21	Gráficos das métricas MARE para os algoritmos implementados	124
4.22	Gráficos das métricas RMSE para os algoritmos implementados	124
4.23	“Cubo hiperespectral” gerado com uso da câmera Specim FX-10	125
4.24	Imagem renderizada produzida pela câmera FX-10 e imagens monocromáticas em diversos comprimentos de onda	126
4.25	Amostra da base de dados de imagens capturadas com o uso da câmera FX-10	126
4.26	Espectros de píxeis em diversas imagens	127
4.27	Imagem obtida com iluminação por lâmpada halógena - sal	128
4.28	Distribuição de potência de uma lâmpada halógena, compensação da corrente de escuro, calibração radiométrica	129
4.29	Espectro de um píxel na região da chama de uma imagem capturada com a câmera Specim FX-10	130
4.30	“Cubo hiperespectral” Specim IQ	130

4.31	Conjunto de imagens RGB construído com o uso da câmera hiperespectral Specim IQ	131
4.32	Imagens de uma vela, RGB e em tons de cinza para diferentes comprimentos de onda	132
4.33	Imagem da chama de uma vela com píxeis designados	133
4.34	Espectro de píxeis de diferentes regiões da chama	133
4.35	Assinatura espectral de emissão de fogo, grama, fumaça, traje de bombeiro e solo vermelho	134
4.36	Imagens em formatos RGB e em tons de cinza relativos a uma operação de Manejo Integrado do Fogo (MIF)	135
4.37	Imagem de incêndio florestal e alguns píxeis selecionados	136
4.38	Espectro de píxeis selecionados de áreas distintas da cena	137
4.39	Curvas de resposta espectral para um filtro <i>Bayer</i>	138
4.40	Imagens hiperespectrais capturadas utilizando uma câmera Specim IQ RGB e em escala de cinza	139
4.41	Processo de síntese <i>Ridge</i> com imagens RGB por projeção com filtro <i>Bayer</i> . . .	140
4.42	Imagens hiperespectrais sintetizadas usando o método de Regressão	141
4.43	Comparação de espectros de píxeis selecionados em diferentes regiões da cena - síntese por <i>Ridge</i>	142
4.44	Processo de síntese de imagens hiperespectrais utilizando um modelo de rede neural do tipo MLP e filtro <i>Bayer</i>	143
4.45	Modelo de rede MLP utilizado para síntese de imagens hiperespectrais	143
4.46	Imagem hiperespectral sintetizada usando uma rede neural do tipo MLP com 12 camadas	144
4.47	Comparação de espectros obtidos a partir de píxeis escolhidos, situados em várias regiões da cena - síntese por MLP	145
4.48	Processo de síntese de imagens hiperespectrais utilizando MLP e <i>K-Means</i> . . .	146
4.49	Imagem hiperespectral sintetizada usando um método que combina agrupamento usando <i>K-Means</i> e <i>MLP</i>	147

4.50	Comparação de espectros obtidos de píxeis localizados em várias regiões da cena - síntese por MLP e <i>minibatch K-Means</i>	147
4.51	Comparação da distribuição da métrica MARE obtida utilizando os quatro processos de síntese hiperespectral.	149
4.52	Comparação da distribuição da métrica RMSE obtida utilizando os três processos de síntese hiperespectral.	150
4.53	Curvas de resposta de um filtro sintetizado com quatro bandas	152
4.54	Arquitetura de rede MLP para síntese de imagens hiperespectrais, a partir de imagens RGB-IR769	152
4.55	Imagens renderizadas RGB e em tons de cinza “sintetizadas” utilizando uma rede MLP de 12 camadas e filtro com 4 componentes	153
4.56	Comparação do espectro de píxeis contendo o pico de emissão de potássio	154
4.57	Amostra de imagens do conjunto de testes para validação do processo MLP-RGB-IR769	154
4.58	Comparação do espectro de píxeis sintetizados por MLP-RGB-IR769	155
4.59	Imagem Specim IQ e sintetizadas por MLP=RGB e MLP-RGB-IR769	156
4.60	Comparação do conteúdo espectral de um píxel da região de fogo	156
4.61	Diagrama em blocos do processo de síntese com base em imagens de uma câmera Panasonic DMC-FZ5	158
4.62	Função de custo na fase de treinamento de quatro arquiteturas MLP com cinco, sete, onze e quinze camadas	159
4.63	Arquitetura de rede neural utilizada na síntese com base em imagens de uma câmera Panasonic DMC-FZ5	160
4.64	Parâmetros da rede neural utilizada na síntese de imagens hiperespectrais	161
4.65	Comparação da função de custo na fase de treinamento a partir de imagens com e sem calibração	162
4.66	MARE médio para arquiteturas de redes MLP com 7, 11 e 15 camadas	162
4.67	MARE de cada imagem do conjunto de testes calculados para arquiteturas MLP com sete, onze e quinze camadas.	162

4.68	Amostras de imagens do conjunto de testes e correspondentes imagens sintetizadas - RGB Specim IQ	163
4.69	Comparação do espectro de píxeis de imagens sintetizadas com base em imagens de uma câmera Panasonic DMC-FZ5	164
4.70	Amostras de imagens do conjunto de testes para o processo utilizando imagens do sensor RGB da câmera Specim IQ	165
4.71	Comparação do espectro de píxeis de imagens obtidas por síntese com base em imagens RGB da câmera Specim IQ	166
A.1	Metadados gerados pelo software de controle da câmera FX-10	179

Lista de Tabelas

3.1	Características da câmera Specim FX-10	89
3.2	Características da câmera Specim IQ	91
3.3	Características da câmera Panasonic DMC-FZ5	92
4.1	Comparação de desempenho dos processos pela métrica MARE, aplicada ao conjunto de testes	120
4.2	Comparação de desempenho dos processos pela métrica RMSE aplicada ao conjunto de testes.	122
4.3	Comparação de desempenho dos processos com base no tempo de processamento em segundos (Hardware: Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz, 196GB RAM)	123
4.4	Análise comparativa dos três processos utilizando a métrica MARE aplicada ao conjunto de testes com 44 imagens.	148
4.5	Análise comparativa dos três processos utilizando a métrica RMSE aplicada ao conjunto de testes com 44 imagens.	149
4.6	MARE do conjunto de validação nas arquiteturas MLP com cinco, sete, onze e quinze camadas. Validação cruzada com 5 partições.	159
4.7	MARE para o conjunto de teste nas arquiteturas MLP com 5, 7, 11 e 15 camadas.	160

Lista de Símbolos

Os símbolos em negrito correspondem a vetores, matrizes ou tensores, os demais são escalares ou funções.

.	Aplicado a escalares, indica uma multiplicação
.	Aplicado a vetores, ou matrizes, indica um produto interno.
A	Ampere - unidade de corrente elétrica
B	<i>Blue</i> - Cor azul de um <i>píxel RGB</i>
B_z	Largura de banda em <i>Hertz</i>
C	Capacitância em um detector CMOS ou CCD
c	Velocidade da luz no vácuo cujo valor é 299.792.458 m/s
col	Número de colunas em uma matriz
Cr	Coefficiente de reflexão do material da superfície do detector
Cr_k	Centroide do agrupamento RGB
Ch_k	Centroide do agrupamento hiperespectral
C_x, C_y	Coordenadas de referência do sensor
Dh	Dicionário de assinaturas espectrais
dx	Dimensão espacial, eixo x
dy	Dimensão espacial, eixo y
dz	Dimensão espectral, eixo z
e	Carga do elétron cujo valor é $1,60217663 \times 10^{-19}$ Coulomb
$\mathcal{E}$	Valor esperado
E	Matriz com termos de ruído para modelagem de fatores principais
E_0	Energia do estado fundamental de um átomo não excitado
E_1	Energia do primeiro estado de um átomo excitado
E_i	Energia do estado de ionização i de um átomo
E_f	Energia de um fóton
E_g	Energia do gap em um material semiconductor
f	Frequência em <i>Hertz</i>
F_c	Função de transferência do sistema óptico, filtro <i>Bayer</i> e sensor de imagem
F	Matriz de fatores latentes utilizado na análise de fatores f_j
f_j	Fator latente utilizado na análise de fatores
f_x	Distância focal no eixo x
f_y	Distância focal no eixo y
g_1, g_2	Pesos dos níveis de excitação 1 e 2 de um átomo no modelo de <i>Boltzmann</i>

G	<i>Green - cor verde de um píxel RGB</i>
h	Constante de Planck cujo valor é $6,62607015 \times 10^{-34} \text{ m}^2 \text{ kg/s}$
$\mathcal{H}$	Conjunto de imagens hiperespectrais
H0	Hipótese nula
H1	Hipótese alternativa
$\mathbf{H}_m$	Tensor contendo uma imagem hiperespectral
I_{dark}	Corrente de escuro em um sensor de imagem
$I_{d_{rms}}$	Ruído balístico da corrente de escuro em um sensor de imagem
$I_d(x_p, y_p, \lambda)$	Função que define a intensidade de um píxel no comprimento de onda λ e coordenadas espaciais x_p e y_p de uma imagem de referência de escuro
Ih	Tensor contendo uma imagem hiperespectral
$I_m(x_p, y_p, \lambda)$	Função que define a intensidade de um píxel no comprimento de onda λ e coordenadas espaciais x_p e y_p de uma imagem
I_p	Corrente fotônica em um sensor de imagem
Irgb	Tensor contendo uma imagem RGB
IR	<i>InfraRed - Faixa espectral de uma onda eletromagnética</i>
IR769nm	Banda espectral centrada no comprimento de onda de 769 nm
$I_s(x_p, y_p, \lambda)$	Função que define a intensidade de um píxel do sensor de imagem no comprimento de onda λ e coordenadas espaciais x_p e y_p
$I_w(x_p, y_p, \lambda)$	Função que define a intensidade de um píxel no comprimento de onda λ e coordenadas espaciais x_p e y_p de uma imagem de referência de branco;
k	Constante de <i>Boltzmann</i> cujo valor é $1,380649 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$
K	Unidade de temperatura - <i>Kelvin</i>
KMO	Critério <i>Kaiser-Meyer-Olkin</i> aplicado na análise de fatores
k_1, k_2, k_3	Coefficientes de distorção radial da lente de uma câmera
lin	Número de linhas em uma matriz
L_d	Comprimento de difusão em um material semicondutor de um detector de luz
M	Matriz com termos de <i>offset</i>
M_t	Número de imagens do conjunto de dados usados no treinamento dos algoritmos de síntese espectral
$\mathbf{M}_p$	Matriz para projeção de um ponto no mundo para um ponto no plano do sensor
n_d	Número de termos em uma regressão linear
N_{i+1}	Número de átomos no estado de ionização $i + 1$
N_i	Número de átomos no estado inicial de ionização i
n_e	Densidade de elétrons
n_s	Número de amostras na fórmula de esfericidade de <i>Bartlett</i>
m_e	Massa do elétron cujo valor é $9,1093837 \times 10^{-22}$ microgramas
N_2, N_1	Densidades numéricas de átomos nos estados excitado e não excitado, para modelagem estatística de <i>Boltzmann</i>
n_f	Número de fatores na análise de fatores
n_h	Número de bandas espectrais em uma imagem hiperespectral
n_p	Número de píxeis em uma imagem
n_s	Número de amostras em uma matriz de píxeis

p	Número de variáveis na fórmula de esfericidade de <i>Bartlett</i>
p_1, p_2	Coeficientes de distorção tangencial de uma lente fotográfica
P_{glv}	Intensidade da banda v de um píxel l na imagem de referência (<i>ground truth</i>)
$\mathbf{p}_m$	Um ponto no mundo
P_{slv}	Intensidade da banda v de um píxel l na imagem hiperspectral sintetizada
$\mathbf{p}_s$	Um ponto no plano do sensor de imagem
$pvalue$	Parâmetro estatístico utilizado para confirmar, ou refutar, a hipótese nula nos testes de <i>Kruskal-Wallis</i> e de <i>Shapiro-Wilk</i>
Q	Matriz de pesos usada para gerar os dados observados a partir das variáveis latentes na análise de fatores
q_{kTC}	Componente de ruído devido à incerteza na carga de um capacitor
r	Distância radial do ponto ao centro da imagem
r_{st}	Elemento fora da diagonal na matriz de correlação da equação Kaiser-Meyer-Olkin
R	Matriz contendo dados de um conjunto de imagens RGB
R	<i>Red</i> - cor vermelha de um píxel RGB
$\mathcal{R}$	Conjunto de imagens RGB obtidas por projeção de imagens hiperspectrais
R^2	Coeficiente de determinação
R_c	Matriz de correlação de dados na fórmula de esfericidade de <i>Bartlett</i>
R_d	Reflectância de um detector de fótons
$\mathbf{R}_m$	Matriz contendo uma imagem RGB
R_{eq}	Resistência equivalente em OHM
R_i	Proporção relativa de átomos excitado para átomos não excitados
S	Esfericidade de <i>Bartlett</i>
Sv_j	Valor singular de índice j
T	Temperatura
U	Matriz contendo autovetores
$u_i(T)$	Função de partição para o estado i de ionização
$u_{i+1}(T)$	Função de partição para o estado $i + 1$ de ionização
v	Banda espectral
V	Matriz transposta contendo autovetores
V_{a_i}	Variância explicada acumulada até o valor singular de índice i
V_e	Componente de ruído em um semicondutor (<i>Johnson-Nyquist</i>)
W	Watt, unidade de potência elétrica
(x, y, z)	Coordenadas de um ponto no mundo
X	Matriz que contém os dados observados, para análise de fatores
X_d	Variável independente em uma regressão linear
x_r	Coordenada do píxel corrigida para compensação de distorções radiais
$\mathbf{x}_i$	Vetor de dados observados, na análise de fatores
xs	Coordenada do píxel na imagem
x_t	Coordenada do píxel corrigida para compensação de distorções tangenciais
ws	Coordenada de um ponto na imagem
W_d	Largura da zona de depleção de uma junção PN
Y_p	Variável dependente em uma regressão linear
y_r	Coordenada do píxel corrigida para compensação de distorções radiais

y_s	Coordenada do píxel na imagem
y_t	Coordenada do píxel corrigida para compensação de distorções tangenciais
α_{si}	Coefficiente de absorção do silício
β_0	Coefficiente da equação de regressão linear univariável correspondente ao intercepto
β_1	Coefficiente da equação de regressão linear univariável correspondente à inclinação
β_i	Coefficientes da equação de regressão linear multivariável
ϵ_i	Vetor com termo de ruído gaussiano de média zero
η	Eficiência quântica externa
λ	Comprimento de onda da radiação eletromagnética
μ_i	Vetor de componentes de <i>offset</i>
mu_{lv}	Valor médio dos valores observados, para cálculo do fator de determinação
σ	Densidade de potência espectral
Σ	Matriz diagonal com os valores singulares em uma SVD
τ_{st}	Termo de correlação parcial das variáveis na equação <i>Kaiser-Meyer-Olkin</i>
$\mathbf{N}$	Vetor com distribuição gaussiana para emular ruído de corrente de escuro
$\mathbb{N}$	Vetor com distribuição de <i>Poisson</i> para emular ruído fotônico

Lista de Abreviaturas e Siglas

ADAM	<i>Adaptative Momentum Estimation</i>
APS	<i>Active Pixel Sensor</i>
AVRIS	<i>Airborne Visible/Infrared Imaging Spectrometer</i>
BEMAD	Batalhão de Emergências Ambientais e Resposta a Desastres do Corpo de Bombeiros Militares de Minas Gerais
CNN	<i>Convolutional Neural Network</i>
CCD	<i>Charge-Coupled Device</i>
CMOS	<i>Complementary Metal-Oxide-Semiconductor</i>
CEMIG	Companhia Energética de Minas Gerais
CEFALA	Centro de Estudos em Fala, Acústica, Linguagem e Música
ELU	Exponencial Linear Unit
FA	<i>Factor Analysis</i>
FNN	<i>Feedforward Network - FNN</i>
GaAsP	Fosfeto arseneto de gálio
GAN	Generative Adversarial Network
GaP	Fosfeto de gálio
Ge	Germânio
ICVL	<i>Interdisciplinary Computational Vision Laboratory</i>
InGaAs	Arseneto de gálio e índio
K-SVD	<i>K-Means Singular Value Decomposition</i>
KMC	<i>K-Means Clustering</i>
KTC	Componente de ruído térmico no canal de um transistor MOS
MARE	<i>Mean Absolute Relative Error</i>
MIF	Manejo Integrado do Fogo
MLP	<i>Multilayer Perceptron</i>
MOSFET	<i>Metal-Oxide-Semiconductor Field-Effect Transistor</i>
LWIR	<i>Long Wavelength Infrared</i>
MWIR	<i>Middle Wavelength Infrared</i>
NIR	<i>Near Infrared</i>
NTIRE	<i>New Trends in Image Restoration and Enhancement</i>
LTSM	<i>Long Short-Term Memory</i>
GAN	<i>(Generative Adversarial Network</i>
OMP	<i>Ortogonal Matching Pursuit</i>
PCA	<i>Principal Component Analysis</i>

RBF	<i>Radial Basis Function - RBF</i>
RELU	<i>Rectified Linear Unit</i>
RGB	<i>Red-Green-Blue</i>
RMSE	<i>Root Mean Squared Error</i>
RNN	<i>Recurrent Neural Network -</i>
RR	<i>Ridge Regression</i>
Si	Silício
SGD	<i>Stochastic Gradient Descent</i>
SVD	Single Value Decomposition
SWIR	<i>Short Wavelength Infrared</i>
TIR	<i>Thermal Infrared</i>
UFMG	Universidade Federal de Minas Gerais
VIS	<i>Visible light range</i>
YOLO	<i>You Only Look Once</i>

Lista de Termos

ADAMAX	Variante do otimizador ADAM (<i>Adaptive Momentum</i>) usado em aprendizado de máquina para treinamento de redes neurais
bit	" <i>Binary Digit</i> " (dígito binário) é a unidade fundamental de informação em ciência da computação
CIE-1964	Padrão de representação para sinais tri-estímulos (cores) com correspondência em luminância
<i>Back-end</i>	Em um contexto de sistemas de monitoramento por imagem, refere-se à parte do sistema que lida com o processamento, armazenamento e gerenciamento de dados de imagens
Espectrorradiômetro	Instrumento utilizado para medir a distribuição espectral da radiação de uma fonte de luz, em diferentes comprimentos de onda
Espectro	Refere-se à gama de frequências da radiação eletromagnética, incluindo ondas de rádio, micro-ondas, luz visível, raios-x e radiação gama
Espectrometria	Técnica utilizada para medir as propriedades da radiação eletromagnética, partículas carregadas ou massas atômicas e moleculares
Espectroscopia	Técnica analítica que estuda a interação entre a matéria e a radiação eletromagnética. Essa interação pode revelar informações sobre as propriedades das substâncias, como composição, estrutura molecular, concentração, temperatura e outras características físicas
ENVI	<i>Environment for Visualizing Images</i> - refere-se a um formato de arquivo usado principalmente em sensoriamento remoto e em sistemas de informação geográfica (SIG)
<i>Front-end</i>	Em um sistema de monitoramento por imagem, refere-se à interface com a qual os usuários interagem para visualizar e analisar informações provenientes das câmeras de monitoramento
Imageador	Dispositivo capaz de produzir imagem por meio de arranjo de elementos sensores de luz.
Irradiância	Medida da potência radiante incidente por unidade de área em uma superfície, expressa em watts por metro quadrado (W/m^2)
Junção PN	Junção de camadas de um semicondutor compostas por material dopado com impureza aceitadora (P) e impureza doadora (N)
Keras	API para construção e treinamento de modelos de aprendizado profundo.
<i>LUMO RECORDER</i>	Software de controle e configuração para a câmera Specim FX-10
<i>Manifold</i>	Variedade matemática - É um espaço topológico que se parece localmente

	com um espaço euclidiano nas vizinhanças de cada ponto
<i>Metamerismo</i>	Fenômeno na percepção de cores, quando objetos que parecem idênticos em cor sob uma condição de iluminação e que parecem com cores diferentes sob outra condição. No contexto de sensores de imagem e reprodução de cores, o metamerismo pode surgir quando distribuições espectrais distintas de luz resultam nos mesmos valores RGB
MIF	Manejo Integrado do Fogo. Técnica empregada para controle preventivo de incêndios, aplicado, sobretudo, em parques florestais e áreas de risco de incêndio. O manejo consiste em queima controlada da vegetação, feito pelo corpo de bombeiros e brigadas de combate a incêndios
Norma L1	Soma das magnitudes absolutas dos componentes de um vetor
Norma L2	Soma dos quadrados dos componentes de um vetor
NTIRE	<i>Challenge on Spectral Reconstruction from RGB Images</i> . Uma competição sobre recuperação do espectro eletromagnético a partir de imagens RGB
<i>Píxel</i>	“Picture element” ou “elemento de imagem”. É a menor unidade que compõe uma imagem digital
Radiância	A radiância é uma medida da quantidade de energia radiante emitida ou transmitida por uma fonte em uma determinada direção, e em uma determinada faixa de comprimentos de onda
RAW	Formato de imagem que contém dados não processados, capturados pelo sensor de uma câmera digital, ou outro dispositivo de captura de imagens
Resolução amplitude	Refere-se à menor unidade de um sinal analógico que pode ser percebido ou identificado em um sistema digital, expressa em bits
Resolução espectral	Capacidade de um instrumento, ou de uma técnica, em distinguir entre diferentes comprimentos de onda ou frequências na radiação eletromagnética
Resolução espacial	Capacidade de um sistema, ou instrumento sensor, distinguir entre objetos, ou detalhes espaciais em uma imagem, ou cena
Resolução radiométrica	Refere-se à menor unidade de um sinal eletromagnético pode ser percebido ou identificado por um detector, expressa em bits em sistemas digitais
<i>Slit</i>	Fenda estreita para coletar luz de uma linha específica de visão da cena
<i>Specim</i>	Fabricante de câmeras hiperespectrais
Tensorflow	Biblioteca de código aberto desenvolvida pelo Google Brain para implementação de modelos de IA e aprendizado de máquina

Sumário

1	Introdução	30
1.1	Motivação	31
1.2	Objetivos	34
1.3	Contribuições científicas e tecnológicas	35
1.4	Organização deste Trabalho	36
2	Referencial teórico	39
2.1	Problema malposto, metamerismo e variedade	39
2.2	Dimensionalidade intrínseca dos dados hiperespectrais e estrutura embutida	41
2.3	Sistemas de imageamento	42
2.4	Responsividade	48
2.5	Ruído em sensores de imagem	50
2.6	Calibração radiométrica	52
2.7	Compensação da corrente de escuro e calibração radiométrica	53
2.8	Calibração geométrica	54
2.9	Câmeras RGB, multiespectrais e hiperespectrais	56
2.9.1	Tipos de sensores hiperespectrais	59
2.10	Detecção de fogo e fumaça	61
2.10.1	Decaimento de elétrons em átomos de potássio	63
2.10.2	Equação de <i>Boltzmann</i>	65
2.10.3	Equação de <i>Saha</i>	65

2.11	Métodos de Regressão Linear	66
2.12	Redes Neurais MLP	68
2.13	Base de dados de imagens ICVL	71
2.14	Trabalhos correlatos	72
2.15	Comentários finais	74
3	Metodologia	76
3.1	Análise de fatores	77
3.1.1	Medida da esfericidade pelo método <i>Bartlett</i>	79
3.1.2	Crítério <i>Kaiser-Meyer-Olkin</i> (KMO)	79
3.1.3	Análise de Fatores	79
3.1.4	Variância explicada acumulada	81
3.2	Métricas para avaliação dos processos de síntese hiperespectral	82
3.2.1	Erro Absoluto Médio Relativo (MARE)	82
3.2.2	Erro quadrático médio (RMSE)	83
3.2.3	Coeficiente de Determinação (R^2 score)	83
3.3	Validação cruzada	84
3.4	Processos para síntese de imagens hiperespectrais	85
3.5	Método para comparação do desempenho dos processos de síntese de imagens hiperespectrais	86
3.6	Preprocessamento das imagens da base de dados ICVL	87
3.7	Câmera hiperespectral Specim FX-10	88
3.8	Câmera Specim IQ	90
3.9	Câmera RGB Panasonic DMC-FZ5	91
3.10	Construção de bases de imagens hiperespectrais e RGB	91
3.10.1	Construção da base de dados de imagens hiperespectrais com uso da câmara hiperespectral <i>Specim FX-10</i>	93

3.10.2	Construção da base de dados de imagens com montagem mecânica da Câmera hiperespectral Specim IQ e da câmera Panasonic DMC-FZ5 . . .	94
3.11	Comentários finais	100
4	Resultados e análises	101
4.1	Análise de fatores em um conjunto de dados hiperespectrais	101
4.2	Modelos e algoritmos para recuperação do hiperespectro a partir de imagens RGB da base de dados ICVL	103
4.2.1	Recuperação do hiperespectro por regressão linear	103
4.2.2	Recuperação do hiperespectro por composição de agrupamentos combinado com regressão linear	108
4.2.3	Recuperação do hiperespectro por redes neurais	112
4.2.4	Recuperação do hiperespectro por redes neurais combinado com agrupamentos	116
4.2.5	Análise qualitativa do desempenho dos algoritmos implementados	117
4.2.6	Análise quantitativa do desempenho dos processos pela métrica MARE e RMSE	120
4.3	Análise de espectros com uso da câmera Specim FX-10	125
4.4	Análise de espectros com uso da câmera Specim IQ	129
4.5	Síntese de imagens hiperespectrais com projeção no espaço RGB através de um filtro <i>Bayer</i>	137
4.5.1	Síntese de imagens hiperespectrais pelo método <i>Ridge Regression</i> a partir de imagens RGB obtidas por projeção usando filtro <i>Bayer</i>	138
4.5.2	Rede Neural para síntese hiperespectral a partir de imagens RGB obtidas por projeção usando um filtro <i>Bayer</i>	142
4.5.3	Agrupamento e MLP para síntese hiperespectral a partir de imagens projetadas usando filtro <i>Bayer</i>	145
4.5.4	Análise de desempenho dos modelos	148
4.6	Projeção do espaço hiperespectral no espaço RGB-IR por função de transferência sintetizada com quatro componentes	151
4.7	Síntese de imagens hiperespectrais a partir de uma câmera comercial Panasonic DMC-FZ5	158

4.8	Síntese de imagens hiperespectrais a partir de imagens captadas pelo sensor RGB da câmera Specim IQ	165
4.9	Comentários finais	167
5	Conclusão	168
5.1	Trabalhos futuros	171
	Referências Bibliográficas	172
A	Base de dados de imagens hiperespectrais FX-10	178
B	Base de Dados de imagens hiperespectrais Specim IQ	180
C	Códigos fonte	181

Capítulo 1

Introdução

*A*nualmente, incêndios devastam vastas áreas florestais em todo o mundo, resultando em prejuízos materiais, impactos ambientais, aumento da poluição atmosférica, destruição de ecossistemas inteiros e a perda de vidas humanas e animais. Em 2023, incêndios florestais causaram danos em vastas áreas na Grécia, Itália, Espanha, Reino Unido, Canadá, Austrália, Estados Unidos, Chile, Brasil e em vários outros locais do mundo ([NASA, 2023](#)). As operações de combate a incêndios são onerosas e representam alto risco para os bombeiros e brigadistas. Pesquisadores em todo o mundo dedicam esforços no desenvolvimento de ferramentas para prevenção, identificação e combate a incêndios florestais. A detecção precoce de incêndios florestais tem passado por extensas pesquisas nos últimos anos, reunindo conceitos de várias disciplinas. O uso de imagens hiperespectrais apresenta-se um caminho viável para a detecção de incêndios florestais por propiciar dados com alta resolução espectral e tem sido usados em sistemas satelitais, ou mesmo embarcados em aeronaves, em diversas aplicações, incluindo a detecção de incêndios florestais. Trata-se de equipamentos caros e de manipulação complexa, restringindo seu uso em sistemas de detecção baseados em solo para cobertura em ampla área geográfica.

As perguntas que nortearam essa pesquisa são:

- É possível obter um espectro de radiação eletromagnética com distribuição espectral contínua a partir de apenas três bandas espectrais R, G e B, capturadas por sensores de câmeras convencionais?

- O uso de imagens com alta resolução espectral possibilitam a identificação de assinaturas e marcadores espectrais específicos em incêndios florestais?

- Há uma estrutura subjacente nos dados correspondentes a imagens hiperespectrais que viabilize uma solução para o problema malposto de se obter dados em um espaço de alta dimensionalidade a partir de dados em um espaço de dimensões reduzidas?

Este trabalho apresenta os resultados de estudos sobre imageamento hiperespectral na faixa do espectro eletromagnético de luz visível ao olho humano e do infravermelho próximo. A pesquisa concentra-se no desenvolvimento de técnicas e metodologias para estimar o conteúdo espectral de elementos de cenas a partir de imagens obtidas por câmeras RGB de baixo custo, equipadas com sensores de silício. A aplicação específica desse estudo está na detecção de incêndios florestais, mas encontra uso na prospecção mineral, análise de solo, medicina, agricultura de precisão, entre outras áreas. A pesquisa estabelece as condições necessárias para o êxito dessa síntese.

A pesquisa também aponta para a busca de marcadores com assinaturas espectrais específicas da região da chama em incêndios florestais. Um exemplo é a presença de átomos de potássio no material combustível, em quantidade suficiente para produzir sinal detectável remotamente, identificado na faixa NIR, como consequência do decaimento de elétrons termicamente excitados, presentes na região da chama.

1.1 Motivação

Inúmeras metodologias têm sido exploradas para detecção precoce de incêndios florestais, visando alcançar níveis elevados de exatidão, precisão e robustez, tais como técnicas de visão computacional e processamento de imagens, juntamente com a integração de algoritmos de aprendizado de máquina. Esse esforço conjunto visa fornecer uma solução eficiente e abrangente para o desafio em questão ([Venâncio et al., 2022](#)). Classificadores baseados em algoritmos de aprendizado de máquina, utilizando dados de imagens fotográficas e vídeos, têm alcançado taxas de acertos acima de 80% na identificação de focos de incêndio ([Lisboa et al., 2020](#)).

Um sistema de monitoramento de incêndios com base em satélites equipados com espectror-

radiômetros e câmeras hiperespectrais foi efetivamente utilizado para detectar inúmeras ocorrências de incêndios em todo o mundo (NASA, 2022). Embora essas soluções ofereçam um valor significativo, elas têm uma limitação em termos da resolução espacial da imagem no solo - *Ground Sample Distance* - devido à distância do satélite para a superfície da terra (Qian, 2021), o que restringe sua capacidade de detectar efetivamente incêndios em estágios iniciais. O uso de técnicas de detecção subpixel é uma forma de melhorar a resolução espacial (Xu et al., 2022). Incêndios menores ou localizados podem não ser identificados com devida acurácia. Fatores como fumaça, cobertura de nuvens e condições climáticas adversas podem obstar a visualização por satélites utilizando câmeras hiperespectrais, tornando desafiador, por esse método, localizar incêndios ou avaliar sua intensidade com precisão (Giglio et al., 2018). Sistemas de detecção de incêndios podem, por vezes, classificar erroneamente fontes de calor, resultando em falsos positivos, ou podem deixar de identificar incêndios reais, resultando em falsos negativos. Essas discrepâncias introduzem imprecisões no processo de detecção, como mencionado anteriormente. Além disso, pode haver uma latência entre o início de um incêndio e sua detecção por sistemas satelitais. Esse atraso é especialmente pronunciado em regiões remotas onde o monitoramento contínuo torna-se inviável. O tempo necessário para que satélites cubram a mesma área geográfica é da ordem de horas, ou até mesmo dias, impossibilitando a detecção precoce de incêndios (Veraverbeke et al., 2014).

Veículos Aéreos Não tripulados (VANT) equipados com câmeras RGB e câmeras hiperespectrais oferecem uma solução versátil para apoiar sistemas de monitoramento de incêndios florestais. Essa abordagem fornece imagens de alta resolução espacial com tamanhos de pixel de poucos centímetros, dependendo da altitude do voo. VANTs desempenham um papel importante no mapeamento dos perímetros do fogo e fornecem consciência situacional em tempo real para equipes de combate a incêndios (Allison et al., 2016). No entanto, o uso de VANTs possui limitações: a duração de voo é determinada pela capacidade de carga das baterias, exigindo recarga frequente; condições climáticas adversas, como ventos fortes, chuva, ou neve, podem afetar a aeronavegabilidade e a estabilidade do voo; e sua faixa operacional é relativamente restrita, dificultando o monitoramento de áreas extensas, ou remotas (Akhloufi et al., 2021). Além disso, operar um VANT envolve a adesão a regulamentações da aviação e restrições na ocupação do espaço aéreo, que proíbem voos próximos a aeroportos, e em regiões densamente povoadas (Easa, 2019; FAA, 2020). A possibilidade de interferência de radiofrequência ou perda de sinal representa um risco para a comunicação com o VANT, seja nas funções de te-

lemetria e controle, ou mesmo na transmissão de dados. Garantir a transmissão contínua de dados, em tempo real, entre o VANT e as estações terrestres, também pode ser desafiador (Kim et al., 2022).

Um sistema de monitoramento de incêndios florestais terrestre apresenta-se como uma abordagem complementar aos métodos mencionados anteriormente. Esse sistema pode utilizar sensores em rede (Singh e Singh, 2012) e tecnologia de visão computacional para aprimorar o gerenciamento de incêndios florestais. Uma estratégia possível envolve a instalação de câmeras de vídeo (RGB) em torres de observação, em parques florestais, ou em torres de transmissão de energia elétrica, entre outros locais pertinentes. Essas câmeras podem fornecer imagens em tempo real, com georreferenciamento, que, processadas por algoritmos de aprendizado de máquina, propiciam informações sobre a progressão e o comportamento do fogo (Lisboa et al., 2020). Ao complementar dados de satélites e de VANTs com observações baseadas em solo, tem-se uma solução mais abrangente da dinâmica de incêndios florestais. A baixa resolução temporal dos sistemas de satélite é compensada pela rápida resposta dos sistemas baseados em solo, da ordem de poucos minutos. Essa abordagem foi implementada com sucesso em um projeto colaborativo, envolvendo entidades como CEMIG (Companhia Energética de Minas Gerais), CEFALA-UFMG (Centro de Estudos da Fala, Acústica, Linguagem e Música da Universidade Federal de Minas Gerais), GAIA *Solutions on Demand* e UFVJM (Universidade Federal dos Vales do Jequitinhonha e Mucuri). O projeto resultou na criação de um sistema conhecido como "Apaga o Fogo" (<http://apagaofogo.eco.br>). Esse sistema concentra-se no monitoramento de incêndios florestais e integra várias câmeras de vídeo instaladas em parques florestais na região metropolitana de Belo Horizonte. Ele utiliza uma infraestrutura de comunicação dedicada, uma plataforma de *back-end* para processamento de imagens e uma interface de *front-end* para agregar imagens das câmeras. Servidores de *back-end* hospedam algoritmos para detecção de incêndios, aproveitando as capacidades do algoritmo YOLO-V4 em processos de tomada de decisão (Redmon et al., 2016). Os dados das imagens passam por um processamento em tempo real, utilizando informações espaciais e temporais. Os eventos são enviados como notificações para brigadas de incêndio e corpo de bombeiros locais, e para outros interessados, possivelmente afetados por incêndios florestais. O sistema é projetado para enviar mensagens, transmitindo a probabilidade de eventos de fogo e fumaça retratados nas imagens de vídeo (Venâncio et al., 2022).

Este esforço colaborativo exemplifica um passo pioneiro para mitigar os desafios da detecção precoce de incêndios florestais, demonstrando a fusão de tecnologia, pesquisa e inovação para proteger proativamente o meio ambiente e comunidades contra ameaças de incêndios (NASCIMENTO et al., 2022). O sistema "Apaga o Fogo" visa capacitar as partes interessadas e afetadas, possibilitando uma resposta rápida aos incêndios por meio da detecção precoce. A agilidade na resposta está se mostrando importante não apenas para extinguir com sucesso o próprio fogo, com muito menos recursos, mas também para identificar sua causa e interagir eventualmente com as pessoas que o causaram. Isso ajuda a mudar a cultura local e até os relacionamentos sociais em relação ao fogo. Embora a detecção de fumaça e fogo, usando câmeras RGB como sensores, tenha alcançado uma alta taxa de sucesso, a ocorrência de falsos alarmes pode afetar a escalabilidade da solução. Esses algoritmos podem apresentar erros em situações em que a cor, textura e movimento do elemento na imagem se assemelham ao fogo, ou fumaça. Os algoritmos, geralmente, utilizam informações apenas da faixa visível da luz correspondentes a somente três bandas espectrais (RGB). A fumaça, em muitos casos, torna-se um obstáculo e impede a localização correta dos focos de incêndio, dificultando o combate.

Esse trabalho aborda a obtenção de informações em alta resolução espectral do conteúdo de imagens, tanto na faixa de luz visível (VIS), quanto na faixa do infravermelho próximo (NIR). Essa abordagem visa reduzir erros e limitações dos processos de detecção atuais, permitindo a distinção de assinaturas espectrais específicas dos elementos da cena. Além disso, a detecção de sinais em uma faixa do espectro na qual a fumaça pouco densa é transparente possibilita a localização mais eficiente dos focos de incêndio.

1.2 Objetivos

O objetivo dessa pesquisa é desenvolver uma metodologia para obtenção de imagens hiperespectrais na faixa de luz visível e do infravermelho próximo, a partir de imagens capturadas por câmeras RGB de baixo custo, utilizando algoritmos de inteligência artificial.

Os objetivos específicos são listados a seguir:

I - Analisar e avaliar diversas técnicas de reconstrução espectral para determinar as restrições e limitações dos métodos de síntese de imagens hiperespectrais a partir de imagens RGB.

II - Realizar análise da dimensionalidade de dados de imagens hiperespectrais para inferir sobre a existência de estruturas subjacentes com dimensão intrínseca reduzida nos dados, e que viabilize encontrar uma solução para o problema malposto de mapear o espaço RGB, de baixa dimensionalidade, para o espaço hiperespectral de alta dimensionalidade.

III - Construir uma base de dados de imagens hiperespectrais e correspondentes imagens RGB utilizando as câmeras Specim FX-10, Specim IQ, e câmera Panasonic DMC-FZ5.

IV - Avaliar a viabilidade de identificação do perfil espectral de ondas eletromagnéticas em incêndios florestais, com identificação da banda de emissão por decaimento de elétrons termicamente excitados, com uso de imagens capturadas por câmeras hiperespectrais, bem como as sintetizadas com uso dos processos desenvolvidos.

1.3 Contribuições científicas e tecnológicas

As contribuições científicas e tecnológicas deste trabalho são:

I - O desenvolvimento de uma metodologia para composição de imagens hiperespectrais na faixa do espectro visível e do infravermelho próximo a partir de imagens RGB capturadas com uso de câmeras de baixo custo, e sua aplicação na detecção de incêndios florestais;

II - A construção de uma base de dados de imagens hiperespectrais de cenas com iluminação natural, em diversas condições, contendo ou não fogo e fumaça. As imagens hiperespectrais foram capturadas com uso das câmeras Specim FX-10 e Specim IQ. A base contém as correspondentes imagens RGB capturadas por uma câmera Panasonic DMC-FZ5.

III - A identificação do perfil espectral da emissão irradiada em incêndios florestais, especificamente, a emissão por decaimento de elétrons termicamente excitados, na região da chama.

IV - A definição de um método para identificar a existência de estrutura nos dados hiperespectrais, associados a imagens com iluminação natural e de emissão eletromagnética de regiões com fogo, por meio do uso de análise de fatores, juntamente com as condições necessárias para sua aplicação, e a inferência acerca da dimensionalidade efetiva do espaço dos dados do *manifold* correspondente.

V - Desenvolvimento de métodos para síntese de imagens hiperespectrais baseados em regressão linear, rede neural e agrupamentos.

VI - Estabelecimento do mapeamento de centroides de agrupamentos no espaço RGB para centroides de agrupamentos no espaço hiperespectral, de maneira simples e eficiente, sem a necessidade de uma etapa de busca.

VII - Desenvolvimento de códigos em Python com implementação dos algoritmos, inferências, métricas e processamentos, utilizando biblioteca *Keras* e *framework Tensorflow*.

VIII - Publicação do artigo científico "*Hyperspectral image synthesis from sparse RGB data: a comparative study combining linear regression, multilayer perceptron, and clustering*", no periódico "*Signal, Image and Video Processing*", *Springer Nature*, doi:10.1007/s11760-023-02875-7.

IX - Elaboração de capítulo de livro com o título "*Hyperspectral image synthesis from RGB images applied to wildfire detection*", para publicação no livro "*Computational Intelligence based Hyperspectral Image Analysis*", *Springer Series "Intelligent Systems Reference Library"* (Electronic ISSN: 1868-4408, Print ISSN: 1868-4394)

1.4 Organização deste Trabalho

Essa tese é composta por cinco capítulos. O Capítulo 2 contém um referencial teórico para oferecer suporte ao entendimento da metodologia, dos conceitos e dos métodos implementados, bem como das análises e discussões dos resultados alcançados. Este capítulo apresenta conceitos relacionados a variedade (*manifold*) e metamerismo, sistemas de imageamento, incluindo tipos de sensores e princípios físicos de detecção de energia eletromagnética, componentes de ruído em sensores, filtros, calibração radiométrica e correções geométricas. O capítulo contém uma breve revisão sobre tipos e modelos de regressão linear e de redes neurais, em especial as do tipo perceptron de múltiplas camadas, incluindo os conceitos de função de ativação e otimizadores. Também são descritos, brevemente, os tipos de imageadores hiperespectrais. Uma abordagem sobre a espectroscopia do fogo é apresentada, incluindo a modelagem do processo de excitação e ionização de átomos de potássio em alta temperatura, promovendo elétrons da

região da chama para níveis elevados de energia, com posterior decaimento, e sua relação com a emissão de sinal característico na faixa NIR. O capítulo conclui com a apresentação de trabalhos correlatos, fornecendo uma breve descrição de artigos relacionados à síntese hiperespectral e detecção de incêndios florestais, com comparação dos resultados obtidos nessa pesquisa. O Capítulo 3 apresenta a metodologia utilizada nessa pesquisa, juntamente com conceitos sobre análise de dados. Aborda-se a dimensionalidade intrínseca dos dados e a hipótese da existência de um *manifold* de espaço dimensional reduzido. Especificamente, a análise de fatores é desenvolvida, assim como métodos para avaliação de sua aplicabilidade por meio dos testes de *Bartlett* e *Kaiser-Meyer-Olkin*. São apresentadas as características das câmeras hiperespectrais e RGB utilizadas na coleta de dados, incluindo os procedimentos de construção das bases de dados e a calibração das câmeras. As etapas de pré-processamento das imagens, com normalização e projeção no espaço dos sensores com três e quatro dimensões, também são detalhadas. O capítulo inclui uma descrição do método de validação cruzada, bem como as métricas utilizadas na avaliação dos modelos relativos aos processos de síntese. São apresentados os métodos de comparação de médias de Kruskal-Wallis e de ANOVA, aplicáveis a depender dos dados terem ou não distribuição normal, o que é avaliado pelo método de Teste de *Shapiro-Wilk*. O Capítulo 4 aborda os detalhes das implementações dos algoritmos e processos de síntese, apresentando os resultados obtidos, além de análises e discussões. As implementações seguiram as etapas definidas na metodologia. São descritos quatro processos para síntese hiperespectral com dados RGB: Regressão linear pelo método de *Ridge*, Regressão *Ridge* combinada com agrupamento utilizando o algoritmo *minibatch K-Means*, rede neural do tipo perceptron multicamadas e rede neural combinada com *minibatch K-Means*. Esses processos foram avaliados utilizando as métricas R^2 , MARE (*Mean Absolute Relative Error*) e RMSE (*Root Mean Squared Error*). Os resultados são comparados com os do concurso NTIRE (*New Trends in Image Restoration and Enhancement*). Preliminarmente, foi efetuada a análise de fatores para avaliar a existência de um *manifold* de dimensões reduzidas nos dados de imagens hiperespectrais como condição para a convergência do problema mal posto relacionado à síntese de dados hiperespectrais a partir de dados de imagens RGB. O capítulo aborda o uso de uma câmera hiperespectral Specim, modelo FX-10, para caracterização do espectro de diversos elementos presentes em cenas naturais e em ambiente de laboratório. Descreve-se, também, os processos nos quais se utilizou uma câmera hiperespectral Specim IQ para síntese hiperespectral com base em imagens de um sensor RGB da própria câmera, no formato PNG, pela projeção de dados hiperespectrais no espaço de um

sensor hipotético com três e quatro dimensões. Por fim, o capítulo descreve a síntese hiperespectral a partir de imagens no formato TIF de uma câmera Panasonic DMC - FZ5. Realizou-se a montagem das câmeras em uma estrutura mecânica para garantir precisão na captura das imagens hiperespectrais, pela câmera Specim IQ, e RGB, pela câmera DMC-FZ5, da mesma cena. As distorções geométricas das câmeras foram consideradas durante a etapa de calibração, assim como o recorte das imagens e ajuste de tamanho, garantindo correspondência entre píxeis das imagens. O Capítulo 5 traz as conclusões com um resumo das análises e resultados da pesquisa, assim como as perspectivas e desdobramentos da pesquisa em trabalhos futuros.

Capítulo 2

Referencial teórico

*E*ste capítulo apresenta informações que concorrem para elucidar conceitos-base para o entendimento das técnicas, métodos, algoritmos e processos utilizados no desenvolvimento desse trabalho. São apresentados fundamentos sobre detecção de imagens, tipos de imageadores, componentes de ruído presentes em detectores de imagem, identificação de parâmetros intrínsecos em sistemas de imageamento e projeção de imagens, algoritmos de regressão linear e modelos de redes neurais. São discutidos, brevemente, trabalhos correlatos no que tange aos aspectos que suportam e no que contrastam com o presente trabalho, destacando as contribuições dessa pesquisa.

2.1 Problema malposto, metamerismo e variedade

O desafio de obter informações hiperespectrais a partir de imagens RGB é considerado um problema malposto devido à tarefa de estimar dados multidimensionais, associados a bandas hiperespectrais, a partir de dados confinados a um espaço tridimensional. Isso pode levar potencialmente a um número infinito de soluções. No entanto, uma solução realista pode ser alcançada sob condições e limitações específicas. Uma dessas condições é que os dados correspondentes a assinaturas espectrais devem existir em uma variedade (*manifold*) de dimensões reduzidas situada em um espaço multidimensional (Arad e Ben-Shahar, 2016). É essencial que o número de componentes ou fatores que explicam a variabilidade dos dados seja significativa-

mente menor do que a dimensionalidade do espaço de dados hiperespectrais. Outra condição diz respeito ao metamerismo. Metamerismo refere-se à propriedade de um sensor produzir a mesma resposta a sinais com diferentes distribuições espectrais de potência.

O metamerismo, para o olho humano, é um fenômeno na percepção de cores, que se manifesta quando objetos que parecem idênticos em cor, sob uma condição de iluminação, aparecem com cores diferentes sob outra condição. Este fenômeno se torna proeminente quando a fonte de luz muda, como na transição da luz natural do dia para a iluminação artificial (Foster et al., 2006).

O metamerismo é também observável em sensores de imagem, especialmente na fotografia digital. Sensores de imagem em dispositivos como CCD (*Charge-Coupled Device*), ou sensores CMOS (*Complementary Metal-Oxide-Semiconductor*) captam luz, convertendo-a em sinais digitais para produzir imagens (Theuwissen, 2002). Esses sensores podem exibir metamerismo devido a variações em suas sensibilidades espectrais, em comparação com a visão humana. Dentro do contexto de sensores de imagem e reprodução de cores, o metamerismo pode surgir quando distribuições espectrais distintas de luz resultam em valores RGB semelhantes, ao serem capturadas por um sensor digital. Isso pode introduzir imprecisões de cor nas imagens, pois os valores RGB sozinhos não têm a capacidade de diferenciar as composições espectrais originais das fontes de luz (Foster et al., 2006).

Considere duas fontes de luz diferentes, caracterizadas por distribuições espectrais distintas, denominadas Fonte de Luz A e Fonte de Luz B. Se um sensor digital captura ambas as Fontes de Luz A e B, registrando valores RGB idênticos, não será possível determinar, apenas com os valores RGB, qual fonte de luz foi utilizada para iluminar a cena. Consequentemente, recuperar o espectro original de uma imagem RGB pode não ser viável. No entanto, se a ocorrência de metâmeros for pouco frequente, é provável que a resposta do sensor seja distinta para cada sinal espectral detectado. No contexto de cenas iluminadas por luz natural, a frequência relativa de pares metâmeros varia de 10^{-6} a 10^{-4} , o que é propício para a reconstrução espectral (Foster et al., 2006).

2.2 Dimensionalidade intrínseca dos dados hiperespectrais e estrutura embutida

Dados hiperespectrais referem-se a informações capturadas em inúmeras bandas espectrais estreitas e próximas. Ao contrário de imagens de cores padrão, que geralmente consistem em apenas três bandas (Vermelho, Verde e Azul), dados hiperespectrais podem abranger dezenas ou até centenas de bandas espectrais (Shukla e Kot, 2016). O número de bandas em dados hiperespectrais contribui significativamente para sua complexidade dimensional. Por exemplo, uma imagem hiperespectral com 31 bandas espectrais pode ser representada como um “cubo hiperespectral” com dimensões espaciais de altura e largura, e uma dimensão espectral. Cada píxel neste “cubo” corresponde a uma localização específica na imagem, e contém informações espectrais em diversas bandas. Isso contrasta com imagens RGB convencionais, onde cada píxel contém apenas três valores de cor. À medida que o número de bandas espectrais aumenta, a dimensionalidade dos dados também aumenta.

Lidar com dados de alta dimensionalidade frequentemente envolve técnicas como extração de características, ou redução de dimensionalidade. Esses métodos visam reter as informações mais importantes, enquanto reduzem o número de dimensões, tornando a análise e visualização viáveis.

Como mencionado anteriormente, a tarefa de recuperar informações hiperespectrais a partir de dados RGB é considerada malposta, envolvendo o desafio de mapear de poucas para muitas dimensões. No entanto, no caso de objetos imageados sob condições de iluminação natural, os dados subjacentes estão frequentemente contidos em um *manifold* de baixa dimensionalidade, tornando esse problema viável (Arad e Ben-Shahar, 2016). Para avaliar a dimensionalidade efetiva dos dados, realizou-se um experimento envolvendo a coleta de píxeis de uma série de imagens, submetendo, posteriormente, os dados a uma análise de fatores (Navlani, 2023), usada para obter informações sobre a estrutura subjacente dos dados da imagem (veja Seção 4.1).

2.3 Sistemas de imageamento

A base teórica da análise hiperespectral no contexto da detecção de incêndios florestais está relacionada à radiação eletromagnética e sua interação com diversos materiais, abrangendo tanto o fogo quanto seu ambiente circundante. O espectro eletromagnético compreende uma ampla gama de radiações, desde ondas de rádio até raios gama. A luz visível, apenas uma pequena parte desse espectro, é composta por ondas eletromagnéticas que abrangem uma faixa específica de comprimento de onda. A análise hiperespectral explora não apenas a luz visível ao olho humano, como nas imagens obtidas com uso de câmeras RGB convencionais, mas também regiões além do espectro visível, incluindo o infravermelho próximo (NIR) e o infravermelho distante (LWIR).

Cada material reflete, absorve, transmite, ou emite radiação de maneira específica, em diferentes bandas do espectro eletromagnético. A análise hiperespectral se baseia nas chamadas "assinaturas espectrais", padrões distintos de resposta espectral associados a diferentes materiais. No contexto de incêndios florestais, essas assinaturas podem variar para chamas, fumaça, vegetação e áreas queimadas. A resolução espectral refere-se à capacidade de um sensor em detectar variações na intensidade da radiação em diferentes comprimentos de onda. Sensores hiperespectrais são projetados para capturar ondas eletromagnéticas em dezenas, e mesmo centenas, de bandas espectrais, possibilitando uma melhor discriminação entre diferentes elementos presentes na cena ([Shukla e Kot, 2016](#)).

O comportamento da radiação eletromagnética ao interagir com materiais é fundamental para se determinar suas propriedades físicas. Por exemplo, a refletância se refere à capacidade do material espalhar a luz incidente, enquanto a absorção envolve a captura de fótons da radiação ([Meneses et al., 2019](#)). Em incêndios florestais, áreas em chamas podem exibir padrões específicos de emissão, fornecendo sinais que possibilitam sua detecção, mesmo em presença de fumaça.

Os sistemas de imageamento desempenham a função de captar ondas eletromagnéticas espalhadas, transmitidas ou irradiadas por elementos da cena. Isso é realizado por meio de componentes ópticos, que constituem um conjunto de lentes, espelhos, obturadores e colimadores. Esses componentes têm a finalidade de concentrar e direcionar a luz incidente para um arranjo

de sensores. Estes, por sua vez, são responsáveis por converter fótons em carga elétrica ou corrente elétrica. Existem diversos tipos de sistemas de imageamento, cada um propiciando a formação de imagens a partir da energia eletromagnética em diferentes faixas de comprimento de onda. Nos sistemas de imageamento digital, a carga ou corrente elétrica gerada em cada elemento do sensor é convertida em uma representação digital. Essa representação é comumente definida com resolução radiométrica de 8, 10 ou 12 bits, correspondendo à intensidade do sinal captado.

Os dispositivos CCD utilizam uma estrutura de capacitores MOS (*Metal-Oxide-Semiconductor*), nos quais a carga é gerada pela incidência de fótons sobre um substrato. Esse processo resulta na liberação de elétrons, os quais também podem ser gerados termicamente. Esses elétrons são então capturados em um poço de potencial elétrico. A profundidade desse poço depende da tensão aplicada entre o substrato e um eletrodo de controle, além de outros parâmetros como constante dielétrica do material, nível de dopagem e espessura da camada de óxido. Diversos modelos de arquitetura para imageadores CCD estão relacionados ao processo de construção do capacitor MOS, ao método de leitura da carga dos capacitores, ao arranjo de elementos sensores, entre outros associados aos sistemas de captura de radiação eletromagnética ([Theuwissen, 2002](#)).

Por outro lado, os detectores CMOS (*Complementary Metal-Oxide Semiconductor*) são baseados em uma matriz de fotodiodos ou fototransistores. Cada elemento sensor contém três ou quatro transistores MOSFET (*Metal-Oxide-Semiconductor Field-Effect Transistor*) para a leitura da tensão em um capacitor contendo carga elétrica. Inicialmente, o capacitor é carregado com uma carga de referência. Posteriormente, esse capacitor é descarregado pela corrente de um fotodiodo, gerada pela incidência de fótons sobre o sensor. A diferença de potencial resultante no capacitor proporciona uma estimativa da intensidade de fótons incidentes sobre o elemento sensor ([Sze, 2002](#)).

Em sistemas de imageamento por cor, um arranjo de microfiltros é sobreposto ao conjunto de detectores. Câmeras comerciais, tipicamente, utilizam filtros para separar três bandas espectrais (R, G e B), em um processo para mimetizar a percepção dos sensores no olho humano. Um filtro *Bayer* é um tipo de arranjo de microfiltros, colocado sobre a estrutura de sensores da câmera, para separar a radiação eletromagnética em comprimentos de onda da luz a ser detectada em cada píxel que comporá a imagem capturada na câmera. Um arranjo do tipo *Bayer*

alterna espacialmente filtros para separar os comprimentos de onda referentes à cor vermelha (R - *Red*), verde (G- *Green*) e azul (B-*Blue*). São utilizados arranjos na forma RGBG, BGGR, RGBG, GRGB ou RGGG para composição de cor dos píxeis da imagem. Os filtros ocorrem na proporção de 50% de píxeis com filtros para cor verde, 25% dos píxeis com filtro para azul (B), e 25% para a cor vermelha (R). Essa proporção se deve ao fato de olho humano ser mais sensível ao verde do que a outras cores, o que significa que uma captura de imagem com uma proporção maior de píxeis verdes pode resultar em imagens que parecem mais naturais e perceptualmente exatas para os observadores. A luminância e os detalhes de uma imagem são mais influenciados pela informação verde do que pelas cores vermelha e azul (Shortis et al., 2005). A definição das intensidades para as demais cores de cada píxel - além daquela correspondente ao filtro de cor - é feita por estimativa, utilizando algoritmos de interpolação com base na intensidade das cores dos píxeis vizinhos. Os diferentes tipos de arranjo de filtros afetam como a interpolação de cores dos píxeis se dá, com implicação nas nuances das cores da imagem resultante, e em artefatos que podem aparecer devido ao processamento (Shortis et al., 2005). O detector integra, no tempo, fótons da radiação eletromagnética incidente nos comprimentos de onda correspondentes à faixa de passagem do filtro, com intensidades que dependem da responsividade do material semiconductor utilizado na construção do detector, sendo essa uma função do comprimento de onda (Teich, 1991). A Figura 2.1 ilustra a arquitetura de um sensor de imagem CMOS, contendo um arranjo de detectores sobrepostos por um mosaico de filtros de cor, circuitos analógicos, sinais de temporização e conversor analógico-digital.

A Figura 2.2 apresenta a composição de um sensor do tipo fotodiodo (*APS - Active Pixel Sensor*) com três transistores MOSFET. Uma microlente concentra a luz sobre o detector, após passar por um filtro de cor, que no caso mostrado na figura corresponde à cor vermelha (*Red*). O transistor de *reset* estabelece a carga inicial do capacitor que é posteriormente descarregado por fótons incidentes em um fotodiodo. Por meio do transistor amplificador, é feita a leitura da carga no capacitor. Diversos materiais semicondutores são utilizados na construção dos dispositivos detectores de luz, dependendo da faixa espectral da radiação eletromagnética considerada. Suas propriedades, de forma geral, definem as características do detector. Um fóton, para ser detectado, precisa promover um elétron da banda de valência para a banda de condução, em um material semiconductor. Para isso, a energia do fóton deve ser maior do que a energia do *gap*,

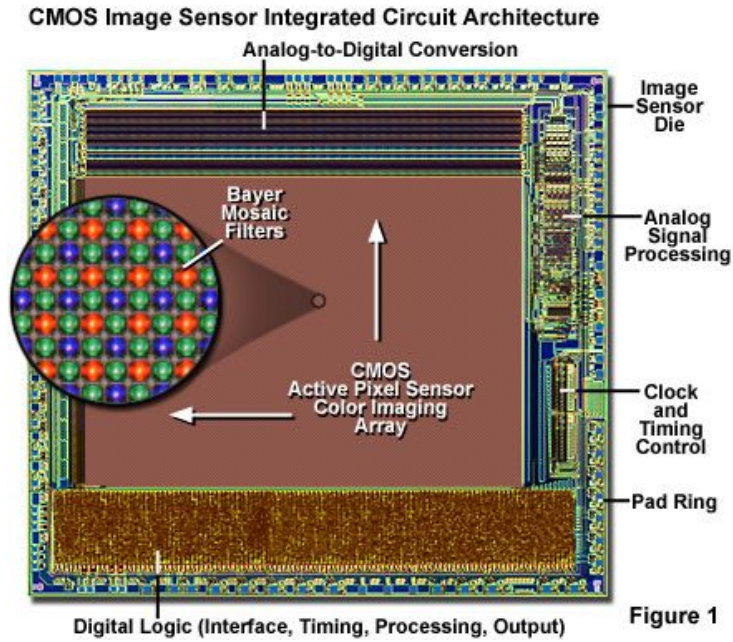


Figure 1

Figura 2.1: Arquitetura de um arranjo de sensores de imagem CMOS. Destacam-se as estruturas do conversor analógico-digital, a etapa de processamento analógico, bem como os sinais de temporização. A figura também apresenta o mosaico de filtros RGB (Teich, 1991).

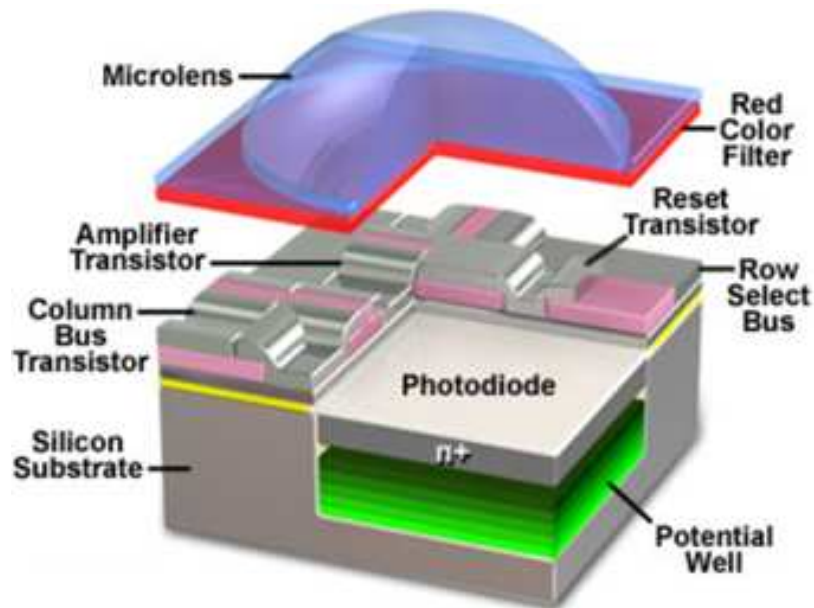


Figura 2.2: Estrutura de um píxel ativo do tipo fotodiodo (APS - Active Pixel Sensor) com três transistores. Uma microlente sobreposta ao filtro Bayer concentra a luz incidente sobre o fotodiodo. O Píxel é considerado ativo pela presença do transistor MOSFET que amplifica o sinal de tensão (Teich, 1991)

conforme ilustrado na Figura 2.3, e estabelecido na equação

$$E_f = \frac{h.c}{\lambda}, \quad (2.1)$$

onde E_f é a energia de um fóton correspondente à onda eletromagnética com comprimento de onda λ ; h é a constante de Planck; e c é a velocidade da luz no vácuo.

A Figura 2.3 ilustra as bandas de valência e condução para o silício e arseneto de gálio, e os respectivos *gaps* (E_g). O arseneto de gálio é um material semiconductor de gap direto, enquanto o silício é de gap indireto.

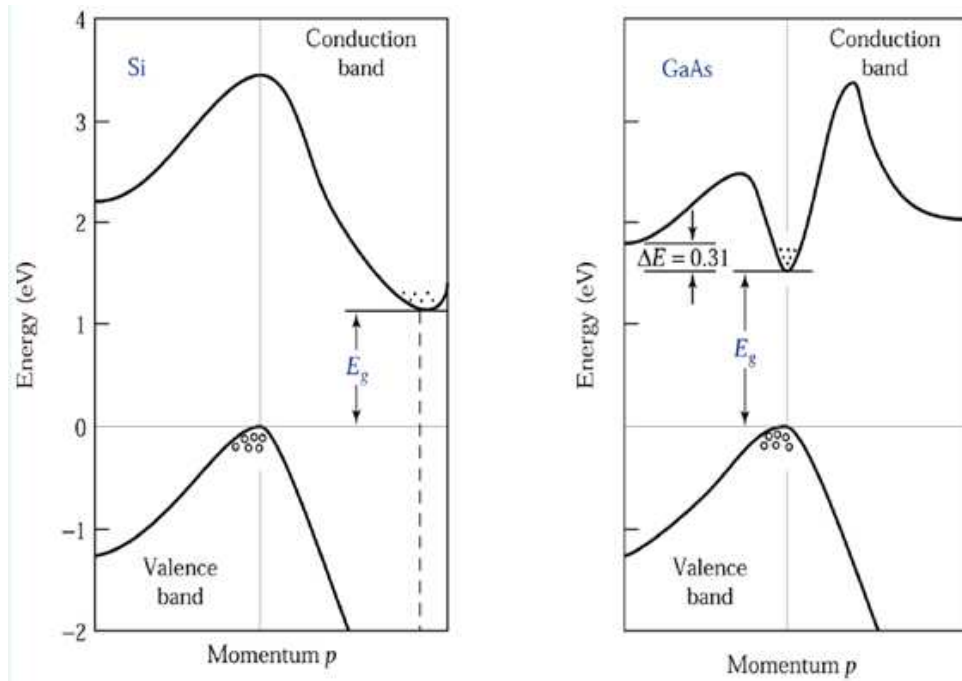


Figura 2.3: Energia do elétron em função do momento linear na banda de valência e na banda de condução do silício (Si) (à direita - gap indireto) e do arseneto de gálio (GaAs) (à esquerda - gap direto). Para que o fóton seja detectado, ele deve promover um elétron da banda de valência para a banda de condução. O comprimento de onda de corte para ejeção de elétrons da banda de valência é dado pela Equação 2.1, fazendo $E_f = E_g$. Para que o fóton seja detectado, o elétron ejetado precisa alcançar o eletrodo do detector antes de recombinar no material semiconductor (Sze, 2002)

O elétron ejetado pode recombinar e retornar à banda de valência e, nesse caso, o fóton que o produziu não será detectado. A recombinação depende do comprimento de difusão (ou recombinação), que é função da mobilidade do elétron (ou da lacuna), da temperatura, do nível de dopagem e do tipo de impureza (doadora ou aceitadora) (Sze, 2002). Para ser detectada, a onda eletromagnética deve penetrar no detector, atravessando a região de depleção até o substrato, onde o fóton promoverá o portador de carga para a banda de condução. Parte da energia da onda incidente é refletida pela superfície do píxel, cuja área útil é somente parte da superfície do sensor (*fill – factor* entre 30 e 80%), enquanto outra parte é transmitida. Uma camada de

óxido de silício colocada sobre o detector reflete a onda espalhada pela superfície do detector de volta para o sensor, aumentando sua eficiência (Teich, 1991).

A eficiência quântica externa η , que também leva em conta a refletância do sistema óptico do sensor, é dada por

$$\eta = (1 - R_d) \left(1 - \frac{e^{-\alpha_{si} W_d}}{(\alpha_{si} L_d + 1)} \right), \quad (2.2)$$

sendo α_{si} o coeficiente de absorção do silício que é função do comprimento de onda da radiação eletromagnética incidente; W_d a largura da zona de depleção da junção PN formada no detector; R_d a refletância do detector e L_d o comprimento de difusão. A Figura 2.4 mostra curvas de coeficiente de absorção em função do comprimento de onda da luz para diversos materiais semicondutores (Sze, 2002).

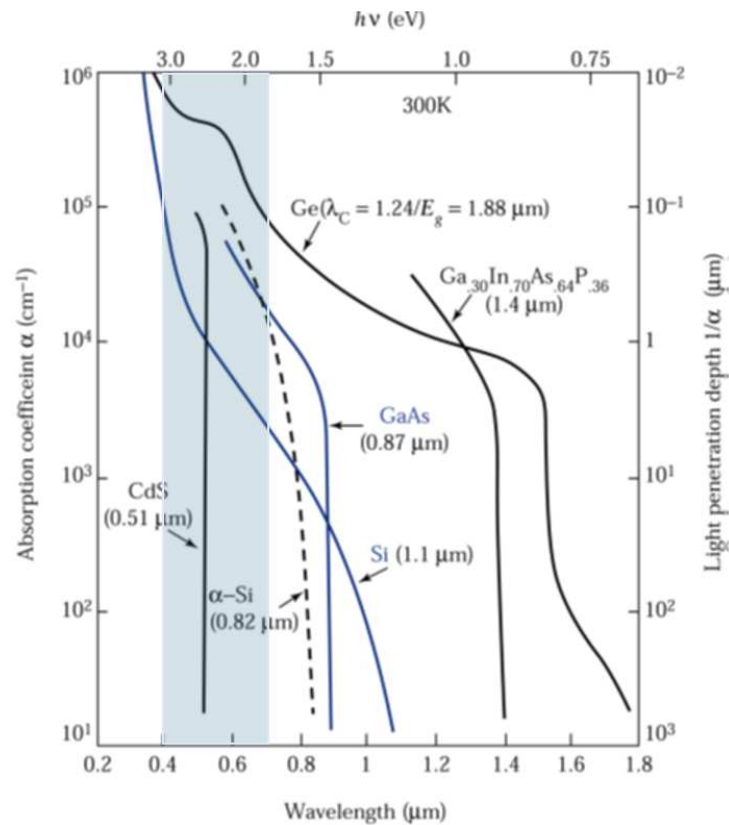


Figura 2.4: Curvas de coeficiente de absorção em função do comprimento de onda da luz para diversos materiais semicondutores. A faixa de comprimentos de onda da luz visível aparece em destaque. A profundidade de penetração é o recíproco do coeficiente de absorção do material. A figura também apresenta o comprimento de onda de corte dos materiais semicondutores - adaptado de (Sze, 2002)

A Figura 2.5 apresenta a curva de profundidade de penetração do silício em função do comprimento de onda. Fótons com maior profundidade de penetração terão maior probabilidade de produzir elétrons detectáveis, a depender da largura da região de depleção. Fótons com comprimento de onda muito pequeno terão pouca penetração, ricocheteando na superfície do detector. Em comprimentos de onda mais longos, a energia do fóton será menor do que a energia do *gap*, não sendo capaz de promover elétrons para a banda de condução.

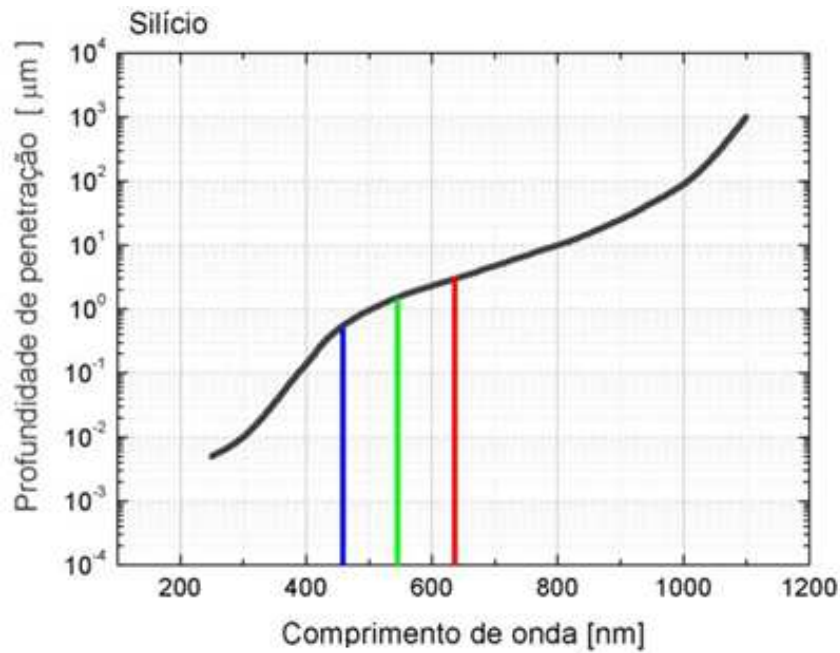


Figura 2.5: Curva de profundidade de penetração em função do comprimento de onda da luz incidente em um detector à base de silício - adaptado de (Sze, 2002). Essa curva corresponde ao recíproco da curva do silício apresentada na Figura 2.4

2.4 Responsividade

A responsividade (A/W) de um detector corresponde à relação entre a corrente elétrica (A - Ampere) produzida e a potência (W - Watt) de luz incidente, sendo função do comprimento de onda, e dependente da tensão de polarização reversa na junção PN, da temperatura, e da eficiência quântica. A responsividade também depende dos tipos de semicondutores que compõem o sensor, da energia do *gap* e outros fatores como nível de dopagem, comprimento de difusão, e dimensões da região de depleção. A largura da região de depleção é modulada pela tensão de polarização dos eletrodos do sensor. A Figura 2.6 mostra curvas de responsividade para diversos materiais semicondutores em função do comprimento de onda do fóton.

Alguns tipos de materiais semicondutores são mais propícios para captura de fótons em comprimentos de onda comparativamente mais curtos, na faixa do espectro visível (VIS - 400 nm a 700 nm), como o silício (Si), fosfeto de gálio (GaP) e fosfeto arseneto de gálio (GaAsP). A resposta do silício se estende até 1100 nm, cobrindo parte da faixa do infravermelho próximo (*NIR – Near infrared* - 700 nm a 1400 nm). Imageadores à base de germânio (Ge) e arseneto de gálio e índio (InGaAs) respondem também a fótons com comprimento de onda na faixa do infravermelho de ondas curtas (*SWIR – Short Wavelength Infrared* - 1400 nm a 3000 nm).

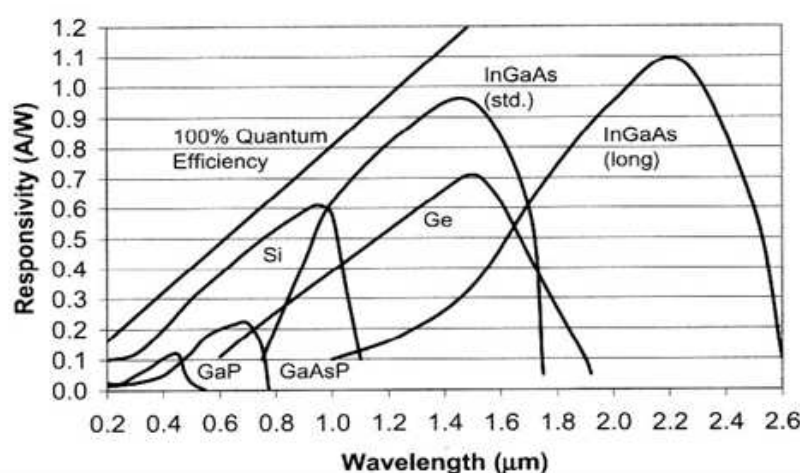


Figura 2.6: Responsividade para diversos materiais semicondutores em função do comprimento de onda. A curva do silício mostra uma responsividade entre 0,2 A/W e 0,6 A/W na faixa de 400 nm a 1000 nm que corresponde à luz visível e ao infravermelho próximo. O Germânio apresenta responsividade semelhante para a faixa final do espectro visível, estendendo-se ao infravermelho próximo (NIR) e ao infravermelho de comprimento de onda curto (SWIR) (Sze, 2002)

Pode-se notar que a responsividade de um dado material semicondutor decresce para comprimentos de onda mais curtos, em função da redução na profundidade de penetração, ao passo que em comprimentos de onda mais longos a energia do fóton é menor, alcançando o comprimento de onda de corte associado à energia do *gap*. O material do sistema óptico do detector também impõe limite ao comprimento de onda do fóton detectável.

Os principais fatores envolvidos na detecção de sinais eletromagnéticos para composição de imagens são os parâmetros que modelam o semicondutor e o sistema óptico, sendo o comprimento de onda do fóton a grandeza física fundamental envolvida na caracterização do processo de detecção (Teich, 1991).

2.5 Ruído em sensores de imagem

O ruído em sensores de imagem do tipo CCD e CMOS está relacionado ao efeito da temperatura nos circuitos eletrônicos do sensor, à natureza estocástica da geração de portadores de carga em um semicondutor, no transporte desses portadores até o eletrodo do detector, com possibilidade de recombinação no percurso, à arquitetura, à tecnologia e processo de fabricação, bem como à etapa de digitalização do sinal analógico do detector. Os principais componentes de ruído em um sensor à base de material semicondutor são descritos nos itens a seguir ([Theuwissen, 2002](#)).

- **Ruído KTC**

Esse componente está associado ao ruído térmico no canal do transistor MOS, durante a carga e descarga do capacitor do detector. Esse fenômeno está relacionado com as flutuações aleatórias na energia dos portadores de carga (elétrons ou lacunas) causadas pela energia térmica. Refere-se à incerteza da carga armazenada no capacitor do detector, carregado via um resistor equivalente ([Theuwissen, 2002](#)). O ruído q_{ktc} (KTC) é dado pela equação

$$q_{ktc} = \frac{\sqrt{kTC}}{e}. \quad (2.3)$$

sendo k a constante de *Boltzmann*, T a temperatura (em *Kelvin*), C a capacitância, e e a carga do elétron.

- **Ruído térmico**

O ruído térmico, também conhecido como ruído de *Johnson-Nyquist*, é uma forma de ruído elétrico que ocorre devido à agitação térmica dos portadores de carga (elétrons) em um condutor elétrico, como resistores, incluindo o canal de um transistor MOS ([Theuwissen, 2002](#)).

A equação do ruído térmico estabelece a variação da tensão eficaz (V_e) como função da temperatura (T), do valor da resistência equivalente (R_{eq}), e da largura de banda (B_z):

$$V_e = \sqrt{4kTR_{eq}B_z}, \quad (2.4)$$

sendo k a constante de *Boltzmann*.

- **Ruído $1/f$ (flicker noise)**

Esse ruído é resultante da flutuação de portadores de carga devido a estados de interface, flutuação de mobilidade, ou ambos. O ruído $1/f$ é principalmente gerado no *driver* da primeira etapa do amplificador de saída. A interação dos estados de interface, localizados na interface $S_i - S_iO_2$ do transistor MOS, com os elétrons no canal de inversão provoca flutuações na tensão de saída da primeira etapa do amplificador. Expressões analíticas para modelagem desse ruído inexistem e são motivo de debate no meio científico (Theuwissen, 2002). Modelamento com expressões semi-empíricas pode ser encontrado em (Tsividis, 1999). Sabe-se que a densidade de potência espectral σ desse componente de ruído é inversamente proporcional à frequência de operação, conforme

$$\sigma = \sqrt{\frac{1}{f}}, \quad (2.5)$$

sendo f a frequência em *Hertz*.

- **Ruído balístico no escuro**

Relaciona-se à flutuação de portadores gerados sem incidência de luz sobre o detector. A geração de corrente de escuro é um processo estocástico e depende da temperatura. Envolve a excitação térmica de portadores de carga através do *gap* de energia do semicondutor. À medida que esses portadores se movem e recombinam, contribuem para a corrente de escuro. A natureza aleatória desse processo leva a flutuações na corrente ao longo do tempo.

O ruído balístico de escuro é devido à variabilidade no número de portadores de corrente de escuro que chegam à saída do dispositivo em um determinado período. Como a geração de corrente de escuro envolve eventos discretos (portadores individuais sendo gerados termicamente), há uma aleatoriedade inerente nos tempos de chegada e no número desses portadores (Theuwissen, 2002). Essa aleatoriedade resulta no ruído balístico de escuro $I_{d_{rms}}$, conforme equação

$$I_{d_{rms}} = \sqrt{2eI_{dark}B_z}, \quad (2.6)$$

onde I_{dark} é a corrente de escuro, e é a carga do elétron, e B_z a largura de banda.

- **Ruído balístico fotônico**

O ruído balístico fotônico refere-se ao fenômeno de flutuações de corrente causadas pela natureza discreta dos fótons em um sistema. Está relacionado ao caráter discreto dos fótons e à natureza quântica da luz. Quando fótons incidentes atingem um fotodetector, eles são absorvidos, gerando pares elétron-lacuna. Como os fótons chegam ao detector de maneira discreta, sua absorção e geração de pares ocorrem de forma não contínua e probabilística. Isso cria flutuações na corrente fotônica que correspondem ao ruído balístico fotônico (Theuwissen, 2002). O ruído balístico fotônico $I_{p_{rms}}$ em um fotodetector pode ser expresso como

$$I_{p_{rms}} = \sqrt{2eI_pB_z}, \quad (2.7)$$

sendo I_p a corrente fotogerada, e a carga do elétron, e B_z a largura de banda.

A intensidade do ruído balístico fotônico associado à corrente de escuro aumenta com o aumento do tempo de exposição e com a temperatura no sensor. A intensidade da corrente de escuro em si determina um *offset* que se soma à corrente do sinal fotônico.

- Outras fontes de incerteza

Há também o ruído de quantização do conversor analógico-digital utilizado para leitura do sensor de imagem e incertezas relacionadas a não uniformidade da resposta do arranjo sensor, com variações na eficiência quântica de píxel para píxel, e no ganho do amplificador de saída do detector, além da ocorrência de contaminantes no sensor (Theuwissen, 2002).

2.6 Calibração radiométrica

A calibração radiométrica visa relacionar a intensidade radiante detectada por uma câmera com a quantidade de luz incidente sobre a cena, para fins de medição do espectro de reflectância. Esse processo melhora a precisão e exatidão das informações obtidas, permitindo comparações consistentes entre diferentes imagens, e a extração de parâmetros físicos significativos. A base teórica da calibração radiométrica envolve conceitos da teoria da radiometria, considerando a resposta espectral da câmera, a sensibilidade espectral do sensor e as características de iluminação da cena imageada. A correção de fatores como reflexões especulares, sombras e variações

temporais na iluminação são importantes para captura de dados de forma confiável (Foster e Amano, 2019). A calibração em laboratório envolve, frequentemente, a utilização de fontes de luz com intensidade e distribuição de potência espectral conhecida, enquanto a calibração no campo pode se basear em condições naturais ou padrões específicos presentes na cena, tais como alvos de referência de branco.

A referência de branco é um elemento para calibração radiométrica que oferece uma superfície uniforme que reflete a luz de forma difusa e com máxima reflectância (Foster e Amano, 2019). Uma referência de branco comumente utilizada é composta por sulfato de bário que apresenta reflectância da ordem de 99% (Moraes et al., 1996).

2.7 Compensação da corrente de escuro e calibração radiométrica

Uma compensação devida à variação da condição de iluminação da cena pode ser alcançada por meio da calibração radiométrica, com identificação da intensidade de píxeis em uma referência de branco.

Uma compensação do *offset* determinado pela corrente de escuro pode ser feita por meio da captura de dados do sensor, sem a incidência de luz (obturador fechado, por exemplo).

A calibração radiométrica e a compensação da corrente de escuro podem ser feitos utilizando a equação

$$I_m(x_p, y_p, \lambda) = \frac{I_s(x_p, y_p, \lambda) - I_d(x_p, y_p, \lambda)}{I_w(x_p, y_p, \lambda) - I_d(x_p, y_p, \lambda)} \quad (2.8)$$

sendo $I_m(x_p, y_p, \lambda)$ a intensidade do píxel na imagem, após compensação da corrente de escuro e correções radiométricas, indexado pelas coordenadas espaciais x_p , y_p e comprimento de onda λ ; $I_s(x_p, y_p, \lambda)$ a intensidade do píxel na imagem conforme lido do sensor de imagem; $I_w(x_p, y_p, \lambda)$ a intensidade do píxel correspondente na imagem de referência de branco; $I_d(x_p, y_p, \lambda)$ a intensidade do píxel correspondente na imagem de referência de escuro (Foster e Amano, 2019).

2.8 Calibração geométrica

A lente do sistema óptico de uma câmera não é perfeita, devido ao processo de fabricação que, tipicamente, produz lentes esféricas, mais fáceis de fabricar do que lentes parabólicas, acarretando distorções na imagem. A precisão e exatidão do processo, por melhor que sejam, ainda são limitadas. Outro fator é o alinhamento mecânico do eixo óptico com o sensor. Distorções radiais são devidas a não uniformidade da lente, ao passo que distorções tangenciais estão associadas ao processo de montagem e alinhamento do sensor da câmera (Bradski e Kaehler, 2008). A distorção radial relaciona-se a píxeis próximos às bordas do sensor serem deslocados, provocando distorções na imagem. A Figura 2.7 apresenta, de forma qualitativa, o efeito da distorção radial, no qual os raios de luz mais distantes do centro da lente são mais curvados do que aqueles mais próximos. A Figura 2.8 apresenta o conceito de distorção tangencial.

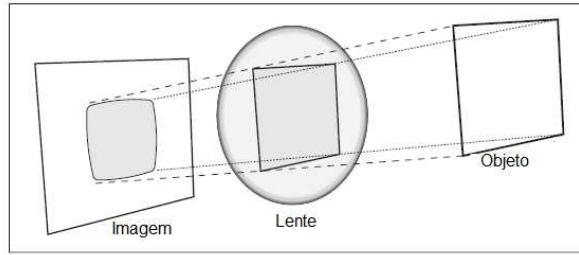


Figura 2.7: Distorção radial (forma qualitativa): raios de luz mais afastados do centro de uma lente simples são curvados em comparação aos raios que passam mais próximos ao centro (Bradski e Kaehler, 2008).

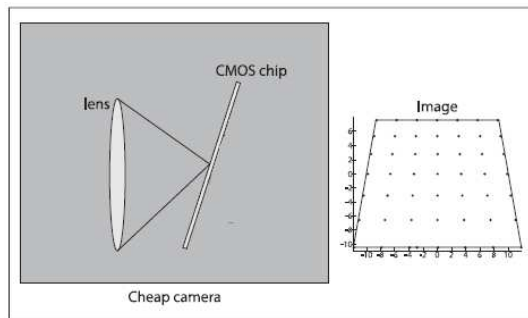


Figura 2.8: Distorção tangencial: ocorre quando a lente não está paralela ao plano da imagem (Bradski e Kaehler, 2008).

Em geral, a localização radial de um ponto no sensor de imagem pode ser definido para compensação de distorções radiais conforme as equações

$$x_r = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) x_s, \quad (2.9)$$

$$y_r = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) y_s. \quad (2.10)$$

sendo k_1 , k_2 e k_3 , coeficientes de distorção radial; x_r e y_r são as coordenadas dos píxeis corrigidos para compensar distorções radiais; x_s e y_s são as coordenadas originais dos píxeis na imagem lida do sensor.

As correções devidas à distorção tangencial são feitas segundo as equações

$$x_t = [2p_1 y_s + p_2 (2x_s^2 + r^2)] + x_s, \quad (2.11)$$

$$y_t = [p_1 (2y_s^2 + r^2) + 2p_2 x_s] + y_s. \quad (2.12)$$

sendo p_1 , p_2 coeficientes de distorção tangencial; r é a distância radial do ponto ao centro da imagem; x_t e y_t são as coordenadas dos píxeis corrigidos para compensar distorções tangenciais; x_s e y_s são as coordenadas originais dos píxeis na imagem lida do sensor (Bradski e Kaehler, 2008).

Um padrão comum usado para calibração geométrica corresponde a um tabuleiro de xadrez. Ao capturar imagens de um padrão, a partir de ângulos diferentes, o software de calibração detecta pontos característicos, como os vértices (*corners*) do padrão de xadrez, estabelece correspondências entre características detectadas nas imagens, usa modelos matemáticos para estimar parâmetros intrínsecos e extrínsecos, minimizando as diferenças entre os pontos observados na imagem e os pontos correspondentes previstos pelo modelo de calibração. Essa diferença é conhecida como erro de reprojeção. A matriz de parâmetros intrínsecos $\mathbf{M}_p$ é dada por

$$\mathbf{M}_p = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (2.13)$$

onde f_x é distância focal no eixo x , f_y é distância focal no eixo y , (C_x, C_y) são as coordenadas de referência do sensor. Essa matriz pode ser usada para projetar um ponto $\mathbf{p}_m$ no mundo com coordenadas (x, y, z) em um ponto $\mathbf{p}_s$ no plano da imagem com coordenadas (x_s, y_s, w_s) , denominada transformação projetiva (Bradski e Kaehler, 2008), conforme equação

$$\begin{bmatrix} x_s \\ y_s \\ w_s \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (2.14)$$

A coordenada w_s é utilizada como artifício matemático na equação matricial, e resulta da projeção do espaço de três dimensões do objeto para o espaço do sensor, com apenas duas dimensões. De forma sintética tem-se que

$$\mathbf{p}_s = \mathbf{M}_p \cdot \mathbf{p}_m. \quad (2.15)$$

Essa equação é utilizada no processo de calibração para obter a matriz $\mathbf{M}_p$, bem como os coeficientes k_1 , k_2 , k_3 , p_1 , e p_2 . Ferramentas de software, como as bibliotecas de código aberto do OpenCV, fornecem funções para a calibração de câmeras. Essas ferramentas, frequentemente, incluem funções para detecção de padrões, correspondência de características e otimização de parâmetros (Bradski e Kaehler, 2008).

2.9 Câmeras RGB, multiespectrais e hiperespectrais

O uso de sensores multiespectrais e hiperespectrais na identificação e coleta de informações acerca de elementos da superfície terrestre é uma realidade há algumas décadas, instalados em sistemas satelitais e aeronaves tripuladas. Também vêm sendo embarcados em VANTs - veículos aéreos não tripulados (drones) - em anos recentes. A combinação de dados obtidos por sensores para compor imagens multiespectrais e hiperespectrais abriu novos campos de aplicação na agricultura de precisão (Thenkabail et al., 2018), (Lowe et al., 2017), diagnóstico médico (Guolan e Baowei, 2014), inspeção de alimentos (Sun, 2010), análise de composição de solos e identificação mineral (Roger e Clark, 2002), mapeamento hidrográfico (Rafaele Furtado et al., 2020), investigação científica criminal (Malkoff e Oliver, 2000), recuperação de documentos (Qureshi et al., 2019), dentre vários.

Câmeras multiespectrais e hiperespectrais possuem sensores com capacidade para detectar uma maior quantidade de bandas e abrangem faixas espectrais além do visível. Tipicamente, câmeras multiespectrais abrangem de 5 a 10 bandas com largura entre 70 nm e 400 nm, em faixas

não contíguas, enquanto câmeras hiperespectrais possuem maior resolução espectral, possibilitando obter dados associados a mais de 500 bandas, com largura típica entre 1 e 10 nm, em faixas contíguas (Shukla e Kot, 2016). A técnica utilizada em câmeras hiperespectrais é conhecida como espectroscopia, espectrometria ou espectrorradiometria. Espectroscopia é o ramo da física que se dedica ao estudo da produção, transmissão e interpretação do espectro eletromagnético. Espectrometria ou espectrorradiometria é derivada da espectrofotometria, a medição de fótons em função do comprimento de onda. Imageamento hiperespectral é um processo para obter um espectro em cada posição de um arranjo espacial, de forma que cada comprimento de onda possa ser usado para produzir uma imagem completa. Dados hiperespectrais possibilitam identificar propriedades dos elementos que compõem o objeto sendo imageado, a partir do conteúdo da radiação eletromagnética emitida, transmitida, refletida e absorvida. A maior resolução espectral fornece informações de detalhes da superfície do objeto e mesmo dos materiais que o compõem. A identificação da composição é feita com base no espectro de absorção e emissão que revelam características físico-químicas do material (Shukla e Kot, 2016). Em câmeras multiespectrais e RGB, estas características são espalhadas nas bandas de detecção, dificultando a identificação. O conteúdo de sinal em cada banda espectral, nos píxeis de imagens RGB (*Red, Green, Blue*), corresponde ao fluxo de fótons em uma faixa de comprimentos de onda definida pela função de espalhamento espectral do sistema óptico (*PSF - Point Spreading Function*) e pela função de transferência do filtro de cor, bem como pela responsividade e limiar de sensibilidade do sensor (Foster e Amano, 2019). As câmeras RGB capturam apenas três canais de cor (vermelho, verde, azul), definindo três bandas espectrais amplas, na faixa da luz visível ao olho humano (VIS), cobrindo de 400 nm a 700 nm, enquanto câmeras hiperespectrais capturam dados em centenas de bandas espectrais estreitas e contíguas, abrangendo a região espectral VIS (*Visual*), NIR (*Near Infrared*) (Foster e Amano, 2019), e além, SWIR (*Short Wavelength InfraRed*), MWIR (*Middle Wavelength InfraRed*) e LWIR (*Long Wavelength InfraRed*) (Signoroni et al., 2019). A Figura 2.9 apresenta as bandas espectrais em imageadores hiperespectrais, geralmente consideradas para detecção de incêndios florestais.

Todo material, incluindo fogo e vegetação, possui uma assinatura espectral específica. Esta assinatura descreve como o material interage com a luz em diferentes comprimentos de onda. Analisando essas assinaturas, pode-se distinguir diferentes materiais com base em suas propriedades. O fogo emite uma combinação de radiação térmica e luz visível. Essa emissão pode ser capturada por sensores RGB, multiespectrais e hiperespectrais. A energia emitida muda

Faixa espectral	Acrônimo	Faixa de comprimento de onda
Visível ao olho humano	VIS	0,4 a 0,7 μm
Infravermelho próximo	NIR	0,75–1,4 μm
Infravermelho de comprimento de ondas curto	SWIR	1,4–3,0 μm
Infravermelho de comprimento de ondas médio	MWIR	3,0–8,0 μm
Infravermelho de comprimento de ondas longo/Infravermelho Termal	LWIR (TIR)	8,0–15,0 μm

Figura 2.9: Bandas espectrais consideradas na detecção de incêndios florestais.

com a temperatura e os tipos de materiais em combustão. A espectroscopia do fogo envolve o estudo das assinaturas espectrais específicas associadas a chamas, fumaça e outros fenômenos relacionados ao fogo (Allison et al., 2016).

O processo de síntese hiperespectral produz um “cubo hiperespectral” a partir de informações espectrais limitadas, como as de imagens RGB. Essa síntese utiliza métodos estatísticos, modelos matemáticos e bibliotecas espectrais conhecidas para prever as propriedades espectrais de cada píxel na cena. Câmeras RGB têm a seu favor o fato de serem fáceis de manusear, amplamente utilizadas e de muito menor custo quando comparadas a câmeras hiperespectrais (Shukla e Kot, 2016). Em imageadores RGB, a energia luminosa é dividida apenas em três bandas amplas, proporcionando boa sensibilidade, enquanto as câmeras hiperespectrais têm sensibilidade reduzida devido à quantidade de energia luminosa que chega ao sistema óptico ser dividida em muitas bandas. Além disso, câmeras hiperespectrais do tipo “*pushbroom*” (varredura de linha) e “*whiskbroom*” (varredura de área) são lentas na aquisição de imagens, em comparação com câmeras RGB, devido ao processo de varredura (Signoroni et al., 2019). Por todas essas razões, há um interesse crescente em desenvolver métodos para incorporar a capacidade de fornecer alta resolução espectral, velocidade e sensibilidade com base em câmeras RGB, por meio de algoritmos de aprendizado de máquina. Isso pode ser alcançado pelo conhecimento do conteúdo espectral de cenários capturados por câmeras hiperespectrais e sua correlação com os mesmos cenários capturados por câmeras RGB. O processo de síntese de imagens hiperespectrais visa recuperar o espectro do sinal original que produziu uma resposta espectral esparsa em um sen-

sor RGB. No contexto da detecção de incêndios florestais, a adoção da análise hiperespectral fornece a capacidade de distinguir a assinatura espectral do fogo de outras fontes de radiação eletromagnética com cores RGB semelhantes. Ao combinar a estimativa hiperespectral com análises espaciais e temporais tradicionais, um sistema de monitoramento de incêndios florestais que utilize essas técnicas pode melhorar a precisão e a acurácia da detecção.

2.9.1 Tipos de sensores hiperespectrais

Diferentes técnicas são empregadas na obtenção de dados espectrais, havendo, em geral, limitações físicas que determinam um compromisso entre resolução espacial, resolução espectral, taxa de amostragem, complexidade computacional, sensibilidade e dimensões do dispositivo sensor. Os sensores hiperespectrais do tipo *whiskbroom* e *pushbroom* exploram o movimento relativo entre o detector e a cena sendo imageada para efetuar a varredura espacial, e são usados especialmente em sensores satelitais. O tipo *whiskbroom* efetua uma varredura de área enquanto o tipo *pushbroom* efetua uma varredura de linha. A captura do espectro é tipicamente feita com uso de prisma ou grade de difração. O tipo *whiskbroom* captura o espectro de um único píxel em um dado instante, enquanto o tipo *pushbroom* obtém o espectro de uma linha inteira da cena, em um dado momento. No tipo *whiskbroom*, a amostragem temporal é muito penalizada em favor de uma melhor resolução espacial. O uso prático desse tipo de sensor na composição de imagens hiperespectrais é muito limitado, uma vez que o tempo para obtenção de uma imagem completa em alta resolução seria da ordem de minutos. O tipo *pushbroom* é mais utilizado do que o *whiskbroom* e possibilita a aquisição de imagens com maior resolução espectral e maior taxa de amostragem temporal. O esquema ilustrado na Figura 2.10 corresponde a um sensor de varredura de linha do tipo *pushbroom* (Signoroni et al., 2019).

Sensores hiperespectrais colecionam dados como um conjunto de imagens. Cada imagem representa uma faixa estreita de comprimentos de onda do espectro eletromagnético. Estas imagens formam uma estrutura tridimensional (“cubo hiperespectral”) onde x e y representam as dimensões espaciais e λ representa a dimensão espectral, conforme pode ser visto na Figura 2.10. Uma fenda (*slit*) restringe a área imageada na varredura de linha. Um difrator espalha os feixes de luz em diferentes posições do sensor. Cada quadro capturado contém dezenas, ou mesmo centenas, de bandas espectrais de uma faixa estreita da cena. Outro esquema de

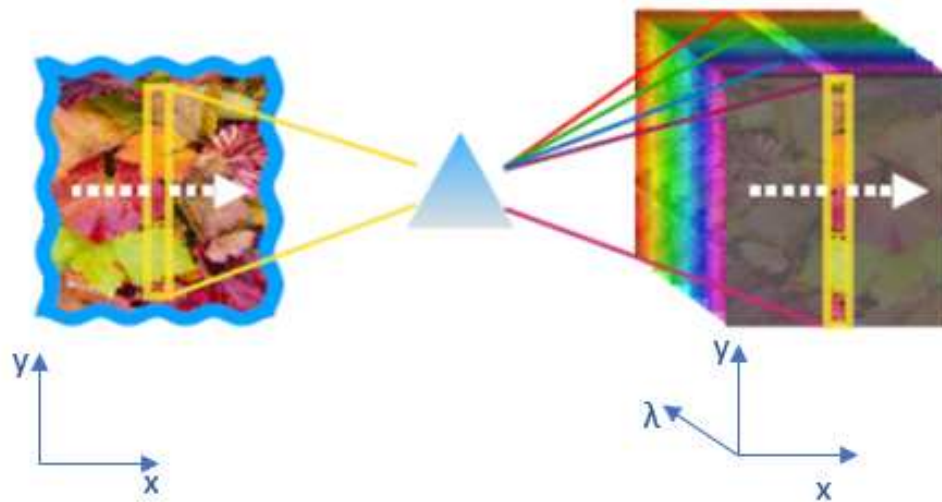


Figura 2.10: Técnicas de sensoriamento hiperespectral com esquema de varredura de linha e composição do “cubo hiperespectral”. O “cubo” possui duas dimensões espaciais (x , y) e uma espectral (λ). Adaptado de (Signoroni et al., 2019).

sensoriamento utiliza aquisição espectral seletiva no tempo. Neste caso, é necessário incorporar um esquema de seleção de comprimento de onda com uso de um filtro acústico-óptico, ou um filtro de cristal líquido sintonizável. Este processo é ilustrado na Figura 2.11. Cada quadro contém a informação de intensidade de luz em um comprimento de onda central.

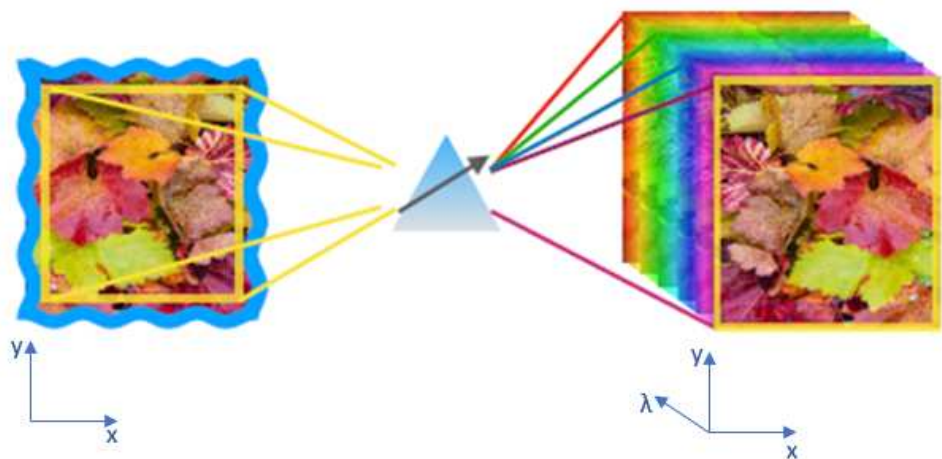


Figura 2.11: Técnicas de sensoriamento hiperespectral com uso de filtro sintonizável do tipo acústico-óptico. O cubo de dados possui duas dimensões espaciais (x , y) e uma espectral (λ). Adaptado de (Signoroni et al., 2019).

O esquema apresentado na Figura 2.12 contém um mosaico de difratores para possibilitar a captura do espectro em cada região da cena, em única tomada. Cada região da cena está associada a um difrator que separa os espectros em diferentes posições no detector.

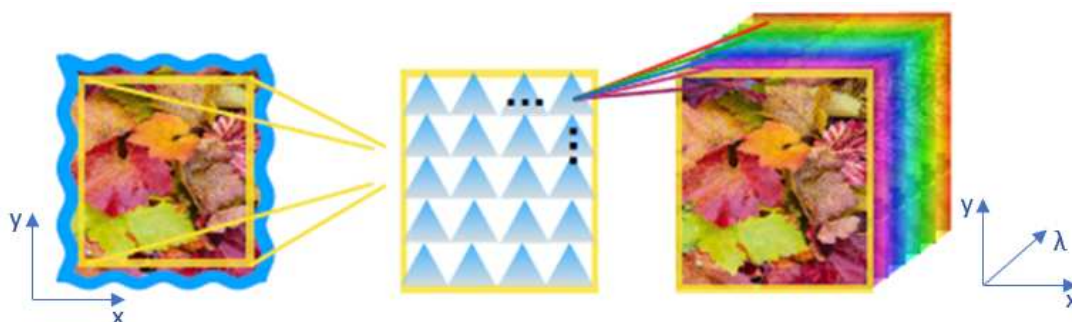


Figura 2.12: Técnicas de sensoriamento hiperespectral com uso de arranjo de difratores. Adaptado de (Signoroni et al., 2019).

2.10 Detecção de fogo e fumaça

Diferentes estratégias vêm sendo utilizadas para detecção de incêndios florestais por imagens de satélites, ou obtidas por câmeras embarcadas em aeronaves tripuladas e não tripuladas. Há técnicas que se baseiam no movimento da chama (*flicker*) e da fumaça (mudança de forma). O calor do fogo produz um sinal de radiação intenso que se destaca da temperatura típica encontrada na superfície. Os processos químicos e os sinais físicos advindos da queima também oferecem oportunidades para aplicação de diversas técnicas e abordagens na detecção (Allison et al., 2016). A matéria emite energia no espectro eletromagnético a depender de sua composição e de sua temperatura. O modelo de emissão de corpo-negro descreve a radiação eletromagnética em função do comprimento de onda, para o corpo em equilíbrio térmico, segundo a lei de Planck (Signoroni et al., 2019). A Figura 2.13 (a) ilustra a radiação de corpo negro para diferentes temperaturas entre 248 K e 323 K. À medida que a temperatura aumenta, a intensidade radiada aumenta e o pico da curva desloca-se para comprimentos de onda menores. Com a temperatura entre 248 K e 323 K, os picos de intensidade da radiação estão localizados entre 8 μm e 12 μm , o que corresponde à faixa TIR (*Thermal Infrared*). Para temperaturas muito elevadas, entre 600 K e 1200 K, conforme Figura 2.13 (b), os picos de emissão deslocam-se para comprimentos de onda entre 3 μm e 5 μm , correspondendo à faixa MWIR. A maioria das técnicas voltadas para detecção de assinatura de calor do fogo são baseadas em sensores TIR e MWIR (Signoroni et al., 2019).

A faixa TIR é transparente para fumaça e possibilita identificar a localização de focos de incêndio, mesmo com fumaça densa. Na faixa TIR, a energia é principalmente determinada pela emissão, sendo baixa a contribuição da energia refletida (Signoroni et al., 2019). A faixa

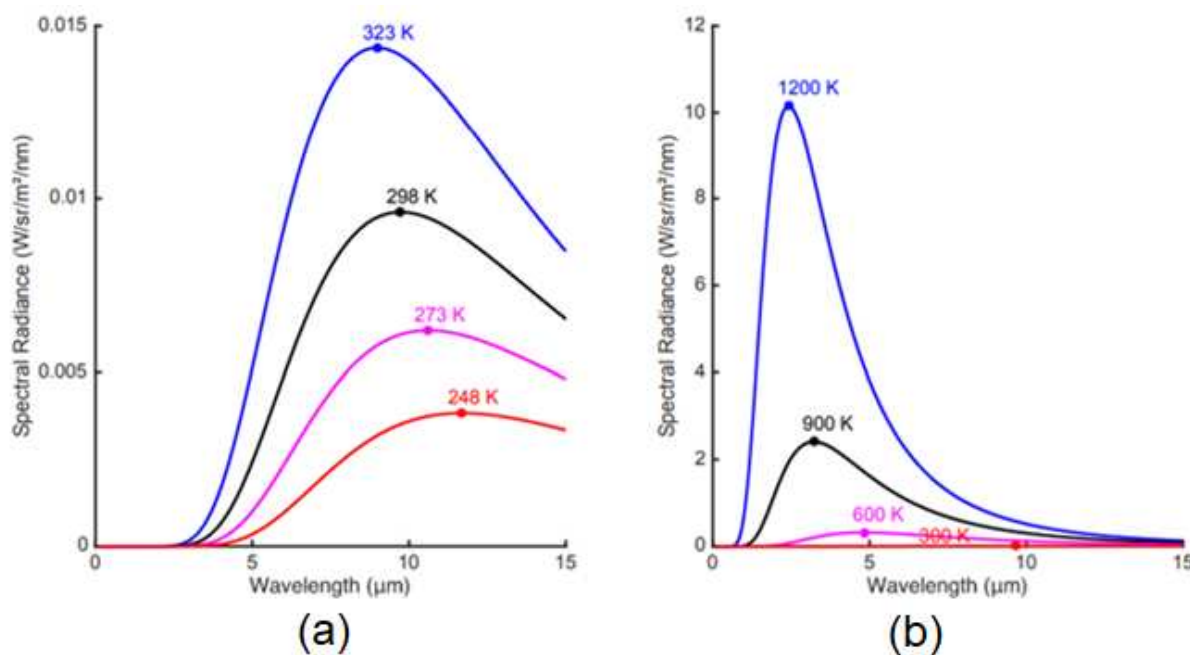


Figura 2.13: (a) Espectro de corpo negro a diferentes temperaturas. (b) Com o aumento da temperatura de 323 K para 1200 K, o pico de radiação desloca-se da faixa entre 8 μm e 12 μm para a faixa entre 3 μm e 5 μm (Signoroni et al., 2019).

MWIR é útil em cenários onde a energia emitida pela chama é alta, em comparação com fontes não associadas ao fogo. A detecção de emissão na faixa MWIR é afetada durante o dia devido à reflexão da luz solar (Signoroni et al., 2019). A detecção baseada na medição de uma única faixa pode conduzir a ambiguidades, sendo preferível medir a energia em diferentes comprimentos de onda. Uma abordagem possível seria identificar a forma do espectro e a localização do pico de energia em relação ao comprimento de onda, obter correlações entre bandas, efetuar correção radiométrica e calcular índices de normalização (Signoroni et al., 2019).

Outra estratégia para detecção de incêndios é baseada na composição da biomassa de florestas, sendo tipicamente 45% de carbono, 5.5% de hidrogênio, 41% de oxigênio, 3.5% de nitrogênio, 7% de potássio, 0.1% de sódio, até 1% de fósforo e até 5% de cálcio (Vodacek et al., 2001).

Potássio é um candidato a ser utilizado para detecção de focos de incêndio, por meio do decaimento de elétrons termicamente excitados na região da chama, com emissão de intenso sinal eletromagnético. Para o potássio, as emissões encontram-se na banda de 766,5 nm e de 769,9 nm, localizadas na faixa NIR (Vodacek et al., 2002). A Figura 2.14 mostra o espectro do fogo em comparação com demais elementos da cena.

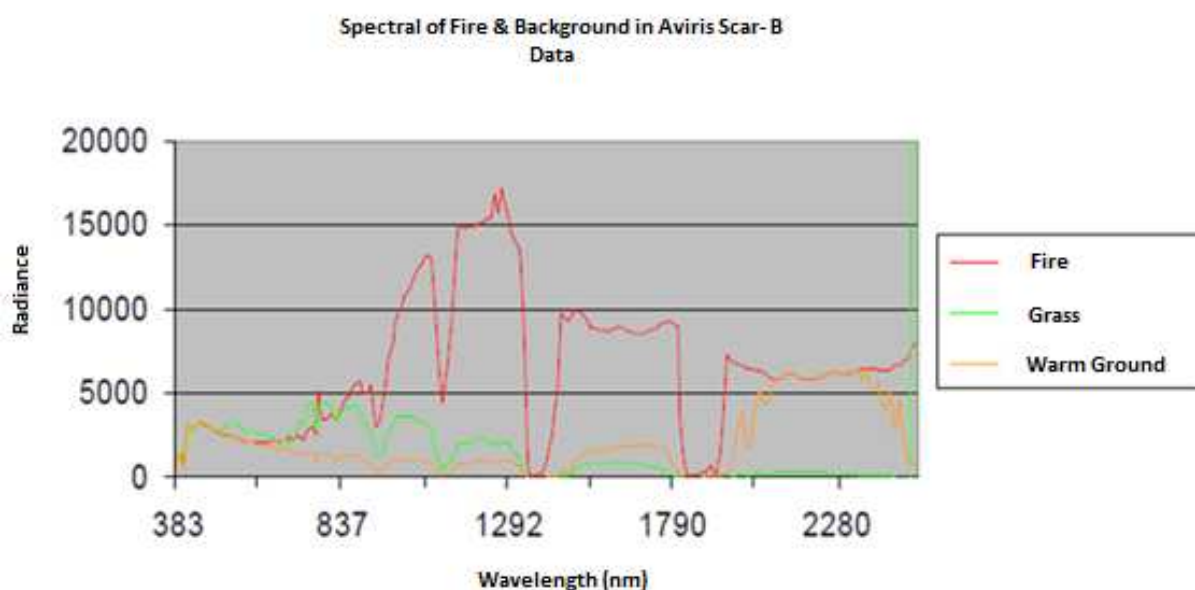


Figura 2.14: Assinatura espectral de fogo, grama e solo aquecido. Observa-se o destaque da intensidade da emissão do fogo na faixa visível da luz, mas sobretudo na faixa do infravermelho de comprimento de ondas longo (Vodacek et al., 2001).

A Figura 2.15 mostra uma imagem de uma área com focos de incêndio e fumaça, feita a partir de satélite. A Figura 2.15 (A) mostra a imagem para o comprimento de onda de 589 nm. Os focos de incêndio são ocultos pela fumaça que não é transparente neste comprimento de onda. Na Figura 2.15 (B) os focos podem ser vistos em uma imagem cujo comprimento de onda corresponde a 770 nm, no qual a fumaça é transparente. A Figura 2.16 (C) mostra uma imagem para o comprimento de onda de 1500 nm, no qual a fumaça também é transparente. A Figura 2.16 (D) mostra uma imagem sintetizada a partir da razão entre as intensidades nos comprimentos de onda 769 nm (emissão do potássio) e 779 nm (banda de normalização). Um filtro de limiar foi aplicado aos dados obtidos através desse cálculo, resultando na imagem da Figura 2.16 (D) (Vodacek, et al., 2021).

2.10.1 Decaimento de elétrons em átomos de potássio

As emissões de potássio estão localizadas na faixa do infravermelho próximo (NIR), especificamente nos comprimentos de onda de 766,5 nm e 769,9 nm (Vodacek et al., 2002). Em baixas temperaturas, o número de átomos de potássio em um estado de energia excitado é estatisticamente baixo. Apenas poucos elétrons encontram-se em níveis quânticos de energia elevada, e esses estão propensos a decair, emitindo luz na faixa do infravermelho próximo. Quando a

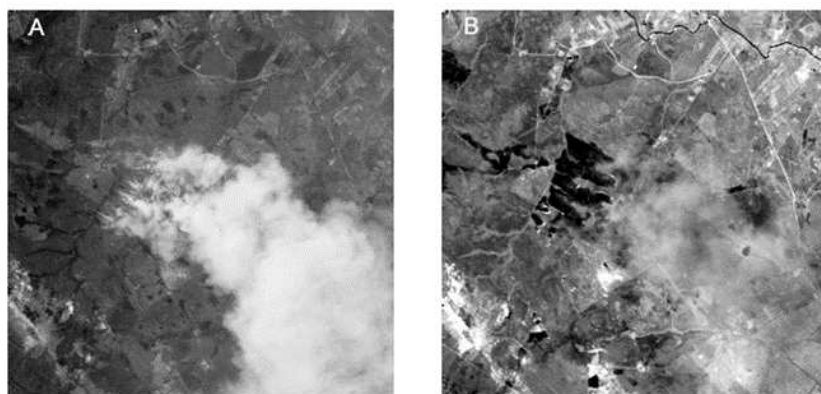


Figura 2.15: Imagens de incêndio em diferentes comprimentos de onda. Imagem em (A) 589 nm e em (B) 770 nm. Note-se que os focos de incêndio na imagem em 770 nm são visíveis, pois neste comprimento de onda a fumaça é transparente (Vodacek, et al., 2021).

vegetação contendo potássio sofre combustão em um incêndio florestal, o calor intenso impulsiona elétrons para níveis quânticos de energia mais alta. Os níveis de energia específicos e as transições que os elétrons sofrem são regidos pelos princípios da mecânica quântica. Essas transições são quantizadas, o que confina os elétrons a ocupar níveis de energia distintos dentro do átomo. À medida que os elétrons em níveis de energia mais alta transitam para níveis mais baixos, liberam energia na forma de fótons. O comprimento de onda desses fótons corresponde à faixa do infravermelho próximo do espectro eletromagnético (Ononye et al., 2005). Este sinal emitido pode servir como uma assinatura para localizar as chamas de um incêndio florestal.

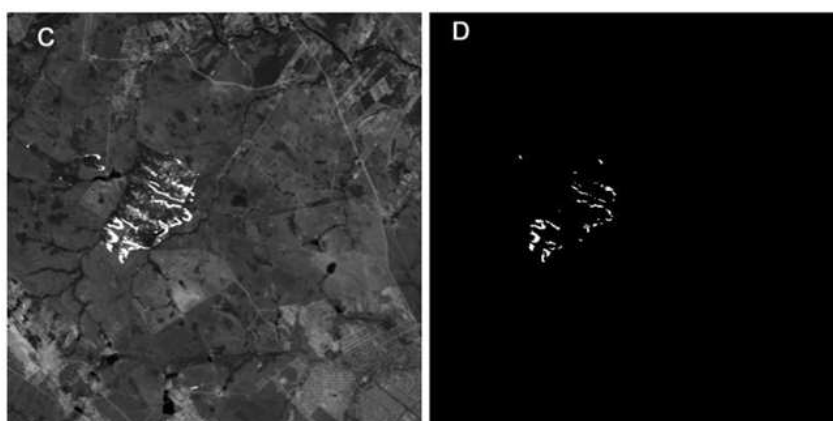


Figura 2.16: Imagens de incêndio em diferentes comprimentos de onda, em (C) 1550 nm e em (D) aplicando filtro 6σ para a razão 769/779 nm. Os focos de incêndio na imagem em 1500 nm são visíveis, pois, neste comprimento de onda, a fumaça é transparente (Vodacek et al., 2001).

2.10.2 Equação de Boltzmann

Átomos podem existir em qualquer estado quântico excitado, cuja população pode ser comparada com a do estado não excitado por meio da estatística de *Boltzmann*, com a condição de não ionização (Carroll e Ostlie, 2006), estimada pela equação

$$R_i = \frac{N_2}{N_1} = \frac{g_2}{g_1} e^{-\left(\frac{E_1 - E_0}{kT}\right)}. \quad (2.16)$$

Essa equação representa a proporção relativa de átomos não excitados, com átomos no primeiro nível de excitação. Sendo g_1 e g_2 os pesos estatísticos dos níveis 1 e 2; N_2 e N_1 são, respectivamente, as densidades numéricas de átomos nos estados excitado e não excitado; k é a constante de Boltzmann; E_1 e E_0 são a energia do primeiro estado excitado e do estado fundamental, respectivamente; T é a temperatura do ambiente no qual os átomos estão inseridos. O estado fundamental é o mais populado, com os estados excitados sendo apenas esparsamente populados. Conforme a temperatura aumenta, a probabilidade de transição do estado fundamental para o primeiro estado excitado e demais estados excitados (de maior energia) também aumenta, dando origem à observada linha espectral de emissão do potássio. A intensidade dessa assinatura espectral incrementa com a população dos estados excitados. Utilizando a equação de *Boltzmann*, é possível prever a variação das intensidades espectrais com a temperatura, bem como a variação das intensidades espectrais com a concentração de átomos de potássio (Ononye et al., 2005).

2.10.3 Equação de Saha

A equação de *Boltzmann* não leva em conta a condição de átomos ionizados. Para isso, utiliza-se a equação de Saha, possibilitando a comparação do número de átomos ionizados, com o número de átomos não ionizados, a uma dada temperatura:

$$R_j = \frac{N_{i+1}}{N_i} = \frac{2u_{i+1}(T)}{n_e u_i(T)} \left(\frac{2\pi m_e kT}{h^2} \right)^{\frac{3}{2}} e^{-\frac{E_i}{kT}}, \quad (2.17)$$

onde N_{i+1} é o número de átomos no estado de ionização $i+1$, e N_i o número de átomos no estado inicial de ionização, i ; $u_{i+1}(T)$ e $u_i(T)$ são funções de partição para íons, respectivamente, nos estados $i+1$ e i . As funções de partição correspondem a soma ponderada das diferentes formas

com que elétrons podem ser distribuídos nos átomos em certo nível de energia. A razão aumenta com a temperatura, mas diminui com o aumento da densidade de elétrons n_e . m_e é a massa do elétron; k é a constante de *Boltzmann*; T é a temperatura, que no contexto de incêndios, reflete a temperatura das chamas; h é a constante de Planck; i , identifica os estados excitados e ionizados do potássio; E_i é a energia de ionização necessária para transitar do nível i para $i+1$.

A equação de Saha descreve a distribuição estatística de partículas carregadas em um gás ionizado, como o plasma em uma estrela. Esta equação é fundamental para entender a ionização de átomos em diferentes estados energéticos, em função da temperatura. A correlação específica com a ionização de átomos de potássio reside na dependência da equação em relação à temperatura e à energia de ionização. Em um incêndio florestal, a temperatura elevada pode resultar em uma população significativa de átomos de potássio em estados excitados, refletindo os estágios iniciais de ionização. Os elétrons resultantes da ionização podem recombinar e, assim, popular níveis energéticos de excitação do átomo de potássio e em seguida decair ao nível fundamental, propiciando a emissão de radiação eletromagnética na faixa NIR (Ononye et al., 2005).

Uma equação adaptada de *Saha-Boltzmann* pode ser usada como uma ferramenta conceitual para modelar a relação entre a temperatura da chama, a ionização de átomos de potássio e a emissão de radiação eletromagnética em determinados comprimentos de onda, como aqueles na faixa do infravermelho próximo, mencionada anteriormente. Essa modelagem contribui para a compreensão das características espectrais observadas em incêndios florestais. Uma emissão amplificada, relacionada à excitação inicial de elétrons dentro dos átomos de potássio, torna-se especialmente pronunciada na faixa de temperatura de 526 °C a 1027 °C. A correlação entre a intensidade da emissão e a temperatura é caracterizada pelas estatísticas de *Saha-Boltzmann* (Ononye et al., 2005).

2.11 Métodos de Regressão Linear

A regressão linear é uma técnica estatística usada para modelar e analisar a relação entre duas ou mais variáveis. O objetivo é encontrar a melhor reta, ou hiperplano, que descreva a relação entre uma variável dependente e uma ou mais variáveis independentes, possibilitando

estimar o valor da variável dependente com base nos valores das variáveis independentes, assim como compreender a relação entre a variável dependente e as variáveis independentes. A aplicação da regressão linear a um conjunto de dados para estimar um modelo de representação pressupõe que a relação entre as variáveis independentes e a variável dependente seja linear, assim como a independência entre as variáveis observadas. Também assume que variância dos erros é constante em todos os níveis das variáveis independentes, assim como a normalidade dos resíduos.

Os tipos de regressão linear são:

- Regressão Linear Simples

Este é o modelo mais básico, onde há apenas uma variável independente X_d e uma variável dependente Y_p . O modelo é expresso pela equação

$$Y_p = \beta_0 + \beta_1 X_d \quad (2.18)$$

sendo β_0 o intercepto e β_1 a inclinação da reta.

- Regressão Linear Múltipla

Este modelo envolve mais de uma variável independente X_{d_i} para prever a variável dependente Y_p . O modelo é expresso pela equação

$$Y_p = \sum_{i=1}^{n_d} (\beta_i X_{d_i}). \quad (2.19)$$

sendo β_i os coeficientes do modelo multivariável.

- Regressão Ridge

O modelo de regressão Ridge é um tipo de técnica de regressão linear que introduz um termo de regularização na equação visando evitar valores excessivamente grandes nos coeficientes, adicionando um termo de penalização com base nas magnitudes dos coeficientes. É utilizado

para controlar a complexidade do modelo e evitar o sobre ajuste (*overfitting*). A penalidade é aplicada sobre a soma dos valores quadráticos dos coeficientes (norma L2) conforme equação

$$Y_p = \sum_{i=1}^{n_d} (\beta_i X_{d_i}) + \alpha \sum_{i=1}^{n_d} (\beta_i^2). \quad (2.20)$$

sendo α o fator de regularização.

- Regressão Lasso (*Least Absolute Shrinkage and Selection Operator*)

Semelhante à regressão Ridge, mas a penalidade é aplicada sobre a soma dos valores absolutos dos coeficientes (norma L1) conforme equação

$$Y_p = \sum_{i=1}^{n_d} (\beta_i X_{d_i}) + \alpha \sum_{i=1}^{n_d} (\|\beta_i\|). \quad (2.21)$$

Informações mais detalhadas e outros modelos de regressão como a polinomial e Elastic Net, podem ser encontrados em ([Gareth et al., 2013](#)).

2.12 Redes Neurais MLP

Uma rede neural de múltiplas camadas (MLP) consiste em camadas de neurônios interconectados, estruturados de maneira a possibilitar a obtenção de padrões complexos a partir dos dados. Essa arquitetura forma um elemento fundamental do aprendizado profundo, frequentemente aplicado a tarefas de classificação e regressão ([Bishop, 2006](#)). O número de neurônios de entrada corresponde às variáveis observáveis no conjunto de dados. Diferentes arquiteturas com diferentes números de camadas ocultas podem ser empregadas. Cada camada oculta acomoda neurônios que processam os dados de entrada da camada e transmitem o resultado do processamento para a camada subsequente. Essas camadas ocultas capacitam a rede a aprender representações hierárquicas dos dados de entrada ([Goodfellow et al., 2016](#)). Funções de ativação, como RELU - *Rectified Linear Unit* - ou ELU - *Exponential Linear Unit* - podem ser empregadas em cada camada oculta para introduzir não linearidade e aprimorar as capacidades de aprendizado da rede. *Frameworks* como *TensorFlow* e *PyTorch* simplificam o processo de

construção, treinamento e implantação de redes neurais. A Figura 2.17 ilustra uma estrutura típica de uma rede neural do tipo perceptron de múltiplas camadas. Essa arquitetura apresenta duas camadas ocultas. No campo das redes de aprendizado profundo encontram-se arquiteturas com dezenas, centenas ou até milhares de camadas ocultas.

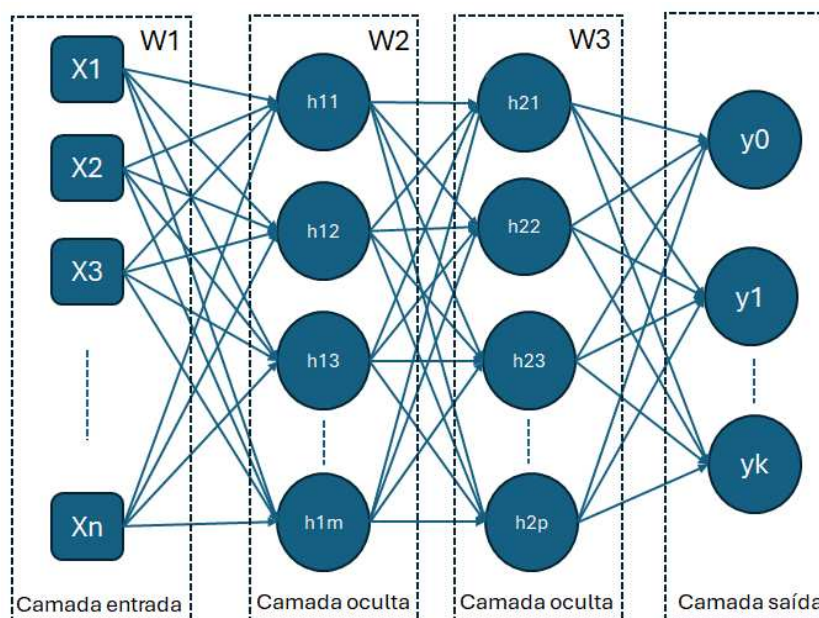


Figura 2.17: Arquitetura de uma rede do tipo multilayer perceptron (MLP) com duas camadas ocultas.

Há diversos tipos de redes neurais aplicáveis a diferentes áreas como reconhecimento de padrões, classificação, regressão, processamento de imagens, detecção de objetos, visão computacional, processamento de linguagem natural (*Natural Language Processing - NLP*), séries temporais, tradução automática, dentre inúmeras. Alguns tipos clássicos são apresentados a seguir:

- Redes Neurais Feedforward (*Feedforward Network - FNN*)

A estrutura mais simples de rede neural. Os dados fluem em uma única direção, da camada de entrada para a camada de saída, passando por uma ou mais camadas ocultas.

- Redes Neurais Convolucionais (*Convolutional Neural Network - CNN*)

Utilizadas principalmente em processamento de imagens e vídeo. Usam camadas convolucionais que aplicam filtros aos dados de entrada, destacando características importantes.

- Redes Neurais Recorrentes (*Recurrent Neural Network - RNN*)

Projetadas para processar dados sequenciais. Elas possuem conexões que formam ciclos, permitindo que informações anteriores influenciem a saída atual.

- *Long Short-Term Memory (LSTM)*:

Essa estrutura de RNN pode aprender dependências de longo prazo, resolvendo o problema de evanescência do gradiente.

- Redes Neurais Generativas Adversárias (*Generative Adversarial Network - GAN*)

Consistem em duas redes neurais, uma geradora e outra discriminadora, que competem entre si. A rede geradora cria dados falsos enquanto a discriminadora tenta distinguir entre dados reais e falsos.

- Redes Neurais de Base Radial (*Radial Basis Function - RBF*):

Usam funções de base radial como funções de ativação. São utilizadas em problemas de classificação onde a decisão de classe é baseada em um centro radial.

Um estudo detalhado sobre redes neurais pode ser encontrado em ([Goodfellow et al., 2016](#)) e em ([Bishop, 2006](#)).

2.13 Base de dados de imagens ICVL

O ICLV – *Interdisciplinary Computing Vision Laboratory, Ben-Gurion University, Israel* – disponibiliza uma base de dados de imagens hiperespectrais (ICVL, 2023), obtidas com a câmara Specim OS Kappa DX4, contendo 201 imagens no formato 1392 linhas por 1300 colunas, e 519 bandas, na faixa de 400 nm a 1000 nm, sendo a resolução espectral de 1,25 nm. Estas imagens estão no formato RAW e informações para decodificação das imagens podem ser encontradas em arquivos no formato ENVI. A base também incorpora arquivos contendo imagens resultantes de subamostragem espectral com resolução de 10 nm, cobrindo somente a faixa de 400 nm a 700 nm, com 31 bandas. As imagens dessa base correspondem a diversos temas de cenas urbanas e rurais, bem como em ambiente de laboratório. A Figura 2.18 apresenta uma amostra das imagens dessa base, renderizadas com projeção no espaço de cores CIE-1964.

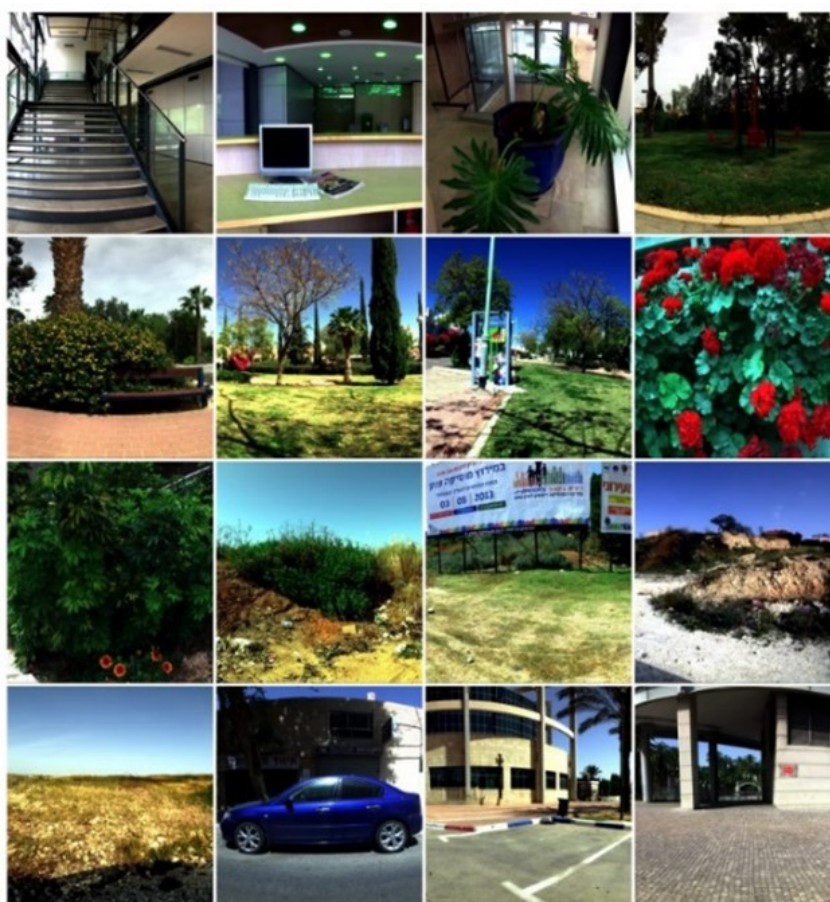


Figura 2.18: Amostra de imagens renderizadas da base ICVL. A base apresenta motivos diversos organizados em grupos como áreas urbanas, rurais, motivos florais e objetos (ICVL, 2023).

2.14 Trabalhos correlatos

Há diversas pesquisas relacionadas à recuperação do hiperespectro tendo por base imagens RGB, utilizando diferentes abordagens. Em (Arad e Ben-Shahar, 2016), os autores utilizaram aprendizado de dicionário para compor uma base com 500 átomos, assim denominados os vetores contendo assinaturas espectrais, impondo restrição de esparsidade de 28 pesos diferentes de zero, por átomo. O dicionário D_h é composto por 1000 píxeis de um conjunto de imagens hiperespectrais da base ICVL, utilizando o algoritmo K-SVD (*K-Means Singular Value Decomposition*) para definir os elementos da base de representação do dicionário. O dicionário foi projetado no espaço do sensor utilizando funções de cores do padrão CIE-1964, por meio de uma matriz $\mathbf{R}^{3 \times 31}$, sendo 3 cores (RGB), e 31 bandas espectrais. A correspondência dos valores de cada píxel da imagem RGB com relação ao vetor-base do dicionário foi definida por meio do algoritmo OMP (*Orthogonal Matching Pursuit*) (Cai e Wang, 2011). O píxel hiperespectral correspondente é definido pela projeção inversa do espaço RGB para o espaço hiperespectral, utilizando a mesma matriz de projeção. Um problema com essa abordagem é que a projeção pelo conjunto de funções CIE-1964 pode não corresponder à resposta do sensor da câmera RGB que capturou a imagem, cujo espectro se deseja estimar.

Em (Akhtar e Mian, 2018) os autores propõem um método de aprendizado de múltiplos processos gaussianos, a partir de *patches* de imagens hiperespectrais, partindo da premissa que regiões da cena contendo materiais, ou elementos similares, possuirão espectros similares. O método consiste na formação de agrupamentos (*clusters*) espaciais espectralmente similares. Os parâmetros dos processos gaussianos são utilizados para definir as correspondências entre os agrupamentos RGB e os agrupamentos hiperespectrais, transformados para o espaço RGB na fase de treinamento. Os *patches* RGB são combinados com os processos gaussianos originais para estimar o espectro da cena. Esses métodos utilizam formação de agrupamentos para construir um dicionário, ou grupo de píxeis nos espaços RGB e hiperespectral. Nessa pesquisa, e conforme será apresentado nos capítulos seguintes, também foram empregados processos de agrupamento, mas com uma abordagem foi mais direta, se comparada com (Arad e Ben-Shahar, 2016) e (Akhtar e Mian, 2018), estabelecendo um mapeamento dos centroides dos agrupamentos no espaço RGB para centroides dos agrupamentos no espaço hiperespectral de maneira simples, sem a necessidade de uma etapa de busca.

Em (Jia et al., 2017), os autores implementam uma técnica de redução de dimensionalidade e mostram que o espectro de cenas naturais reside em um *manifold* de baixa dimensionalidade. A partir deste *manifold* efetuam um processo para mapear o espaço RGB em um espaço de 3 dimensões contido no espaço hiperespectral. Desta forma, o problema de mapeamento de 3 para muitas dimensões fica restrito ao problema de mapeamento de 3 para 3 dimensões. A recuperação do hiperespectro se dá por uma técnica de reconstrução baseada em aprendizado de *manifold*.

No presente trabalho, por sua vez, foi empregada uma técnica de Análise de Fatores a um conjunto de dados de imagens hiperespectrais (Seção 4.1) que também concluiu pela dimensão intrínseca reduzida de representação dos dados no espaço hiperespectral.

Em (Arad et al., 2018) e (Arad et al., 2020) são descritos o *NTIRE - New Trends in Image Restoration and Enhancement, Challenge on Spectral Reconstruction from RGB Images*, edições 2018 e 2020, respectivamente. O desafio estabelece parâmetros e metodologias para comparação de diversos processos desenvolvidos pelos participantes. Na edição de 2018 foram definidas duas categorias “*Clean*” e “*real world*”. A categoria “*Clean*” estabeleceu a recuperação do hiperespectro a partir de imagens geradas utilizando funções de cores do padrão CIE-1964. Na categoria “*real world*” o desafio consistiu na recuperação do hiperespectro a partir de imagens no formato JPEG de 8 bits, aplicando uma curva do sensor de uma câmera desconhecida. As soluções apresentadas foram baseadas em redes neurais convolucionais e redes residuais densas. No concurso de 2020 também foram definidas duas categorias “*Clean*” e “*real world*”. A categoria “*Clean*” manteve as características do concurso de 2018, enquanto a categoria “*real world*” definiu o seguinte processamento:

- Aplicação de uma função de resposta de uma câmera Canon D40;
- Subamostragem dos canais RGB para compor uma imagem – mosaico do tipo *Bayer RGGB*;
- Simulação de ruído da câmera (*Poisson shot noise and normally distributed dark noise*);
- Aplicação de processo de demosaicagem utilizando biblioteca *OPENCV*, produzindo uma imagem RGB;
- Armazenamento da imagem no formato JPEG de 8 bits.

A metodologia de testes também incluiu a variação do brilho das imagens em 50% e 100%. Também na edição de 2020, as soluções foram baseadas em arquiteturas de redes neurais convolucionais e uso das imagens da base ICVL. As métricas e métodos de avaliação da edição NTIRE 2020 foram utilizadas como referências nesse trabalho. No presente trabalho, além da base ICVL, também foi construída e utilizada uma base de dados de imagens próprias com conteúdo de fogo e fumaça. Quatro métodos foram desenvolvidos para síntese de imagens hiperespectrais a partir de imagens RGB da base ICVL, produzindo imagens com 31 bandas espectrais. O NTIRE teve como foco a síntese de imagens hiperespectrais abrangendo a faixa de luz visível. Nessa pesquisa o escopo foi ampliado para também abarcar a faixa do infravermelho próximo até o limite de 1000 nm, disponibilizado pela câmera Specim IQ, uma das câmeras hiperespectrais utilizadas na pesquisa. Nesse caso, as imagens capturadas possuem 204 bandas espectrais.

Em (Ononye et al., 2005) é investigado o uso de assinatura remotamente detectável associada às linhas intensas de emissão de potássio (K) nos comprimentos de onda de 766,5 nm e 769,9 nm. Espectros ao nível do solo de um experimento de ignição ilustram o sinal de alto contraste proporcionado pela emissão de potássio, a partir de dados de imagem coletados por meio do sensor AVIRIS (*Airborne Visible InfraRed Imaging Spectrometer* - Espectrômetro Infravermelho Visível Aerotransportado). Esses dados foram coletados em um incêndio ocorrido em 1995 no Brasil.

No presente trabalho utilizou-se uma câmera hiperespectral e filtros para projeção das imagens dessa câmera em um espaço com quatro dimensões, por meio de uma matriz, produzindo um tensor com quatro bandas espectrais. Essas imagens foram utilizadas para treinamento e testes de um processo de síntese espectral no qual uma linha de emissão do átomo de potássio pôde ser identificada, além da detecção direta pelas câmeras hiperespectrais Specim FX-10 e Specim IQ usadas nos experimentos.

2.15 Comentários finais

Este capítulo estabeleceu os fundamentos teóricos essenciais para o entendimento da metodologia e análise de dados aplicadas nessa pesquisa, conforme detalhado no Capítulo 3. No

âmbito desses fundamentos, foram abordados tópicos como detecção de imagens, tipos de imageadores hiperespectrais, procedimentos relacionados à calibração geométrica e radiométrica, componentes de ruído em sistema de imageamento, métodos de regressão linear e tipos de redes neurais, bem como a espectroscopia do fogo. Em particular, este capítulo apresentou o embasamento teórico que respalda a identificação, pelas câmeras hiperespectrais utilizadas na pesquisa, do pico de emissão resultante do decaimento de elétrons em átomos de potássio, conforme detalhado no Capítulo 4. Este capítulo também trouxe uma análise crítica de trabalhos correlatos no tema síntese hiperespectral a partir de dados de imagens RGB, situando o presente trabalho no estado da arte no tema.

Capítulo 3

Metodologia

A metodologia utilizada no desenvolvimento dessa pesquisa compreende um conjunto de etapas que se inicia com a identificação da dimensionalidade intrínseca dos dados de imagens hiperespectrais, a fim de determinar, parcialmente, a viabilidade de se resolver o problema malposto referente à síntese de dados em um espaço de dimensionalidade superior, com espectro contínuo, a partir de dados contidos em um espaço de dimensões reduzidas e espectralmente esparsos. A existência de uma estrutura subjacente aos dados (*manifold*) e a dimensão intrínseca do espaço de representação foram inferidos a partir da identificação de fatores latentes que explicam a variância dos dados. Para isso, foi utilizada uma base de dados de imagens hiperespectrais disponibilizada pelo ICVL (*Interdisciplinary Computing Vision Laboratory*) ([ICVL, 2023](#)).

Seguiu-se uma etapa de avaliação da viabilidade de síntese de imagens hiperespectrais a partir de imagens RGB, por meio de algoritmos básicos como regressão, agrupamento e rede neural. A etapa seguinte envolveu a construção de uma base própria de dados de imagens e o estudo de assinaturas espectrais de elementos, em cenas diversas, com uso de uma câmera hiperespectral Specim FX-10.

Uma nova etapa iniciou-se com a composição de uma base de dados própria de imagens hiperespectrais e RGB, utilizando uma câmera Specim IQ, com captura simultânea da mesma cena também por uma câmera RGB Panasonic DMC-FZ5. Com o uso dessa base, foram desenvolvidos experimentos e processos de síntese de imagens hiperespectrais com base em imagens

RGB da câmera DMC-FZ5, assim como do sensor RGB que equipa a câmera Specim IQ. Foram avaliados processos que envolvem a projeção de imagens hiperespectrais em espaços de cores, com uso de uma matriz de funções representativas da resposta do sensor de imagem e de filtros RGB, com três componentes de cor, assim como a projeção em um espaço com quatro dimensões por meio de uma matriz que incorporou uma banda NIR, em adição às bandas RGB.

A síntese de imagens hiperespectrais com uso da câmera Panasonic DMC-FZ5 correspondeu a uma nova etapa que requereu procedimentos de calibração geométrica das câmeras e o recorte das imagens para haver correspondência nos píxeis, necessária ao treinamento dos algoritmos.

A avaliação dos resultados de cada etapa valeu-se do uso de métricas tipicamente utilizadas na verificação do desempenho de algoritmos de inteligência artificial e específicos para tratamento de imagens, especificamente as métricas RMSE, MARE e R^2 score, conforme apresentado na Seção 3.2.

A Figura 3.1 mostra um infográfico referente à metodologia de pesquisa utilizada nesse trabalho, estruturada em etapas. Ao lado de cada etapa são listados os resultados produzidos.

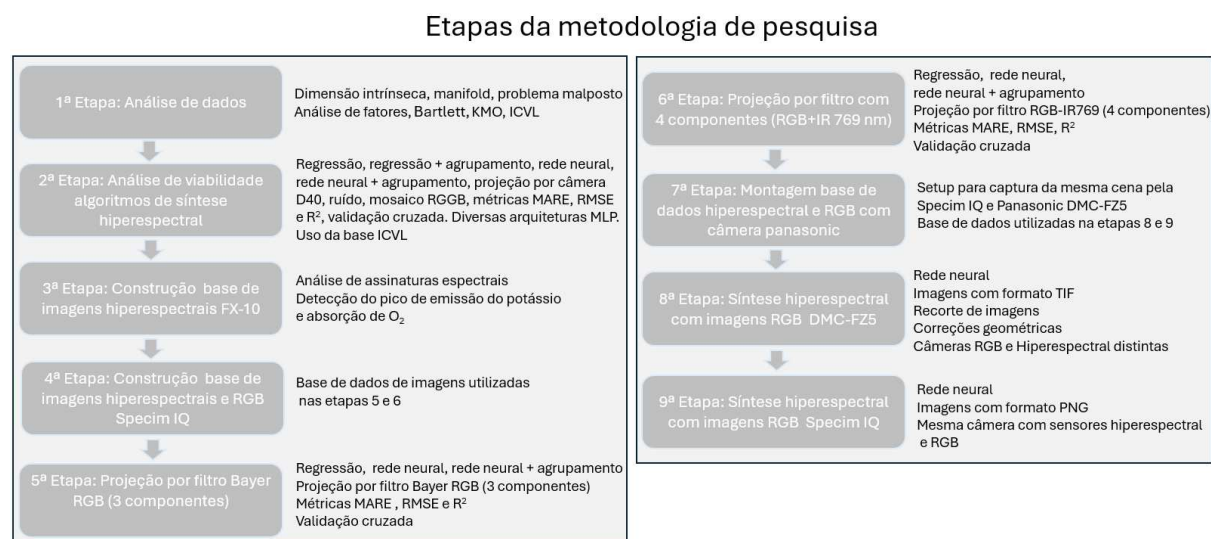


Figura 3.1: Infográfico com as etapas da metodologia de pesquisa.

3.1 Análise de fatores

Dados hiperespectrais referem-se a informações capturadas em numerosas bandas espectrais contíguas e estreitas, e podem abranger dezenas ou até centenas de bandas espectrais, definindo

sua complexidade dimensional ([Signoroni et al., 2019](#)). Por outro lado, imagens RGB consistem em apenas três bandas. À medida que o número de bandas espectrais aumenta, a dimensão dos dados também aumenta.

Lidar com dados de alta dimensionalidade, frequentemente, envolve técnicas como extração de características, ou redução dimensional. Esses métodos visam reter as informações mais importantes enquanto reduzem o número de dimensões, tornando a análise e visualização viáveis.

Conforme mencionado anteriormente, a tarefa de recuperar informações hiperespectrais a partir de dados RGB é considerada malposta, envolvendo o desafio de mapear de poucas para muitas dimensões. No entanto, no caso de objetos capturados sob condições de iluminação natural, os dados subjacentes, muitas vezes, estão contidos em um espaço de dimensão intrínseca reduzida, tornando esse problema viável ([Jia et al., 2017](#)).

A análise de fatores é uma técnica estatística que busca identificar as variáveis latentes que explicam os dados observados. O modelo representa as variáveis observadas como uma combinação linear de um conjunto de fatores latentes não observados ([Navlani, 2023](#)). O principal objetivo da análise de fatores é identificar o menor número de fatores latentes (ou dimensões) que explicam as correlações ou covariações observadas no conjunto de dados. A Análise de Fatores (FA - *Factor Analysis*) e Análise em Componentes Principais (PCA - *Principal Component Analysis*) são técnicas estatísticas relacionadas, mas distintas, sendo a análise de fatores mais geral. A PCA visa projetar variáveis do espaço de representação original em um novo espaço de representação composto por vetores ortogonais chamados componentes principais. Essas componentes principais são combinações lineares das variáveis originais e são escolhidas para capturar a máxima variância nos dados ([Goodfellow et al., 2016](#)). Na análise de fatores, o foco está em identificar os fatores comuns que explicam as correlações entre variáveis observadas, e assume que existem fatores não observáveis (latentes) que influenciam as variáveis observadas.

Para efetuar a análise de fatores de um conjunto de dados, preliminarmente, verifica-se sua aplicabilidade por meio do teste de esfericidade de *Bartlett* e pelo critério *Kaiser-Meyer-Olkin* ([Navlani, 2023](#)).

3.1.1 Medida da esfericidade pelo método *Bartlett*

O Teste da esfericidade de Bartlett é utilizado para verificar se as variáveis em um conjunto de dados estão correlacionadas, comparando a matriz de correlação dos dados observados com a matriz identidade que representa variáveis descorrelacionadas. A fórmula de Bartlett (Navlani, 2023) é utilizada para calcular a esfericidade S^2 e corresponde à equação

$$S^2 = -(n_s - 1 - \frac{2p + 5}{6}) \ln(|\det R_c|), \quad (3.1)$$

onde n_s é o número de amostras, p é o número de variáveis, e R_c é a matriz de correlação dos dados.

3.1.2 Critério *Kaiser-Meyer-Olkin* (KMO)

Essa medida avalia se a análise de fatores é aplicável. É uma medida estatística que representa o grau no qual cada variável observada é predita, sem erros, por outras variáveis do conjunto de dados. A formulação para o critério KMO é dado por

$$KMO = \frac{\sum \sum_{s \neq t} r_{st}^2}{\sum \sum_{s \neq t} r_{st}^2 + \sum \sum_{s \neq t} \tau_{st}^2}, \quad (3.2)$$

onde r_{st} corresponde ao elemento fora da diagonal principal da matriz de correlação R_c , e τ_{st} corresponde à correlação parcial das variáveis (Navlani, 2023).

3.1.3 Análise de Fatores

A Equação 3.3 descreve um modelo em que cada vetor de dados observado $\mathbf{x}_i$ é uma combinação linear de fatores latentes $\mathbf{f}_j$, somados a um termo de ruído gaussiano com média zero ϵ_i e a um componente de *offset* μ_i . A matriz de pesos $\mathbf{Q}$ é usada para gerar os dados observados, a partir das variáveis latentes subjacentes:

$$\mathbf{x}_i = \mathbf{Q}\mathbf{f}_j + \epsilon_i + \mu_i. \quad (3.3)$$

Considerando $\mathbf{x}_i$ como os vetores coluna da matriz $\mathbf{X}$ que contém os dados observados, $\mathbf{F}$ como uma matriz com fatores $\mathbf{f}_j$, $\mathbf{E}$ como a matriz com termos de ruído, e $\mathbf{M}$ como uma matriz com termos de *offset*, a decomposição de $\mathbf{X}$ pode ser definida como (Johnson e Dean, 2007):

$$\mathbf{X} = \mathbf{QF} + \mathbf{E} + \mathbf{M}. \quad (3.4)$$

Considerando que $\mathbf{f}_j$ e ϵ_i são variáveis aleatórias e assumindo que

$$\mathcal{E}(\mathbf{F}) = 0, \quad (3.5)$$

$$COV(\mathbf{F}) = \mathcal{E}(\mathbf{FF}^T) = \mathbf{I}, \quad (3.6)$$

$$\mathcal{E}(\mathbf{FE}^T) = 0, \quad (3.7)$$

$$\mathcal{E}(\mathbf{E}) = 0, \quad (3.8)$$

$$COV(\mathbf{E}) = \mathcal{E}(\mathbf{EE}^T) = \psi, \quad (3.9)$$

sendo ψ uma matriz diagonal com fatores específicos relacionados à componente de ruído, a matriz de covariância pode-se ser assim calculada:

$$(\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^T = (\mathbf{QF} + \mathbf{E})(\mathbf{QF} + \mathbf{E})^T = (\mathbf{QF} + \mathbf{E})(\mathbf{QF})^T + \mathbf{E}^T \quad (3.10)$$

$$= (\mathbf{QF}(\mathbf{QF})^T + \mathbf{E}(\mathbf{QF})^T + \mathbf{QFE}^T + \mathbf{EE}^T) \quad (3.11)$$

$$\Sigma = COV(\mathbf{X}) = \mathcal{E}((\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^T) = \mathbf{Q}\mathcal{E}(\mathbf{FF}^T)\mathbf{Q}^T + \mathcal{E}(\mathbf{EF}^T)\mathbf{Q}^T + \mathbf{Q}\mathcal{E}(\mathbf{FE}^T) + \mathcal{E}(\mathbf{EE}^T) \quad (3.12)$$

$$= \mathbf{QQ}^T + \psi. \quad (3.13)$$

A estimativa da matriz $\mathbf{Q}$ e da matriz $\mathbf{F}$ pode ser feita usando fatoração por SVD (*Single Value Decomposition*) (Goodfellow et al., 2016). Tem-se

$$(\mathbf{X} - \mathbf{M}) = \mathbf{U}\Sigma\mathbf{V}^T, \quad (3.14)$$

sendo,

$$\mathbf{U} = (\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^T, \mathbf{V} = (\mathbf{X} - \mathbf{M})^T(\mathbf{X} - \mathbf{M}). \quad (3.15)$$

$\mathbf{U}$ é uma matriz ($n_s \times n_s$) cujas colunas são os autovetores esquerdos de $(\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^T$, Σ é uma matriz diagonal ($n_s \times p$) com valores singulares de $(\mathbf{X} - \mathbf{M})$. $\mathbf{V}^T$ é uma matriz ($p \times p$) cujas linhas são os autovetores direitos de $(\mathbf{X} - \mathbf{M})^T(\mathbf{X} - \mathbf{M})$.

Os dados observáveis são compostos por uma matriz $\mathbf{F}$ de fatores comuns, ponderados pela matriz $\mathbf{Q}$, e por uma matriz $\mathbf{E}$ de fatores específicos correspondentes ao ruído. A matriz $\mathbf{F}$ pode ser obtida por matrizes $\mathbf{U}$, Σ , e $\mathbf{V}$ incompletas contendo elementos referentes aos n_f fatores com os maiores valores singulares (Johnson e Dean, 2007)

$$\mathbf{F} = \mathbf{U}_{n_f} \Sigma_{n_f}^{\frac{1}{2}}. \quad (3.16)$$

A matriz $\mathbf{Q}$ pode ser calculada como

$$\mathbf{Q} = \mathbf{V}_{n_f}^T \Sigma_{n_f}^{\frac{1}{2}}. \quad (3.17)$$

3.1.4 Variância explicada acumulada

A variância explicada acumulada corresponde à variabilidade nos dados que pode ser explicada por um certo número de variáveis latentes. A variância explicada acumulada até o fator i é dada pela soma dos valores singulares dos fatores até i , normalizado pela soma de todos os valores singulares. A variância explicada acumulada pode ser calculada por

$$V_{a_i} = \frac{\sum_{j=1}^i S v_j}{\sum_{j=1} S v_j}. \quad (3.18)$$

3.2 Métricas para avaliação dos processos de síntese hiperespectral

3.2.1 Erro Absoluto Médio Relativo (MARE)

O valor médio do erro absoluto relativo (MARE - *Mean Absolute Relative Error*) pode ser empregado para avaliar o desempenho de métodos de síntese hiperespectral. Trata-se de uma métrica estatística comumente utilizada para avaliar a exatidão e o desempenho de um modelo preditivo, especialmente no campo da análise de dados. O MARE estima a magnitude média dos valores absolutos relativos dos erros entre os valores previstos e os valores observados (Botchkarev, 2019), conforme

$$\text{MARE} = \frac{1}{n_p n_h} \sum_l \sum_v \left| \frac{P_{g_{lv}} - P_{s_{lv}}}{P_{g_{lv}}} \right|, \quad (3.19)$$

onde $P_{g_{lv}}$ e $P_{s_{lv}}$ representam o valor de intensidade da banda v de um píxel l na imagem de referência (*ground truth*) e na imagem hiperespectral sintetizada, respectivamente; n_p refere-se ao número de píxeis da imagem, e n_h denota o número de bandas espectrais.

O MARE é calculado como a média dos valores absolutos relativos dos erros. O MARE fornece uma medida da discrepância relativa média entre valores previstos e valores reais, levando em consideração a escala dos valores reais. É particularmente útil ao comparar o desempenho de modelos preditivos em conjuntos de dados, ou variáveis, com magnitudes distintas, pois padroniza a medição do erro. Valores menores de MARE indicam maior exatidão do modelo, refletindo uma correspondência mais próxima entre previsões e valores reais. No entanto, como qualquer métrica, o MARE deve ser interpretado no contexto do problema específico e do domínio para avaliar completamente o desempenho do modelo. O MARE captura as diferenças locais de píxel para cada banda espectral, possibilitando comparações em diversas regiões de uma cena imageada. Além disso, essa métrica pode servir como função de custo durante a fase de treinamento, auxiliando na determinação do melhor modelo para o processo.

3.2.2 Erro quadrático médio (RMSE)

Outra métrica utilizada na avaliação dos modelos é o erro quadrático médio (RMSE - *Root Mean Squared Error*), que quantifica estatisticamente a magnitude média dos erros quadráticos entre valores previstos, utilizando um modelo treinado, e os valores observados. Ele é especialmente sensível a erros de maior magnitude devido ao cálculo do quadrado do erro. (Botchkarev, 2019).

$$\text{RMSE} = \sqrt{\frac{1}{n_p n_h} \sum_l \sum_v (Pg_{lv} - Ps_{lv})^2}. \quad (3.20)$$

Um valor menor de RMSE indica maior exatidão do modelo, pois denota uma diferença média menor entre previsões e valores reais. O RMSE é uma métrica útil para avaliar o desempenho geral de um modelo e comparar o desempenho preditivo de diferentes modelos. No contexto da síntese de imagens hiperespectrais, o RMSE pode ser usado para avaliar o desempenho geral do método, levando em consideração toda a imagem.

3.2.3 Coeficiente de Determinação (R² score)

Para avaliar modelos lineares, como a regressão *Ridge*, uma métrica aplicável é o Coeficiente de determinação (R² score). Ele mede a proporção da variância na variável dependente que pode ser prevista a partir das variáveis independentes em uma tarefa de regressão. O coeficiente de determinação indica quão efetivamente as previsões do modelo de regressão se alinham com os valores observados reais da variável dependente. Ele varia de 0 a 1, onde um valor de R² igual a zero indica que o modelo falha em explicar qualquer variabilidade na variável dependente. Um valor de R² igual a 1 significa que o modelo explica perfeitamente a variabilidade na variável dependente. Um valor de R² entre 0 e 1 fornece informações sobre a qualidade do ajuste alcançado pelo modelo de regressão. Valores mais altos de R² indicam que uma maior proporção da variabilidade na variável dependente é explicável pelas variáveis independentes empregadas no modelo (Gareth et al., 2013). O coeficiente de determinação é calculado pelas equações

$$R^2 = 1 - \frac{\sum_l \sum_v (Pg_{lv} - Ps_{lv})^2}{\sum_l \sum_v (Pg_{lv} - \mu_{lv}^2)}, \quad (3.21)$$

$$\mu_{lv} = \frac{1}{n_p n_h} \sum_l \sum_v Pg_{lv}. \quad (3.22)$$

3.3 Validação cruzada

A validação cruzada é uma técnica utilizada em aprendizado de máquina para avaliar o desempenho de modelos de maneira robusta e confiável, propiciando uma estimativa mais precisa de como um modelo se comportará ao lidar com novos dados não utilizados durante o treinamento. A abordagem clássica de validação cruzada envolve a divisão do conjunto de dados em partições (*folds*) distintas. O modelo é treinado com várias partições do conjunto de dados. Um dos métodos mais comuns de validação cruzada é o *k-fold*. Nesse método, o conjunto de dados é dividido em k_f partições, e o modelo é treinado k_f vezes, cada vez usando $k_f - 1$ partições para treinamento, e uma partição para validação. A validação cruzada é utilizada para avaliar o *overfitting*, no qual um modelo se ajusta excessivamente aos dados de treinamento, mas tem um desempenho inferior com novos dados de teste. Ao avaliar o modelo em diferentes subconjuntos de dados, a validação cruzada fornece uma estimativa da sua capacidade de generalização (Gareth et al., 2013). A Figura 3.2 ilustra a divisão dos dados em cinco partições.

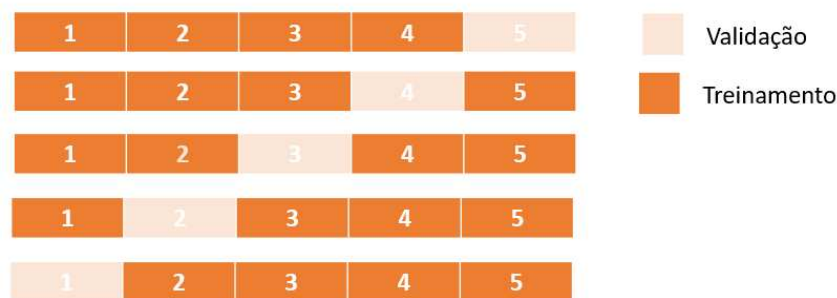


Figura 3.2: Método de validação cruzada com $k_f = 5$ (*5-fold cross-validation*).

Nesse trabalho os dados de imagens foram divididos em conjuntos de treinamento, validação e testes. Nas etapas iniciais do desenvolvimento, utilizou-se 10 partições ($k_f = 10$). A dispersão de valores entre experimentos nas diferentes partições mostrou-se reduzida. Dessa forma, optou-se pela validação cruzada com 5 partições nas etapas posteriores, reduzindo o número de experimentos ($k_f = 5$).

A avaliação final do desempenho dos algoritmos e processos de síntese foi feita com um conjunto de dados de testes não vistos pelos algoritmos. Os conjuntos de imagens das bases foram divididos na proporção de 80% para treinamento e validação, e 20% para testes. Em alguns experimentos específicos, a proporção dos conjuntos foi um pouco diferente em função da quantidade de imagens disponível.

3.4 Processos para síntese de imagens hiperespectrais

Na metodologia empregada neste trabalho, foram definidos quatro processos para a síntese de imagens hiperespectrais com base em dados espectralmente esparsos presentes em imagens RGB. Esses processos resultam da combinação de três técnicas fundamentais: regressão linear, agrupamento e redes neurais. Os quatro processos avaliados são: regressão linear pelo método Ridge, regressão linear Ridge combinada com agrupamento, redes neurais do tipo perceptron multicamadas e redes neurais combinadas com agrupamento. Esses processos foram selecionados para abranger diferentes níveis de complexidade.

O método de regressão Ridge é o mais simples de implementar e possui o menor custo computacional. Espera-se que ele ofereça o menor tempo de processamento, mas também uma menor acurácia em comparação aos outros métodos avaliados. A técnica de agrupamento será avaliada em combinação com o método de regressão e com redes neurais, visando a redução do custo computacional, ao considerar que a síntese será realizada com base apenas nos centróides dos agrupamentos. Para o método baseado em redes neurais, espera-se alcançar a maior acurácia, embora com um custo computacional mais elevado.

Diversos modelos de regressão serão avaliados, considerando diferentes valores para o fator de regularização. A técnica de agrupamento utilizará o algoritmo *Minibatch K-Means*, devido ao seu menor custo computacional, avaliada para diferentes tamanhos de agrupamentos, o que implica em uma diversidade na distribuição espectral.

Diferentes arquiteturas de redes neurais do tipo *Multilayer Perceptron* serão consideradas na etapa de implementação, com variações no número de camadas e neurônios em cada camada, no tipo de função de ativação, no uso ou não de camadas de regularização, na taxa de aprendizado e no tipo de otimizador. Também serão avaliadas as influências da função de custo, do tamanho do lote e do número de épocas na acurácia alcançada na fase de treinamento

3.5 Método para comparação do desempenho dos processos de síntese de imagens hiperespectrais

A avaliação do desempenho dos processos de síntese desenvolvidos neste trabalho se dará, de forma qualitativa, a partir de aspectos como nitidez, contraste e definição de cores das imagens sintetizadas em comparação com as correspondentes imagens hiperespectrais de referência (*ground truth*) do conjunto de testes. A avaliação do desempenho também será feita de forma quantitativa utilizando as métricas descritas na seção 3.2. Os valores médios, máximos e mínimos das métricas MARE e RMSE são determinados para cada processo de síntese hiperespectral, considerando as imagens do conjunto de teste.

Um teste de comparação de médias possibilitará averiguar se há uma distinção significativa da distribuição de valores MARE e RMSE entre os processos de síntese. Em uma etapa preliminar, será utilizado o Teste de *Shapiro-Wilk* ([Shapiro e Wilk, 1965](#)) para determinar se os valores dessas métricas aplicadas ao conjunto de imagens de teste seguem, ou não, uma distribuição normal e, assim, definir qual método de comparação de médias, ou medianas, deve ser utilizado. Se a distribuição for normal, o Teste de ANOVA ([Ott e Longnecker, 2015](#)) é aplicável, caso contrário, utilizar-se-á o Teste Kruskal-Wallis ([Kruskal e Wallis, 1952](#)). O valor do parâmetro *pvalue* obtido no Teste de Shapiro-Wilk será utilizado para determinar se a hipótese nula que estabelece que os dados seguem uma distribuição normal pode ser rejeitada, ou não.

Hipóteses do Teste de *Shapiro-Wilk*

- Hipótese nula (H_0): A amostra vem de uma distribuição normal.
- Hipótese alternativa (H_1): A amostra não vem de uma distribuição normal.

No Teste de Kruskal-Wallis e no Teste de ANOVA, o valor do parâmetro *pvalue* indica a probabilidade de que as diferenças observadas entre os grupos tenham ocorrido por acaso, sob a hipótese nula de que todas as amostras provêm da mesma distribuição.

- Hipótese nula (H_0): Todas as amostras provêm da mesma distribuição, ou seja, não há diferença significativa entre as medianas dos grupos.
- Hipótese alternativa (H_1): Pelo menos uma das amostras provém de uma distribuição

diferente, ou seja, há diferença significativa entre as medianas dos grupos.

3.6 Preprocessamento das imagens da base de dados ICVL

A análise de viabilidade e desempenho de algoritmos de recuperação do hiperespectro a partir de dados contidos em imagens RGB será feita, preliminarmente, utilizando a base de dados de imagens disponibilizada no ICVL (ICVL, 2023), seguindo o processo proposto pelo NTIRE 2018/2020 (Arad et al., 2018, 2020). O processo consiste em projetar as imagens hiperespectrais no espaço do sensor hipotético de uma câmera RGB, conforme a função de transferência do sistema óptico, características responsivas do sensor, arranjo de filtros de cor e adição de ruído com distribuição normal, para emular o ruído balístico de escuro, assim como ruído com distribuição de *Poisson* para emular o ruído fotônico. A resposta do sistema de imagem de uma câmera Canon D40 será utilizada para projeção das imagens hiperespectrais da base.

A imagem hiperespectral é tida como referência (*ground truth*) do espectro de luz da cena. Uma representação funcional de um arranjo de filtros *Bayer* RGGB é utilizada para emular o processamento na câmera D40, com uma proporção de 50% de píxeis na cor verde, 25% de píxeis na cor vermelha e 25% de píxeis na cor azul. Um total de 183 imagens dessa base, todas no formato (1392 x 1300 x 31), foram convertidas para o formato (1392 x 1300 x 3), por meio da projeção no espaço do sensor. Tanto as imagens hiperespectrais quanto as correspondentes imagens RGB foram normalizadas para conter píxeis com valores entre 0 e 1.

As imagens do conjunto de treinamento, validação e teste são submetidas a uma etapa de processamento como mostrado na Figura 3.3. A etapa de processamento das imagens hiperespectrais consiste da normalização, na qual se divide o tensor contendo a imagem por 4095, valor esse que corresponde à amplitude máxima em 12 bits (formato das imagens da base ICVL). Uma abordagem alternativa seria subtrair cada imagem de seu valor médio e dividir pela variância. Essas imagens são a base (*ground truth*) para treinamento dos algoritmos descritos nas Seções 4.2.1, 4.2.2, 4.2.3 e 4.2.4. As imagens hiperespectrais são projetadas no espaço do sensor RGB por meio de uma matriz $\mathbf{F}_c$ que representa a função de transferência do sistema óptico, dos filtros *Bayer* e do sensor de imagem da câmera Canon D40. $\mathbf{F}_c$ depende do comprimento de onda λ (radiação eletromagnética). A matriz de projeção contendo a resposta da

câmera D40 e a função para emulação do arranjo de filtros podem ser encontradas no *link* indicado no Apêndice C. Para emular a corrente de escuro e o ruído fotônico adicionou-se uma componente de ruído, com distribuição gaussiana $\mathbf{N}$, e outra componente $\mathbb{J}$, com distribuição de Poisson, conforme

$$\mathbf{I}_{rgb} = \mathbf{F}_c(\lambda)\mathbf{I}_h + \mathbf{N} + \mathbb{J}. \quad (3.23)$$



Figura 3.3: Etapa de pré-processamento no qual as imagens hiperespectrais são projetadas no espaço do sensor (RGB), sendo adicionado ruído devido à corrente de escuro e ruído fotônico.

3.7 Câmera hiperespectral Specim FX-10

A câmera hiperespectral FX-10 é do tipo *line-scan* com resolução espacial de 1024 píxeis por linha, resolução radiométrica de 12 bits, e resolução espectral de 1,33 nm na faixa de 400 nm a 1000 nm. A Tabela 3.1 apresenta algumas características da câmera. O número de linhas depende do ângulo e da velocidade de varredura, bem como da taxa de quadros, configuráveis por meio da interface do software controle da câmera. Algumas imagens possuem 300 linhas, outras 3900 linhas. A câmera FX-10 foi utilizada, inicialmente, para caracterizar espectros de elementos de cenas diversas, valendo-se de iluminação natural. Como visto no Capítulo 2, o espectro da radiação eletromagnética do sol que alcança a atmosfera inferior, abrange ampla faixa do espectro eletromagnético, incluindo a de luz visível (VIS) e do infravermelho próximo (NIR), sendo essa a faixa de captura, tanto da câmera Specim FX-10, quanto da câmera Specim IQ. A Figura 3.4 mostra uma imagem da câmera Specim FX-10, assim como a unidade rotor RS10, utilizada para efetuar a varredura da cena.

A Figura 3.5 apresenta um conjunto de imagens de fogo, monocromáticas, em diversos

Tabela 3.1: Características da câmera Specim FX-10

Faixa espectral	400 nm - 1000 nm
Resolução Espectral	1,33 nm
Bandas espectrais	448
Resolução espacial	1024 píxeis por linha
Resolução radiométrica	12 bits
Amostragem temporal	330 fps
Dimensões do píxel	19,9 x 9,97 μm
Resolução radiométrica	12 bits
<i>Full field of view</i> (FOV)	31 x 31 deg
<i>F-number</i>	$F/1,7$
Relação Sinal-Ruído	600:1
Dimensões	150 x 85 x 71 mm

**Figura 3.4:** Câmera Hiperespectral FX-10: 448 bandas espectrais, faixa de 400 nm a 1000 nm, alojada sobre a unidade rotor RS-10.

comprimentos de onda, tanto na faixa visível (VIS), quanto do infravermelho próximo (NIR), capturadas com uso da câmera Specim FX-10. Nas imagens das bandas 466,96 nm e 533,93 nm a chama não está visível, por corresponder à região do azul e do verde do espectro eletromagnético, respectivamente. As bandas 769,64 nm, 801,79 nm e 935,71 nm correspondem a faixas espectrais não visíveis ao olho humano. Observa-se que na faixa 935,71 nm a chama se destaca em relação ao restante da imagem. A segmentação de imagens por comprimento de onda,

utilizando imagens hiperespectrais, apresenta-se como técnica viável. A Figura 3.5 também apresenta uma imagem RGB renderizada. Essa renderização consiste do preenchimento dos canais R, G e B da imagem com os conteúdos das bandas hiperespectrais 650 nm, 550 nm e 450 nm, respectivamente. A câmera FX-10 foi utilizada para construção de uma base de dados (*dataset*) de imagens hiperespectrais cuja descrição está contida no Apêndice A.

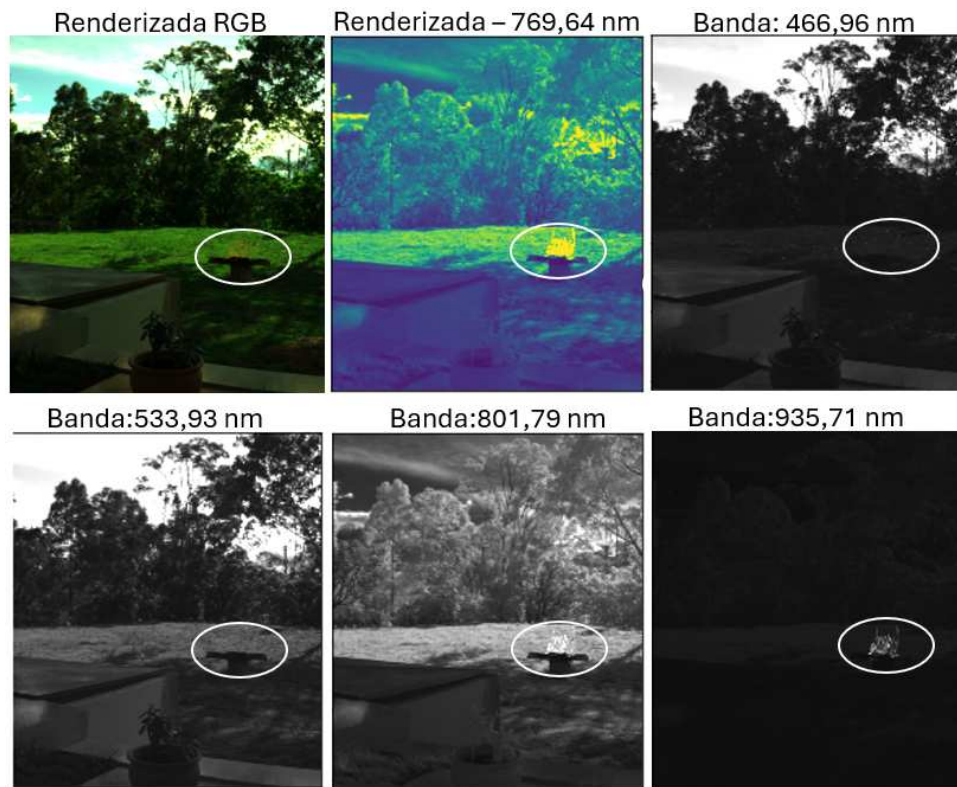


Figura 3.5: Imagens monocromáticas de fogo em diversos comprimentos de onda (VIS e NIR), capturadas com uso da câmera Specim FX-10. A figura apresenta uma imagem RGB renderizada com a combinação das bandas 650 nm, 550 nm e 450 nm.

3.8 Câmera Specim IQ

A câmera hiperespectral Specim IQ também é do tipo *line-scan (push-broom)*, com dispositivo de rotação interno. Possui resolução espacial de 512 x 512 píxeis, resolução radiométrica de 12 bits e resolução espectral de 2,94 nm. A Tabela 3.2 contém as principais características dessa câmera. A câmera é capaz de capturar sinais eletromagnéticos na faixa de 400 nm a 1000 nm, abrangendo o espectro de luz visível (VIS) e do infravermelho próximo (NIR).

A Figura 3.6 ilustra uma amostra da base de dados de imagens construídas com o uso da

câmera Specim IQ, descrita no Apêndice B.

Tabela 3.2: Características da câmera Specim IQ

Faixa espectral	400 nm - 1000 nm
Resolução Espectral	2,94 nm
Bandas espectrais	204
Resolução espacial	512 x 512 píxeis
Amostragem temporal	330 fps
Dimensões do píxel	17,58 x 17,58 μm
Resolução radiométrica	12 bits
Relação Sinal-Ruído	400:1
<i>Full field of view</i> (FOV)	31 x 31 deg
<i>F-number</i>	F/1,7
<i>View Finder Camera</i> (RGB)	5M píxeis
<i>Focus Camera</i>	1.3M píxeis
<i>QE peak</i>	> 45 %
Tipo de sensor	CMOS
Dimensões	207 x 91 x 74 mm

3.9 Câmera RGB Panasonic DMC-FZ5

Em uma nova etapa da pesquisa, foi avaliado o uso dos processos de recuperação do espectro de uma cena, a partir de imagens capturadas com uma câmera comercial de baixo custo. A câmera utilizada foi o modelo Panasonic DMC-FZ5. A Tabela 3.3 apresenta os dados dessa câmera. Trata-se de uma câmera compacta, com lente de 35mm e sensor CCD que provê uma resolução espacial 2560 x 1920.

3.10 Construção de bases de imagens hiperespectrais e RGB

No desenvolvimento desse trabalho, e como etapa da metodologia de pesquisa, foram construídas bases de dados de imagens hiperespectrais utilizando as câmeras Specim FX-10 e Specim IQ. As bases também contêm as correspondentes imagens RGB. No caso das imagens RGB relativas à base de imagens produzidas com a câmera Specim FX-10, as mesmas foram obtidas por renderização, utilizando as bandas centradas nos comprimentos de onda de 650 nm,

Tabela 3.3: Características da câmera Panasonic DMC-FZ5

Megapixels	5MP
Resolução máxima de imagem	2560 x 1920
Resolução máxima de vídeo	320 x 240
Resolução radiométrica	8 bits
Tipo de sensor de imagem	CCD
Dimensão do sensor de imagem	1/2,5"
Taxa de quadros	30 fps
Zoom digital	4 x
Zoom óptico	12 x
Comprimento focal (equivalente a lente 35mm)	36 - 432 mm
Velocidade do obturador	8 - 1/2000s
Dimensões	108 x 84,8 x 68,4 mm
Peso	326 g

550 nm e 450 nm para compor os conteúdos dos canais R, G e B da imagem. Em relação às imagens RGB correspondentes à base de imagens produzidas com a câmera Specim IQ, essas estão disponíveis de duas formas: imagens em formato PNG, produzidas pelo sensor RGB da própria câmera Specim IQ; imagens em formato TIF, produzidas por uma câmera Panasonic DMC-FZ5, com correspondência das cenas imageadas.

Para a construção da base de imagens, foram feitas incursões em áreas urbanas e rurais para captura de imagens com motivos diversos. Em especial, imagens de fogo e fumaça foram obtidas em eventos de queimas prescritas, em operações conduzidas pelo BEMAD (Batalhão de Emergências Ambientais e Resposta a Desastres) do Corpo de Bombeiros Militar de Minas Gérias, em cooperação com pessoal das brigadas de incêndio em parques florestais. Os eventos denominados Manejo Integrado do Fogo – MIF – são utilizados como ações preventivas de incêndios florestais, por meio da queima controlada de áreas designadas pela administração dos parques. Esses eventos também são utilizados para treinamento de brigadistas e de soldados do Corpo de Bombeiros. A Figura 3.6 contém imagens da ação dos bombeiros e brigadistas no Parque Estadual Serra Verde, localizado na região metropolitana de Belo Horizonte. Esse evento foi umas das várias oportunidades em que foi possível participar do MIF, e utilizar a câmera Specim IQ para capturar imagens de fogo e fumaça.



Figura 3.6: Operação de queima prescrita (MIF - Manejo Integrado do Fogo) – Parque Estadual Serra Verde. Amostra da base de dados de imagens de fogo e fumaça com uso da câmera Specim IQ. Ação executada pelo BEMAD (Batalhão de Emergências Ambientais e Resposta a Desastres), do Corpo de Bombeiros Militares de Minas Gerais.

3.10.1 Construção da base de dados de imagens hiperespectrais com uso da câmera hiperespectral *Specim FX-10*

Foi construída uma base de imagens hiperespectrais utilizando a câmera *Specim FX-10* e o software de configuração e controle *LUMO RECORDER*. A base contém imagens com diferentes formatos, tendo sido capturadas com diversas configurações da câmera, referentes ao ângulo e velocidade de varredura, ao tempo de integração, à taxa de quadros, agrupamento espectral (*spectral binning*) e agrupamento espacial (*spatial binning*). As condições de iluminação também foram diversas, incluindo iluminação natural e artificial com lâmpada fluorescente e lâmpada halógena. Os objetos imageados, sobretudo o fogo, foram capturados a diferentes distâncias. As imagens naturais foram obtidas em diferentes horários, sempre diurno. A base contém 339 imagens de elementos diversos como Leds, lâmpadas, velas, lamparina, frutas, sal de cozinha, fogo, fumaça, vapor, vegetação, dentre outros. As características e conteúdos da base estão descritos no Apêndice [A](#).

3.10.2 Construção da base de dados de imagens com montagem mecânica da Câmera hiperespectral Specim IQ e da câmera Panasonic DMC-FZ5

As câmeras Specim IQ e Panasonic DMC-FZ5 foram alojadas em uma estrutura mecânica que possibilitou fixar o alinhamento entre os eixos ópticos das câmeras, assim como as distâncias entre os seus centros de imagens. A Figura 3.7 apresenta a estrutura mecânica utilizada para fixação das câmeras e a Figura 3.8 apresenta a montagem da estrutura em um tripé. Com



Figura 3.7: Estrutura mecânica para fixação das câmeras Specim IQ e Panasonic DMC-FZ5. A posição e orientação das câmeras se manteve fixa no processo de captura de imagens para composição da base de dados. Em (d) o detalhe da colocação de um filtro de infravermelho.

essa estrutura, foram feitas incursões a campo para captura de imagens em diferentes condições de iluminação natural e com diversos motivos, incluindo imagens de fogo e fumaça, a fim de possibilitar a composição da base de dados para treinamento, validação e teste. A base é composta de 224 imagens RGB no formato TIF, capturadas com uso da câmera Panasonic DMC-FZ5, e as correspondentes imagens hiperespectrais tomadas da mesma cena com uso da câmera Specim IQ. A base também contém 224 imagens RGB no formato .png capturadas com a câmera Specim IQ. A base é composta por um conjunto de pastas para cada imagem e um conjunto de arquivos em cada pasta. Os arquivos contêm imagens .raw da cena capturada, assim como referência de branco e referência de escuro, esse obtido com o obturador da câmera fechado. Um arquivo de metadados em formato .XML contém informações sobre data, horário,



Figura 3.8: Estrutura mecânica para fixação das câmeras Specim e Panasonic montada em um tripé. Essa configuração foi utilizada nas incursões a campo.

e coordenadas geográficas referentes à localização da câmera SPECIM IQ. Um arquivo *.HDR* contém informações sobre os comprimentos de onda de cada canal da imagem, taxa de quadros e o tempo de integração configurado para captura da imagem. As imagens foram capturadas em diferentes horários, sempre diurnos. Uma vez que o campo de visão, distância focal e resolução das câmeras são diferentes, foi elaborado um procedimento de calibração das câmeras para correção de distorções geométricas, assim como o recorte das imagens para corresponderem à mesma cena imageada. O processo de calibração para distorções geométricas foi implementado seguindo o procedimento descrito na Seção 2.8. Um total de 22 imagens de um padrão de tabuleiro de xadrez foram utilizadas para estimar a matriz de parâmetros intrínsecos e o vetor de coeficientes para correção das distorções geométricas das câmeras Specim IQ e Panasonic DMC-FZ5.

A Figura 3.9 (a) mostra um conjunto de imagens de um tabuleiro de xadrez capturadas pela câmera Specim IQ. A Figura 3.9 (b) apresenta a matriz de parâmetros intrínsecos e o vetor para correção das distorções geométricas da câmera Specim IQ, estimados com o processo de calibração. A primeira linha da matriz corresponde aos valores da distância focal f_x , o valor zero

e a coordenada C_x de referência do sensor. A segunda linha da matriz contém o valor zero, a distância focal f_y , e a coordenada C_y . A terceira linha contém os valores (0,0,1), necessária para projeção do espaço 3D no plano 2D do sensor (veja Equação 2.13 na Seção 2.8). O vetor contém os coeficientes k_1, k_2, k_3, p_1 , e p_2 para correções de distorções radiais e tangenciais (veja Equações 2.9 e 2.11 na Seção 2.8). A Figura também mostra (em c e d), de forma qualitativa, o efeito da calibração para correção geométrica em uma imagem obtida com essa câmera. Observe a diferença na extensão do telhado em destaque na figura.



(a)

```
In [7]: mtxSpecim
Out[7]:
array([[965.43931963,  0., 242.53717902],
       [ 0., 963.81320607, 285.19313532],
       [ 0.,  0., 1.]])

In [9]: distSpecim
Out[9]:
array([-2.63556129e-01, -4.36976717e-01, 1.43417959e-03, -2.92914371e-03, 6.14146959e+00])
```

(b)



(c)



(d)

Figura 3.9: Em (a), conjunto de imagens de tabuleiro de xadrez para calibração da câmera Specim IQ. Em (b), matriz de parâmetros intrínsecos e o vetor com coeficientes para correção das distorções geométricas da câmera Specim IQ. Em (c), imagem capturada pela câmera Specim sem calibração. Em (d), efeito da calibração na imagem capturada.

A Figura 3.10 (a) mostra um conjunto de imagens de um tabuleiro de xadrez capturadas pela Panasonic DMC-FZ5. Em (b), a matriz de parâmetros intrínsecos e o vetor para correção das distorções geométricas da câmera, estimados com o processo de calibração. Em (c) e (d), o efeito da calibração para correção geométrica em uma imagem obtida com essa câmera. Observe o deslocamento da parede em destaque na figura.



(a)

```

In [3]: mtxPana
Out[3]:
array([[2.76674992e+03, 0.00000000e+00, 1.24910410e+03],
       [0.00000000e+00, 2.76573040e+03, 9.80764871e+02],
       [0.00000000e+00, 0.00000000e+00, 1.00000000e+00]])

In [5]: distPana
Out[5]: array([-3.500e-01, -1.000e-01, 1.790e-03, -2.079e-02, 2.440e+00])

```

(b)



(c)

(d)

Figura 3.10: Em (a), um conjunto de imagens de tabuleiro de xadrez para calibração da câmera Panasonic DMC-FZ5. Em (b), a matriz de parâmetros intrínsecos e o vetor com coeficientes para correção das distorções geométricas da câmera Panasonic DMC-FZ5. Em (c), imagem capturada pela câmera Panasonic sem calibração. Em (d), efeito da calibração na imagem capturada.

A Figura 3.11 mostra a etapa de pré-processamento das imagens tanto da câmera RGB (a) quanto da câmera Specim IQ (b). Essa etapa envolveu a calibração das imagens para correção de distorções geométricas, conforme descrito no Capítulo 2, utilizando a matriz de parâmetros intrínsecos. A câmera RGB produz imagens no formato 2048 x 1536, ao passo que a Câmera Specim IQ produz imagens no formato 512 x 512. O pré-processamento das imagens RGB produzidas pela câmera DMZ-FZ5 inclui uma etapa de ajuste de resolução da imagem para o formato 512 x 512. As imagens de ambas as câmeras foram normalizadas para valores entre 0 e 1. O campo de visão das duas câmeras é diferente, o que implica na necessidade de efetuar um recorte na imagem de cada câmera, reduzindo ao mesmo campo.

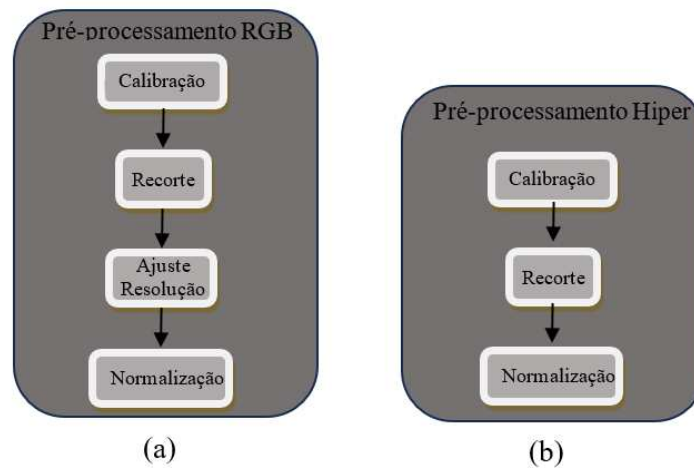


Figura 3.11: Etapa de pré-processamento das imagens das câmeras RGB (a) e Specim IQ (b).

A Figura 3.12 apresenta as imagens obtidas por meio das câmeras Panasonic e Specim, após serem processadas por meio de recorte, calibração e redução de tamanho (resolução). Após o recorte e redimensionamento, as imagens processadas foram ajustadas para o formato (263, 302, 3), para a câmera Panasonic, e formato (263, 302, 204) para a câmera Specim. Essa imagem é renderizada com a combinação das bandas 450 nm, 550 nm e 650 nm.

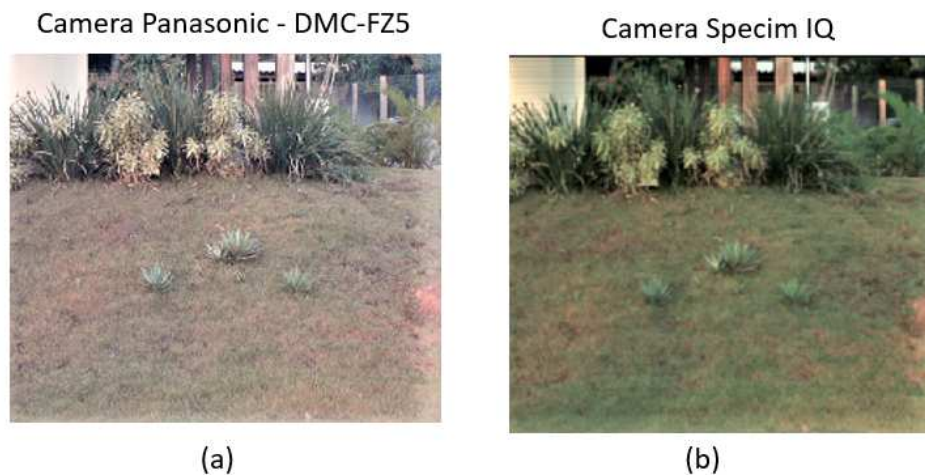


Figura 3.12: Amostra de imagem da base obtida com a montagem das câmeras Panasonic DMC-FZ5 e Specim IQ com eixos alinhados. (a) Imagem Panasonic processada com recorte e redução de resolução, formato (263, 302, 3). (b) Imagem Specim processada com recorte, formato (263, 302, 204). Essa é uma imagem renderizada.

Considerando que o objeto imageado se comporta como superfície lambertiana (Foster et al., 2006), a distribuição de potência espectral que alcança as duas câmeras é semelhante, dada a sua proximidade, não sendo necessário retificar as imagens, em cada câmera, reduzindo ao mesmo referencial. As sensibilidades e responsividade das câmeras são, por si só, diferentes,

entretanto, o algoritmo foi capaz de sintetizar imagens com distribuição de potência espectral semelhante, como será demonstrado nos resultados experimentais.

3.11 Comentários finais

Esse capítulo descreveu a metodologia empregada nessa pesquisa, abrangendo os processos de síntese hiperespectral, a base ICVL bem como as características e o uso das câmeras hiperespectrais Specim FX-10 e Specim IQ, além da câmera RGB Panasonic DMC-FZ5, utilizadas para compor as bases de dados de imagens empregadas nos treinamentos dos algoritmos, e nos processos de síntese de imagens hiperespectrais, como será detalhado no Capítulo 4. Os procedimentos de calibração geométrica foram aplicados como etapas necessárias à construção da base de dados de imagens, envolvendo o uso combinado da câmera Specim IQ e DMC-FZ5, juntamente com passos adicionais de recorte e ajuste do tamanho das imagens. A formulação da análise de fatores foi demonstrada, tendo sido apresentados os testes de esfericidade e o critério KMO. Os métodos de comparação de médias por meio dos testes de ANOVA e Kruskal-Wallis foram apresentados, assim como as métricas para avaliação dos processos de síntese. O próximo capítulo apresenta a utilização da metodologia aqui descrita, bem como das bases de imagens, juntamente com as implementações dos algoritmos, os resultados obtidos e as principais discussões da pesquisa.

Capítulo 4

Resultados e análises

E

sse capítulo apresenta os principais resultados alcançados na pesquisa e suas análises, seguindo a metodologia apresentada no Capítulo 3.

4.1 Análise de fatores em um conjunto de dados hiperespectrais

A análise de fatores foi utilizada para avaliar a existência de um *manifold* de dimensões reduzidas nos dados de imagens hiperespectrais como condição para a convergência do problema malposto relacionado à síntese de dados hiperespectrais a partir de dados de imagens RGB. Para realizar a análise de fatores, foi composto um conjunto com 120 imagens hiperespectrais da base ICVL (Seção 2.13), na forma de uma matriz $\mathbf{X}$ contendo 10^8 píxeis com 31 bandas espectrais. A análise de fatores, assim como os testes para sua aplicação, estão descritos na Seção 3.1

- Teste de esfericidade de Bartlett

O teste de esfericidade de Bartlett está descrito na Seção 3.1.1. De acordo com (Navlani, 2023), a análise de fatores pode ser aplicada a um conjunto de dados quando a esfericidade for menor do que 0,05. O teste de esfericidade de Bartlett para a matriz $\mathbf{X}$ produziu um valor para a esfericidade $< 0,01$. O valor obtido indica que análise de fatores é aplicável ao conjunto de

dados.

- Critério KMO

O critério *Kaiser-Meyer-Olkin*, descrito na Seção 3.1.2, é usado para avaliar se a análise fatorial é aplicável ou não a um conjunto de dados e diz respeito à adequação da amostragem dos mesmos. Segundo (Navlani, 2023), um valor de KMO maior do que 0,6 indica uma amostragem adequada. O valor de KMO obtido para a matriz $\mathbf{X}$ foi de 0,97.

A Figura 4.1 apresenta os autovalores para cada fator latente do conjunto de dados representados pela matriz $\mathbf{X}$. Observa-se que a partir do fator 5 os autovalores são nulos, o que corrobora a hipótese de um espaço de representação com dimensões reduzidas e a existência de uma estrutura subjacente nos dados (*manifold*). Interessante notar que, de fato, apenas as três primeiras componentes possuem autovalores significativos. Esse resultado indica potencial existência de solução para o problema malposto referente à síntese de dados em um espaço de dimensionalidade superior (hiperespectral) a partir de dados contidos em um espaço de dimensões reduzidas (RGB).

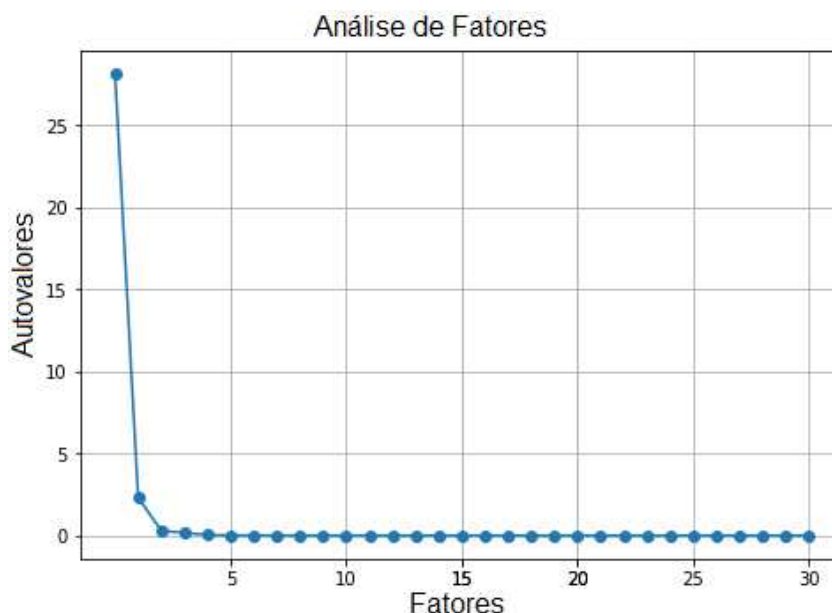


Figura 4.1: Autovalores em função dos fatores principais (variáveis latentes). O gráfico indica que apenas três ou quatro componentes são significativos.

Adicionalmente, calculou-se, para a matriz $\mathbf{X}$, a variância acumulada cujos resultados são apresentados no gráfico da Figura 4.2. Apenas três fatores explicam mais de 99% da variância dos dados, indicando a relevância dessas variáveis latentes para representação dos dados

observados. Esses resultados permitem inferir a existência de uma estrutura nos dados com di-

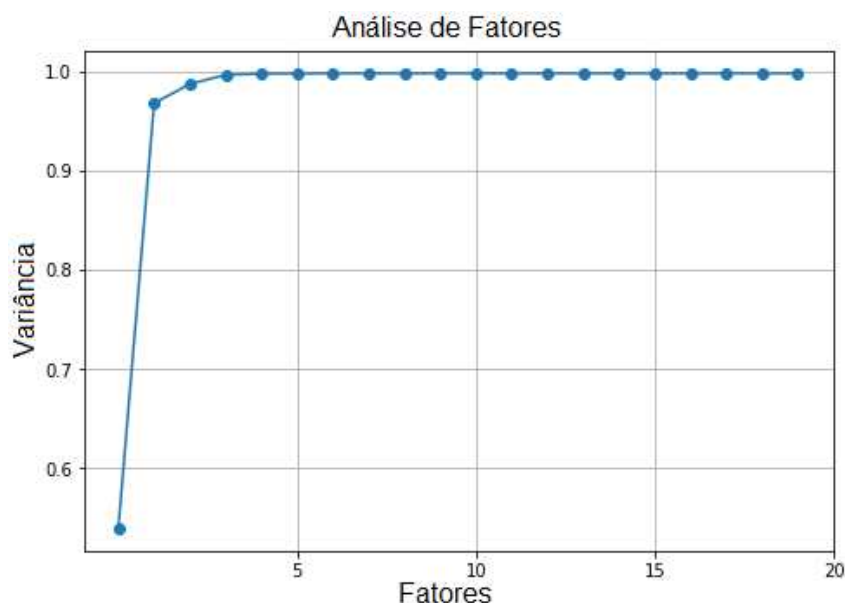


Figura 4.2: Variância acumulada em função dos fatores associados a variáveis latentes. Apenas três fatores correspondem a mais de 99% da variância.

mensões significativamente reduzidas em comparação com o espaço das variáveis observáveis.

4.2 Modelos e algoritmos para recuperação do hiperespectro a partir de imagens RGB da base de dados ICVL

4.2.1 Recuperação do hiperespectro por regressão linear

Um método simples para recuperação do conteúdo hiperespectral contido em uma imagem RGB consiste no uso de regressão linear. O modelo de regressão *Ridge*, frequentemente referido como "modelo *Ridge*", é um tipo de técnica de regressão linear que é particularmente útil para lidar com problemas de multicolinearidade e *overfitting* na modelagem preditiva (veja Seção 2.11). Ele introduz um termo de regularização como uma penalização com base nas magnitudes dos coeficientes. Ao incorporar esse termo de penalização na função de custo, a regressão *Ridge* visa encontrar um equilíbrio entre ajustar os dados o mais próximo possível (como na regressão linear comum) e manter os coeficientes do modelo sob controle. Essa regularização ajuda a mitigar o risco de o modelo ser excessivamente sensível ao ruído presente nos

dados de treinamento (Gareth et al., 2013). A intensidade da ação de regularização na regressão *Ridge* é controlada por um hiperparâmetro α (fator de regularização). Um valor mais alto de α aumenta a ação de regularização e leva a valores reduzidos dos coeficientes. À medida que α se aproxima de zero, a regressão *Ridge* se torna semelhante à regressão linear comum. Os modelos *Ridge* são comumente usados quando se lida com conjuntos de dados que têm características correlacionadas, ou quando há a possibilidade de *overfitting*. Eles são especialmente eficazes quando há muitas características no conjunto de dados, pois ajudam a estabilizar o modelo e melhorar seu desempenho de generalização em dados não vistos (Gareth et al., 2013).

Para implementação e avaliação do método, foi utilizada a base ICVL para compor um conjunto $\mathcal{H}$ contendo M_t imagens hiperespectrais com 31 bandas espectrais, projetadas no espaço de cores do sensor, conforme descrito na seção 3.6, obtendo-se um conjunto $\mathcal{R}$ de imagens correspondentes no espaço RGB, com três dimensões. Cada imagem do conjunto $\mathcal{R}$ é convertida em um arranjo matricial $\mathbf{R}_m \in \mathbf{R}^{(lin \cdot col) \times 3}$, com três canais de cor, sendo *lin* o número de linhas e *col* o número de colunas.

Cada elemento do arranjo contém três componentes, referentes aos canais de cor R, G e B. Da mesma forma, cada imagem do conjunto $\mathcal{H}$ é convertida em um arranjo matricial $\mathbf{H}_m \in \mathbf{R}^{(lin \cdot col) \times 31}$. Cada elemento do arranjo contém 31 componentes, correspondentes às 31 bandas da imagem hiperespectral. Um método de regressão por mínimos quadrados é então utilizado para estimar uma matriz $\mathbf{W} \in \mathbf{R}^{3 \times 31}$ que minimiza a função de custo L :

$$L = \frac{1}{M_t} \sum_{m=1}^{M_t} ((\mathbf{R}_m \mathbf{W} - \mathbf{H}_m)^2 + \alpha \mathbf{W}^2). \quad (4.1)$$

Um estimador do valor médio sobre as M_t imagens de cada conjunto é calculado para o quadrado do produto interno $\mathbf{R}_m \mathbf{W}$, do qual se subtrai $\mathbf{H}_m$, sendo m o índice da imagem no conjunto. A Equação 4.1 corresponde ao modelo *Ridge Regression* que utiliza um fator de regularização α aplicado à norma L2 da matriz de coeficientes $\mathbf{W}$.

Foi utilizado um conjunto de treinamento e validação composto por 120 imagens hiperespectrais contendo 31 bandas. O conjunto de testes corresponde a 63 imagens com essa mesma estrutura. Nesse caso, a proporção de imagens para treinamento/validação e testes foi de 65% e 35%, respectivamente. As imagens hiperespectrais (*ground truth*) da base possuem resolução em amplitude de 12 bits, tendo sido normalizadas para obter a intensidade de cada píxel entre

0 e 1. Utilizou-se representação em ponto flutuante de 32 bits. Além da projeção pela curva de resposta do sensor, foi adicionado ruído para emular a corrente de escuro e ruído impulsivo com distribuição de *Poisson*, conforme descrito na Seção 3.6.

A Figura 4.3 apresenta o diagrama do processo. Os blocos na cor laranja correspondem às etapas de treinamento, incluindo o estágio de préprocessamento descrito na Seção 3.6. Demais blocos na cor azul referem-se às etapas de inferência, nas quais imagens do conjunto de testes são usadas para avaliar o desempenho do processo. O modelo extraído na etapa de treinamento foi utilizado para sintetizar imagens hiperespectrais a partir de imagens RGB, sendo comparadas com as imagens de referência (*ground truth*). O desempenho do algoritmo foi avaliado com base no Coeficiente de Determinação (R^2 score) e nas métricas MARE e RMSE (veja Seção 3.2). O tempo de execução também foi levado em conta como métrica de desempenho do algoritmo, usando uma plataforma de *hardware* de referência. A função de custo foi minimizada utilizando o algoritmo gradiente estocástico descendente (SGD - *Stochastic Gradient Descent*) (Goodfellow et al., 2016), com número de iterações igual a 50, e tolerância de 10^{-4} como condição de convergência. Utilizou-se validação cruzada e o coeficiente de determinação R^2 para avaliação dos diversos modelos associados a diferentes valores para o fator de regularização. O conjunto de treinamento foi dividido em dez partições (*10-fold Cross Validation*). Nove partições foram utilizados para treinamento e uma partição foi utilizada para validação e estimativa do erro na fase de testes.

A Figura 4.4 (a) apresenta o uma comparação do desempenho, segundo o coeficiente de determinação R^2 , para valores do fator α ajustados entre 10^{-3} e 10^5 , havendo uma curva com cor distinta para cada valor do fator. As curvas apresentam silhuetas semelhantes e se sobrepõem conforme pode ser visto. O valor R^2 corresponde à média obtida para o conjunto de validação, em cada partição dos dados. As partições 2 e 9 apresentam os menores valores para a métrica R^2 em todos os modelos, o que indica uma particularidade das imagens desses subconjuntos e não implicam na escolha do modelo. As curvas indicam que o fator de regularização tem pouca influência no desempenho do modelo. Isso se deve à grande quantidade de dados (píxeis) usada no treinamento, o que reduz a probabilidade de *overfitting*. Sendo assim, definiu-se o valor de $\alpha = 1$ para os testes posteriores, mantendo um modelo com complexidade intermediária nos demais experimentos.

A Figura 4.4 (b) mostra o resultado da métrica R^2 para os conjuntos de treinamento, vali-

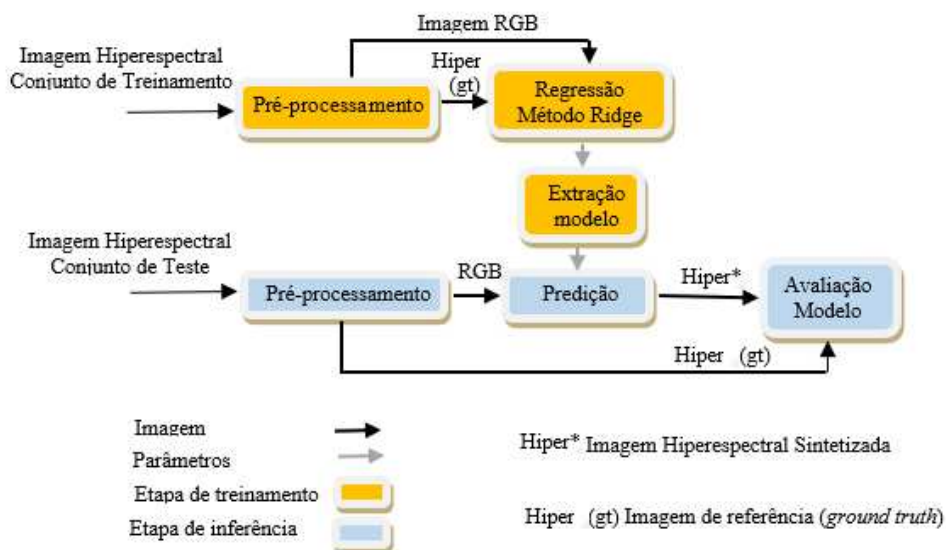


Figura 4.3: Diagrama de processo para recuperação do hiperespectro a partir de imagens RGB. As imagens hiperespectrais são projetadas no espaço RGB do sensor, sendo adicionado ruído devido à corrente de escuro e ruído impulsivo com distribuição de *Poisson*. As imagens hiperespectrais de referência (*ground truth*) são comparadas com as imagens hiperespectrais sintetizadas pelo método *Ridge Regression*.

dação e teste, para diferentes partições, segundo o método de validação cruzada. O valor do coeficiente R^2 foi calculado para cada subgrupo (partição) e para um valor de $\alpha = 1$. Não houve discrepâncias significativas nessa métrica, em cada conjunto utilizado nas etapas de treinamento, validação e teste, indicando que cada modelo representa bem a estrutura dos dados, e que a etapa de inferência aponta para uma correta generalização do modelo.

A Figura 4.4 (c) mostra a comparação do desempenho dos modelos de regressão, apurado para os conjuntos de treinamento, validação e teste, para três valores do fator de regularização α . O gráfico apresenta o valor médio do coeficiente de determinação para as partições dos conjuntos de treinamento e validação, assim como o valor médio de R^2 para as imagens do conjunto de teste. Observa-se que o desempenho, segundo a métrica R^2 , varia pouco com a alteração do modelo. O resultado para o conjunto de testes (inferência) apresenta desempenho ligeiramente inferior, o que é condizente com o esperado para processos de aprendizado de padrões.

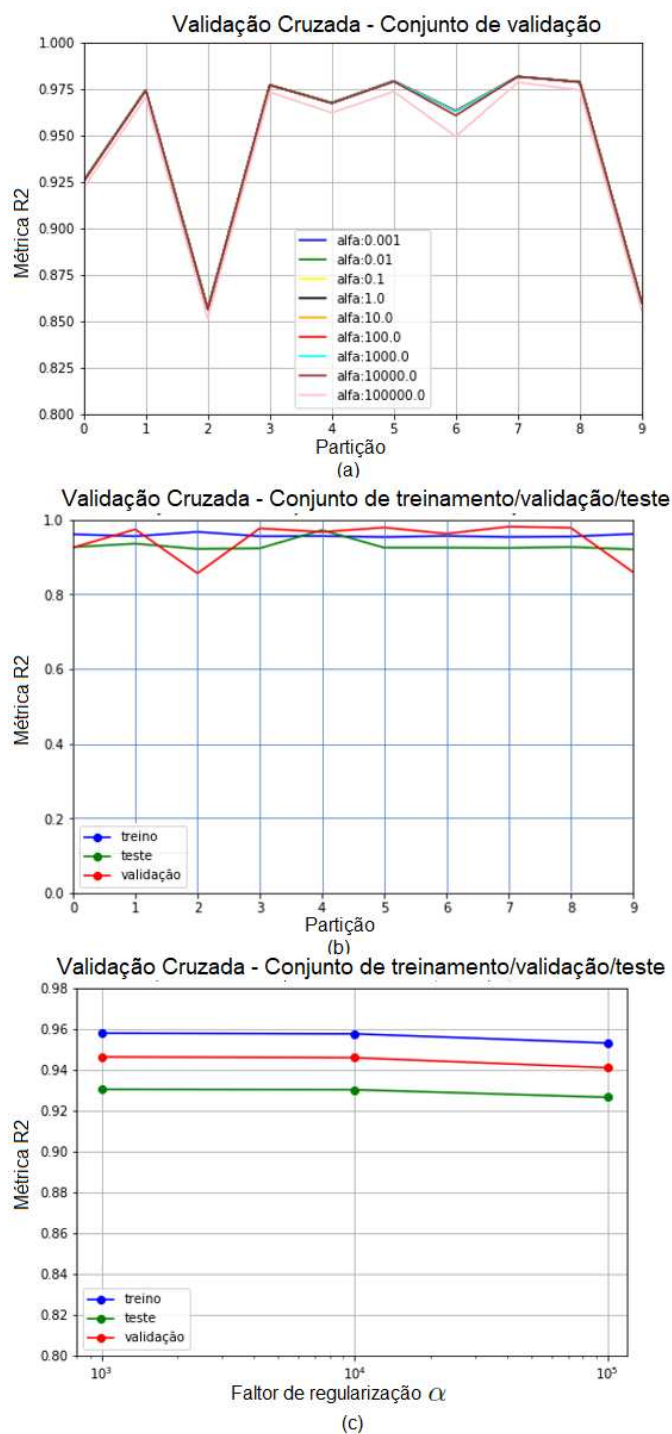


Figura 4.4: Em (a), métrica R^2 para cada combinação do conjunto de validação e fator de regularização. Em (b), métrica R^2 para cada combinação dos conjuntos de treinamento, validação e teste ($\alpha = 1$). Em (c), um gráfico com o valor médio para a métrica R^2 , em função do fator de regularização α .

4.2.2 Recuperação do hiperespectro por composição de agrupamentos combinado com regressão linear

A utilização de algoritmos para composição de agrupamentos (*clusters*) possibilita reunir píxeis com correlação espectral, considerando os píxeis como pontos em um espaço euclidiano multidimensional. Os píxeis das imagens hiperespectrais da base de dados ICVL utilizada nesse estudo são representados em um espaço com 31 dimensões associadas às bandas espectrais das imagens. Os píxeis RGB, por sua vez, são representados em um espaço tridimensional. Nos agrupamentos de pontos com assinatura espectral semelhante, há uma maior probabilidade de se obter uma representação que resida em um *manifold* de dimensões reduzidas, comparativamente às dimensões do espaço em que está inserido, levando-se em conta a variabilidade dos dados (Arad e Ben-Shahar, 2016). Também nesses agrupamentos, espera-se uma menor frequência de metâmeros, o que favorece a obtenção de uma correspondência única entre pontos do espaço RGB e pontos do espaço hiperespectral. Assumindo que pontos que representam píxeis que estão próximos no espaço RGB também estarão próximos no espaço hiperespectral, pode-se propor um método para recuperação da assinatura espectral a partir de imagens RGB por meio da composição de agrupamentos, sendo aplicado um algoritmo de regressão linear aos centroides de cada agrupamento, considerando-os como representativos da assinatura espectral desse agrupamento. Assim, a regressão é feita em um número de pontos consideravelmente menor, apenas os que correspondem aos centroides, otimizando o processamento.

A Figura 4.5 ilustra o esquema do processo. Um conjunto de imagens hiperespectrais foi projetado no espaço do sensor, conforme descrito na seção 3.6, sendo adicionado ruído para maior realismo e verificação da robustez do método, seguido de normalização das intensidades, obtendo-se um conjunto de imagens RGB. Esses dois conjuntos foram utilizados para treinar o algoritmo *Ridge Regression* para estimar o hiperespectro de píxeis centroides RGB. Dessa forma, obtém-se um modelo aplicado para compor uma representação hiperespectral dos píxeis-centroides definidos para o mesmo conjunto RGB, por meio do algoritmo *minibatch K-means* (Gareth et al., 2013). A avaliação do desempenho do modelo foi feita com as métricas R^2 , MARE e RMSE (Seção 3.2).

Os dados para treinamento do algoritmo compõem um conjunto com as mesmas 120 imagens hiperespectrais utilizadas para treinamento do método descrito na seção 4.2.1, para fins de

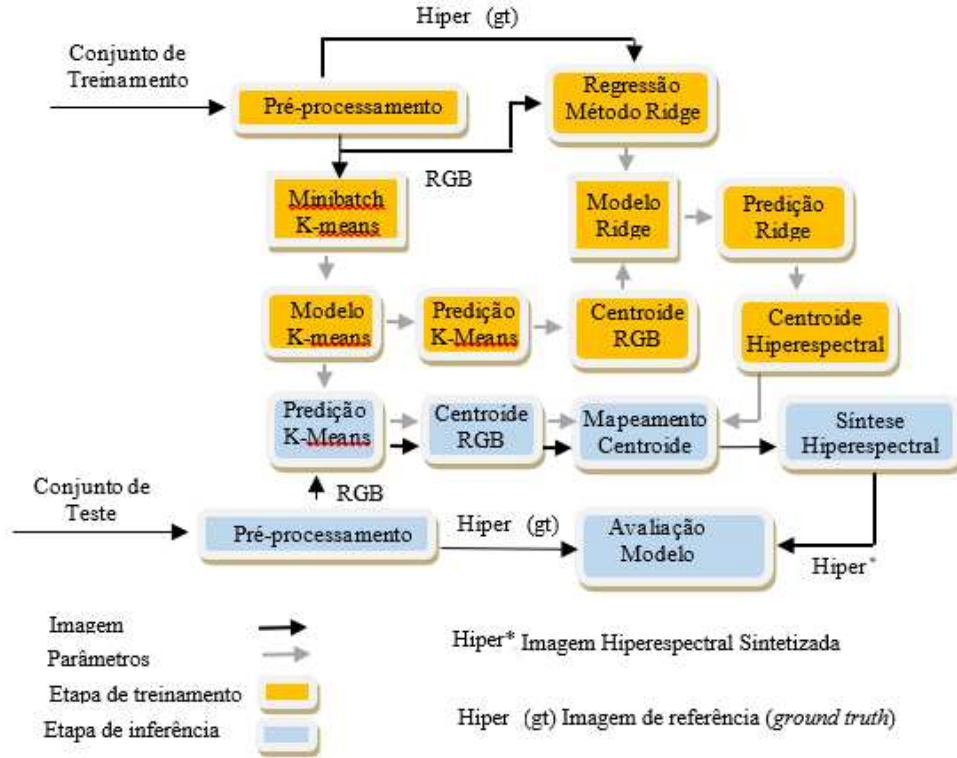


Figura 4.5: Diagrama de processo para recuperação do hiperespectro a partir de imagens RGB. O algoritmo *Ridge Regression* foi utilizado para obter as assinaturas espectrais dos centroides RGB, definidos para o conjunto de imagens RGB de treinamento, pela composição de agrupamentos por meio do algoritmo *minibatch K-Means*. As etapas em azul, no diagrama, correspondem à fase de inferência do processo.

comparação. O conjunto foi convertido em uma matriz $D_h \in \mathbf{R}^{n_s \times 31}$, onde

$$n_s = M_t.lin.col, \quad (4.2)$$

sendo n_s o número de píxeis na matriz, M_t o número de imagens utilizadas, lin o número de linhas, col é o número de colunas, sendo 31 bandas espectrais. Estas mesmas imagens, após etapa de pré-processamento, produziram a matriz $D_r \in \mathbf{R}^{n_s \times 3}$. Os agrupamentos RGB foram definidos usando o algoritmo *minibatch K-Means*, tomando os valores R, G e B de cada píxel como coordenadas no espaço RGB. Para cada agrupamento do espaço RGB, o algoritmo *minibatch K-Means* define um centroide, Cr_k , que minimiza a variância intra agrupamento. Este centroide corresponde a um píxel que representa o valor médio dos componentes R, G e B do agrupamento. Este píxel foi mapeado para o espaço hiperespectral (31 componentes espectrais) usando o modelo de *Ridge*, calculando o píxel centroide hiperespectral Ch_k para o agrupamento. Assim, apenas o espectro de cada centroide hiperespectral é estimado, usando regressão, ao invés vez de toda a imagem, reduzindo o custo computacional. A imagem hi-

perespectral sintetizada é obtida reconstruindo a imagem RGB, conforme os agrupamentos de píxeis e o espectro do píxel centroide hiperespectral correspondente.

Na etapa de inferência, cada imagem RGB do conjunto de teste foi submetida ao modelo *mini-batch K-Means* obtido na etapa de treinamento, para definir o agrupamento ao qual cada píxel pertence. A síntese da imagem hiperespectral foi realizada mapeando os centroides dos agrupamentos RGB (Cr_k) para os centroides dos agrupamentos hiperespectrais (Ch_k), conforme mostrado na Figura 4.6. Cada píxel em um agrupamento é mapeado para a mesma distribuição espectral correspondente ao centroide do espaço hiperespectral. O desempenho para diferentes números de agrupamentos foi avaliado usando validação cruzada com 10 partições (*10-fold*) do conjunto de dados.

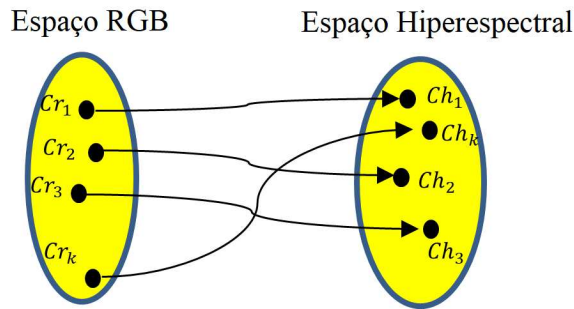


Figura 4.6: Mapeamento dos centroides do espaço RGB (Cr_k) para o espaço hiperespectral por meio da síntese dos píxeis centroides hiperespectrais (Ch_k)

A Figura 4.7 (a) mostra o valor médio do coeficiente de determinação para o conjunto de validação, em cada partição. O valor apresentado corresponde à média obtida para as imagens do conjunto de validação. A Figura 4.7 (b) mostra o valor médio do coeficiente de determinação (R^2 score) para o conjunto de validação em função do número de agrupamentos, usando validação cruzada. Foi calculada a média considerando as imagens de cada partição do conjunto de dados contendo 12 imagens. Observa-se uma melhora no desempenho, segundo essa métrica, à medida que o número de agrupamentos aumenta. Acima de 3.000 agrupamentos, a melhora no desempenho é mínima, com aumento significativo no custo computacional. Esse comportamento é um indício de que as distribuições espectrais para o conjunto de dados é limitada no que tange à sua diversidade. Sendo assim, definiu-se o número de agrupamentos igual a 1000 como um valor de compromisso entre acurácia e custo computacional. Utilizou-se um fator de regularização $\alpha = 1$ no modelo *Ridge* para avaliação dos modelos. A Figura 4.8 (a) mostra o valor médio da métrica MARE do conjunto de validação, para cada partição, utilizando valida-

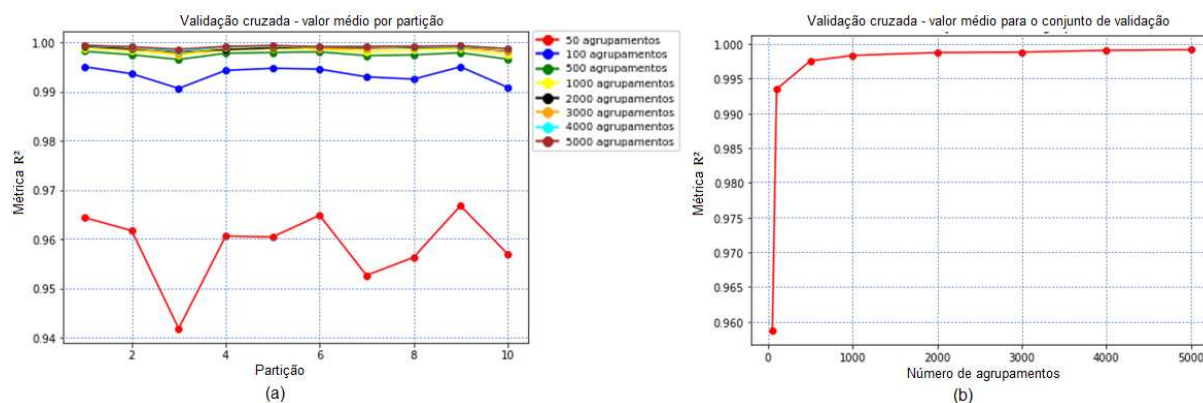


Figura 4.7: Em (a), o valor médio do coeficiente de determinação para o conjunto de validação, em cada partição. Em (b), o desempenho do modelo em função do número de agrupamentos, avaliados usando o valor médio do coeficiente de determinação (R^2 Score), calculado para 10 partições. Aumentar o número de agrupamentos melhora o desempenho, mas aumenta significativamente o custo computacional.

ção cruzada com 10 partições. O valor apresentado corresponde ao valor médio das imagens do conjunto de validação (12 imagens), usando *10-fold cross-validation*. Diversos modelos com diferentes números de agrupamentos foram avaliados. A Figura 4.8 (b) apresenta o valor médio da métrica MARE do conjunto de validação, em função do número de agrupamentos. O comportamento assintótico da curva é semelhante ao apresentado com a métrica R^2 . A Figura 4.9

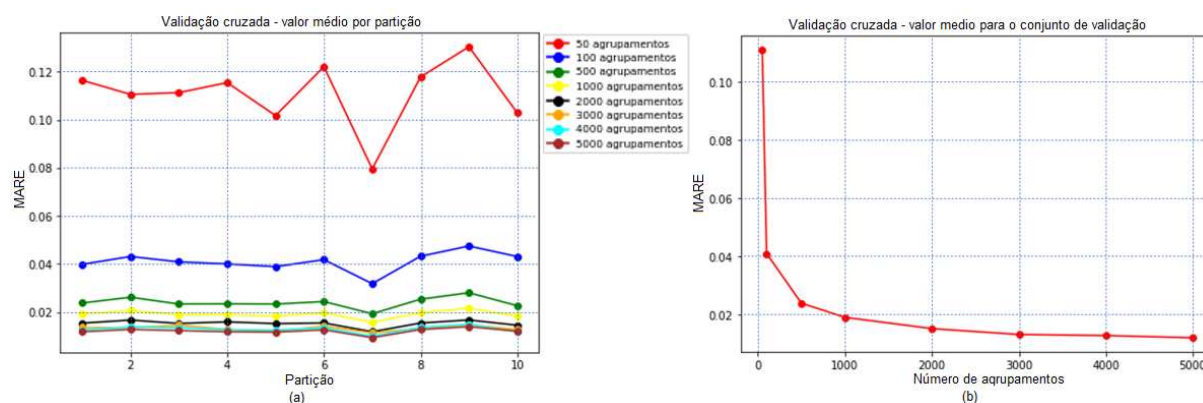


Figura 4.8: Valor médio da métrica MARE, calculado para cada partição do conjunto de treinamento e de validação.

mostra o valor médio da métrica MARE para o conjunto de validação (curva em vermelho) e o conjunto de teste (curva em azul) em função do número de agrupamentos. O valor apresentado em cada ponto da curva corresponde à média do MARE das imagens do conjunto de validação e do conjunto de testes, usando *10-fold cross-validation*. Observa-se que o desempenho, segundo essa métrica, na etapa de validação de modelos e na etapa de inferência são congruentes, indicando a correta generalização do modelo.

4.2.3 Recuperação do hiperespectro por redes neurais

Outra abordagem investigada para a resolução do problema de recuperação do hiperespectro a partir de imagens RGB utiliza redes neurais do tipo *multilayer perceptron* (Seção 2.12.

Da mesma forma que os processos anteriores, as imagens para treinamento da MLP foram pré-processadas por meio de normalização, projeção no espaço RGB e adição de ruído para emular maior realismo. Os parâmetros da rede foram definidos por meio de ensaios diversos. Os dados para treinamento foram os mesmos utilizados nos processos descritos na Seção 4.2.1 e na Seção 4.2.2. A Figura 4.10 apresenta o diagrama em blocos com as etapas desse processo.

Vários experimentos foram conduzidos com diferentes arquiteturas, variando a quantidade de camadas e de neurônios em cada estrutura de rede. Além disso, foram testadas funções de ativação ELU e RELU, funções de custo por erro médio quadrático (RMSE) e por erro absoluto médio relativo (MARE), bem como algoritmos otimizadores SGD e ADAMAX (Goodfellow et al., 2016).

A Figura 4.11 apresenta modelos de arquitetura com duas e três camadas. Variações da arquitetura de três camadas foram utilizadas para comparar o desempenho dos modelos, modificando o tamanho de lote (número de amostras utilizadas em cada iteração de treinamento do

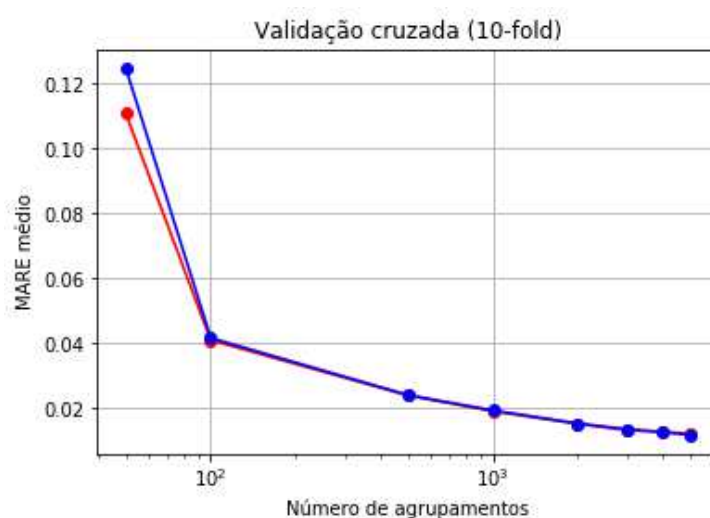


Figura 4.9: Valor médio da métrica MARE, calculado em função do número de agrupamentos. A curva em vermelho representa os valores para o conjunto de validação, a curva em azul corresponde aos valores do conjunto de teste.

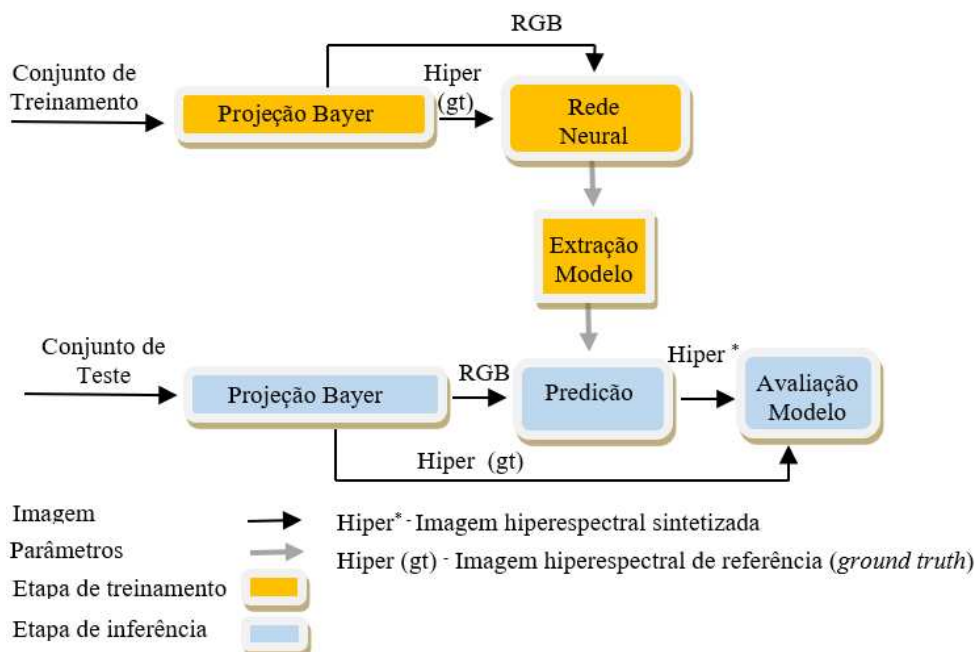


Figura 4.10: Diagrama de processo contendo uma rede neural treinada para obter as assinaturas espectrais de pixels a partir dos valores RGB. O processo inclui normalização das imagens e adição de ruído, emulando o sensor de uma câmera fotográfica. A projeção no espaço RGB utilizou a curva de resposta da câmera Canon 40D.

algoritmo MLP), função de ativação e tipo de otimizador. A taxa de aprendizado (*lr* - learning rate) foi definida em 0,01 por mostrar-se como um valor de compromisso entre convergência suave e tempo de execução. A métrica MARE foi utilizada para avaliar os modelos.

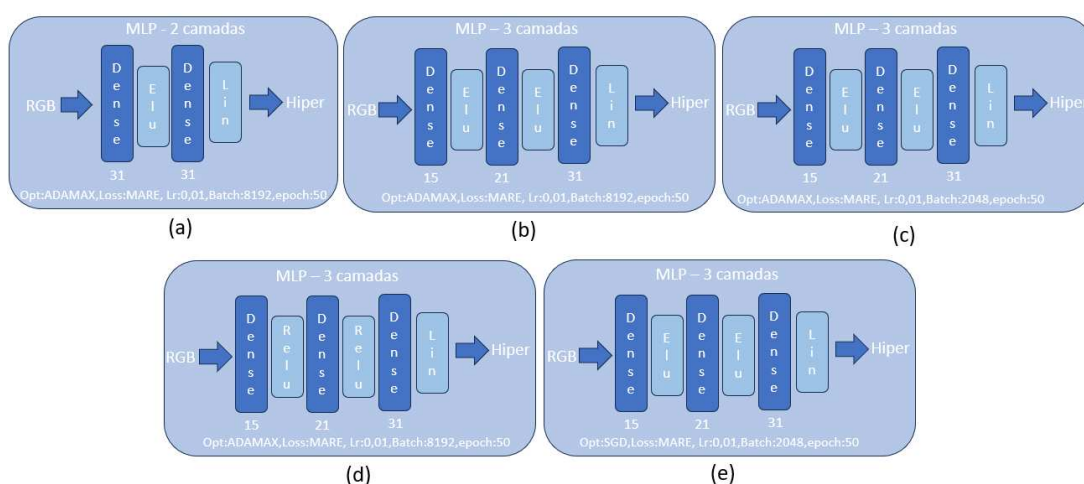


Figura 4.11: Arquitetura de redes MLP com duas e três camadas e variações nos parâmetros e configurações, que incluem o tipo de otimizador, função de ativação e tamanho do lote (*batch*). Os números junto às camadas “Dense” indicam a quantidade de neurônios. Em (a), uma rede com duas camadas, otimizador (*Opt*) ADAMAX, função de ativação ELU, lote (*batch*) de 8192, função de custo (*loss*) MARE, e taxa de aprendizado (*lr* - learning rate) de 0,01. Em (b), (c), (d) e (e), as redes são de três camadas. Em (c), o tamanho do lote foi de 2048. Em (d), modificou-se a função de ativação para RELU. Em (e) modificou-se o otimizador para SGD.

A Figura 4.12 mostra modelos de arquitetura com quatro, cinco e sete camadas.

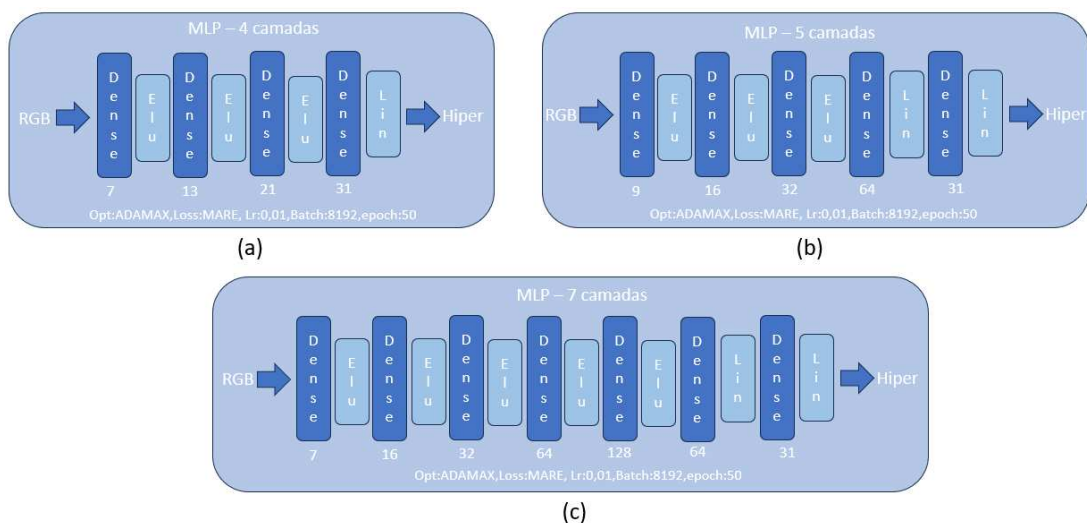


Figura 4.12: Arquitetura de redes MLP com quatro (a), cinco (b) e sete (c) camadas. Em todas as arquiteturas utilizou-se otimizador (*Opr*) ADAMAX, função de ativação ELU, lote (*batch*) de 8192, função de custo (*loss*) MARE, e taxa de aprendizado (*lr -learning rate*) de 0,01. O treinamento foi feito em 50 épocas (*epoch*). O número de neurônios em cada camada variou nas diversas arquiteturas.

A Figura 4.13 contém modelos de nove e onze camadas. Nesses casos, foi acrescida uma etapa *dropout* para regularização e prevenção de sobreajuste (*overfitting*) (Goodfellow et al., 2016). O método de validação cruzada (*10-fold*) foi utilizado para avaliar o desempenho dos

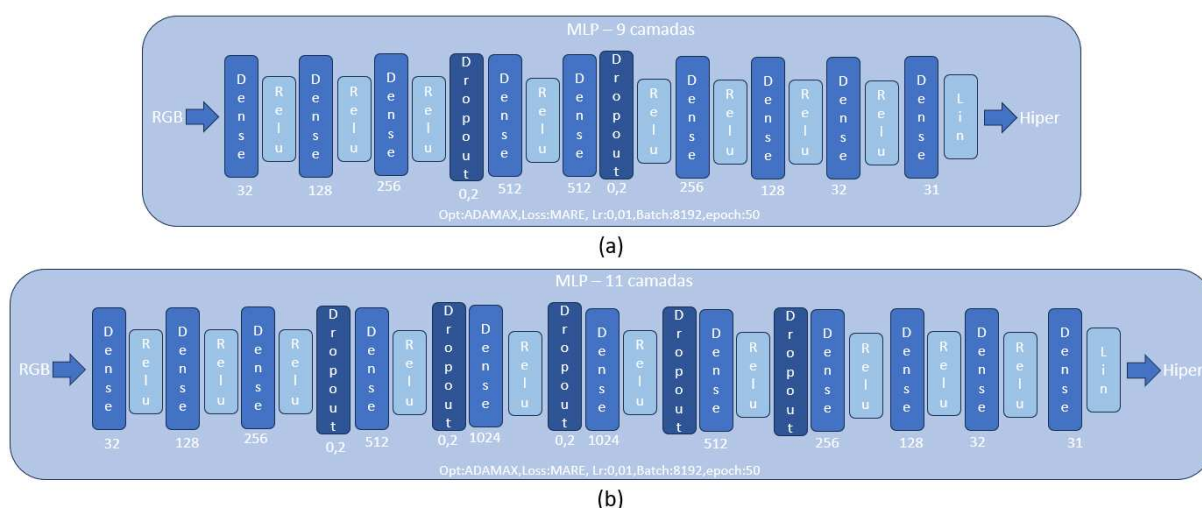


Figura 4.13: Arquitetura de redes MLP com nove (a) e onze (b) camadas. Nessas arquiteturas mais complexas foi acrescida uma função de *dropout* para prevenir *overfitting*. O valor de *dropout* em 0,2 indica que 20% dos neurônios puderam ser seletivamente desativados.

modelos, sendo definida uma taxa de aprendizado de 0,01. Esse valor representou um compromisso entre velocidade de convergência e suavidade na amortização da função de custo. Cada

modelo foi treinado com 50 épocas. A função de perda exibe uma convergência assintótica para um valor estável, mesmo para um número de épocas menor. As curvas de custo, correspondentes a esses experimentos, são apresentadas na Figura 4.14. O tamanho do lote não influenciou o desempenho da rede. O otimizador ADAMAX apresentou melhor desempenho, em relação ao SGD, assim como a função RELU, em relação à função de ativação ELU.

O acréscimo de camadas e neurônios melhora o desempenho, entretanto acarreta aumento substancial no custo computacional. Uma arquitetura com sete e nove camadas propiciou desempenho semelhante, segundo a métrica MARE, enquanto uma arquitetura com onze camadas melhorou o desempenho em 10%. A métrica MARE foi utilizada como função de custo na etapa de treinamento da rede. A partir destes experimentos, foi escolhida uma arquitetura com

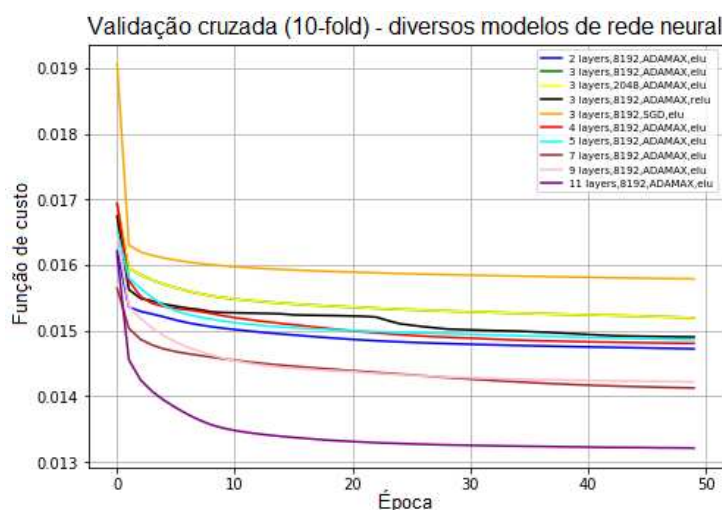


Figura 4.14: Curvas da função de custo para várias arquiteturas de redes neurais, tamanho do lote, função de ativação e tipo de otimizador. A validação cruzada com 10 partições e a métrica MARE foram usadas para avaliar os modelos.

onze camadas, por representar um compromisso entre acurácia e custo computacional. O melhor resultado foi obtido com a função RELU (*Rectified Linear Unit*) nas camadas iniciais e uma função de ativação linear na camada de saída. Como o objetivo era obter uma imagem com 31 bandas espectrais, esta camada foi configurada com 31 neurônios.

4.2.4 Recuperação do hiperespectro por redes neurais combinado com agrupamentos

Nesta proposta de arquitetura foi utilizada uma combinação do algoritmo *minibatch K-Means* para composição de agrupamentos e uma rede neural do tipo MLP de onze camadas. A formação de agrupamentos se deu da mesma forma como descrito na Seção 4.2.2, utilizando a mesma base de dados de treinamento e teste. Neste caso, uma rede neural foi utilizada para compor uma representação hiperespectral dos píxeis-centroides. Na etapa de verificação da qualidade do método, cada imagem do conjunto de testes foi submetida ao modelo *minibatch K-Means* para definir a qual agrupamento cada píxel da imagem RGB pertence, substituindo o centroide RGB pelo centroide hiperespectral do mesmo agrupamento. A Figura 4.15 ilustra o diagrama do processo.

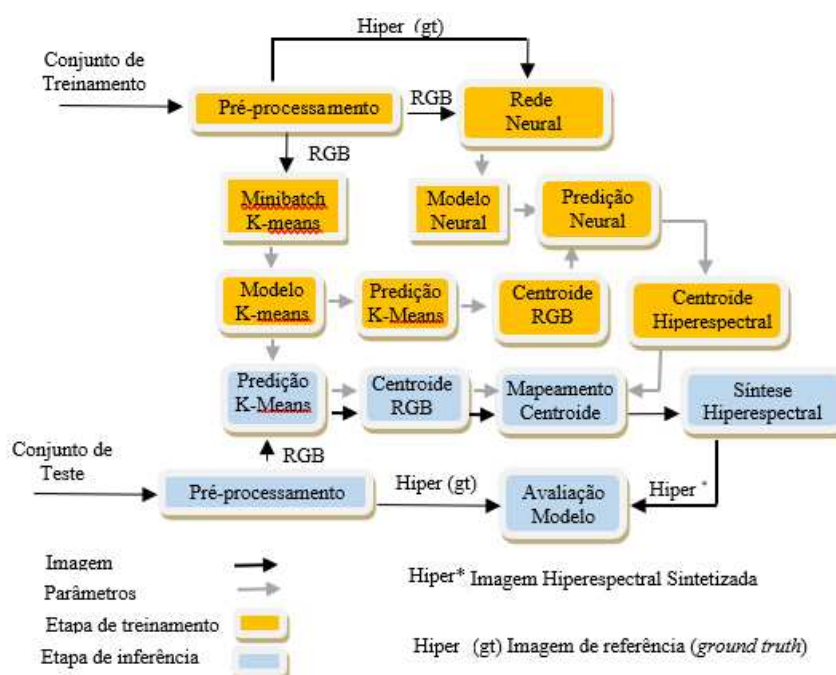


Figura 4.15: Diagrama de processo contendo as etapas de processamento das imagens RGB e hiperespectrais, o algoritmo *minibatch K-Means* para formação de agrupamentos e uma rede neural treinada para obter as assinaturas espectrais dos píxeis centroides. O processo inclui normalização das imagens e adição de ruído. As etapas em azul no diagrama correspondem à fase de inferência do processo. A projeção no espaço RGB utilizou a curva de resposta da câmera Canon D40.

4.2.5 Análise qualitativa do desempenho dos algoritmos implementados

A Figura 4.16 (i) apresenta uma imagem hiperespectral do conjunto de teste, renderizada a partir da composição das bandas espectrais 650 nm, 550 nm e 450 nm. Essas bandas correspondem, respectivamente, ao comprimento de onda central para as cores vermelha (R), verde (G) e azul (B). Para essa base de dados, a largura de cada banda espectral é de 10 nm. As Figuras 4.16 (ii), (iii), (iv), e (v) correspondem às imagens hiperespectrais sintetizadas, respectivamente, pelos processos utilizando o algoritmo *Ridge Regression* (RR), *Ridge Regression* combinado com *K-Means* (RR/KMC), *multilayer perceptron* (MLP) e MLP combinado com *K-Means* (MLP/KMC). As imagens sintetizadas também foram renderizadas a partir das mesmas bandas. De uma forma geral, os processos recuperaram a estrutura espacial da imagem de forma correta, com boa relação de contraste, sem artefatos ou distorções. Do ponto de vista qualitativo, não há distinção significativa de desempenho entre os processos.



Figura 4.16: Em (i), uma amostra de uma imagem hiperespectral do conjunto de teste, renderizada a partir da composição das bandas espectrais 650 nm, 550 nm e 450 nm. Em (ii), (iii), (iv) e (v) a correspondente imagem hiperespectral estimada em cada método, respectivamente, *Ridge*, *Ridge/K-Means*, MLP e MLP/*K-Means*.

A Figura 4.17 (a) apresenta imagens originais (*ground truth*) em tons de cinza para cada uma das 31 bandas espectrais na faixa de 400 nm a 700 nm. As Figuras 4.17 (b), (c), (d) e (e), correspondem, respectivamente, às imagens estimadas pelos processos utilizando o algoritmo *Ridge Regression*, *Ridge Regression* combinado com *K-Means*, MLP e MLP combinado com o algoritmo *K-Means*. O contraste e nitidez varia nas imagens em cada comprimento de onda, sendo menor nos comprimentos na faixa da cor azul (B), em todos os processos de síntese e mesmo na imagem original. Isso se deve à menor intensidade de sinal nessa faixa, produzindo uma menor relação sinal-ruído. O processo MLP apresentou o melhor desempenho, e até melhor que a imagem original em termos de contraste.

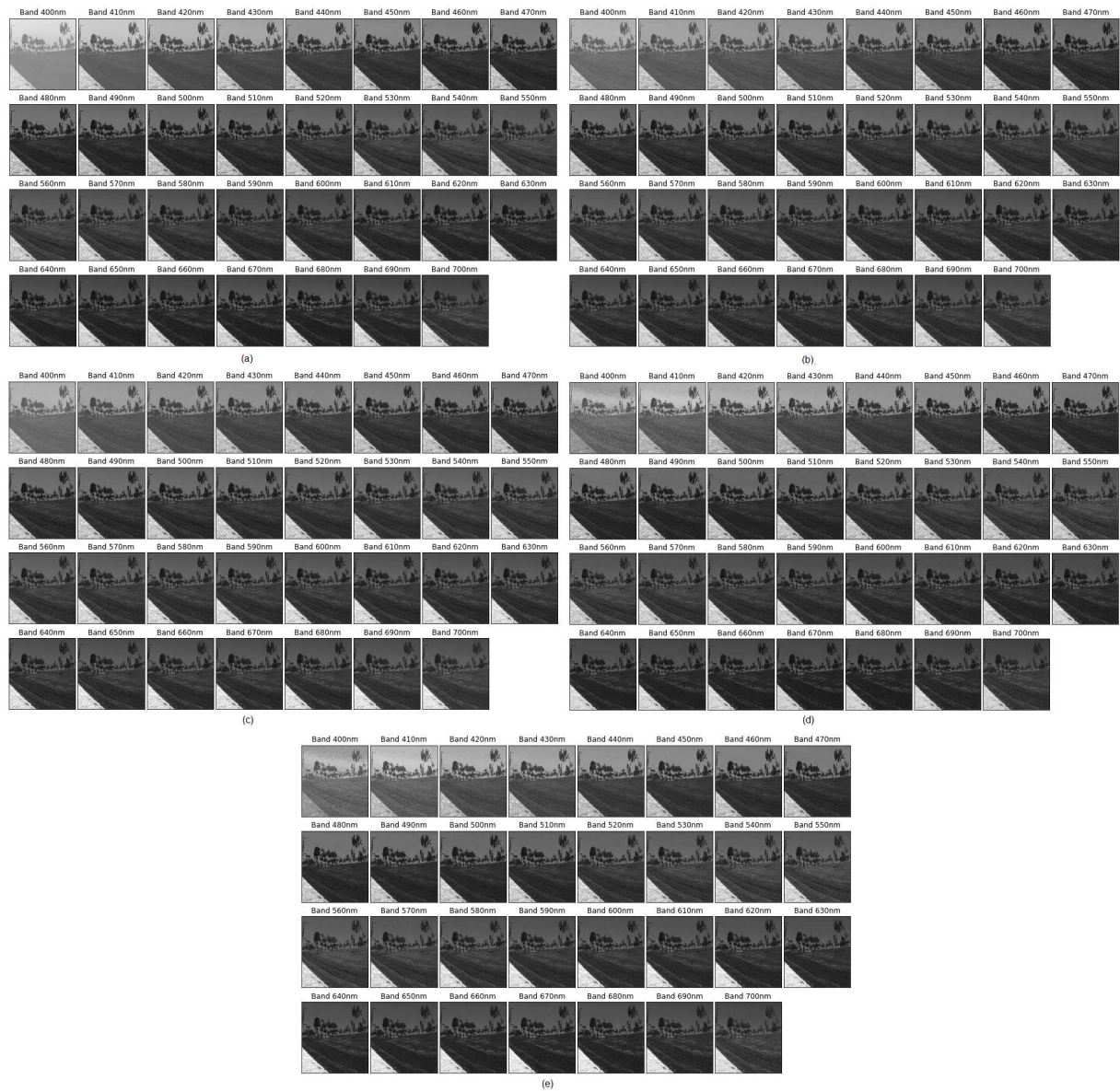


Figura 4.17: Figura (a) Amostra de imagem hiperespectral do conjunto de teste, em tons de cinza, para 31 bandas espectrais. Em (b), (c), (d) e (e) as correspondentes imagens em tons de cinza, para as 31 bandas espectrais, conforme estimadas em cada método, respectivamente, *Ridge* (RR), *Ridge/K-Means* (RR/KMC), MLP e MLP combinado com *K-Means* (MLP/KMC).

Foram selecionados píxeis, por diversidade de cor e aspecto na imagem, para terem seu espectro comparado em cada um dos quatro processos, conforme apresentado na Figura 4.16. A Figura 4.18 apresenta as assinaturas espectrais desses píxeis na imagem sintetizada, utilizando os quatro processos. O registro da curva na cor azul corresponde à intensidade do píxel em função do comprimento de onda, na faixa de 400 nm a 700 nm, referente à imagem original (*ground truth*). A curva em amarelo representa a estimativa feita com uso do algoritmo de regressão *Ridge* (RR). A curva na cor preta representa a estimativa feita com uso combinado dos

algoritmos de regressão linear e *K-Means* (RR/KMC). A curva em verde representa a estimativa feita com uso da rede neural (MLP). A curva em laranja representa a estimativa feita com uso combinado dos algoritmos MLP e *K-Means* (MLP/KMC). Como pode ser visto na figura, todos os algoritmos foram bem sucedidos em estimar a distribuições espectrais dos píxeis “a”, “b” e “c” (marcados na Figura 4.16). No entanto, somente os processos MLP e MLP/KMC, por corresponderem a modelos de maior complexidade referentes a redes neurais, foram bem sucedidos em acompanhar a silhueta espectral do píxel “d” em suas variações menos suaves.

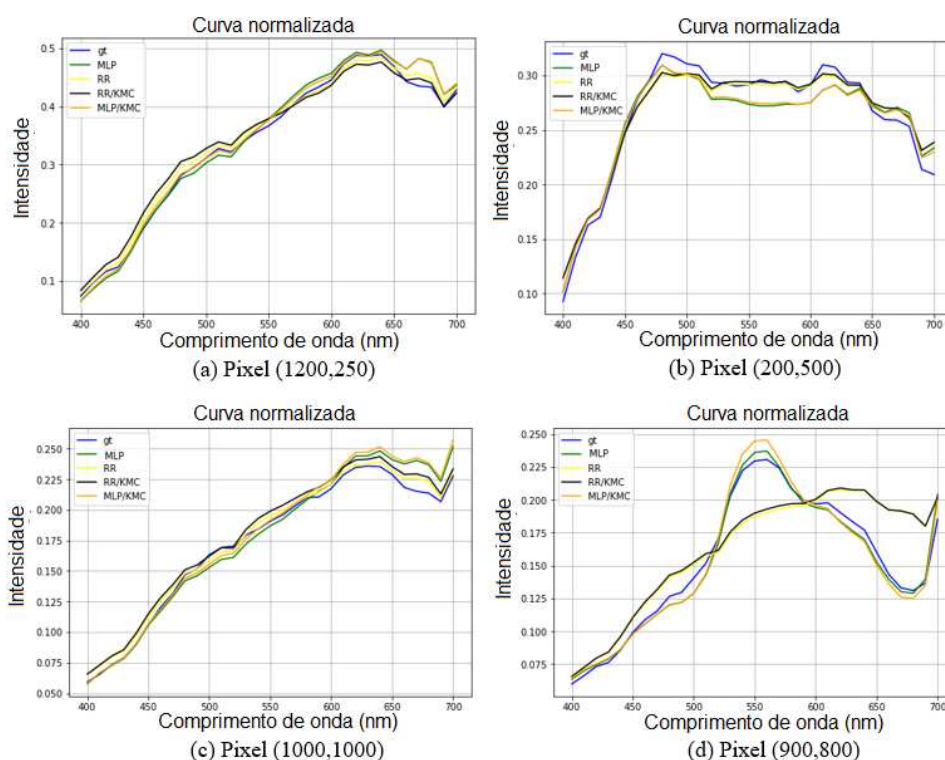


Figura 4.18: Comparação das assinaturas espectrais de píxeis da imagem original e das imagens estimadas utilizando os quatro processos.

4.2.6 Análise quantitativa do desempenho dos processos pela métrica MARE e RMSE

Uma análise quantitativa, comparativa do desempenho dos processos de síntese hiperespectral, foi realizada com uso de 63 imagens do conjunto de teste, considerando as métricas MARE e RMSE. A Tabela 4.1 resume os dados comparativos da métrica MARE para avaliação de desempenho dos algoritmos, nessa etapa do trabalho. O melhor desempenho foi alcançado no processo utilizando rede neural do tipo MLP, tanto pela métrica MARE quanto RMSE. Considerando o valor médio do MARE, a rede neural apresentou um desempenho 3,5% melhor do que o processo MLP/KMC, 18,3% melhor do que o processo RR/KMC e 18,2% melhor do que o processo RR.

Tabela 4.1: Comparação de desempenho dos processos pela métrica MARE, aplicada ao conjunto de testes

Processo	RR	RR/KMC	MLP	MLP/KMC
Média	0,0888	0,0889	0,0726	0,0752
Desvio padrão	0,0795	0,0798	0,0811	0,0831
Mínimo	0,0435	0,0435	0,0269	0,0282
Máximo	0,5399	0,5436	0,5049	0,5049

A Figura 4.19 apresenta a comparação da dispersão de valores da métrica MARE para os quatro processos de síntese, considerando as 63 imagens do conjunto de testes. O gráfico do tipo *boxplot* mostra o valor da mediana e os valores limites da distribuição, assim como os valores discrepantes de algumas imagens do conjunto (*outliers*). Pode-se notar que os processos MLP e MLP/KMC possuem menor dispersão de valores em relação ao valor da mediana. Os gráficos mostram que o processo MLP tem melhor desempenho, seguido pelo processo MLP/KMC. Os processos RR e RR/KMC tiveram desempenho semelhante, segundo o gráfico, e corroborado pelos dados apresentados na Tabela 4.1.

Um teste de comparação de médias foi aplicado para averiguar se há uma distinção significativa da distribuição de valores MARE entre os processos. Conforme apresentado na Seção 3.5, preliminarmente, aplica-se o Teste de *Shapiro-Wilk* para avaliar se os dados seguem uma distribuição normal. Aplicando o Teste de *Shapiro-Wilk* aos dados de MARE que produziram a

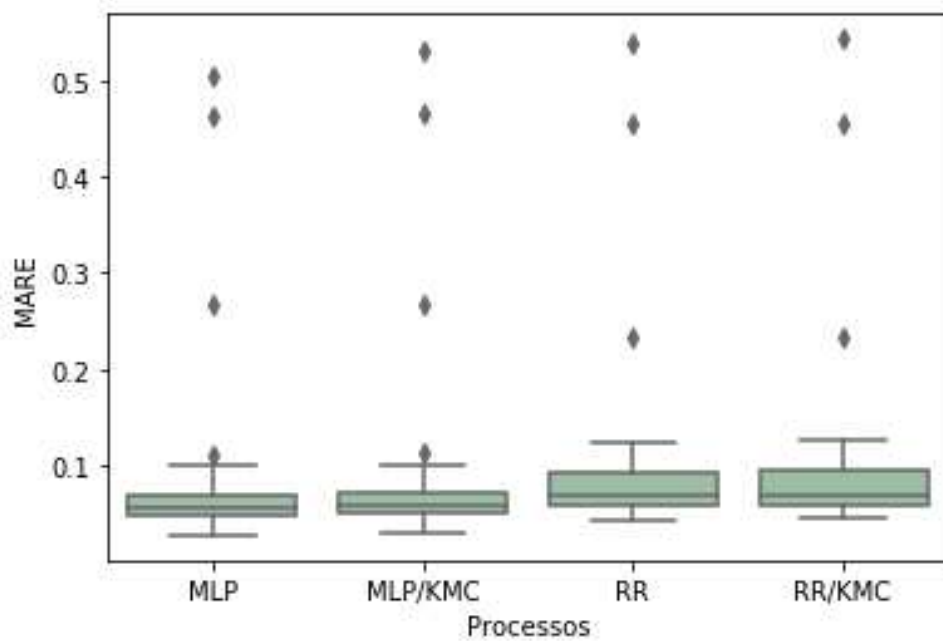


Figura 4.19: Comparação da distribuição da métrica MARE obtida utilizando os quatro processos de síntese hiperespectral.

Tabela 4.1 e as distribuições mostradas na Figura 4.19, obteve-se

$$pvalue = 1,7 \times 10^{-27}. \quad (4.3)$$

Com isso, rejeita-se a Hipótese nula (critério: $pvalue < 0,05$), ou seja, os dados não vem de uma distribuição normal. Sendo assim, optou-se pelo método de *Kruskal-Wallis* para comparar os valores das medianas dos valores de MARE dos quatro processos, ao invés do Teste ANOVA que pressupõe que os dados são gaussianos.

No Teste de *Kruskal-Wallis*, o valor de $pvalue$ indica a probabilidade de que as diferenças observadas entre os grupos tenham ocorrido por acaso, sob a hipótese nula de que todas as amostras provêm da mesma distribuição.

Aplicando o Teste de *Kruskal-Wallis* aos dados de MARE obteve-se

$$pvalue = 1,6 \times 10^{-6}. \quad (4.4)$$

Dessa forma, rejeita-se a hipótese nula (critério: $pvalue < 0,05$). Isso indica que há evidências suficientes para afirmar que pelo menos um dos grupos tem uma mediana diferente dos outros grupos.

A Tabela 4.2 contém dados referentes à métrica RMSE. A tabela apresenta os valores mínimo, médio e máximo associados ao processamento das imagens do conjunto de teste. Considerando o valor médio do RMSE, a rede neural apresentou um desempenho 4,6% melhor que o processo MLP/KMC, 15,4% melhor do que o processo RR/KMC e 14,2% melhor que o processo RR.

Tabela 4.2: Comparação de desempenho dos processos pela métrica RMSE aplicada ao conjunto de testes.

Processo	RR	RR/KMC	MLP	MLP/KMC
Média	0,0218	0,0221	0,0187	0,0196
Desvio padrão	0,0135	0,0136	0,0140	0,0144
Mínimo	0,0074	0,0075	0,0054	0,0059
Máximo	0,0847	0,0859	0,0806	0,0864

A Figura 4.20 apresenta a comparação da distribuição da métrica RMSE para os quatro processos de síntese, considerando as 63 imagens hiperespectrais do conjunto de testes. O gráfico mostra a mediana e a dispersão de valores em cada processo. De forma diversa ao comportamento para a métrica MARE, o processo MLP/KMC mostra uma dispersão maior em relação do processo MLP e se assemelha aos processos RR e RR/KMC. Isso indica que a métrica MARE é mais eficaz em diferenciar o desempenho entre os processos.

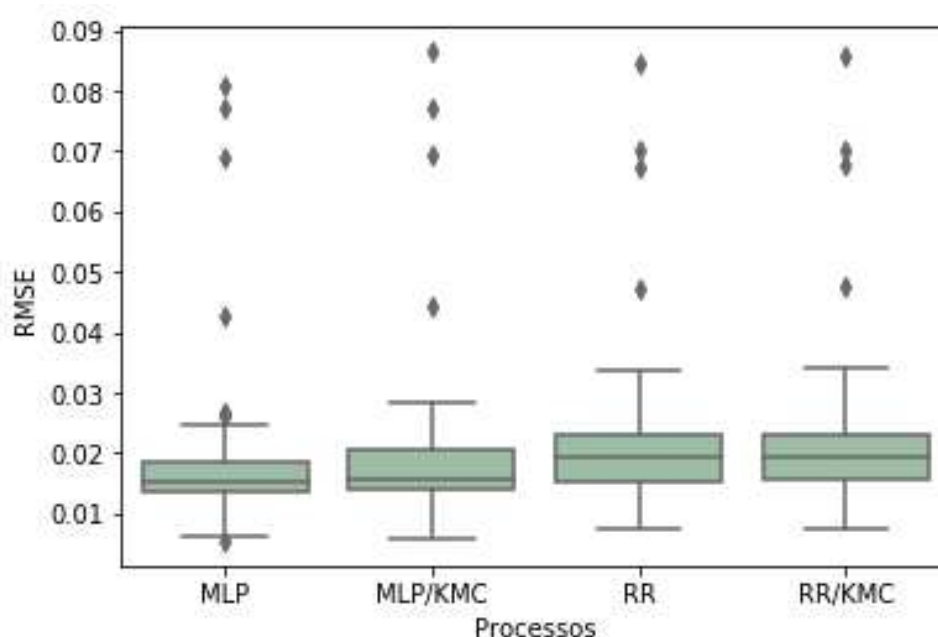


Figura 4.20: Comparação da distribuição da métrica RMSE obtida utilizando os quatro processos de síntese hiperespectral.

Um teste de comparação de médias foi aplicado para averiguar se há uma distinção significativa da distribuição de valores RMSE entre os processos.

Aplicando o Teste de *Shapiro-Wilk* aos dados de RMSE que produziram a Tabela 4.2 e as distribuições mostradas na Figura 4.20, obteve-se

$$pvalue = 5,0 \times 10^{-23}. \quad (4.5)$$

Com esse resultado, rejeita-se a Hipótese nula (critério: $pvalue < 0,05$) de que os dados de RMSE dos processos vem de uma distribuição normal. Sendo assim, aplicou-se o Teste de *Kruskal-Wallis* aos dados de RMSE obtendo

$$pvalue = 2,1 \times 10^{-3}. \quad (4.6)$$

Isso indica que os dados vem de distribuições diferentes para cada processo. A Tabela 4.3 apresenta dados de desempenho quanto ao tempo de processamento dos algoritmos. O teste utilizou um método de medição que leva em conta o tempo que o processo associado ao programa ficou em modo *sleep*, ou seja, este tempo não é contabilizado. O valor da tabela corresponde à média de 10 medições.

Tabela 4.3: Comparação de desempenho dos processos com base no tempo de processamento em segundos (Hardware: Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz, 196GB RAM)

Processo	RR	RR/KMC	MLP	MLP/KMC
Média	0,47	14,5	219,9	13,8
Desvio padrão	0,1	0,8	4,3	1,0

O algoritmo RR é 468 vezes mais rápido do que MLP e 29,4 vezes mais rápido do que o processo MLP/KMC, sendo ainda 30,9 vezes mais rápido do que o processo RR/KMC. Os processos que utilizam agrupamentos podem ser ainda otimizados para tempo de processamento pela redução no número de agrupamentos, que foi de 1000 na presente versão, com implicações nas métricas MARE e RMSE.

A Figura 4.21 apresenta os gráficos da métrica MARE para os processo descritos nessa seção, para cada imagem do conjunto de testes. A Figura 4.22 apresenta os gráficos da mé-

trica RMSE. Poucas imagens do conjunto de testes apresentaram valores muito discrepantes (*outliers*) do valor médio MARE e RMSE, sendo as mesmas em todos os processos.

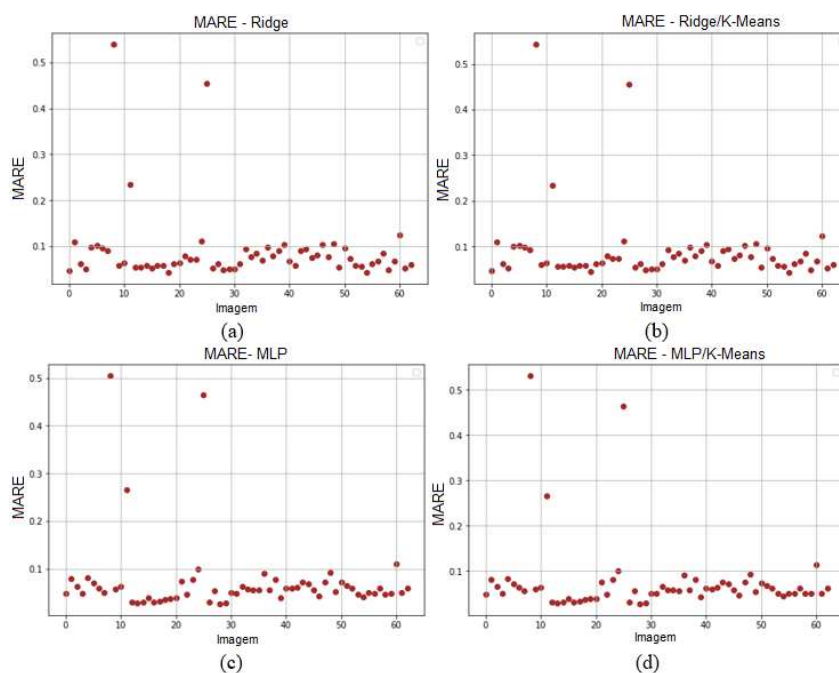


Figura 4.21: Gráficos das métricas MARE para os algoritmos implementados. Em (a) desempenho do algoritmo *Ridge*. Em (b) desempenho do método combinado *K-Means* e *Ridge*. Em (c) desempenho da MLP. Em (d) desempenho do método combinado *K-Means* e MLP.

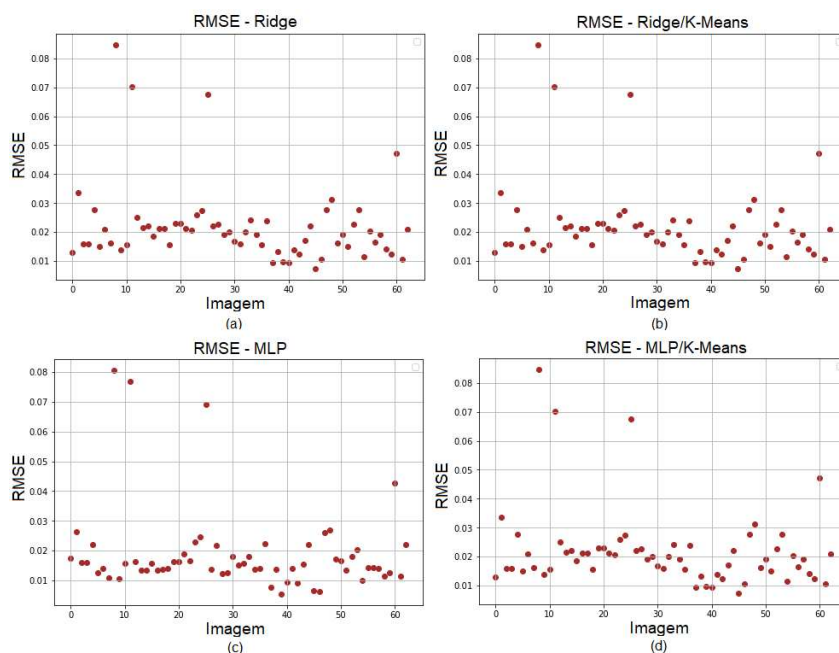


Figura 4.22: Gráficos das métricas RMSE para os algoritmos implementados. Em (a) desempenho do algoritmo *Ridge*. Em (b) desempenho do método combinado *K-Means* e *Ridge*. Em (c) desempenho da MLP. Em (d) desempenho do método combinado *K-Means* e MLP.

4.3 Análise de espectros com uso da câmera Specim FX-10

A câmera hiperspectral Specim FX-10 produz 448 bandas espectrais. A Seção 3.7 apresenta as principais características dessa câmera. A Figura 4.23 apresenta o “cubo hiperspectral” produzido por essa câmera, contendo 448 imagens monocromáticas, na faixa de comprimentos de onda de 400 nm a 1000 nm, com amostragem de 1,33nm.

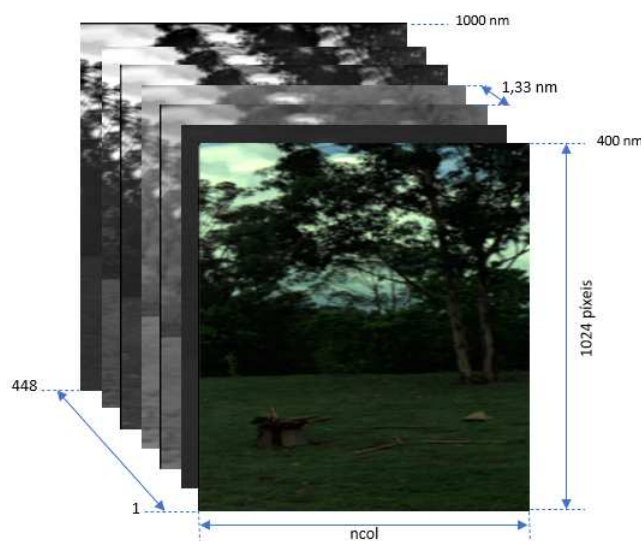


Figura 4.23: “Cubo hiperspectral” gerado com uso da câmera Specim FX-10 contendo 448 imagens monocromáticas, correspondentes aos “canais de cor” da câmera. As bandas espectrais são estreitas com largura de 1,33 nm, cobrindo a faixa de 400 nm a 1000 nm. As imagens possuem 1024 píxeis por coluna, sendo o número de píxeis por linha dependente do ângulo de varredura configurado na câmera.

A Figura 4.24 (b) contém um excerto desse “Cubo” com imagens monocromáticas na faixa de luz visível e na faixa do infravermelho próximo. A Figura 4.24 (a) apresenta, na parte superior, uma renderização com a combinação das bandas 450 nm, 550 nm e 650 nm, correspondentes, respectivamente, ao centro da faixa do azul, verde e vermelho. Na parte inferior da figura, encontra-se uma imagem disponibilizada no conjunto de arquivos produzidos pela câmera, a partir do software aplicativo “*LUMO RECORDER*”. A banda de 769,64 nm tem especial interesse para a detecção de incêndios florestais por possibilitar o seu uso como marcador de emissão do potássio. A imagem na banda 950 nm apresenta alto contraste para o sinal de fogo. Isso indica um potencial de segmentação de imagens com base no comprimento e onda.

A Figura 4.25 mostra um conjunto de imagens extraído da base de dados composta com o uso dessa câmera. Em (i) imagens obtidas de cenas com iluminação pela luz do sol. Em (ii),

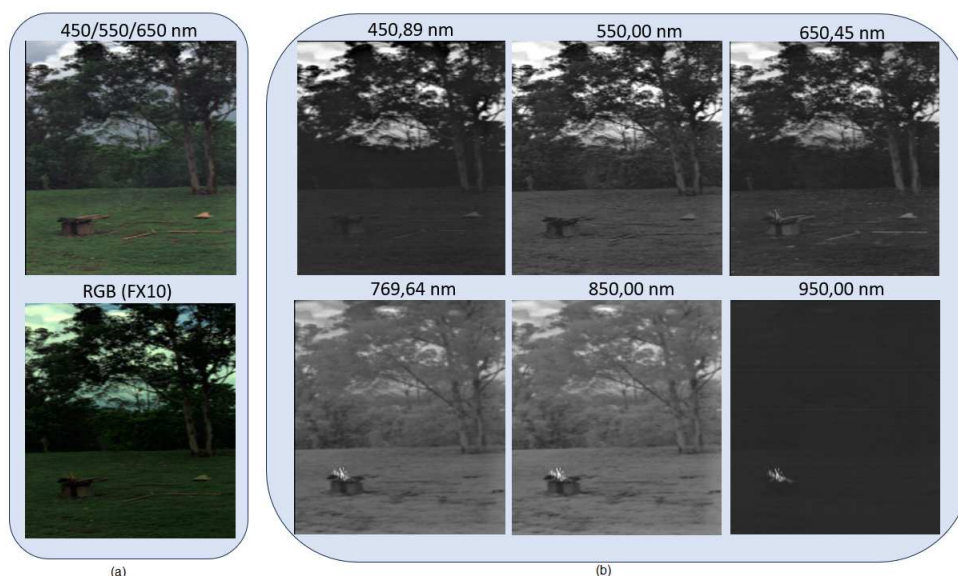


Figura 4.24: Em (a), parte superior, uma imagem renderizada com a combinação das bandas 450 nm, 550 nm e 650 nm. Na parte inferior, uma imagem renderizada produzida pelo software de controle da câmera Specim FX-10. Em (b), imagens monocromáticas correspondentes aos canais centrados nos comprimentos de onda indicados em cada imagem.

imagens produzidas em laboratório utilizando iluminação com lâmpadas halógenas.

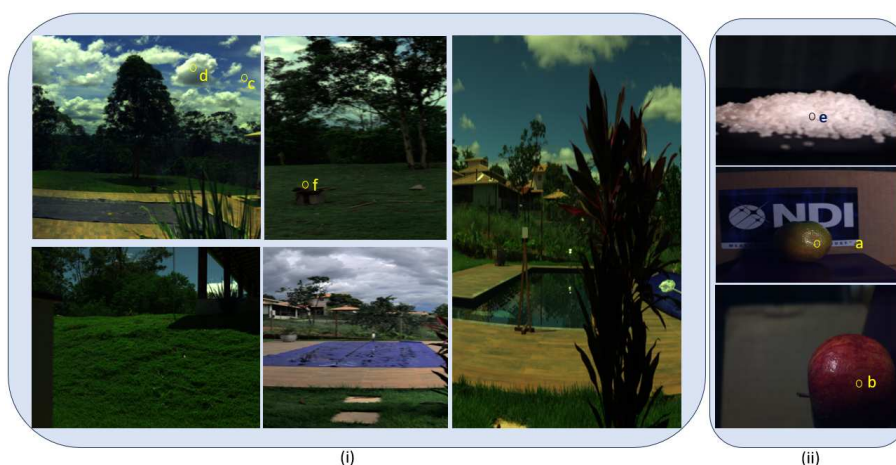


Figura 4.25: Amostra da base de dados de imagens capturadas com o uso da câmera FX-10. As imagens foram obtidas com iluminação natural pela luz do sol (i) e artificial por lâmpada halógena (ii). Os pontos dos centros dos círculos correspondem a píxeis cujos espectros são apresentados na Figura 4.26.

A Figura 4.26 apresenta espectros de píxeis relacionados a diversos elementos imageados com uso da câmera Specim FX-10. Em (a) o espectro do píxel correspondente ao ponto “a” na imagem da Figura 4.26 (imagem de uma laranja). Em (b), o espectro do píxel referente ao ponto “b” (imagem de uma maçã). Essas imagens foram iluminadas com lâmpada halógena, cujo espectro é apresentado na Figura 4.28 (a). Observe o maior conteúdo de energia na região do verde (em torno de 550 nm) do píxel “a” em comparação com o píxel “b”, que apresenta

maior energia relativa no vermelho (em torno de 650 nm). A distribuição de potência acima de 700 nm é semelhante nos dois píxeis e acompanha o perfil de distribuição de potência da lâmpada halógena nessa região do espectro. Em (c), o espectro do píxel “c”, de um ponto em uma imagem de uma nuvem, e em (d) o píxel de uma região da imagem relativa ao azul do céu. A silhueta da curva se deve ao perfil espectral do sol, das características de reflexão e absorção do objeto, assim como as características de transmissão da atmosfera, assim como da resposta espectral do sensor da câmera. O forte pico de absorção no comprimento de onda em torno de 760nm presente nos píxeis (c) e (d), não está presente nos píxeis (a) e (b) por serem capturadas com iluminação artificial e obtidas em laboratório. Essa absorção se deve à excitação da molécula de oxigênio pela onda eletromagnética que atravessa a atmosfera.

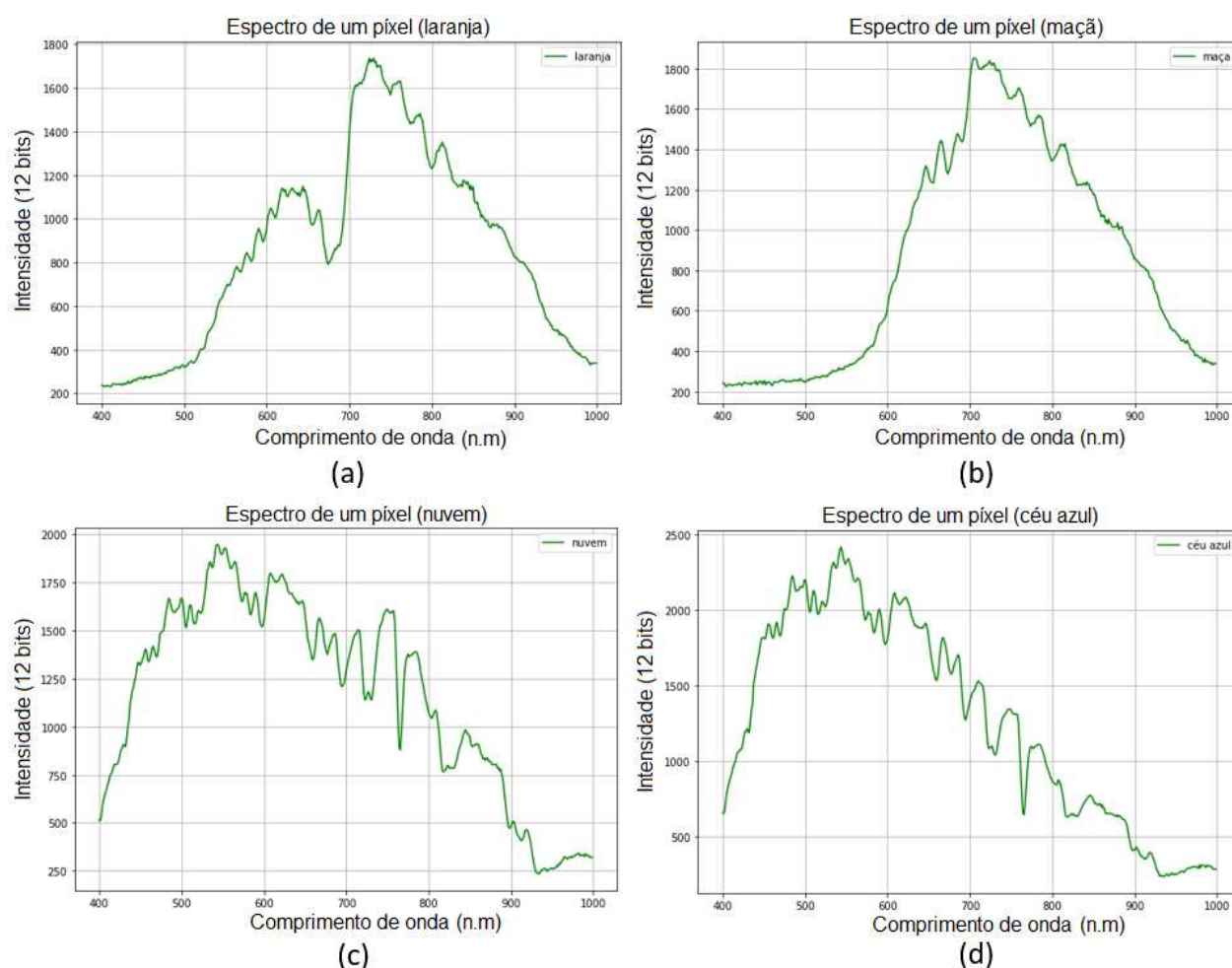


Figura 4.26: Espectros de píxeis em diversas imagens apresentadas na Figura 4.26. Em (a) o espectro do píxel “a” (laranja). Em (b), o espectro do píxel “b” (maçã). Essas imagens foram iluminadas com lâmpada halógena. Em (c), o espectro do píxel “c” (nuvem) e em (d) o píxel do céu azul.

A Figura 4.27 apresenta imagens obtidas com a câmera Specim FX-10 e renderizadas combinando o conteúdo espectral das bandas centradas em 450 nm, 550 nm e 650 nm, referentes a uma amostra de sal, iluminado por lâmpada halógena. As bandas possuem largura de 1,33 nm. Em (a), imagem original RGB renderizada. Em (b), imagem monocromática no comprimento de onda de 769,64 nm. Esse comprimento de onda está relacionado à emissão por decaimento de elétrons em átomos de potássio na região da chama de incêndios florestais. Em (c), imagem renderizada após ser processada com subtração de corrente de escuro. Em (d), imagem renderizada após ser processada por calibração radiométrica utilizando a curva de distribuição espectral da lâmpada halógena apresentada na Figura 4.28 (a). Em (e), uma imagem renderizada após ser processada por subtração da corrente de escuro e calibração radiométrica. A calibração radiométrica e compensação da corrente de escuro foram feitas conforme descrito na Seção 2.7.

Ainda não se percebe, visualmente, o efeito da compensação e calibração nas imagens. Entretanto, o efeito da compensação altera significativamente a distribuição de potência espectral do píxel, conforme pode ser visto na Figura 4.28 (c).



Figura 4.27: Imagem obtida com iluminação por lâmpada halógena. Em (a), imagem renderizada combinando o conteúdo espectral das bandas 450 nm, 550 nm e 650 nm. Em (b), imagem no comprimento de onda de 769,64 nm. Em (c), imagem renderizada após ser processada com subtração de corrente de escuro. Em (d), imagem renderizada após ser processada por calibração radiométrica. Em (e), imagem renderizada após ser processada por subtração da corrente de escuro e calibração radiométrica

A Figura 4.28 (a) mostra a distribuição de potência por comprimento de onda da lâmpada halógena utilizada para iluminação em ambiente de laboratório. O espectro foi obtido do valor médio de um conjunto de píxeis de uma referência de branco, composta por uma superfície coberta com sulfato de bário, por meio da câmera Specim FX-10. Essa fonte foi utilizada para captura da imagem apresentada na Figura 4.27 (Sal grosso) e das imagens da Figura 4.25 (ii), conforme dito. Na Figura 4.28 (b) é apresentado o espectro do píxel “e” da Figura 4.25, sendo a

curva em vermelho correspondente ao espectro sem compensação da corrente de escuro e sem calibração radiométrica. A curva em azul apresenta a compensação da corrente de escuro, conforme descrito na Seção 2.7. Observa-se o deslocamento para baixo da curva azul. As flutuações apresentadas em (a) correspondem às aquelas presentes em (b), na mesma figura, ou seja, são devidas à fonte de luz. Em (c), o resultado da aplicação da calibração radiométrica e compensação da corrente de escuro, segundo a curva de distribuição espectral da fonte, e conforme descrito na Seção 2.7 e modelado pela Equação 2.8. Observa-se uma mudança significativa no perfil espectral, após compensação da variação de amplitude, por comprimento de onda, da fonte de radiação eletromagnética. Isso atesta a importância da calibração para fins de espectrometria e espectroscopia.

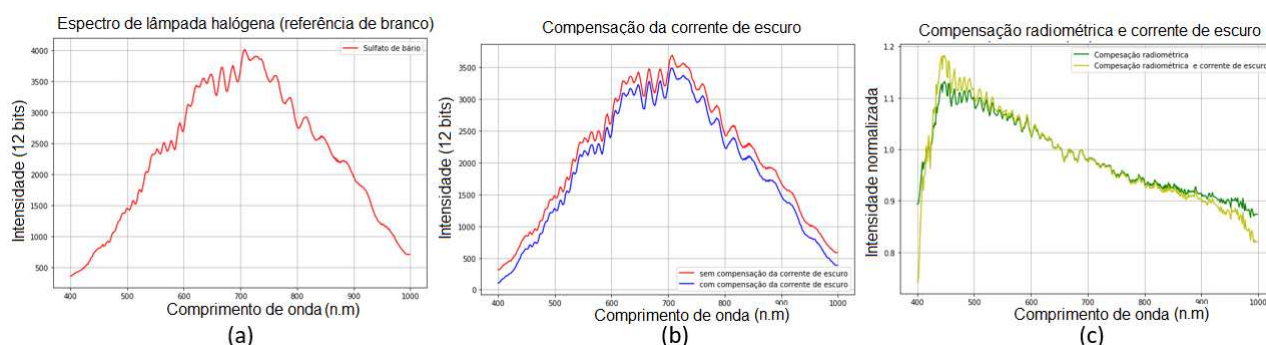


Figura 4.28: Em (a) é apresentada a distribuição de potência por comprimento de onda de uma lâmpada halógena, utilizada para iluminação do objeto imageado, cujo espectro de um píxel é mostrado em (b) e (c). Em (b), o espectro do píxel “e” da Figura 4.25. A curva em vermelho correspondente ao espectro sem compensação da corrente de escuro, e sem calibração radiométrica. A curva em azul apresenta a compensação da corrente de escuro. Em (c), o resultado da aplicação da calibração radiométrica e compensação da corrente de escuro.

A Figura 4.29 mostra o espectro de um píxel da região de fogo, correspondente ao ponto “f” da Figura 4.25. Observa-se um forte sinal no comprimento de onda de 769,64 nm, associado ao decaimento de elétrons termicamente excitados em átomos de potássio, conforme descrito na Seção 2.10.1 e modelado pela Equação 2.16 e pela Equação 2.17.

4.4 Análise de espectros com uso da câmera Specim IQ

A câmera Specim IQ possui um sensor hiperespectral que abrange 204 bandas espectrais na faixa de comprimentos de onda de 400 nm a 1000 nm, tendo cada banda uma largura de



Figura 4.29: Espectro de um píxel na região da chama de uma imagem capturada com a câmera Specim FX-10. O pico de emissão em átomos de potássio destaca-se na figura, em 769,64 nm.

2,94 nm. Esta ampla faixa cobre tanto o espectro visível, observável pelo olho humano, quanto a região do infravermelho próximo. Além disso, a Specim IQ possui um sensor RGB. Ambos sensores produzem imagens com uma resolução espacial de 512 x 512 píxeis. A Figura 4.30 apresenta uma coleção de 204 imagens que formam um “cubo hiperespectral” obtido com a câmera Specim IQ. Cada imagem do cubo corresponde à radiação eletromagnética capturada em uma estreita faixa espectral.

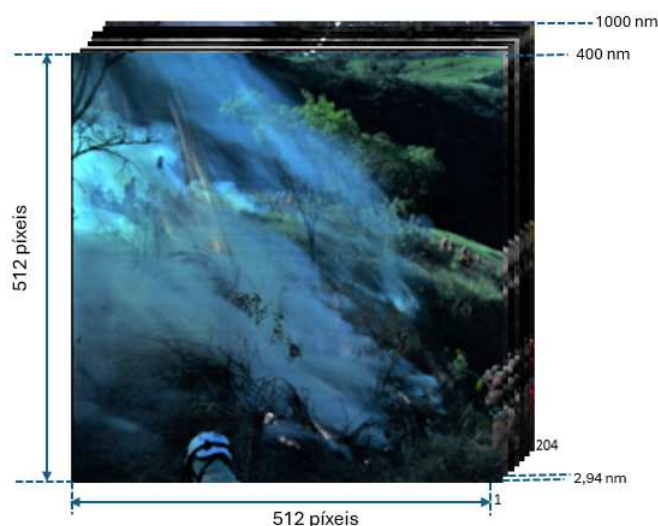


Figura 4.30: “Cubo hiperespectral” Specim IQ consistindo em 204 imagens, cada uma correspondendo à energia eletromagnética em uma banda espectral estreita, com largura de 2,94 nm, cobrindo a faixa de 400 nm a 1000 nm. As imagens possuem o formato de 512 x 512 píxeis.

Um conjunto de dados de imagens hiperespectrais foi construído para o treinamento dos

algoritmos projetados para sintetizar imagens hiperespectrais a partir de imagens RGB com o uso dessa câmera. A compilação de imagens abrangeu cenas diversas, incluindo casos de incêndios florestais, que são o principal foco dessa pesquisa. Uma amostra da base de dados é apresentada na Figura 4.31. A base é composta por um total de 224 imagens.



Figura 4.31: Exemplo de conjunto de imagens construído com o uso do sensor RGB da câmera hiperespectral Specim IQ. O conjunto de dados também contém as imagens correspondentes captadas com o sensor hiperespectral dessa câmera, cada uma com 204 canais espectrais.

A Figura 4.32 apresenta uma comparação entre uma imagem em formato RGB de uma vela e imagens em tons de cinza, capturadas em vários comprimentos de onda, usando o sensor hiperespectral da câmera Specim IQ. Na figura, à medida que se desloca para comprimentos de onda maiores, apenas a chama permanece discernível. A Figura 4.32 (a) mostra, em sua parte superior, uma imagem RGB capturada com uso do sensor RGB da câmera Specim IQ. A figura exibe, na parte inferior, uma imagem renderizada combinando os canais em 450 nm, 550 nm e 650 nm. A Figura 4.32 (b) exibe uma série de imagens em tons de cinza. Vale ressaltar que esta imagem foi obtida com o menor tempo de integração da câmera, para evitar saturação. Devido às condições de iluminação restritas, a imagem em formato RGB apresenta um viés perceptível em direção ao azul. A Figura 4.33 retrata uma chama de uma vela, mostrando píxeis que

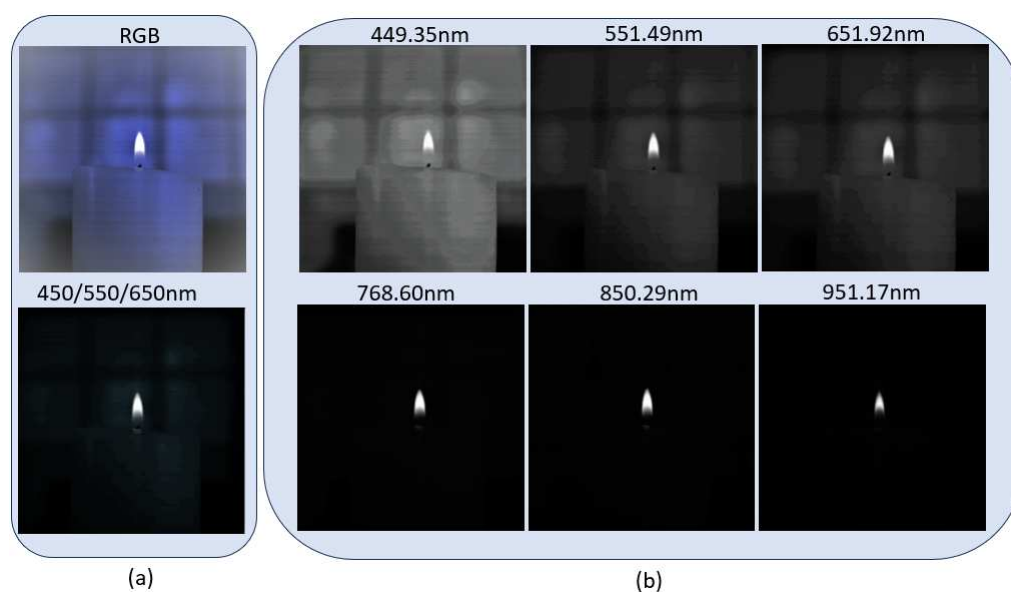


Figura 4.32: Em (a), imagens de uma vela na representação RGB (parte superior) e renderizada (parte inferior) combinando canais em 450 nm, 550 nm e 650 nm. Em (b), imagens em tons de cinza, em diferentes comprimentos de onda.

representam diferentes pontos dentro e fora da chama. As curvas espectrais correspondentes a esses píxeis são apresentadas na Figura 4.34. Na ilustração, a região “a” é caracterizada por uma temperatura relativamente menor (indicada por menor intensidade do píxel), em comparação com as regiões “b” e “c”, que exibem temperaturas mais altas (com intensidade de píxel correspondentemente mais elevada). A região “d”, posicionada fora da chama, exibe intensidade reduzida, indicativa de sua baixa temperatura. Note-se que, nesse caso, não há o pico de emissão associado ao potássio, por não ser queima de material vegetal.

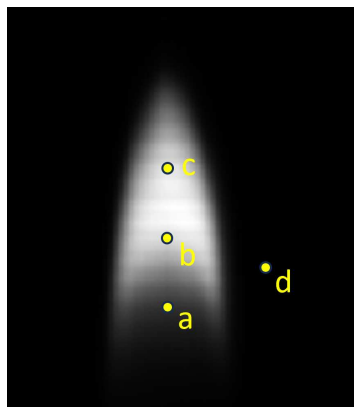


Figura 4.33: A imagem retrata a chama de uma vela com píxeis designados, indicando pontos com distintas características em regiões dentro e fora da chama.

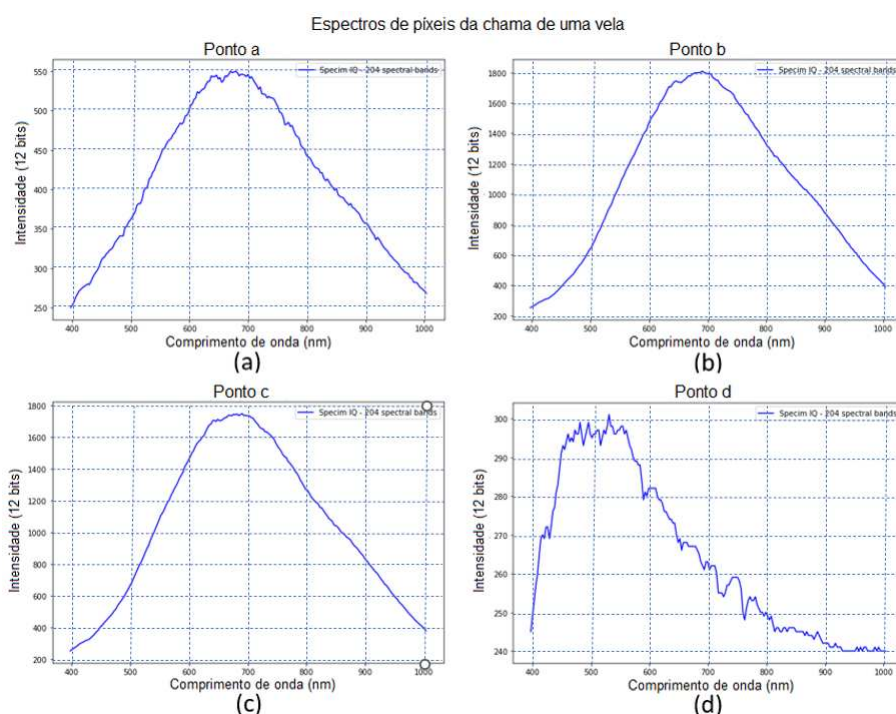


Figura 4.34: Espectro de píxeis de diferentes regiões da chama. A região “a” possui menor intensidade (temperatura), se comparada às regiões de maior intensidade (temperatura mais alta), “b” e “c”. A região “d”, situada fora da chama, exibe intensidade reduzida, condizente com sua temperatura mais baixa, e exibindo uma distribuição espectral diferente da região da chama da vela.

A Figura 4.35 apresenta assinaturas espectrais de vários elementos na cena capturada em uma queima prescrita (MIF), incluindo fogo, grama, fumaça, traje de bombeiro e solo vermelho. A imagem foi normalizada por meio da divisão do tensor da imagem por 4095 (valor máximo de 12bits do formato das imagens da câmera Specim IQ). O espectro de reflectância do traje de bombeiro exibe um aumento de intensidade relativa na faixa acima de 600 nm. A intensidade das emissões de fogo é especialmente proeminente em torno de 769 nm, devido à emissão por

decaimento de elétrons em átomos de potássio, na região da chama.

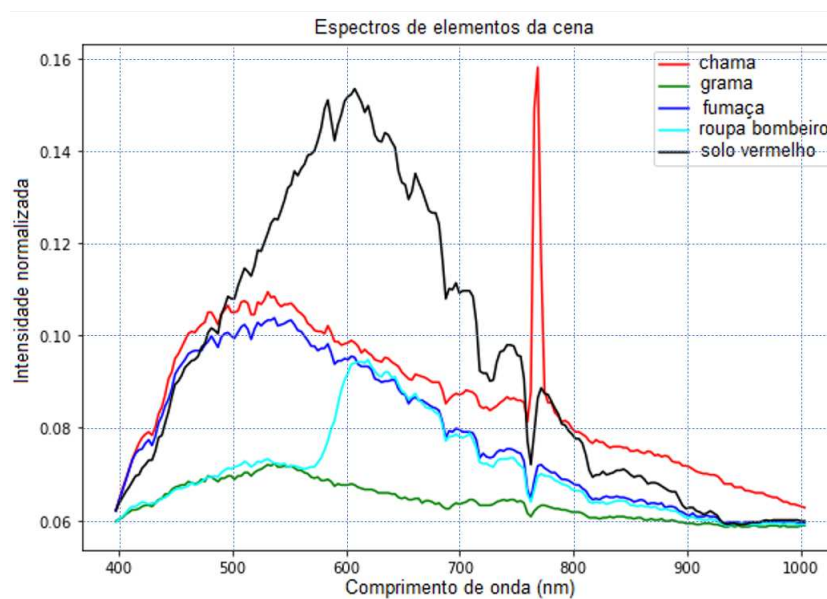


Figura 4.35: Assinatura espectral de emissão de fogo, grama, fumaça, traje de bombeiro e solo vermelho. A intensidade da emissão de fogo é proeminente na faixa de luz visível, mas particularmente intensa na banda 769 nm.

A Figura 4.36 (a) exibe uma cena de incêndio florestal captada pelo sensor RGB da câmera Specim IQ, e em (b) imagens em tons de cinza capturadas em vários comprimentos de onda. A Figura 4.36 (a) também apresenta uma imagem renderizada sintetizada combinando canais em 450 nm, 550 nm e 650 nm. Observe que a chama permanece indistinguível dentro da faixa do espectro visível. No entanto, a chama torna-se perceptível a partir 768,60 nm, apesar do obstáculo causado pela fumaça. Na banda 951,17 nm, somente a chama é discernível na cena, uma imagem com alto contraste e de fácil segmentação.

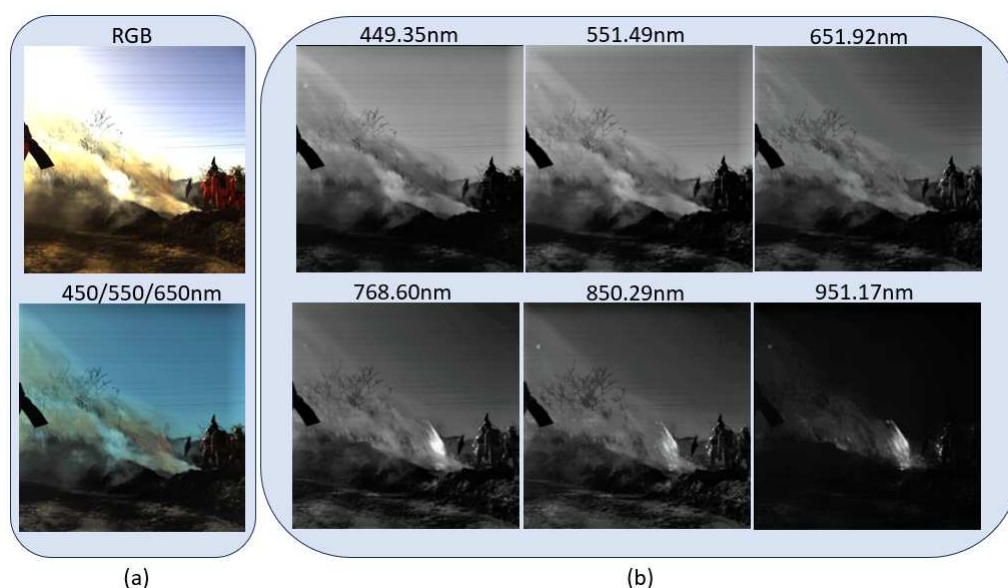


Figura 4.36: Na parte superior de (a) uma imagem captada pelo sensor da câmera Specim IQ. Na parte inferior de (a) uma imagem renderizada. Em (b) imagens em tons de cinza captadas pelo sensor hiperspectral da câmera Specim IQ. As imagens são relativas a uma cena de Manejo Integrado do Fogo (MIF). A chama é imperceptível na da faixa visível do espectro. Na banda de 768,60 nm, no entanto, chama torna-se discernível, apesar da obstrução causada pela fumaça.

Isso indica que câmeras com alta resolução espectral tem o potencial de serem utilizadas em sistemas de detecção de incêndios florestais, possibilitando a redução ou até mesmo a eliminação de erros de detecção presentes em sistemas que utilizam somente detecção com base em imagens RGB.

A transparência à fumaça pouco densa também possibilita a melhor localização do fogo e melhor diagnóstico situacional no que se refere ao planejamento das orientações e ações de combate, assim como na predição do curso dos incêndios.

Na Figura 4.37, um conjunto de pontos identificam regiões da cena, cujos espectros (dos pixels) são apresentados na Figura 4.38. O pixel “a” está na região da chama, o pixel “b”

corresponde à área com fumaça, o píxel “c” está na região azul do céu, e o píxel “d” está situado no traje de bombeiro. O espectro do píxel “a” (região da chama) apresenta pico de

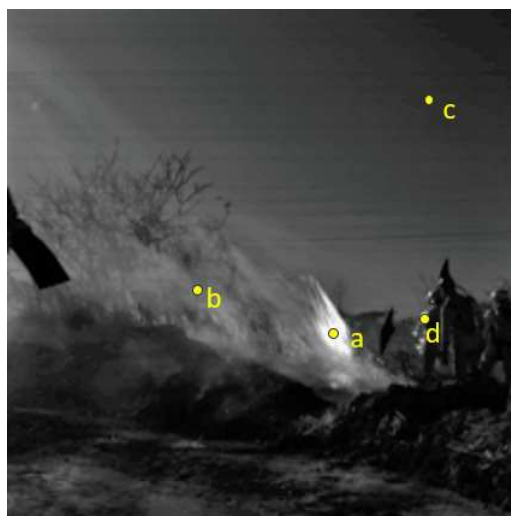


Figura 4.37: Imagem de incêndio florestal e alguns píxeis selecionados de pontos distintos da imagem. Píxel “a” está na região do fogo, píxel “b” na área com fumaça, o píxel “c” está na região azul do céu, e o píxel “d” está situado no traje de bombeiro.

emissão em torno de 769 nm. Esse efeito corresponde à emissão resultante do decaimento de elétrons excitados nos átomos de potássio presentes na região da chama. A fumaça pouco densa do incêndio florestal exhibe transparência nesse comprimento de onda específico.

Neste cenário de iluminação pela luz do sol, a fumaça parece atuar como uma superfície refletora difusa (*lambertiana*). O espectro do céu azul revela uma distribuição de intensidade mais alta nos comprimentos de onda mais curtos, o que está conforme as expectativas. Note-se que os píxeis “c” e “d” exibem baixa intensidade em comparação com as outras regiões amostradas.

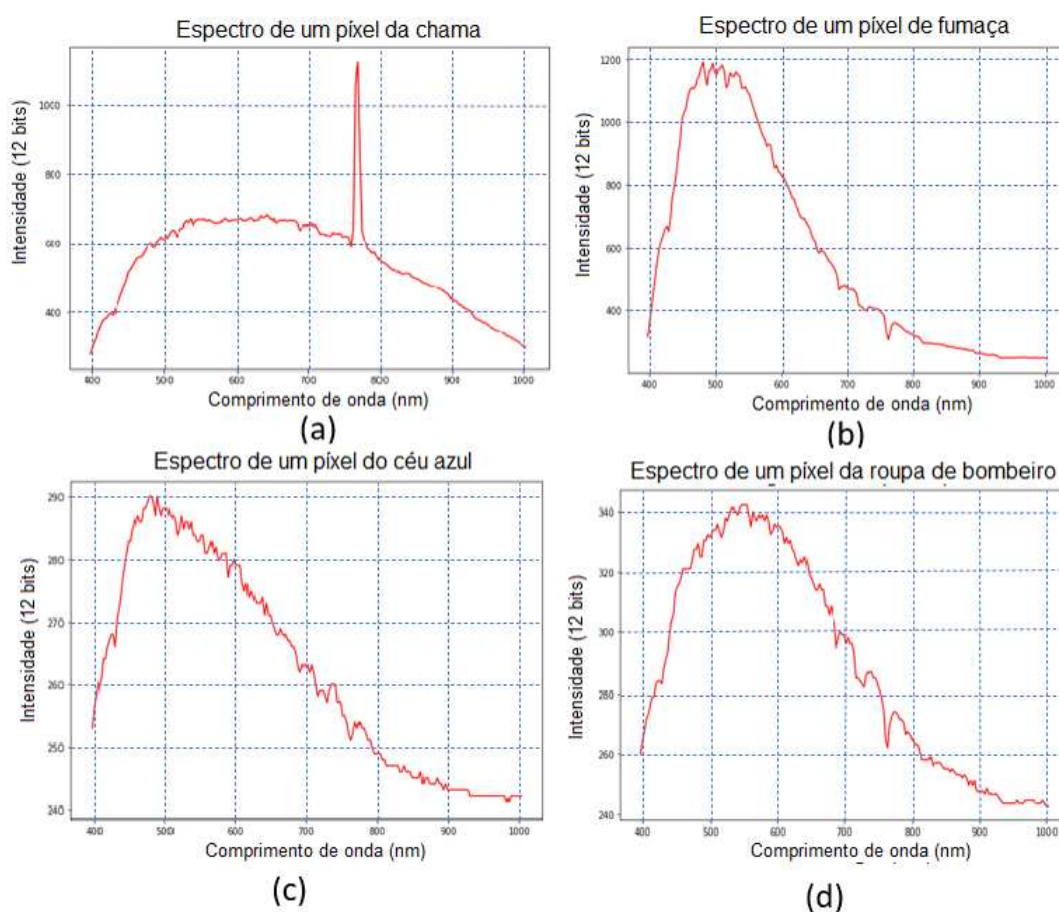


Figura 4.38: Espectro de píxeis selecionados de áreas distintas da cena. A intensidade é apresentada em uma escala de 12 bits (valor máximo é 4095).

4.5 Síntese de imagens hiperespectrais com projeção no espaço RGB através de um filtro *Bayer*

Para simplificar o processo de treinamento a partir de imagens hiperespectrais produzidas pela câmera Specim IQ, foi construído um conjunto de imagens RGB utilizando um filtro *Bayer* para projeção das imagens hiperespectrais, cobrindo o espectro de comprimentos de onda de 400 nm a 1000 nm. Esse filtro é representado como uma matriz de coeficientes, alinhando-se com a resposta do filtro para cada banda espectral da câmera Specim IQ. Esse filtro mapeia informações do espaço de dados hiperespectral, multidimensional, para o espaço RGB, tridimensional, correspondente a câmeras convencionais. A Figura 4.39 apresenta a resposta do filtro Bayer RGB utilizado para projeção das imagens hiperespectrais no espaço RGB. Note-se que a resposta se estende à faixa do infravermelho próximo.

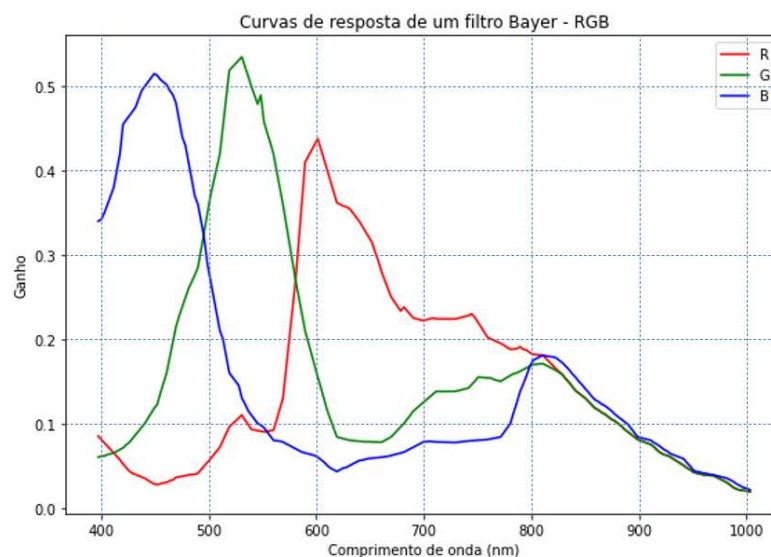


Figura 4.39: Curvas de resposta espectral para um filtro *Bayer* RGB, na faixa de comprimentos de onda de 400 nm a 1000 nm (Flasseur et al., 2017).

A Figura 4.40 (i) apresenta, em sua parte superior, uma imagem captada com o sensor RGB da câmera Specim IQ, acompanhada de suas contrapartes captadas pelo sensor hiperespectral dessa câmera e mostradas em escala de cinza, para cada canal, em comprimentos de onda específicos (ii). Uma imagem renderizada a partir das bandas centradas em 450 nm, 550 nm e 650 nm também é mostrada na parte inferior de (i). Essa figura serve de base para a comparação dos métodos de síntese de imagens hiperespectrais utilizados nesse trabalho, como detalhado nas seções subsequentes. Os píxeis marcados na imagem RGB (pontos “a”, “b”, “c” e “d”) terão seus espectros estimados e comparados com o espectro de referência obtido através do sensor hiperespectral da câmera Specim IQ (considerado o valor de referência - *ground truth*). Ao observar a figura, torna-se evidente que a chama é significativamente menos discernível nos comprimentos de onda de 449,35nm (na região espectral azul) e 551,49 nm (na região espectral verde). Por outro lado, é observado um aumento gradual na transparência da fumaça, à medida que o comprimento de onda se desloca em direção à região do infravermelho próximo.

4.5.1 Síntese de imagens hiperespectrais pelo método *Ridge Regression* a partir de imagens RGB obtidas por projeção usando filtro *Bayer*

Esta seção descreve a síntese de imagens hiperespectrais pelo método *Ridge* de regressão linear, a partir de imagens RGB da base de dados construída pela projeção, no espaço RGB, de

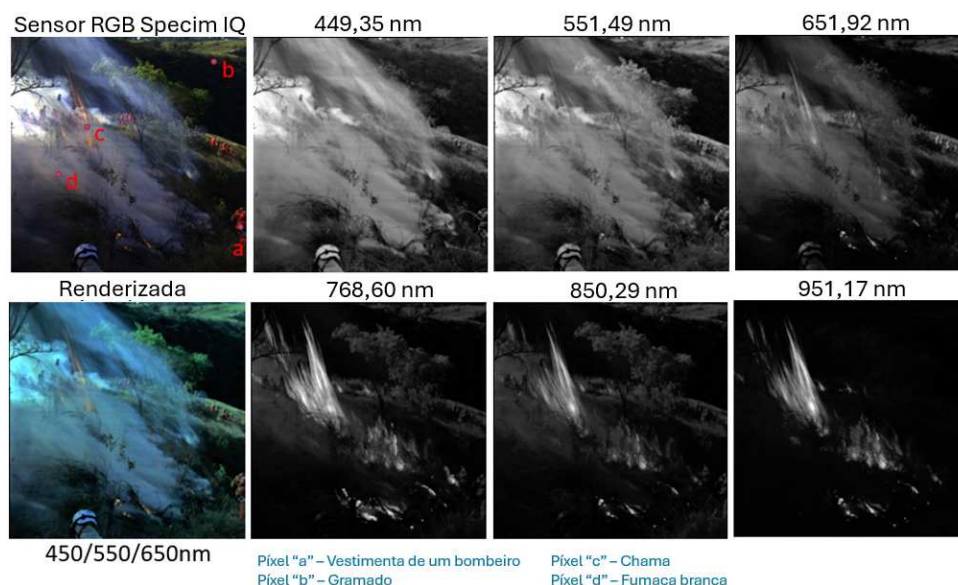


Figura 4.40: Imagens hiperespectrais capturadas utilizando uma câmera Specim IQ. Na parte superior de (i), uma imagem capturada pelo sensor RGB da câmera Specim IQ. Os pontos na figura indicam os píxeis de diversos elementos da cena, cujos espectros serão sintetizados utilizando os métodos descritos nas seções seguintes. Em (ii), imagens capturadas pelo sensor hiperespectral da câmera Specim IQ, representadas em escala de cinza, em diversos comprimentos de onda. Incluída na parte inferior de (i) uma imagem renderizada composta pela combinação das bandas espectrais centradas em 450 nm, 550 nm e 650 nm. Cada banda possui a largura de 2,94 nm.

imagens hiperespectrais capturadas com uso da câmera Specim IQ, utilizando uma matriz de projeção construída com funções de resposta espectral de um filtro *Bayer*. O processo é ilustrado na Figura 4.41. Os blocos na cor laranja correspondem à etapa de treinamento e os de cor azul à fase de inferência (testes). Foram utilizadas 180 imagens para treinamento, escolhidas de forma aleatória do conjunto de 224 imagens da base. As imagens restantes foram utilizadas para avaliação da qualidade do modelo na etapa de testes. Diferentemente do processo descrito na Seção 4.2.1, que produziu imagens hiperespectrais sintetizadas com 31 bandas, restrita à faixa de luz visível, com comprimentos de onda entre 400 nm e 700 nm, o processo da Figura 4.41, assim como os demais descritos nas seções subsequentes, sintetiza imagens com 204 bandas espectrais, na faixa de 400 nm a 1000 nm, abrangendo a faixa de luz visível ao olho humano (VIS) e também a do infravermelho próximo (NIR).

Na parte superior da Figura 4.42 (a), tem-se uma imagem do conjunto de testes capturada pelo sensor RGB da câmera Specim IQ. Na parte inferior da figura, uma imagem sintetizada por meio de um modelo *Ridge* e renderizada com base nos conteúdos presentes nas bandas 450 nm (B), 550 nm (G) e 650 nm (R). A Figura 4.42 (b) mostra imagens em escala de cinza sinteti-

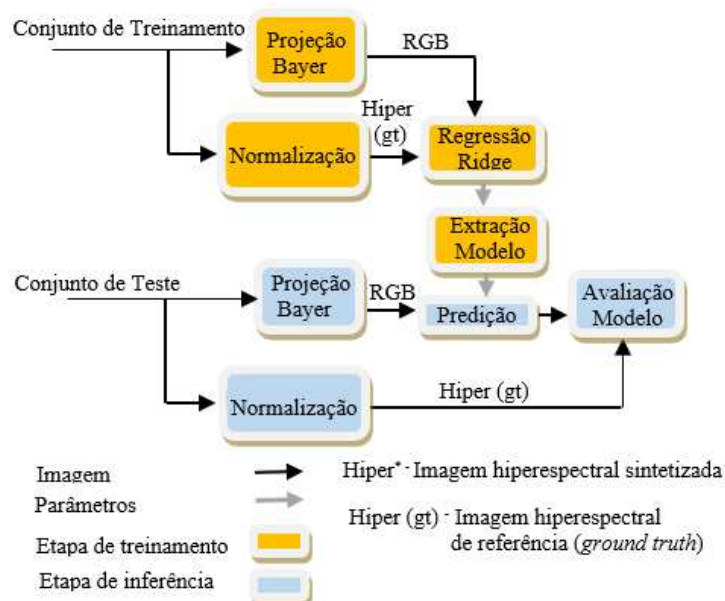


Figura 4.41: Processo de síntese de imagens hiperespectrais utilizando um modelo *Ridge* e um conjunto de imagens obtidas pela projeção de imagens hiperespectrais (*ground truth*) no espaço RGB usando um filtro Bayer, cujas curvas são apresentadas na Figura 4.39.

zadas pelo modelo *Ridge* em comprimentos de onda específicos. Com esse modelo, foi gerada uma coleção de 204 canais de "cor", abrangendo vários comprimentos de onda, na faixa entre 400 nm e 1000 nm, utilizando apenas três valores (R, G, B) de cada píxel. Ao comparar com a Figura 4.40 (*ground truth*), vê-se que essa abordagem foi efetiva na síntese hiperespectral, enquanto preservou a estrutura espacial da imagem. Note-se também que a síntese foi capaz de preservar características de discriminação de conteúdos por comprimento de onda, produzindo a transparência da fumaça, ou realce da chama, na faixa NIR. A Figura 4.43 mostra os espectros dos píxeis situados em diferentes regiões da cena, já designados como pontos "a", "b", "c" e "d" na parte superior da Figura 4.40. Esses pontos correspondem a diferentes elementos da cena, com distribuição espectral distinta. A curva na cor azul corresponde aos dados espectrais captados pela câmera Specim IQ. A curva em vermelho representa os dados espectrais sintetizados. Em particular, a síntese do espectro do píxel "d", referente a uma região com fumaça, apresentou uma excelente concordância entre as curvas, em uma avaliação qualitativa. Nos píxeis "a" e "b", a curva sintetizada acompanhou a silhueta da curva Specim IQ, porém com discrepâncias em algumas faixas espectrais, por exemplo, a faixa entre 700 nm e 800 nm, no píxel "b". Interessante notar que o pico de absorção, devido à excitação da molécula de O_2 , próximo a 760 nm, está presente nos espectros de todos os píxeis. Entretanto, o pico de emissão, em torno de 769 nm, não foi corretamente sintetizado conforme pode ser visto nas curvas

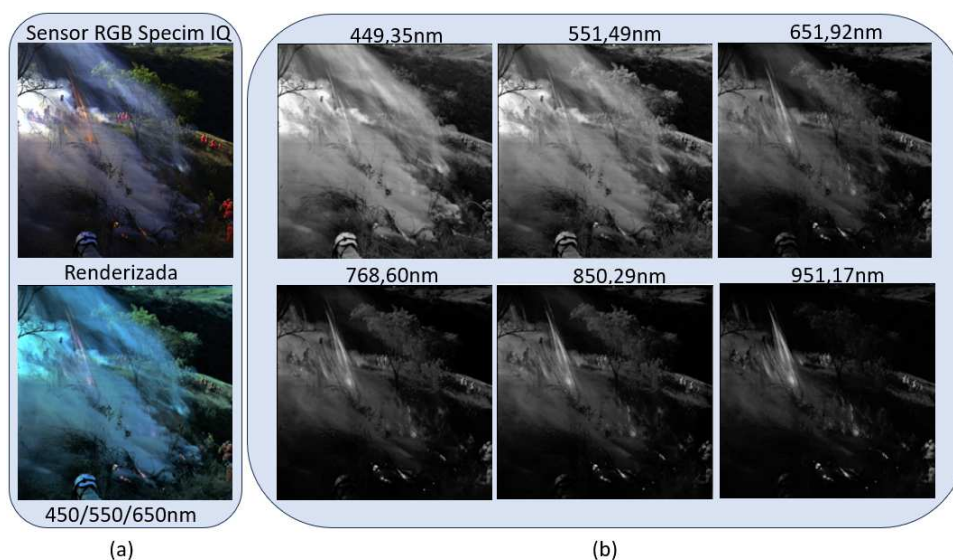


Figura 4.42: Imagens hiperespectrais sintetizadas usando o método de Regressão *Ridge*. Este algoritmo propiciou a geração de 204 canais espectrais, abrangendo centenas de comprimentos de onda. Essa síntese foi alcançada usando apenas os valores R, G e B de píxeis individuais. Na parte superior de (a), uma imagem RGB capturada pelo sensor RGB da câmera Specim IQ. Na parte inferior de (a) uma imagem sintetizada pelo modelo *Ridge*, renderizada a partir dos canais sintetizados em 450 nm, 550 nm e 650 nm. Em (b), imagens sintetizadas em vários comprimentos de onda, em tons de cinza.

do píxel “c”. Isso se deve, possivelmente, à suavização do modelo de síntese usando a projeção Bayer em relação a essa variação específica na curva sintetizada.

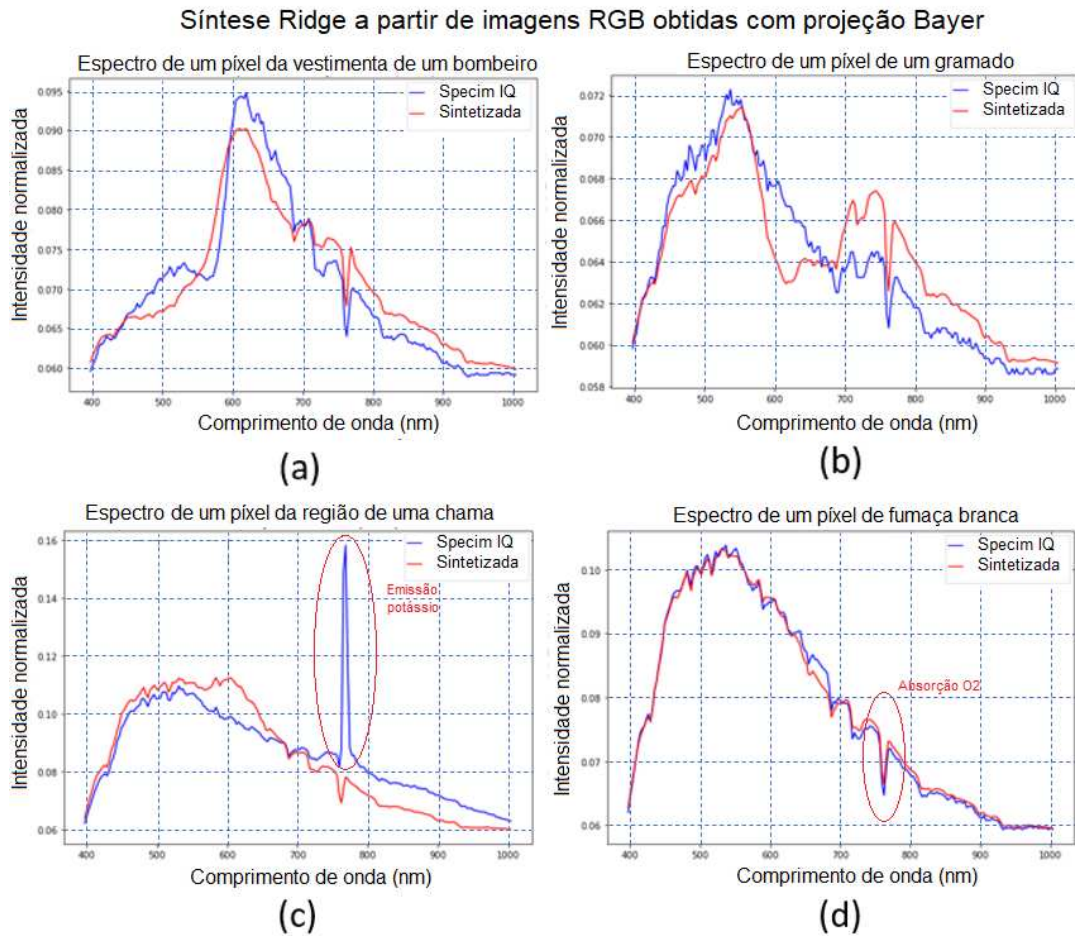


Figura 4.43: Comparação de espectros de pixels selecionados em diferentes regiões da cena, conforme identificado na Figura 4.40. A curva em azul representa o espectro medido usando o sensor de imagem hiperespectral Specim IQ, enquanto a curva na cor vermelha corresponde à intensidade do espectro estimada, usando o método de regressão *Ridge*. Em destaque, pico de absorção de O_2 e pico de emissão de potássio.

4.5.2 Rede Neural para síntese hiperespectral a partir de imagens RGB obtidas por projeção usando um filtro *Bayer*

Um processo utilizando uma rede MLP de doze camadas (*Dense*) foi utilizado para sintetizar imagens hiperespectrais, a partir de imagens RGB da base de dados construída pela projeção no espaço RGB de imagens hiperespectrais capturadas com uso da câmera Specim IQ, por meio de uma matriz de projeção, conforme apresentado na Seção 4.5.1. A matriz de projeção pode ser encontrada no *link* indicado no Apêndice C. A Figura 4.44 ilustra o diagrama desse processo e a Figura 4.45 apresenta a arquitetura MLP utilizada para síntese. A escolha da arquitetura se deu com base nos experimentos descritos na Seção 4.2.3.

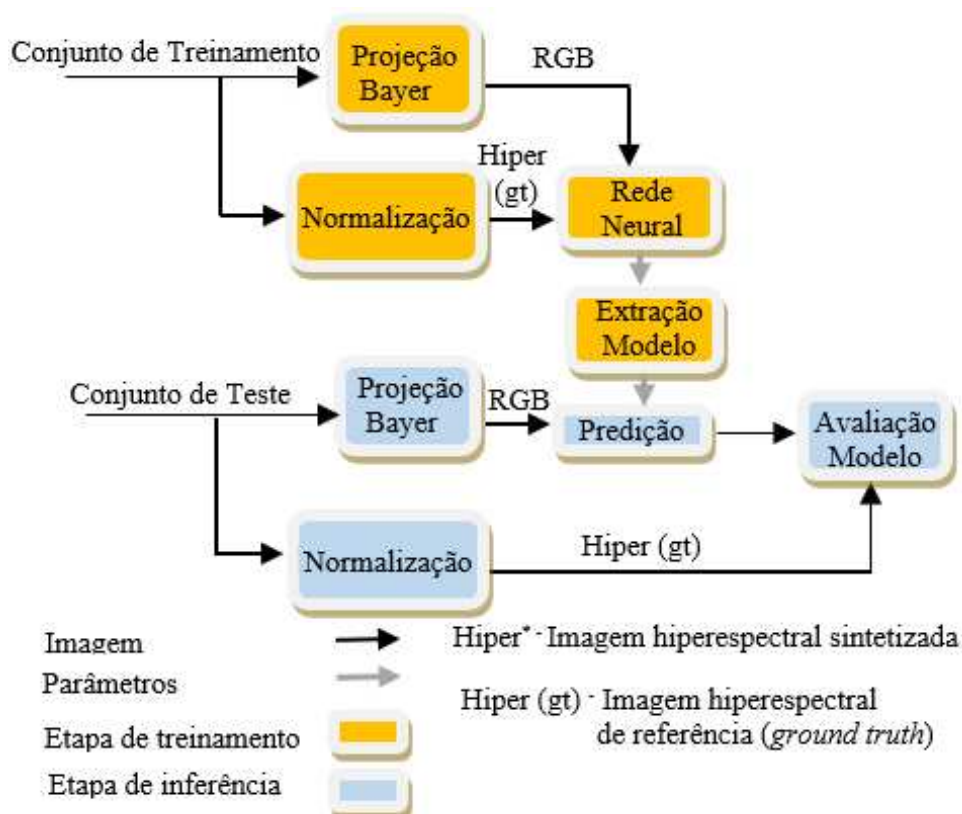


Figura 4.44: Processo de síntese de imagens hiperespectrais utilizando um modelo de rede neural do tipo MLP, e projeção de imagens hiperespectrais no espaço RGB por meio de um filtro *Bayer*.

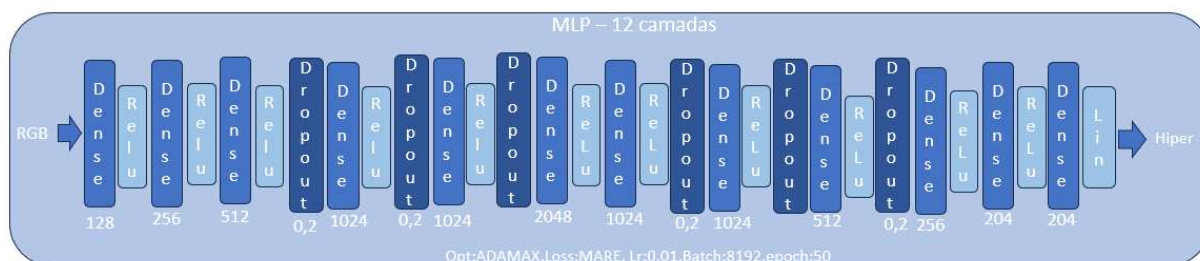


Figura 4.45: Modelo de rede MLP utilizado para síntese de imagens hiperespectrais. O treinamento do modelo se deu com uso de imagens RGB obtidas pela projeção de imagens hiperespectrais no espaço RGB com uso de um filtro *Bayer*.

A Figura 4.46 (a) apresenta uma imagem RGB (parte superior da figura) do conjunto de testes capturada usando o sensor RGB da câmera Specim IQ e uma imagem sintetizada (parte inferior da figura), gerada usando a rede MLP da Figura 4.45. Essa rede neural produziu um “cubo hiperespectral” com 204 bandas. A intensidade de cada banda, em cada píxel, foi determinada a partir apenas dos valores R, G e B de cada píxel da imagem RGB. A Figura 4.46 (b) ilustra uma amostra de imagens sintetizadas, em escala de cinza, correspondentes a alguns comprimentos de onda. Esse processo apresenta um resultado superior em comparação com o

processo utilizando regressão *Ridge*, em especial o realce da chama e a melhora na nitidez e contraste. Na Seção 4.5.4 é feita uma análise comparativa, quantitativa, dos processos, usando as métricas MARE e RMSE.

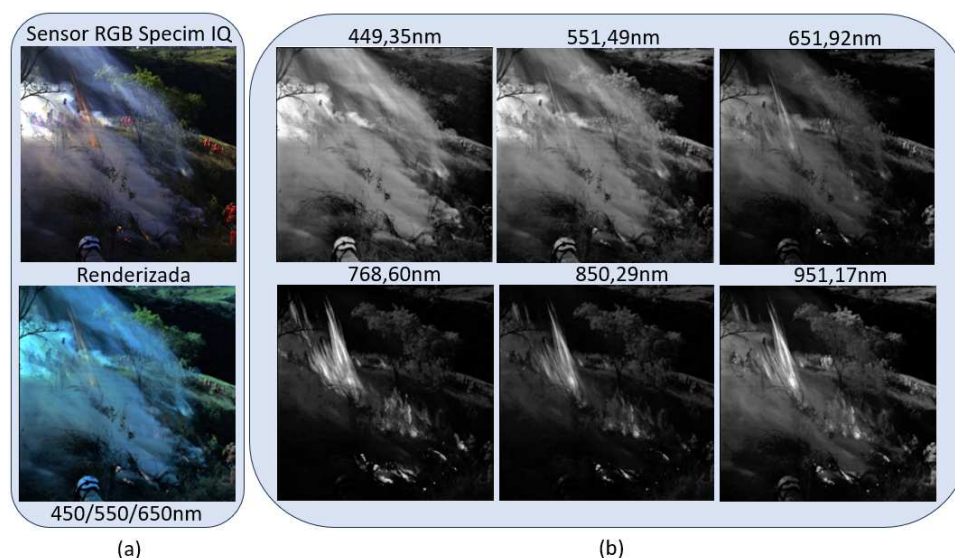


Figura 4.46: Imagem hiperespectral sintetizada usando uma rede neural do tipo MLP com doze camadas. Na parte superior de (a), imagem *ground truth*, na parte inferior, uma imagem gerada combinando canais em 450 nm, 550 nm e 650 nm. A figura (b) apresenta uma seleção de imagens em escala de cinza, extraídas do cubo hiperespectral. Cada imagem corresponde a um comprimento de onda distinto.

Foi realizada uma comparação de espectros de uma imagem sintetizada e a correspondente imagem (*ground truth*) do conjunto de testes capturada pelo sensor hiperespectral da câmera Specim IQ. Essa comparação envolveu píxeis localizados em regiões distintas da cena, especificamente marcados como pontos “a”, “b”, “c”, e “d”, na Figura 4.40. Ao realizar uma comparação qualitativa entre os resultados obtidos usando o método *Ridge* e a rede neural de múltiplas camadas (MLP), nota-se, ao observar a curva do píxel “b” (região do gramado), no intervalo de 700 nm a 900 nm, que a MLP demonstra uma capacidade superior de seguir as mudanças na silhueta da curva espectral. Embora possa haver ligeiras variações na intensidade, a MLP mantém uma silhueta que se assemelha à referência (consulte a Figura 4.47). Por outro lado, o processo *Ridge* tem melhor comportamento, em comparação com a MLP, para o píxel “a” (vestimenta do bombeiro). Esse processo também não foi eficaz em estimar o pico de emissão em átomos de potássio termicamente excitados (píxel “c” - região da chama).

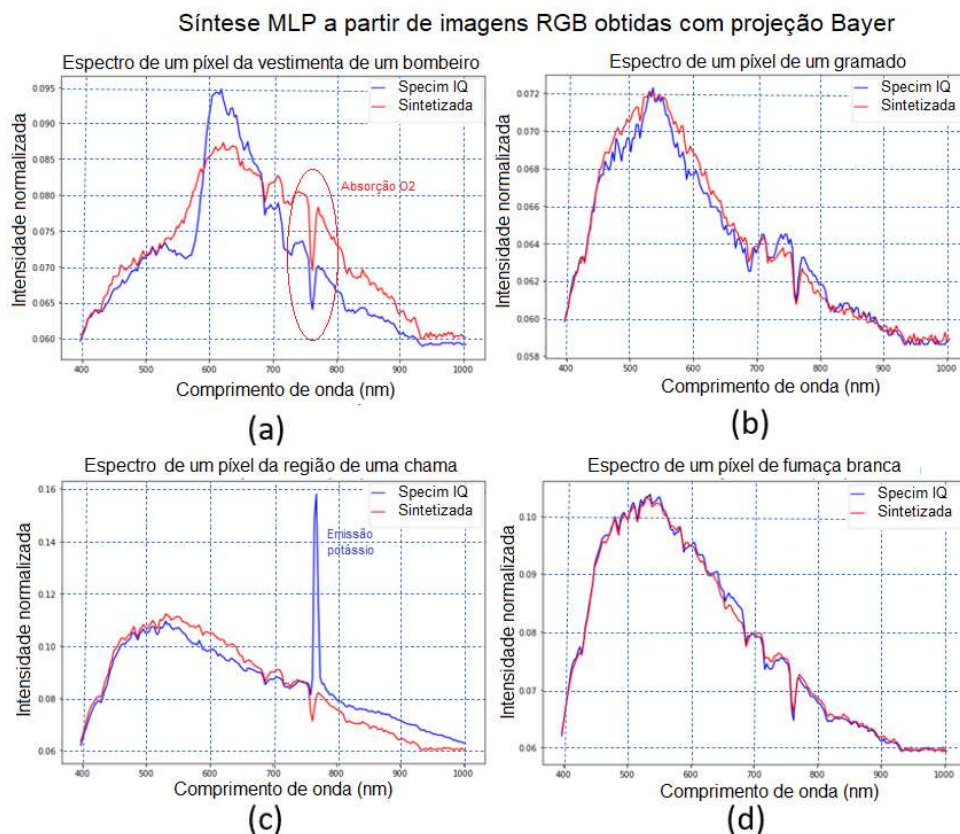


Figura 4.47: A figura fornece uma comparação de espectros obtidos a partir de píxeis situados em diferentes regiões da cena. A curva azul ilustra o espectro medido (*ground truth*) usando o sensor de imagem hiperespectral Specim IQ, enquanto a curva vermelha representa a intensidade estimada pela MLP ao longo da faixa de comprimento de onda de 400 nm a 1000 nm.

4.5.3 Agrupamento e MLP para síntese hiperespectral a partir de imagens projetadas usando filtro *Bayer*

De forma semelhante ao apresentado na Seção 4.2.4, utilizou-se um processo de síntese de imagens hiperespectrais pela combinação dos algoritmos *minibatch K-Means* e MLP. O treinamento dos algoritmos foi feito com imagens hiperespectrais da base construída nesse trabalho, projetadas no espaço RGB por meio de uma matriz com função de resposta espectral segundo o filtro *Bayer*. Foi definido o número de 3000 agrupamentos e a mesma arquitetura MLP descrita na subseção 4.5.2. A Figura 4.48 mostra o diagrama em blocos do processo. O modelo de rede neural é definido na etapa de treinamento e utilizado para estimar o conteúdo espectral dos píxeis centroides de agrupamentos RGB. Os agrupamentos RGB são estabelecidos por meio do algoritmo *minibatch K-Means*. As etapas são as mesmas descritas na Seção 4.2.4, o que difere é como as imagens RGB foram geradas, nesse caso, utilizando a projeção por filtro *Bayer*.

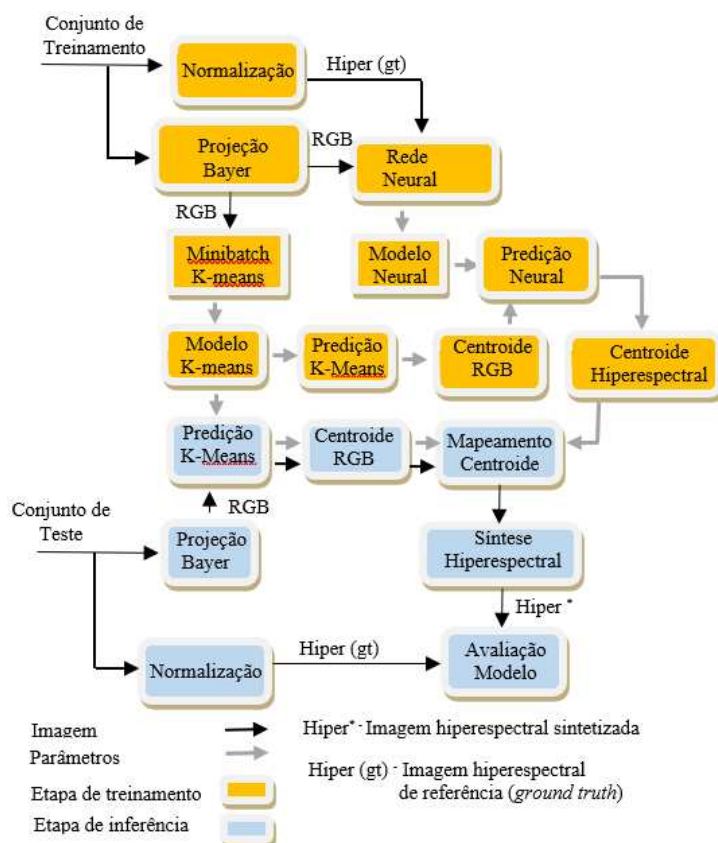


Figura 4.48: Processo de síntese de imagens hiperespectrais utilizando um modelo de rede neural do tipo MLP, combinado com o algoritmo *minibatch K-Means*, aplicado a um conjunto de imagens obtidas por meio da projeção no espaço RGB, através de um filtro *Bayer*.

A Figura 4.49 representa a imagem hiperespectral sintetizada com 204 bandas espectrais, obtida da mesma imagem RGB de teste utilizada nos métodos anteriores. Esse processo também produziu imagens com realce do fogo em relação à fumaça, qualitativamente semelhante ao resultado com a MLP, e sabidamente com menor custo computacional.

Assim como nas técnicas *Ridge* e MLP anteriormente descritas, procedeu-se à comparação das curvas espectrais dos píxeis marcados como pontos “a”, “b”, “c”, e “d” na Figura 4.40. O desempenho assemelha-se ao das metodologias MLP e *Ridge* para os píxeis “c” (chama), e “d” (fumaça), exibindo uma resposta mais suave para os píxeis “a” (vestimenta do bombeiro) e “b” (gramado). O método *Ridge* apresenta uma resposta superior para o píxel “a” em comparação com o método de agrupamento e MLP combinados. Essa observação pode ser feita examinando as curvas de cada processo na faixa de 500 nm a 650 nm, como visto na Figura 4.50.

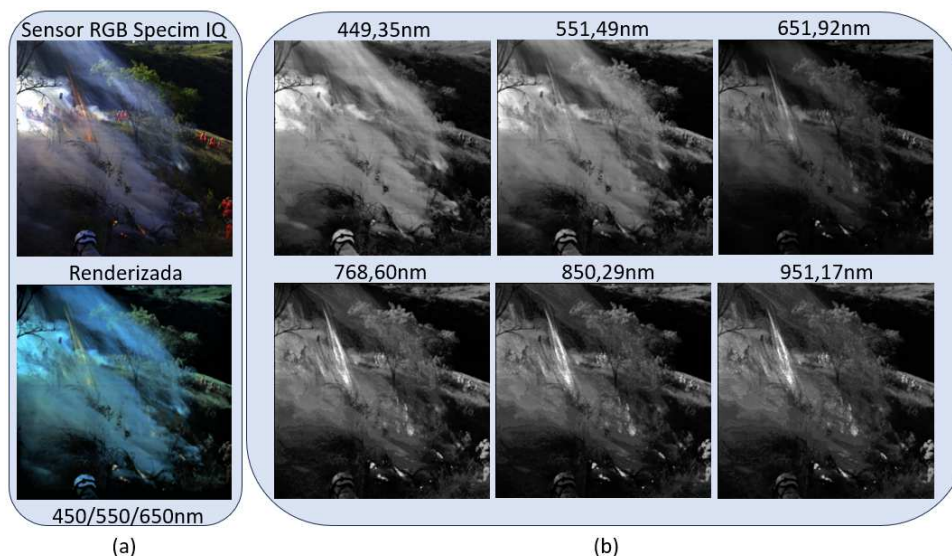


Figura 4.49: Imagem hiperespectral sintetizada usando um método que combina agrupamento usando *minibatch K-Means* e *MLP*. A figura (b) apresenta uma seleção de imagens em escala de cinza, extraídas do cubo hiperespectral. Cada imagem corresponde a um comprimento de onda distinto. Na parte superior de (a), imagem *ground truth*, na parte inferior imagem sintetizada e renderizada combinando canais em 450 nm, 550 nm e 650 nm.

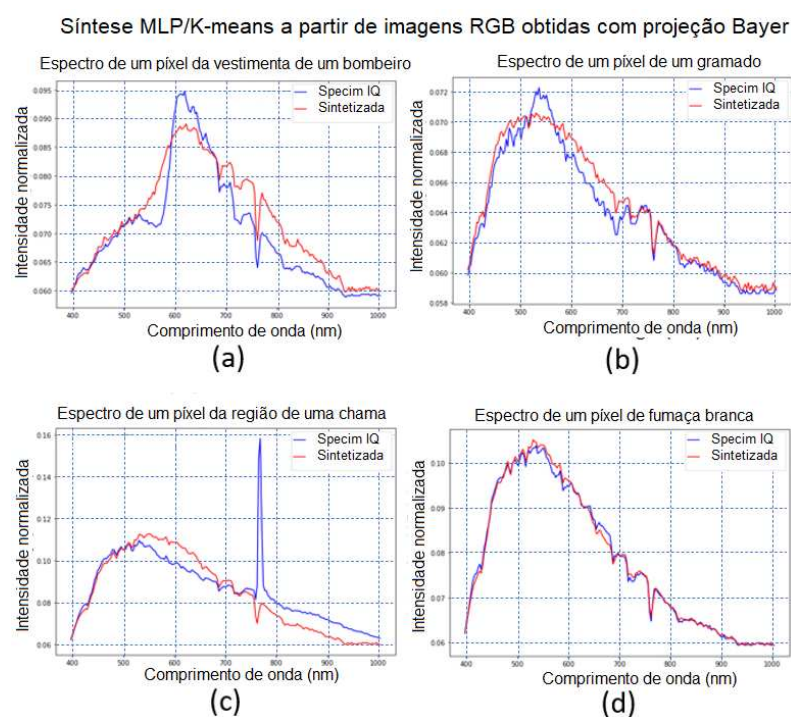


Figura 4.50: A figura fornece uma comparação de espectros obtidos de pixels localizados em várias regiões da cena. A curva azul ilustra o espectro medido usando o sensor de imagem hiperespectral Specim IQ, enquanto a curva vermelha representa a intensidade estimada, usando *minibatch K-Means* e *MLP*.

4.5.4 Análise de desempenho dos modelos

Uma comparação quantitativa entre os três métodos foi feita com as métricas MARE e RMSE para um conjunto de testes contendo 44 imagens da base construída com uso da câmera Specim IQ (veja Apêndice B). A Tabela 4.4 apresenta os resultados para a métrica MARE. O conjunto de testes consiste em imagens hiperespectrais projetadas no espaço RGB, utilizando o mesmo processo descrito na Seção 3.6, onde uma matriz é empregada para emular a resposta de um filtro Bayer. Conforme mostrado na Tabela 4.4, a MLP apresentou o melhor desempenho, sendo 18% melhor que o processo de MLP/KMC e 35% melhor do que o processo RR. Uma vez que o método MLP/KMC é aplicado apenas aos píxeis correspondentes aos centroides do agrupamento, há, nesse método, significativa redução do custo computacional e aumento da velocidade de processamento, em comparação com MLP. O método mais rápido foi o baseado na regressão *Ridge*, entretanto, foi o de menor exatidão.

Tabela 4.4: Análise comparativa dos três processos utilizando a métrica MARE aplicada ao conjunto de testes com 44 imagens.

Processo	RR	MLP	MLP/KMC
Média	0,0260	0,0168	0,0205
Desvio padrão	$6,2 \times 10^{-5}$	$4,3 \times 10^{-5}$	$4,9 \times 10^{-5}$
Mínimo	0,0110	0,0059	0,0104
Máximo	0,0571	0,0401	0,0462

Um teste de comparação de médias foi aplicado para averiguar se há uma distinção significativa da distribuição de valores MARE entre os processos (Veja Seção 3.5). Aplicando o Teste de *Shapiro-Wilk* aos dados de MARE que produziram a Tabela 4.5 e as distribuições mostradas na Figura 4.51, obteve-se

$$pvalue = 1,7 \times 10^{-5}. \quad (4.7)$$

Com esse resultado, rejeita-se a Hipótese nula (critério: $pvalue < 0,05$) para o teste de *Shapiro-Wilk*, de que os dados de MARE dos processos vem de uma distribuição normal. Sendo assim, aplicou-se o Teste de *Kruskal-Wallis* aos dados de MARE obtendo

$$pvalue = 2,3 \times 10^{-8}. \quad (4.8)$$

Esse resultado indica que os dados vem de distribuições diferentes para cada processo. A Fi-

gura 4.51 apresenta a comparação da distribuição da métrica MARE para os três processos de síntese, considerando as 44 imagens do conjunto de testes.

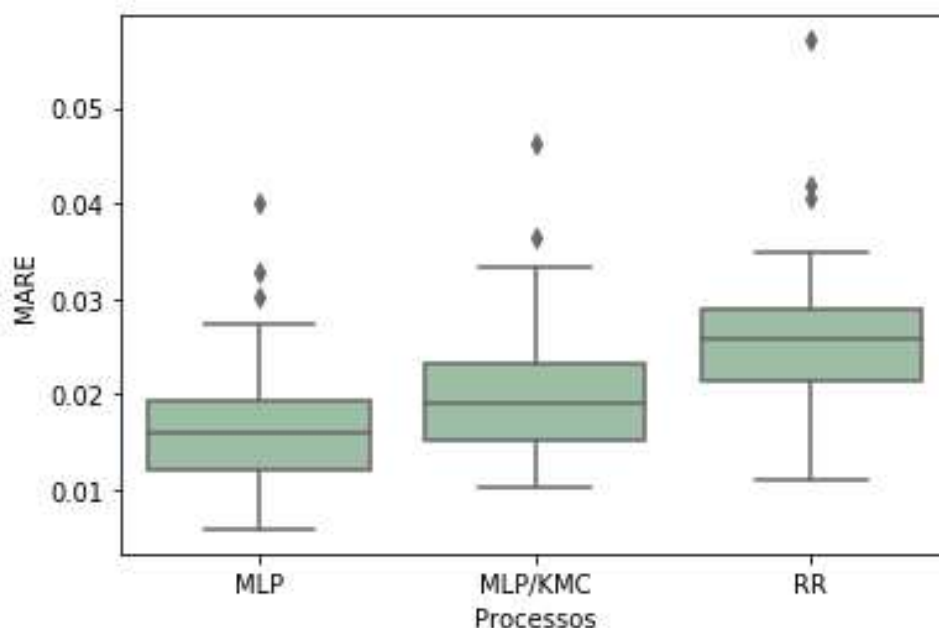


Figura 4.51: Comparação da distribuição da métrica MARE obtida utilizando os quatro processos de síntese hiperespectral.

A Tabela 4.5 apresenta os resultados para a métrica RMSE, na qual se pode notar que o processo MLP apresentou um desempenho 23,5 % melhor que o processo de MLP/KMC e 25 % melhor do que o processo RR.

Tabela 4.5: Análise comparativa dos três processos utilizando a métrica RMSE aplicada ao conjunto de testes com 44 imagens.

Processo	RR	MLP	MLP/KMC
Média	0,0036	0,0027	0,0034
Desvio padrão	$2,5 \times 10^{-6}$	$1,7 \times 10^{-6}$	$2,7 \times 10^{-6}$
Mínimo	0,0015	0,0009	0,0014
Máximo	0,0107	0,0083	0,0113

Um teste de comparação de médias foi aplicado para averiguar se há uma distinção significativa da distribuição de valores RMSE entre os processos. Aplicando o Teste de *Shapiro-Wilk* aos dados de RMSE que produziram a Tabela 4.5 e as distribuições mostradas na Figura 4.52, obteve-se

$$pvalue = 1,6 \times 10^{-10}. \quad (4.9)$$

Com esse resultado, rejeita-se a Hipótese nula (critério: $pvalue < 0,05$) para o teste de *Shapiro-Wilk*, de que os dados de RMSE dos processos vem de uma distribuição normal. Sendo assim, aplicou-se o Teste de *Kruskal-Wallis* aos dados de RMSE obtendo

$$pvalue = 6,8 \times 10^{-3}. \quad (4.10)$$

Esse resultado indica que os dados vem de distribuições diferentes para cada processo. A Figura 4.52 apresenta a comparação da distribuição da métrica RMSE para os três processos de síntese, considerando as 44 imagens do conjunto de testes.

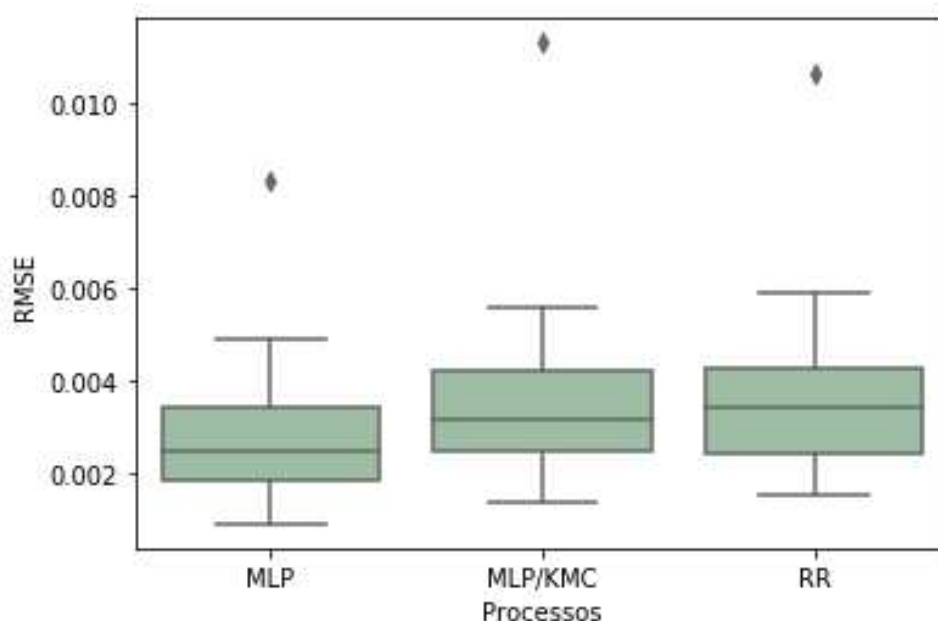


Figura 4.52: Comparação da distribuição da métrica RMSE obtida utilizando os três processos de síntese hiperespectral.

Como pode ser observado nas Figuras 4.43 (c), 4.47 (c) e 4.50 (c), nenhum dos processos logrou a correta recuperação do perfil espectral na região de fogo, no que tange especificamente ao pico de emissão oriundo do decaimento de elétrons em átomos de potássio (766,5 nm e 769,9 nm) (Vodacek et al., 2002). Entretanto, todas as curvas exibiram a absorção atmosférica devido à molécula de oxigênio, que se localiza no comprimento de onda de 760 nm. Uma hipótese é que a suavização efetuada pelos algoritmos na interpolação reduziu a resolução espectral do processo de síntese. A seção seguinte traz um método de síntese baseado em imagens hiperespectrais projetadas no espaço de um sensor hipotético com quatro dimensões, correspondentes às bandas espectrais R, G, B com a adição de uma banda estreita na faixa do

infravermelho próximo (NIR). Isso possibilitou a identificação do pico de emissão de potássio em imagens com presença de fogo tendo vegetação como material combustível.

4.6 Projeção do espaço hiperespectral no espaço RGB-IR por função de transferência sintetizada com quatro componentes

Um experimento foi conduzido no qual foi feita uma síntese de um filtro para projeção das imagens hiperespectrais em um espaço de quatro dimensões, correspondentes às três componentes de um filtro *Bayer* RGB somada a um filtro passa-banda centrado em 769 nm. O objetivo desse experimento foi verificar a viabilidade do uso de um filtro para realçar o espectro de emissão por decaimento de elétrons em átomos de potássio. A Figura 4.53 apresenta a resposta do filtro sintetizado. Esse filtro foi utilizado para emular a resposta de uma câmera RGB com adição de filtro IR de banda estreita. As imagens hiperespectrais capturadas pela câmera Specim foram projetadas em um espaço de quatro dimensões do sensor, utilizando uma matriz para emular a resposta do filtro. Essa matriz está disponível no link indicado no Apêndice C. As imagens resultantes foram utilizadas para treinar uma rede MLP de doze camadas, conforme Figura 4.54.

Na etapa de testes, as imagens do conjunto de testes foram projetadas no mesmo espaço de quatro dimensões. A Figura 4.55 contém uma amostras do conjunto de testes utilizado para validação do processo de síntese. O conjunto é o mesmo utilizado nos processos descritos na Seção 4.5. A figura mostra em (i), (iii) e (v) imagens renderizadas pela composição das bandas 450 nm, 550 nm e 650 nm, sendo “Specim IQ” a imagem de referência (*ground truth*) e a correspondente imagem “Sintetizada”. Em (ii), (iv) e (vi) imagens em tons de cinza na banda 769 nm. Os pontos “a”, “b”, “c” e “d”, marcados na figura, indicam píxeis cujos espectros são apresentados na Figura 4.56. São apresentadas assinaturas espectrais dos píxeis “a” (céu azul), “b” (região da chama), “c” (região de fumaça) e “d” (solo nu) indicados na Figura 4.55, obtidos com o processo de síntese hiperespectral, e as correspondentes capturadas pelo sensor hiperespectral da câmera Specim IQ. Observa-se uma forte correspondência entre as curvas produzidas com dados da câmera Specim IQ e dos sintetizados pela rede MLP.

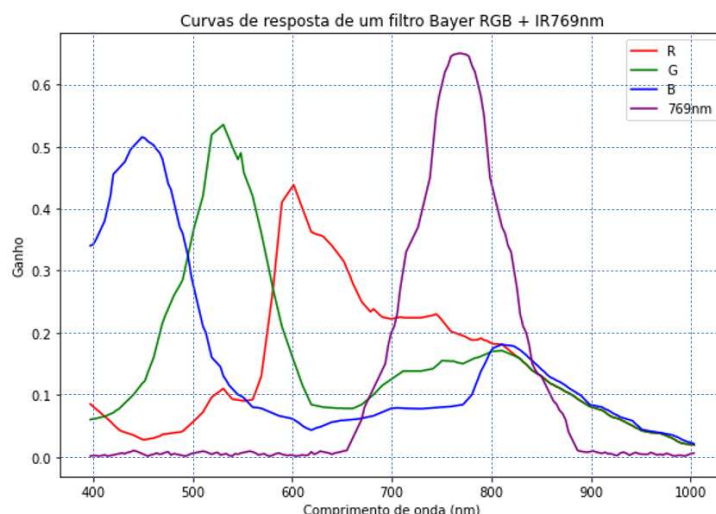


Figura 4.53: Curvas de resposta de um filtro sintetizado com quatro bandas espectrais R, G, B e 769 nm. Esse filtro foi utilizado para projetar as imagens hiperespectrais da câmera Specim em um espaço de quatro dimensões, para posterior treinamento dos algoritmos de síntese.

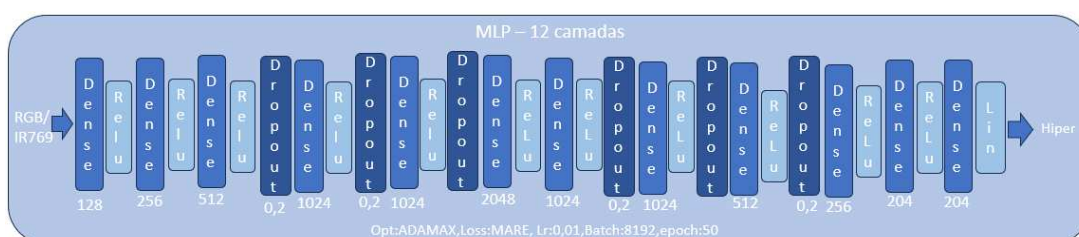


Figura 4.54: Arquitetura de rede MLP para síntese de imagens hiperespectrais, a partir de imagens RGB-IR769, obtidas pela projeção de imagens hiperespectrais em um espaço de quatro dimensões (R, G, B, IR769)

Observa-se, também, a correta identificação da absorção atmosférica devido à excitação da molécula de oxigênio, assim como o pico de emissão oriundo do decaimento de elétrons termicamente excitados presentes na região da chama (píxel “b”). O solo nu, com cor avermelhada, costuma gerar falso-positivo em algoritmos de detecção de incêndios florestais. Conforme pode ser visto pela curva do píxel “d”, a silhueta da assinatura espectral oferece oportunidade para diferenciação do espectro do fogo, reduzindo potencialmente a probabilidade de confusão na classificação. O espectro do píxel “a” mostra maiores intensidades entre 450 nm e 500 nm em direção à banda espectral da cor azul (B).

A Figura 4.57 contém amostras de imagens do conjunto de testes com marcação dos píxeis “a” (céu azul), “b” (vestimenta de um bombeiro), “c” (fumaça) e “d” (fogo), cujos espectros são apresentados na Figura 4.58. As curvas dos píxeis sintetizados “a”, “b” e “c” possuem excelente correspondência com as curvas de referência (Specim IQ).

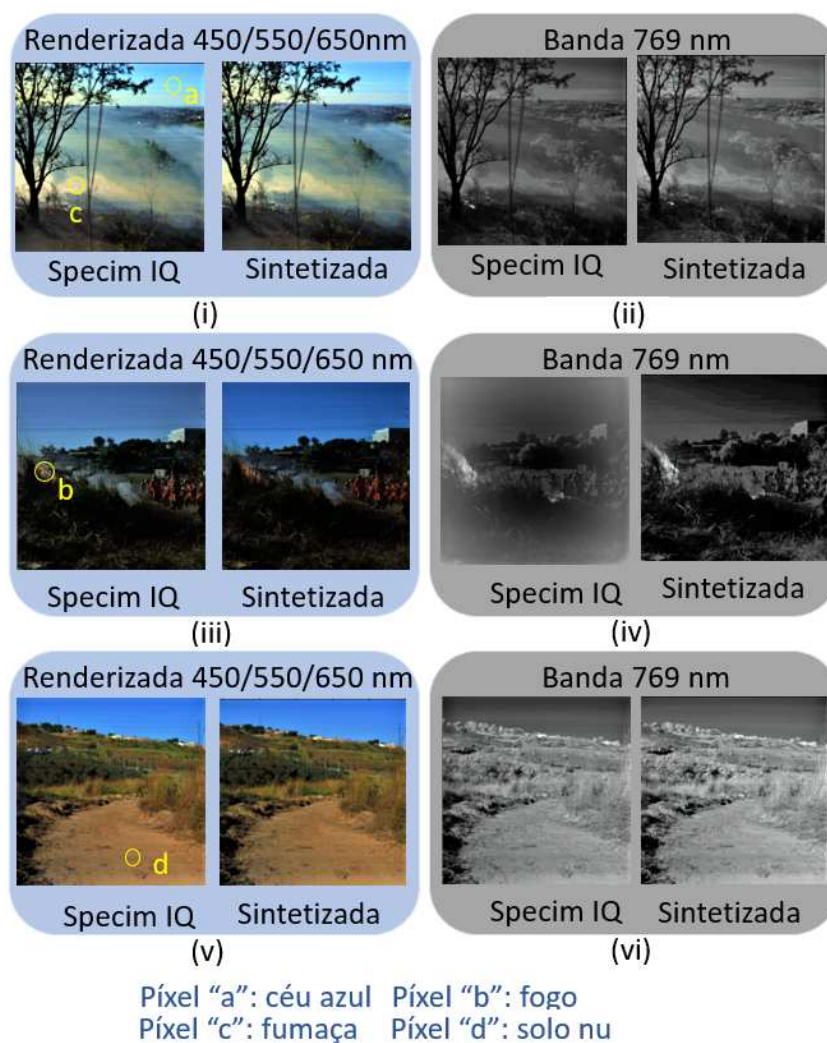


Figura 4.55: Em (i), (iii) e (v) imagens renderizadas pela composição das bandas 450 nm, 550 nm e 650 nm, sendo “Specim IQ” a imagem de referência (*ground truth*) e a correspondente imagem “Sintetizada” utilizando uma rede MLP de 12 camadas, ambas renderizadas. Em (ii), (iv) e (vi) imagens em tons de cinza na banda 769 nm. Os pontos “a”, “b”, “c” e “d” marcados na figura indicam píxeis cujos espectros são apresentados na Figura 4.56

A curva do píxel “d” apresenta maiores discrepâncias na faixa NIR.

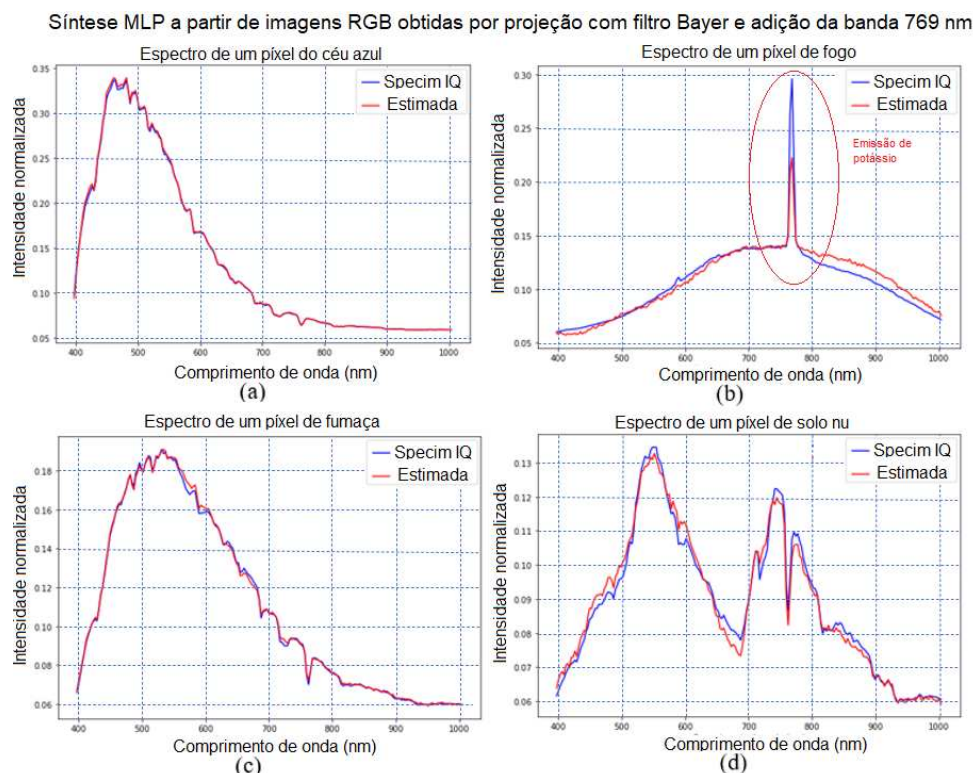


Figura 4.56: Comparação do conteúdo espectral de píxeis de diversas imagens sintetizadas a partir de imagens contendo canais RGB e um quarto canal correspondente à banda de 769 nm (MLP-RGB-IR769). Observa-se a correta síntese do pico de emissão por decaimento de elétrons em átomos de potássio e o pico de absorção devido à excitação da molécula de oxigênio, assim como perfis de assinaturas espectrais de diferentes elementos das cenas.

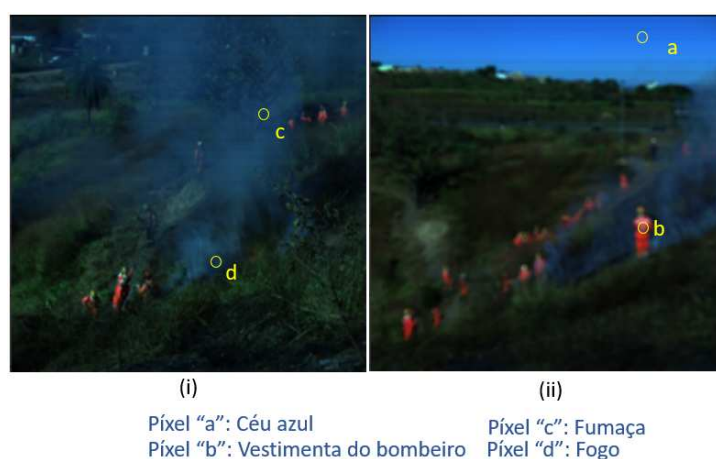


Figura 4.57: Amostra de imagens do conjunto de testes para validação do processo MLP-RGB-IR769, com marcação de píxeis em diferentes regiões da cena. Os espectros desses píxeis são apresentados na Figura 4.58

A Figura 4.59 mostra em (i) uma imagem capturada pelo sensor hiperespectral da câmera Specim IQ (*ground truth*), sendo a mesma imagem apresentada na Figura 4.40 e utilizada na avaliação dos processos de síntese descritos na Seção 4.5. O píxel "a" corresponde à região de

Síntese MLP a partir de imagens RGB obtidas por projeção com filtro Bayer e adição da banda 769 nm

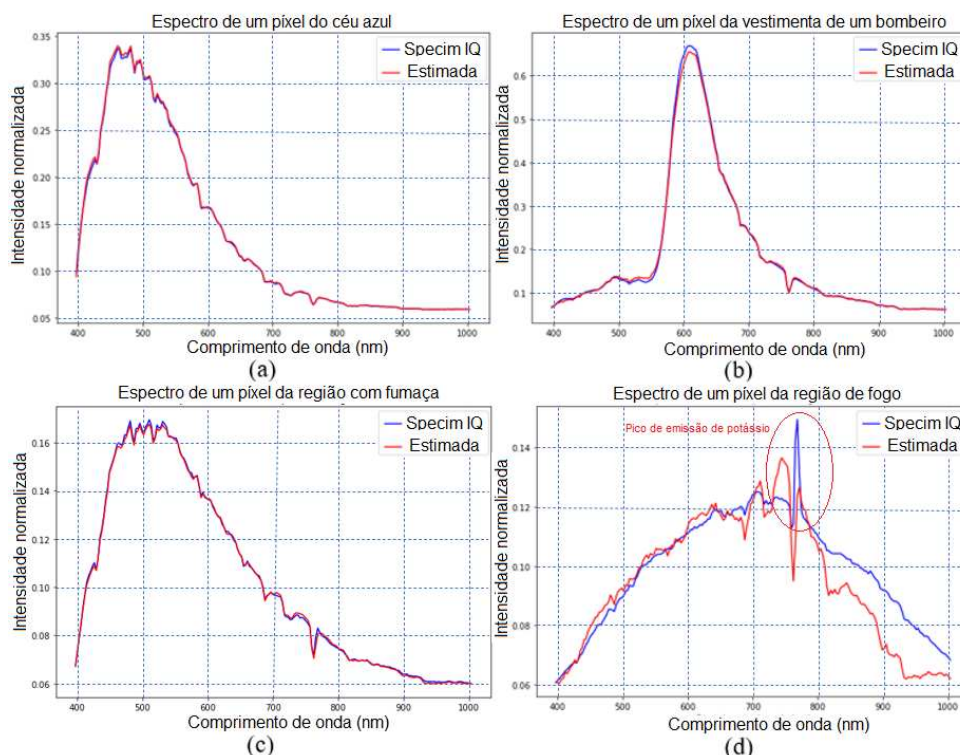


Figura 4.58: Comparação do conteúdo espectral de píxeis em imagens sintetizadas pelo processo MLP-RGB-IR769. Observa-se a correta síntese do pico de emissão por decaimento de elétrons em átomos de potássio e o pico de absorção devido à excitação da molécula de oxigênio.

fogo, sendo o mesmo usado na comparação dos processos apresentados nessa seção. A figura também mostra a correspondente imagem sintetizada (“Sintetizada RGB”) pelo processo MLP com projeção utilizando um filtro Bayer, MLP-RGB, conforme descrito na Seção 4.5.2, assim como a imagem sintetizada (“Sintetizada RGB-IR”) pelo processo MLP-RGB-IR769. Em (ii), imagem original Specim IQ, e sintetizadas por MLP-RGB (“Sintetizada RGB”) e por MLP-RGB-IR769 (“Sintetizada RGB-IR”), todas em tons de cinza na banda 769 nm.

A Figura 4.60 mostra uma comparação das assinaturas espectrais do píxel “a” (região de fogo) da Figura 4.59, conforme sintetizado pelo processo MLP-RGB pelo processo MLP-RGB-IR769. A comparação também se dá com a curva do mesmo píxel capturada com uso do sensor hiperespectral da câmera Specim IQ (*ground truth*). Como pode ser observado, o processo MLP-RGB-IR769 logrou êxito em sintetizar o pico de emissão em átomos de potássio termicamente excitados, o que não ocorreu com o processo MLP-RGB.

Como pode ser visto, esse resultado indica a viabilidade do uso de filtros na faixa NIR como

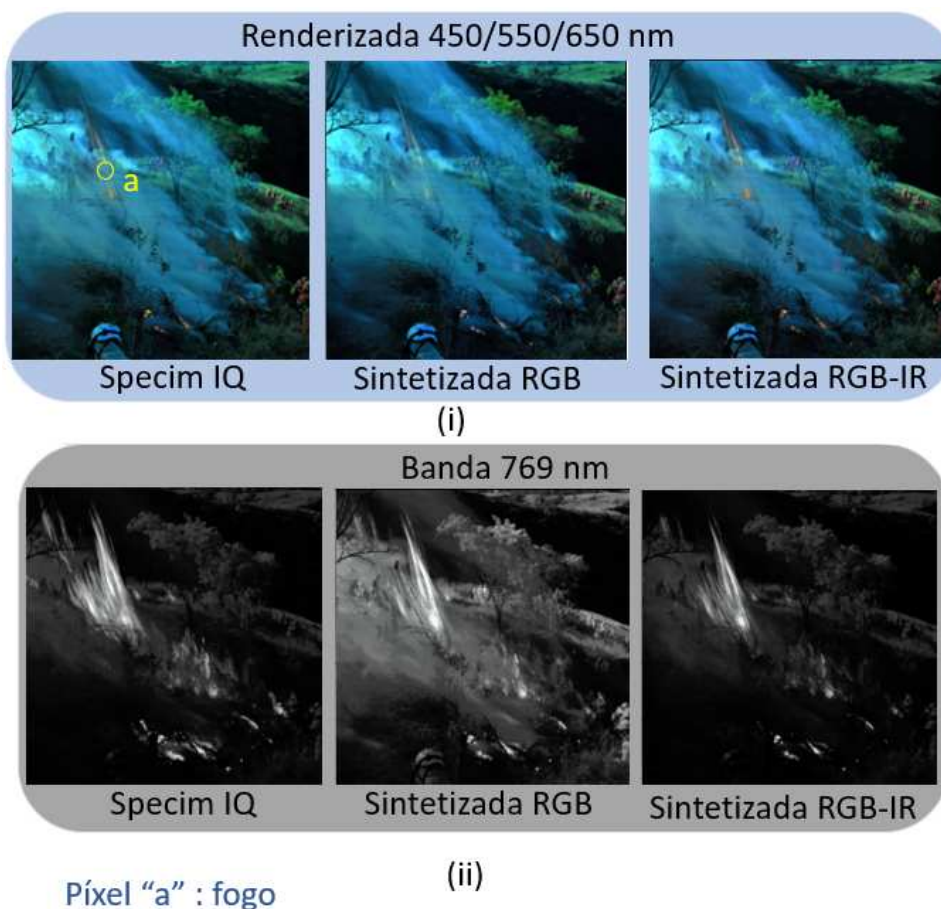


Figura 4.59: Em (i), imagens, *ground truth* (Specim IQ) e sintetizadas por MLP-RGB (“Sintetizada RGB”) e por MLP-RGB-IR769 (“Sintetizada RGB-IR”). Em (ii), imagens em tons de cinza na banda 769 nm, *ground truth* (Specim IQ) e sintetizadas.

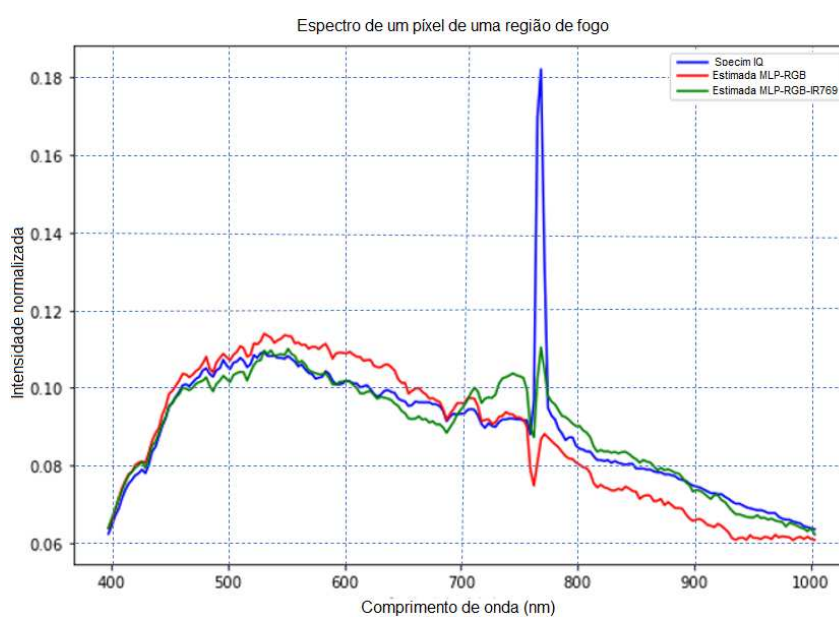


Figura 4.60: Comparação do conteúdo espectral de um píxel da região de fogo. Observa-se a correta síntese do pico de emissão por decaimento de elétrons em átomos de potássio pelo processo MLP-RGB-IR769. O processo MLP-RGB não identificou o pico de emissão.

forma de prover informações espectrais complementares para síntese mais acurada de imagens hiperespectrais. As imagens apresentadas na Figura 4.59 (ii) mostram excelente contraste com o plano de fundo e um sinal intenso na região de fogo.

Sistemas de detecção de incêndios florestais podem ser concebidos com base no uso de câmeras RGB comuns e de baixo custo, acrescidas de filtro NIR de banda estreita, associados a algoritmos de síntese hiperespectral. A capacidade de detectar o pico de emissão por decaimento em átomos de potássio, assim como de outras substâncias com assinaturas espectrais características, devem ser levadas em conta como uma forma de rastrear focos de incêndio, assim como na redução de erros de detecção de incêndios florestais.

Essa seção e as anteriores nesse capítulo descreveram métodos e resultados de síntese hiperespectral a partir de imagens RGB e RGB com adição de uma banda centrada em 769 nm, obtidas com uso de funções de projeção. As seções seguintes descrevem métodos e resultados de síntese hiperespectral a partir de imagens obtidas diretamente de sensores RGB de câmeras, tanto no formato PNG quanto no formato TIF.

4.7 Síntese de imagens hiperespectrais a partir de uma câmera comercial Panasonic DMC-FZ5

Seguindo a metodologia descrita no Capítulo 3, em uma nova etapa da pesquisa, foi avaliado um processo de recuperação do espectro de uma cena a partir de imagens RGB capturadas com uma câmera comercial de baixo custo. A câmera utilizada foi o modelo Panasonic DMC-FZ5. A Figura 4.61 mostra o processo para síntese de imagens hiperespectrais. Os blocos na cor laranja correspondem às etapas de treinamento e aqueles na cor azul às etapas de inferência. As imagens de treinamento e teste, tanto da câmera Specim, quanto da câmera Panasonic foram submetidas a uma etapa de pré-processamento, conforme apresentado na Seção 3.11.

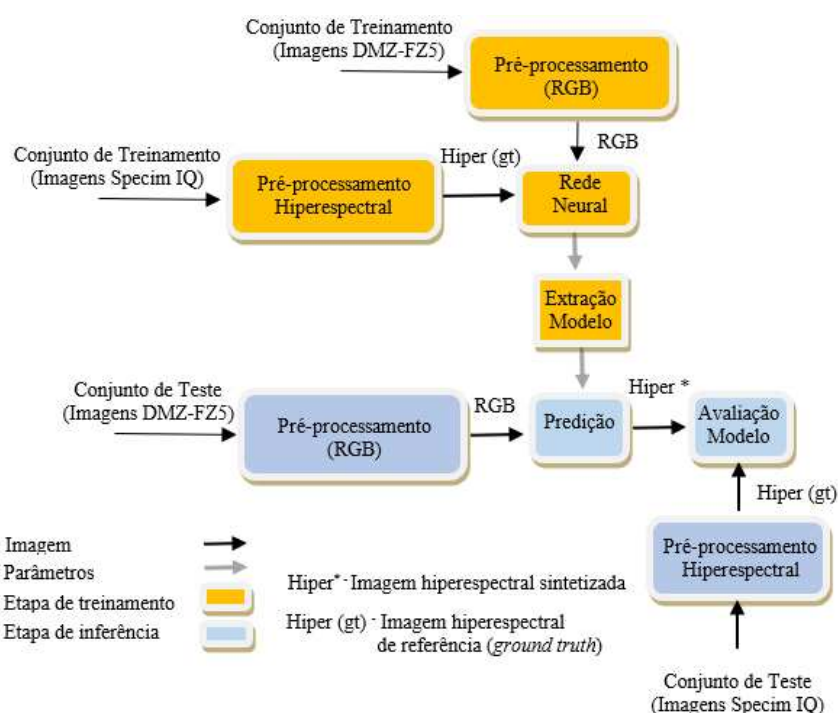


Figura 4.61: Diagrama em blocos das etapas do processo utilizado para síntese de imagens hiperespectrais com base em imagens RGB de uma câmera Panasonic DMC-FZ5 e imagens captadas pelo sensor hiperespectral de uma câmera Specim IQ.

Foram avaliadas arquiteturas de redes MLP com cinco, sete, onze e quinze camadas e testadas funções de ativação ELU e RELU, com otimizador ADAMAX e taxa de aprendizado de 0,01, em 50 épocas. As redes foram treinadas com lotes de 2048 amostras. A Figura 4.62 apresenta as curvas da função de custo com métrica MARE para os modelos MLP testados, utilizando validação cruzada (*5-fold cross-validation*). Os modelos representados pelas diferentes arquiteturas foram treinados cinco vezes, com diferentes conjuntos de treinamento e validação.

Como pode ser visto, para cada modelo, a curva MARE converge assintoticamente para um valor mínimo que difere ligeiramente em cada treinamento. Esse comportamento é o mesmo para as diversas arquiteturas MLP. No modelo com cinco camadas, o valor mínimo da curva de custo (MARE), na fase de treinamento, variou entre 0,0815 e 0,0857. No modelo com sete camadas, a variação foi de 0,0807 a 0,0847. No modelo com onze camadas, a variação foi de 0,0809 a 0,0849, e no modelo com quinze camadas, os valores ficaram entre 0,0810 e 0,0851. A diferença entre o melhor e o pior desempenho foi de apenas 0,98%, considerando o menor valor alcançado pela função de custo, em cada modelo.

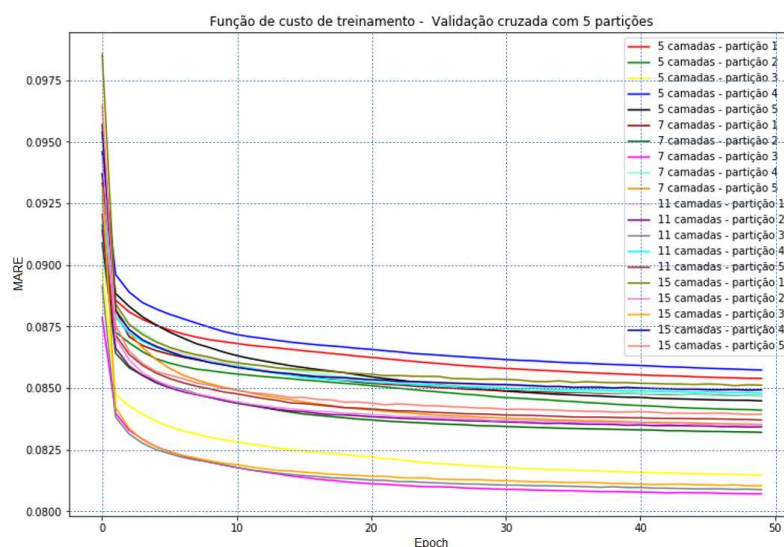


Figura 4.62: Função de custo na fase de treinamento de quatro arquiteturas MLP com cinco, sete, onze e quinze camadas, utilizando validação cruzada com 5 partições. Observa-se um desempenho semelhante para as arquiteturas, com uma diferença de 0,98% no valor mínimo alcançado pela função de custo nas diversas arquiteturas.

A Tabela 4.6 resume os resultados obtidos nos experimentos com as diferentes arquiteturas para o conjunto de validação. A tabela apresenta os valores mínimo, médio, máximo, assim como o desvio padrão. Os valores foram definidos para partições sobre um total de 180 imagens, usando validação cruzada com 5 partições.

Tabela 4.6: MARE do conjunto de validação nas arquiteturas MLP com cinco, sete, onze e quinze camadas. Validação cruzada com 5 partições.

MARE	5 camadas	7 camadas	11 camadas	15 camadas
Máximo	0,2436	0,2406	0,2397	0,2406
Mínimo	0,0336	0,0321	0,0331	0,0334
Médio	0,0909	0,0914	0,0918	0,0917
Desvio padrão	0,0394	0,0392	0,0393	0,0392

A Tabela 4.7 resume os resultados obtidos nos experimentos com as diferentes arquiteturas para o conjunto de testes. Os valores foram calculados para um total de 44 imagens. A rede

Tabela 4.7: MARE para o conjunto de teste nas arquiteturas MLP com 5, 7, 11 e 15 camadas.

MARE	5 camadas	7 camadas	11 camadas	15 camadas
Máximo	0,2164	0,2151	0,2166	0,2165
Mínimo	0,0344	0,0346	0,0375	0,03656
Médio	0,0996	0,0998	0,1002	0,0998
Desvio padrão	0,0454	0,0446	0,0452	0,0449

com melhor desempenho foi uma MLP com 7 camadas, mas com valores muito próximos das demais. Dada a semelhança de desempenho dos modelos, optou-se por uma arquitetura com 15 camadas para prover maior capacidade à rede neural (o controle de *overfitting* é feito com uso da camada *dropout*), uma vez que neste estágio da pesquisa o tempo de processamento e complexidade computacional não é o mais relevante e sim a avaliação do desempenho na tarefa de síntese hiperespectral. A Figura 4.63 apresenta a arquitetura da rede neural considerada nessa etapa. A arquitetura utiliza função de ativação “RELU” nas camadas internas e uma função “Linear” na camada de saída. Os blocos “*dropout*” implementam a ação de regularização para controle da complexidade do modelo e prevenção do sobreajuste (*overfitting*).

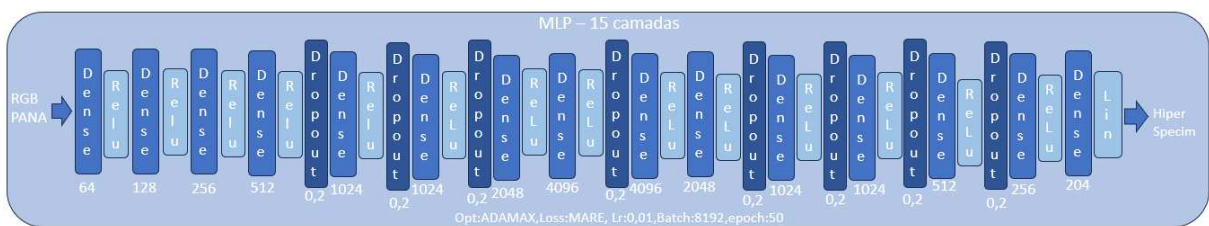


Figura 4.63: Arquitetura de rede neural utilizada na síntese de imagens hiperespectrais a partir de imagens RGB de uma câmera Panasonic DMC-FZ5.

A Figura 4.64 mostra o sumário dos parâmetros e sequência de camadas da arquitetura correspondente à Figura 4.63.

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 64)	256
dense_1 (Dense)	(None, 128)	8320
dense_2 (Dense)	(None, 256)	33024
dense_3 (Dense)	(None, 512)	131584
dropout (Dropout)	(None, 512)	0
dense_4 (Dense)	(None, 1024)	525312
dropout_1 (Dropout)	(None, 1024)	0
dense_5 (Dense)	(None, 1024)	1049600
dropout_2 (Dropout)	(None, 1024)	0
dense_6 (Dense)	(None, 2048)	2099200
dense_7 (Dense)	(None, 4096)	8392704
dropout_3 (Dropout)	(None, 4096)	0
dense_8 (Dense)	(None, 4096)	16781312
dense_9 (Dense)	(None, 2048)	8390656
dropout_4 (Dropout)	(None, 2048)	0
dense_10 (Dense)	(None, 1024)	2098176
dropout_5 (Dropout)	(None, 1024)	0
dense_11 (Dense)	(None, 1024)	1049600
dropout_6 (Dropout)	(None, 1024)	0
dense_12 (Dense)	(None, 512)	524800
dropout_7 (Dropout)	(None, 512)	0
dense_13 (Dense)	(None, 256)	131328
dense_14 (Dense)	(None, 204)	52428

```

Total params: 41,268,300
Trainable params: 41,268,300
Non-trainable params: 0

```

Figura 4.64: Sumário dos parâmetros da rede neural utilizada na síntese de imagens hiperespectrais a partir de imagens RGB de uma câmera Panasonic DMC-FZ5.

A Figura 4.65 mostra uma comparação da função de custo na fase de treinamento para imagens submetidas ao processo de calibração correspondente à cor rosa, com a função de custo referente às mesmas imagens, na cor violeta, sem passar pelo processo de calibração. A calibração das imagens provê uma melhora de 1% no desempenho do processo. A Figura 4.66 apresenta o MARE médio calculado para cada imagem do conjunto de validação. Os valores correspondem à média para cada imagem do conjunto obtida no cômputo do MARE em cada iteração da validação cruzada, com diferentes partições. A partição do conjunto de validação contém 36 imagens. Observa-se a proximidade dos valores para as diferentes arquiteturas MLP.

A Figura 4.67 apresenta o MARE de cada imagem do conjunto de testes, para as arquiteturas MLP com sete, onze e quinze camadas. O conjunto de testes contém 44 imagens. Observa-se a proximidade dos valores para as diferentes arquiteturas MLP.

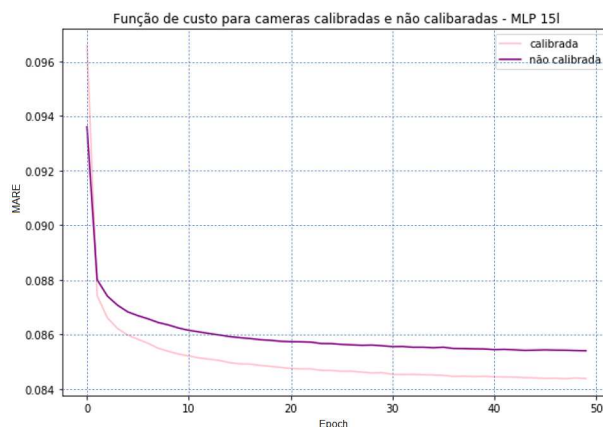


Figura 4.65: Comparação da função de custo na fase de treinamento em uma rede neural de 15 camadas, para imagens submetidas ao processo de calibração (cor rosa), e sem calibração (cor violeta). A melhora obtida é de somente 1%.

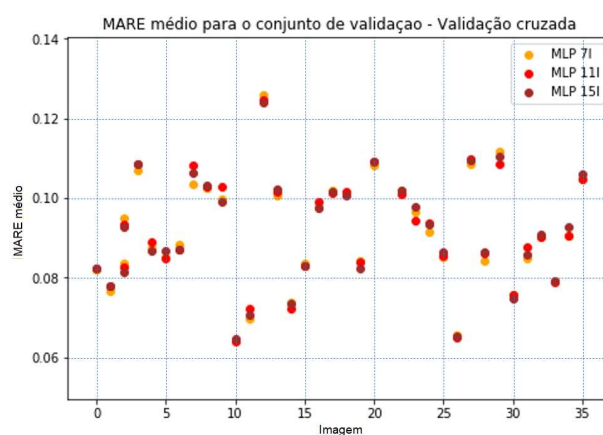


Figura 4.66: MARE médio do conjunto de imagens das partições na validação cruzada. Os valores correspondem a arquiteturas de redes MLP com sete, onze e quinze camadas.

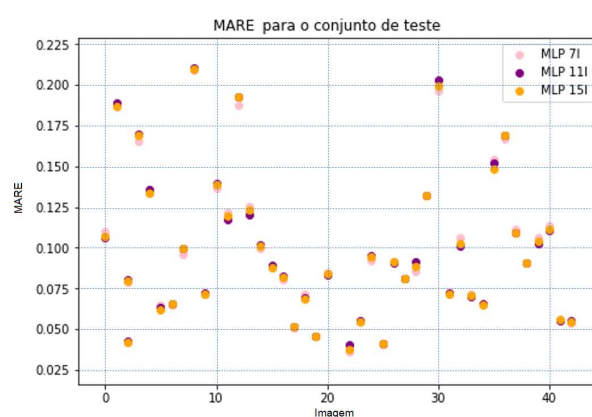


Figura 4.67: MARE de cada imagem do conjunto de testes calculados para arquiteturas MLP com sete, onze e quinze camadas.

A Figura 4.68 apresenta amostras de imagens do conjunto de testes e correspondentes imagens sintetizadas. Todas as imagens são renderizadas utilizando a composição das bandas em 450 m, 550 nm e 650 nm. Em (i), (iii) e (v) imagens capturadas pelo sensor hiperespectral da

câmera Specim IQ. Os círculos identificados como “a”, “b”, “c”, “d”, “e”, “f”, “g”, “h” marcam píxeis cujos espectros estão contidos na Figura 4.69. Em (ii), (iv) e (vi), imagens sintetizadas pela MLP.

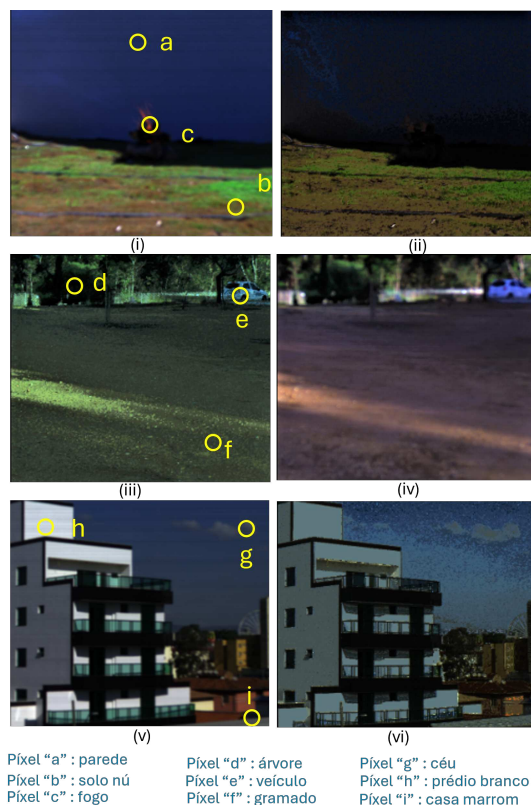


Figura 4.68: Amostras de imagens do conjunto de testes e correspondentes imagens sintetizadas. O espectro dos píxeis marcados com círculos nas imagens são apresentados na Figura 4.69.

A Figura 4.69 mostra a comparação de píxeis de imagens hiperespectrais sintetizadas e *ground truth* do conjunto de testes. Os píxeis considerados nessa análise estão marcados na Figura 4.68. A curva na cor vermelha corresponde à distribuição espectral de píxeis da imagem sintetizada. A curva na cor azul refere-se à distribuição espectral de píxeis de imagens das mesmas cenas capturadas com uso da câmera Specim IQ (*ground truth*). Os píxeis “a”, “d”, e “e” apresentam boa correspondência, tanto na intensidade, quanto no perfil (silhueta) de distribuição espectral. Os demais píxeis apresentam divergência na intensidade em alguns comprimentos de onda, mas uma boa semelhança na forma da curva de distribuição espectral. As intensidades divergem mais na faixa visível. Na faixa NIR a correspondência é mais próxima. O píxel “c” está localizado na região contendo fogo na cena. Observa-se o pico de emissão devido ao decaimento de elétrons em átomos de potássio localizados na região da chama, conforme capturado pela câmera Specim IQ. Essa característica não foi sintetizada pelo processo

Síntese hiperespectral a partir de imagens RGB de uma câmera Panasonic DMC-FZ5

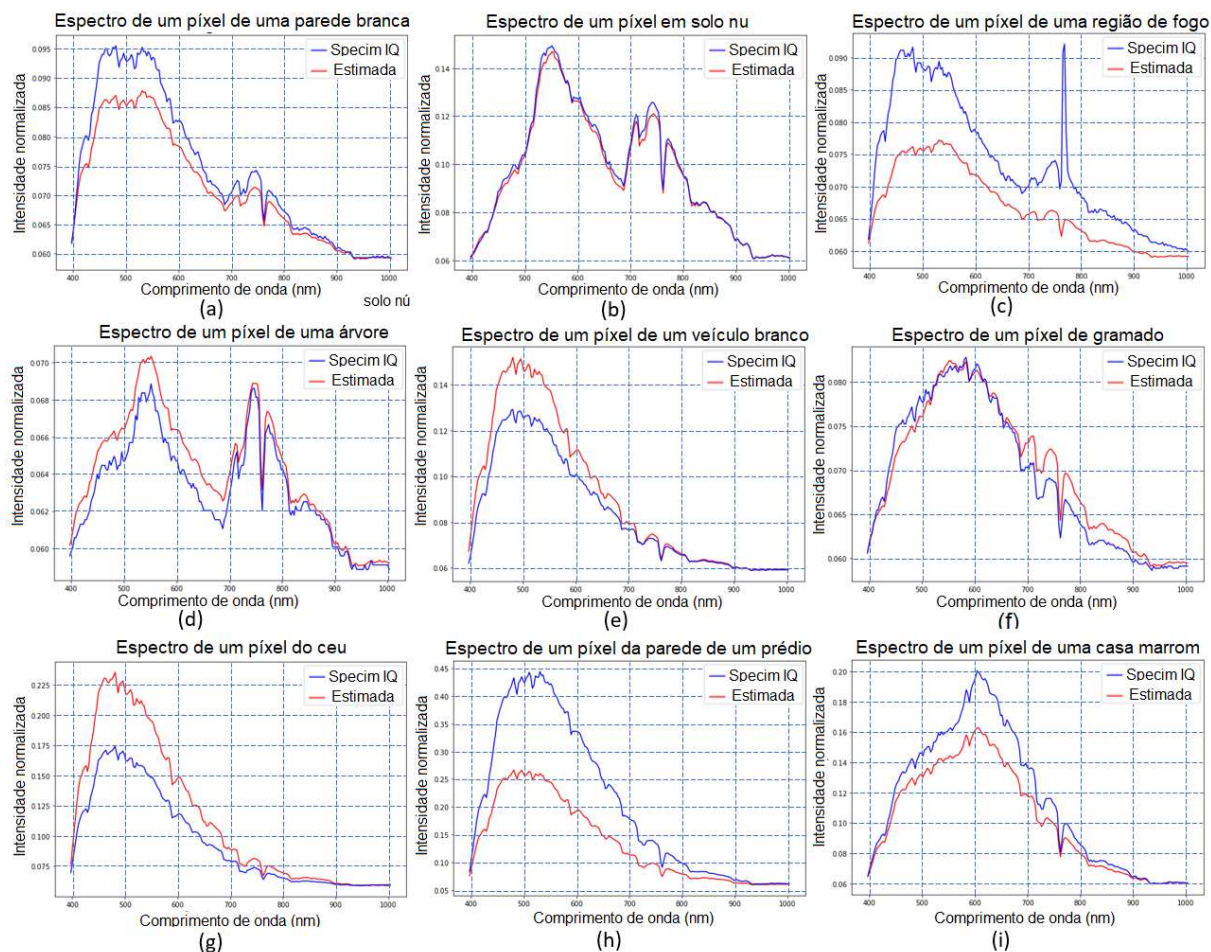


Figura 4.69: Comparação do espectro de pixels de três imagens do conjunto de testes, obtidos com o processo de síntese com base em imagens RGB produzidas com uso da câmera Panasonic DMC-FZ5 e os correspondentes capturados pela câmera Specim IQ (*ground truth*).

da rede neural. O pico de absorção devido à molécula de oxigênio esteve presente em todos os pixels das imagens sintetizadas.

Esses resultados indicam que o processo foi efetivo em sintetizar o perfil espectral de diversos pixels, mas os valores de MARE foram maiores em comparação com os métodos baseados em projeção do conteúdo espectral em um espaço de cores. Os resultados podem ser melhorados utilizando outras formas de normalização baseados na subtração do valor médio e na divisão pela variância. Esse método pode ser útil em etapas de extração de características baseados em conteúdo espectral, por exemplo, a razão entre intensidades de bandas espectrais específicas, tanto na faixa do espectro visível como na do infravermelho próximo. O processo de síntese de imagens hiperespectrais a partir de imagens de uma câmera RGB de baixo custo apresenta-

se com viável e produziu resultados consistentes em um conjunto de teste com diversidade de motivos, incluindo fogo e fumaça. Os resultados podem ser aprimorados para encontrar uso efetivo na detecção de incêndios florestais. Um caminho possível seria a adição de uma banda espectral na faixa NIR, conforme evidenciado na Seção 4.6

4.8 Síntese de imagens hiperespectrais a partir de imagens captadas pelo sensor RGB da câmera Specim IQ

Um novo experimento consistiu do uso das imagens da base de dados produzidas pelo sensor RGB embutido na câmera Specim IQ, cujas imagens são no formato PNG (512 x 512 píxeis). As cenas imageadas são as mesmas capturadas pelo sensor hiperespectral da câmera. A síntese de imagens hiperespectrais foi feita utilizando a mesma arquitetura MLP e mesmos parâmetros definidos no processo descrito na Seção 4.7 e conforme Figura 4.63. As etapas de pré-processamento envolvendo calibração geométrica e recorte das imagens não foram necessárias, uma vez que as câmeras possuem o mesmo campo de captura de imagem e mesma resolução espacial.

A Figura 4.70 contém amostras de imagens do conjunto de testes utilizadas para avaliar o desempenho do processo de síntese. A Figura 4.71 apresenta uma comparação dos espec-

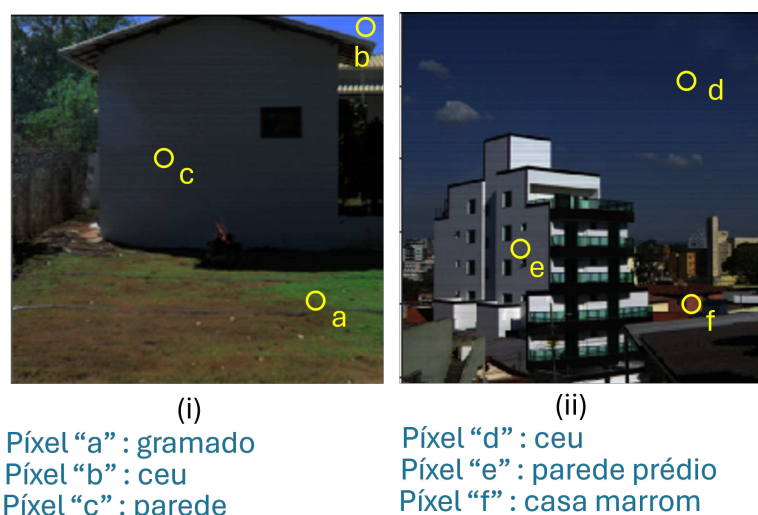


Figura 4.70: Amostras de imagens do conjunto de testes para o processo utilizando imagens do sensor RGB da câmera Specim IQ. Os círculos marcam os píxeis cujos espectros são apresentados na Figura 4.71.

tros de alguns píxeis de imagens hiperespectrais sintetizadas. Os píxeis avaliados são aqueles

indicados na Figura 4.70. A curva na cor vermelha corresponde à distribuição espectral de píxeis da imagem sintetizada. A curva na cor azul refere-se à distribuição espectral de píxeis de imagens das mesmas cenas capturadas com uso do sensor hiperespectral da câmera Specim IQ (*ground truth*). Os píxeis “a”, “d”, e “e” apresentam boa correspondência, tanto na intensidade, quanto no perfil de distribuição espectral (silhueta). Os demais píxeis apresentam divergência na intensidade em alguns comprimentos de onda, mas uma boa correspondência com o perfil de distribuição espectral comparado com as curvas dos píxeis capturados pelo sensor hiperespectral da câmera Specim IQ.

O resultado é similar ao obtido com o uso da câmera Panasonic DMC-FZ5. Isso indica que a metodologia de calibração, recorte e ajuste na resolução das imagens utilizado no processo com uso dessa câmera não introduziu perdas significativas na acurácia do processo de síntese e corrobora a avaliação da viabilidade de uso de uma câmera RGB de baixo custo, tendo a vantagem de ter imagens com maior resolução espacial.

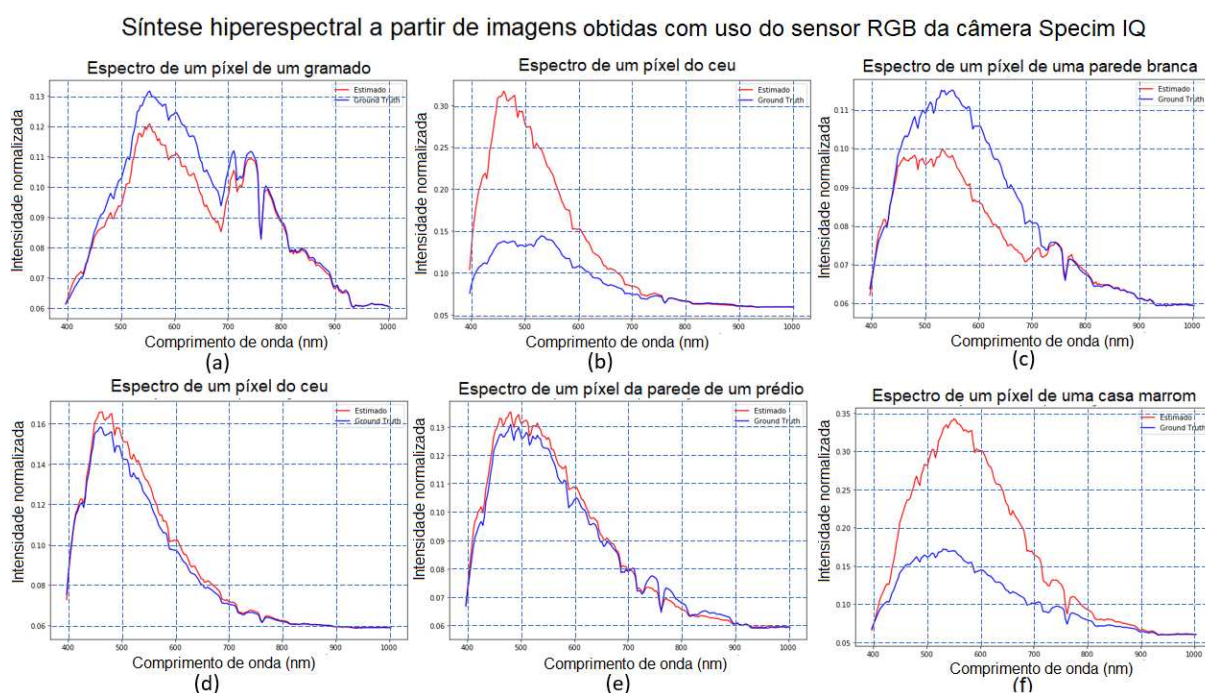


Figura 4.71: Comparação do espectro de píxeis de duas imagens do conjunto de testes, obtidos com o processo de síntese com base em imagens RGB produzidas com uso do sensor RGB embutido na câmera Specim IQ, e os correspondentes capturados pelo sensor hiperespectral da mesma câmera (*ground truth*).

4.9 Comentários finais

Neste capítulo, foram apresentadas diversas implementações, seguindo a metodologia delineada para a pesquisa. A etapa preliminar envolveu a implementação de quatro processos de síntese, permitindo concluir sobre a viabilidade de sintetizar imagens hiperespectrais a partir de imagens RGB, obtidas por meio da projeção em um espaço tridimensional. A síntese, nesse caso, foi restrita à faixa visível, com apenas 31 componentes espectrais. Diversos modelos de regressão, agrupamento e redes neurais foram testados, possibilitando a seleção dos modelos mais promissores. Essa etapa serviu para identificar que apenas três processos agregavam valor e seriam utilizados nas etapas seguintes.

Na etapa seguinte, foram utilizadas bases de imagens hiperespectrais construídas com as câmeras Specim FX-10 e Specim IQ. Nessa fase, o processo de síntese abrangeu uma faixa mais ampla, incluindo a luz visível e o infravermelho próximo. O número de componentes espectrais foi ampliado para 204 bandas espectrais, propiciando maior resolução espectral. A projeção no espaço do sensor foi realizada em três e quatro dimensões. A inclusão de uma dimensão adicional, na banda de 769 nm, mostrou um aprimoramento na síntese. Foram avaliados apenas os processos baseados em regressão linear, *Ridge Regression*, rede neural do tipo MLP e MLP combinada com agrupamento por *minibatch K-Means*. As análises e discussões foram fundamentadas em aspectos qualitativos e quantitativos, utilizando as métricas e métodos de comparação definidos no Capítulo 3.

Por fim, apresentou-se o processo de síntese a partir de imagens RGB produzidas por câmeras RGB, sendo uma da Specim IQ, em formato PNG, e outra por uma câmera Panasonic DMC-FZ5, em formato TIF. Com o uso da câmera RGB Specim, não foi necessária a etapa de calibração e ajustes requeridos para a câmera Panasonic. Nessa etapa, foram empregadas somente arquiteturas de redes neurais, devido ao seu melhor desempenho. O próximo capítulo trará uma proposta de novos desenvolvimento e desdobramentos da pesquisa, bem como as conclusões da tese.

Capítulo 5

Conclusão

*E*ste trabalho apresentou os resultados de uma pesquisa desenvolvida em etapas seguindo uma metodologia onde se empregou, em um primeiro momento, uma base de dados de imagens hiperespectrais disponibilizada pelo ICVL para implementar quatro processos de síntese de imagens hiperespectrais com base em dados espectrais esparsos associados a imagens RGB. Uma avaliação preliminar dos dados, utilizando análise de fatores, sugeriu a presença de uma estrutura nos dados residindo em uma variedade de dimensão reduzida (*manifold*). Essa constatação permitiu encontrar uma solução para o problema malposto de mapear do espaço RGB para o espaço de dados hiperespectrais, utilizando quatro processos de síntese hiperespectral. Esses processos foram selecionados, indo de uma regressão simples, que obteve os melhores resultados em termos de tempo de execução, a uma abordagem de menor custo computacional (em comparação com MLP) que incorpora agrupamento. Avançou-se, então, para uma solução baseada em rede neural MLP, seguindo o mesmo caminho em direção a uma solução com redução de custo computacional, baseada em agrupamento. Neste caso, houve uma melhora significativa no desempenho. O processo RR exibiu o melhor desempenho em relação ao tempo de execução, enquanto o MLP alcançou os melhores resultados em termos da métrica MARE (veja Tabelas 4.1, 4.2 e 4.3). O processo RR/KMC apresentou os piores resultados em todas as métricas. O processo MLP/KMC demonstrou tempos de execução significativamente melhores, entregando resultados comparáveis em termos da métrica MARE em referência ao processo baseado somente em MLP. Os resultados de cada processo foram comparados entre si e com o estado da arte estabelecido no concurso NTIRE – 2020. A

comparação indica que, mesmo ao empregar algoritmos mais simples, foram alcançadas soluções em uma ordem de magnitude similar, considerando a métrica MARE. A hipótese de que pontos representando píxeis próximos no espaço RGB também serão próximos no espaço hiperespectral foi validada, utilizando o conjunto de dados cujas imagens foram adquiridas sob condições de iluminação por luz natural, conforme evidenciado pelos resultados apresentados na Seção 4.2.4. Consequentemente, permitiu implementar e testar uma abordagem direta para mapear dados do espaço RGB para o espaço hiperespectral, sem a necessidade de qualquer mecanismo de busca, o que significa menor custo computacional.

Existe potencial para aprimorar o desempenho da regressão linear ao empregar funções de custo associadas a métricas mais adequadas para capturar o comportamento local da distribuição de píxeis, como MARE. A técnica de agrupamento mostra-se promissora para reduzir os custos computacionais quando combinado com técnicas que oferecem maior acurácia, como MLP, mas que demandam maior complexidade computacional. Nessa primeira etapa da metodologia de pesquisa, a síntese ficou restrita à faixa de luz visível, com imagens contendo 31 componentes espectrais. Essa etapa serviu para identificar que apenas três processos agregavam valor e seriam utilizados nas etapas seguintes.

Seguindo a metodologia, foi construída uma base de dados de imagens hiperespectrais utilizando as câmeras Specim FX-10 e Specim IQ. As imagens obtidas com a câmera FX-10 foram utilizadas na caracterização de perfis espectrais de diversos elementos nas cenas imageadas. Em particular, foi possível identificar o pico de emissão na banda 769 nm, oriundo do decaimento de elétrons, termicamente excitados, em átomos de potássio, na região da chama, assim como o pico de absorção atmosférica na banda de 760 nm, devido à excitação das moléculas de oxigênio. Essas mesmas características foram identificadas nos dados captados com a câmera Specim IQ. Foram treinados três processos para sintetizar imagens hiperespectrais, a partir de imagens RGB obtidas pela projeção no espaço de três dimensões de um sensor hipotético, utilizando curvas de um filtro *Bayer*. Nessa etapa, a síntese abrangeu uma faixa mais ampla, incluindo a luz visível (VIS) e o infravermelho próximo (NIR), e valeu-se dos processos RR, RR/KMC e MLP. O número de componentes espectrais foi ampliado para 204, propiciando maior resolução espectral. Também foi feita a projeção no espaço do sensor com quatro dimensões, com a inclusão de uma dimensão adicional referente à banda centrada em 769 nm. Os resultados com o filtro de quatro componentes mostrou aprimoramento da síntese e possibilitou a identificação

do pico de emissão de potássio, o que não foi possível no método com apenas três componentes. As análises e discussões foram fundamentadas em aspectos qualitativos e quantitativos, utilizando as métricas MARE, RMSE e R^2 .

Por fim, desenvolveu-se um processo de síntese a partir de imagens RGB produzidas por sensores RGB de duas câmeras. Uma dos sensores RGB equipa a própria câmera Specim IQ, e gera imagens no formato PNG. O outro corresponde ao sensor de uma câmera Panasonic DMC-FZ5 que gera imagens no formato TIF. Nessa etapa, foram empregadas somente arquiteturas de redes neurais, devido ao seu melhor desempenho, já identificado nas etapas anteriores. A captura de imagens se deu com uso de um *setup* mecânico para abrigar as câmeras Specim IQ e DMZ-FZ5, possibilitando a captura da mesma cena. Os resultados apresentaram um pior desempenho se comparado aos métodos de projeção por filtros, uma vez que incorpora outras variáveis no processo, como diferenças no arranjo óptico, distorções, ruído, distorções geométricas, sensibilidade. No entanto, a silhueta das curvas de referência (Specim IQ - *Ground Truth*) e sintetizadas apresentam semelhança em sua distribuição de potência espectral, diferindo, em geral, na intensidade. Isso indica que a forma da assinatura espectral é preservada, o que viabiliza seu uso na identificação de elementos da cena, como incêndios florestais.

Vale ressaltar que certos comportamentos das curvas espectrais, caracterizados por mudanças abruptas, apresentam maior desafio. Notavelmente, os métodos encontraram dificuldades em identificar o pico em 769 nm oriundo do decaimento de elétrons em átomos de potássio termicamente excitados na região da chama. Este aspecto específico sugere necessários refinamentos nos processos e procedimentos, e mesmo nas técnicas de síntese utilizadas.

Ao empregar uma distribuição espectral contínua adquirida por meio desses métodos, espera-se aprimorar a acurácia na identificação de incêndios, melhorando assim o desempenho geral de algoritmos de detecção. Essa abordagem pode servir como uma etapa de extração de características para fases subsequentes de classificação, especialmente em sistemas de detecção de incêndios que utilizam câmeras RGB de baixo custo.

Esses métodos têm o potencial de serem usados em uma ampla gama de aplicações, incluindo agricultura de precisão, medicina, mineralogia e outros. A tecnologia moderna nos proporciona meios de usar informações de imagens hiperespectrais em situações do mundo real.

5.1 Trabalhos futuros

Uma etapa que se apresenta de forma imediata é o uso do processo de síntese em sistemas de detecção de incêndio que utilizam câmeras RGB, com base nas assinaturas espectrais sintetizadas. Uma versão de algoritmos de detecção de objetos, como o YOLO, poderia ser desenvolvida baseado em imagens hiperespectrais sintetizadas. Aprimoramentos para redução do tempo de processamento serão necessários, nesse caso, como podas nas redes neurais, otimização de algoritmos de agrupamentos, baseados tipicidade e excentricidade, e melhoria nos algoritmos de regressão, utilizando função de custo baseada na métrica MARE. Outra via de pesquisa está relacionada a explorar novas arquiteturas de redes como as convolucionais, residuais e generativas. Uma abordagem por séries temporais associada a imagens de vídeo hiperespectrais também se apresenta como caminho para desenvolvimentos futuros. O uso da emissão de potássio na faixa NIR como marcador espectral deve ser explorado como forma de redução de erros de detecção nos sistemas de monitoramento de incêndios florestais. A transparência da fumaça, pouco densa, na banda espectral específica, é um facilitador na detecção precoce de incêndios florestais. Uma possibilidade seria utilizar filtros passa-banda, centrados na faixa de 769 nm, em câmeras convencionais e eventualmente, sem o uso de lentes, o que pode favorecer a detecção por não haver restrição de resolução espectral do sistema óptico.

Referências Bibliográficas

- Akhloufi, M. A., Couturier, A., e Castro, N. A. (2021). Unmanned Aerial Vehicles for Wildland Fires: Sensing, Perception, Cooperation and Assistance. *Drones*, 5(1).
- Akhtar, N. e Mian, A. S. (2018). Hyperspectral Recovery from RGB Images Using Gaussian Processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42:100–113.
- Allison, R. S., Johnston, J. M., Craig, G., e Jennings, S. (2016). Airborne Optical and Thermal Remote Sensing for Wildfire Detection and Monitoring. *Sensors*, 16(8).
- Arad, B. e Ben-Shahar, O. (2016). Sparse Recovery of Hyperspectral Signal from Natural RGB Images. In Leibe, B., Matas, J., Sebe, N., e Welling, M., editors, *Computer Vision – ECCV 2016*, pages 19–34. Springer International Publishing.
- Arad, B., Ben-Shahar, O., Timofte, R., Van Gool, L., Zhang, L., Yang, M.-H., Xiong, Z., Chen, C., Shi, Z., Liu, D., Wu, F., Lanaras, C., Galliani, S., Schindler, K., Stiebel, T., Koppers, S., Seltsam, P., Zhou, R., El Helou, M., Lahoud, F., Shahpaski, M., Zheng, K., Gao, L., Zhang, B., Cui, X., Yu, H., Can, Y. B., Alvarez-Gila, A., Van de Weijer, J., Garrote, E., Galdran, A., Sharma, M., Koundinya, S., Upadhyay, A., Manekar, R., Mukhopadhyay, R., Sharma, H., Chaudhury, S., Nagasubramanian, K., Ghosal, S., Singh, A. K., Singh, A., Ganapathysubramanian, B., e Sarkar, S. (2018). NTIRE 2018 Challenge on Spectral Reconstruction from RGB Images. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1042–1049.
- Arad, B., Timofte, R., Ben-Shahar, O., Lin, Y.-T., Finlayson, G., Givati, S., Li, J., Wu, C., Song, R., Li, Y., Liu, F., Lang, Z., Wei, W., Zhang, L., Nie, J., Zhao, Y., Po, L.-M., Yan, Q., Liu, W., Lin, T., Kim, Y., Shin, C., Rho, K., Kim, S., ZHU, Z., HOU, J., Sun, H., Ren, J., Fang, Z., Yan, Y., Peng, H., Chen, X., Zhao, J., Stiebel, T., Koppers, S., Merhof, D., Gupta, H., Mitra,

- K., Fubara, B. J., Sedky, M., Dyke, D., Banerjee, A., Palrecha, A., Sabarinathan, S., Uma, K., Vinothini, D. S., Sathya Bama, B., e Md Mansoor Roomi, S. M. (2020). NTIRE 2020 Challenge on Spectral Reconstruction from an RGB Image. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1806–1822.
- Bishop, C. (2006). *Pattern Recognition and Machine Learning*. New York, Springer-Verlag, ISBN 978-1-4939-3843-8.
- Botchkarev, A. (2019). A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management*, 14:45–79.
- Bradski, G. e Kaehler, A. (2008). *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly, ISBN 978-0-596-51613-0.
- Cai, T. e Wang, L. (2011). Orthogonal Matching Pursuit for Sparse Signal Recovery With Noise. *IEEE Transactions on Information Theory*, 57:4680–4688.
- Carroll, B. e Ostlie, D. (2006). *An Introduction to Modern Astrophysics and Cosmology*. San Francisco, Pearson, Addison-Wesley, ISBN 0-8053-0402-9.
- Easa (2019). Easy Access Rules for Unmanned Aircraft Systems. <https://www.easa.europa.eu/en/document-library/easy-access-rules/easy-access-rules-unmanned-aircraft-systems-regulations-eu>.
- FAA (2020). FAA - Concept of operation V2.0 Unmanned Aircraft System (UAS) Traffic Management. https://www.faa.gov/sites/faa.gov/files/2022-08/UTM_ConOps_v2.pdf.
- Flasseur, O., Fournier, C., Verrier, N., Denis, L., Jolivet, F., Cazier, A., e Lepine, T. (2017). Self-calibration for lensless color microscopy. *Applied Optics*, 56:F189–F199.
- Foster, D. e Amano, K. (2019). Hyperspectral Imaging in Color Vision Research: Tutorial. *Journal of the Optical Society of America A*, 36:606.
- Foster, D., Amano, K., Nascimento, S., e Foster, M. (2006). Frequency of Metamerism in Natural Scenes. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 23:2359–72.

- Gareth, J., Witten, D., Hastie, T., e Tibshirani, R. (2013). *An Introduction to Statistical Learning*. New York, Springer, ISBN 978-1-4614-7138-7.
- Giglio, L., Hall, J. V., e Justice, C. O. (2018). MODIS Collection 6 Active Fire Product User's Guide, Revision B, . https://modis-fire.umd.edu/files/MODIS_C6_Fire_User_Guide_B.pdf.
- Goodfellow, I., Bengio, Y., e Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Guolan, L. e Baowei, F. (2014). Medical Hyperspectral Imaging: a Review. *Journal of Biomedical Optics*, 19.
- ICVL (2023). Hyperspectral Image Dataset. <http://icvl.cs.bgu.ac.il/hyperspectral>.
- Jia, Y., Zheng, Y., Gu, L., Subpa-Asa, A., Lam, A., Sato, Y., e Sato, I. (2017). From rgb to spectrum for natural scenes via manifold-based mapping. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 4715–4723.
- Johnson, R. A. e Dean, W. W. (2007). *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall, ISBN 978-0-13-187715-3.
- Kim, S., Euibum, L., Hong, I., e Yook, J. (2022). Review of Intentional Electromagnetic Interference on UAV Sensor Modules and Experimental Study. *Sensors*, 22(6). Publisher Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland.
- Kruskal, W. H. e Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260):583–621.
- Lisboa, A., Venâncio, P., Neto, A., Magalhães, H., Do Nascimento, C., e Yehia, H. C. (2020). Monitoramento de Incêndios por Meio de Redes e Linhas Aéreas de Energia Elétrica. *Revista da Universidade Federal de Minas Gerais*, 26:66–81.
- Lowe, A., Harrison, N., e French, A. (2017). Hyperspectral Image Analysis Techniques for the Detection and Classification of the Early Onset of Plant Disease and Stress. *Plant Methods*, 13.

- Malkoff, D. B. e Oliver, W. R. (2000). Hyperspectral imaging applied to forensic medicine. In *BiOS*, volume 3920, pages 108–116. SPIE.
- Meneses, P. R., Almeida, T. d., e Baptista, G. M. d. M. (2019). *Reflectância dos Materiais Terrestres: Análise e Interpretação*. Oficina de Textos, ISBN 978-85-7975-301-5.
- Moraes, E. C., Nogueira, J. L. M., Ênio Salvatore Carmine Esposito, de Brito Bastos, E. J., e de Faria Freire, M. L. (1996). *Comparação entre placas de referência de sulfato de bário (pintada e prensada)*, pages 927–931. Anais VIII Simpósio Brasileiro de Sensoriamento Remoto, Salvador.
- NASA, E. O. (2022). Fire information for resource management system (firms). <https://www.earthdata.nasa.gov/learn/find-data/near-real-time/firms>.
- NASA, E. O. (2023). 2023 fire season in the northern hemisphere. <https://earthobservatory.nasa.gov/images/event/151148/2023-fire-season-in-the-northern-hemisphere>.
- NASCIMENTO, C. A. M., LISBOA, A. C., Yehia, H. C., MAGALHÃES, H. A., FOLLADOR NETO, A., BARBOSA, A. V., VENANCIO, P. V. A. B., REZENDE, T. M., MAGALHAES, A. H., CAMPOS, R. J., MELO, M. A. S., CABELO, G. S., LIMA, D. Q., e SOUZA, M. R. S. (2022). Wildfire Detection System Using Artificial Intelligence with the Collaboration of the Web Society. In *Cigre Session Paris, SC-B2*, volume 10138. CIGRE - Conseil International des Grands Réseaux Electriques.
- Navlani, A. (2023). Introduction to Factor Analysis in Python. <https://www.datacamp.com/tutorial/introduction-factor-analysis>.
- Ononye, A., Vodacek, A., e Kremens, R. (2005). Improved fire temperature estimation using constrained spectral unmixing. In *Remote Sensing for Field Users. Am. Soc. Photogram. Remote Sens., pp. CD-ROM, proc. 10th Biennial USDA Forest Service Remote Sensing Applications Conference. Salt Lake City, UT*.
- Ott, L. R. e Longnecker, M. (2015). *An Introduction to Statistical Methods and Data Analysis*. Cengage Learning, ISBN 978-1305269477.
- Qian, S.-E. (2021). Hyperspectral satellites, evolution, and development history. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:7032–7056.

- Qureshi, R., Uzair, M., Khurshid, K., e Yan, H. (2019). Hyperspectral document image processing: Applications, challenges and future prospects. *Pattern Recognition*, 90:12–22.
- Rafaele Furtado, L., Moraes, R., e Rodrigues, R. (2020). Análise espacial de uma pequena bacia hidrográfica na amazônia. *Journal of Hyperspectral Remote Sensing*, 10:95.
- Redmon, J., Divvala, S., Girshick, R., e Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788.
- Roger e Clark (2002). *Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy*. New York, John Wiley and Sons, pages 3–58.
- Shapiro, S. S. e Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4):591–611.
- Shortis, M., Seager, J., Harvey, E., e Robson, S. (2005). Influence of Bayer Filters on the Quality of Photogrammetric Measurement. *Videometrics VIII*, 5665:164–171.
- Shukla, A. e Kot, R. (2016). An Overview of Hyperspectral Remote Sensing and its applications in various Disciplines. *IRA-International Journal of Applied Sciences (ISSN 2455-4499)*, 5:85.
- Signoroni, A., Savardi, M., Baronio, A., e Benini, S. (2019). Deep Learning Meets Hyperspectral Image Analysis: A Multidisciplinary Review. *Journal of Imaging*, 5.
- Singh, A. e Singh, H. (2012). Forest Fire Detection through Wireless Sensor Network using Type-2 Fuzzy System. *International Journal of Computer Applications*, 52:19–23.
- Sun, D. W. (2010). *Hyperspectral Imaging for Food Quality Analysis and Control*. San Diego, Academic Press, ISBN 978-0-12-374753-2.
- Sze, S. M. (2002). *Semiconductor Devices Physics and Tech*. New York, John Wiley & Sons, ISBN 978-0470-53794-7.
- Teich, B. S. A. M. (1991). *Fundamentals of Photonics*. New York, John Wiley & Sons, ISBN: 9781119506874.

- Thenkabail, P., Lyon, J., e Huete, A. (2018). *Hyperspectral Remote Sensing of Vegetation*. Florida, CRC Press, ISBN 9780429192180.
- Theuwissen, A. J. P. (2002). *Solid-State Imaging with Charge-Coupled Devices*. New York, Kluwer Academic Publishers, ISBN 978-0792334569.
- Tsividis, Y. (1999). *Operation and Modeling of the MOS Transistor*. EngineeringPro collection. Oxford University Press, ISBN 9780195170146.
- Venâncio, P. V. A. B., Lisboa, A. C., e Barbosa, A. V. (2022). An Automatic Fire Detection System Based on Deep Convolutional Neural Networks for Low-power Resource-Constrained Devices. *Neural Computing and Applications*, 34:15349–15368.
- Veraverbeke, S., Sedano, F., Hook, S., Randerson, J., Jin, Y., e Rogers, B. (2014). Mapping the Daily Progression of Large Wildland Fires Using MODIS Active Fire Data. *International Journal of Wildland Fire*, 23:655–667.
- Vodacek, A., Kremens, R., Fordham, A., Luisi, D., Schott, J., VanGorden, S., e Latham, D. (2001). Spectral Features of Biomass Fires. *Digital Imaging and Remote Sensing Laboratory*. https://www.cis.rit.edu/~rlkpci/PDFs/OPTONe_SPectral.PDF.
- Vodacek, A., Kremens, R., Fordham, A., Vangorden, S., Luisi, D., Schott, J., e Latham, D. (2002). Remote Optical Detection of Biomass Burning Using a Potassium Emission Signature. *International Journal of Remote Sensing - INT J REMOTE SENS*, 23:2721–2726.
- Xu, H., Zhang, G., Zhou, Z., Zhou, X., e Zhou, C. (2022). Forest Fire Monitoring and Positioning Improvement at Subpixel Level: Application to Himawari-8 Fire Products. *Remote Sensing*, 14:2460.

Apêndice A

Base de dados de imagens hiperespectrais FX-10

A base de dados de imagens hiperespectrais FX-10 foi construída com uso da câmera Specim modelo FX-10, sendo composta por 339 imagens de motivos diversos como vegetação do cerrado, eucalipto, fogo, fumaça, frutas, vela, lamparina, sal e referência de branco. As imagens possuem formatos diferentes, uma vez que a câmera, sendo do tipo *line-scan* possibilita definir o ângulo de varredura da cena, o tempo de exposição e a velocidade de varredura. A altura da imagem é fixa com 1024 píxeis. A câmera é controlada e configurada por um software denominado *LUMORECORDER*. A cada imagem capturada, corresponde um conjunto de arquivos armazenados em pastas específicas. Um arquivo *.hdr* contém a descrição dos parâmetros utilizados na captura da imagem, em formato ENVI. Esse arquivo é utilizado para leitura da imagem *.RAW*. Além de arquivos com dados da cena, há também arquivos com a referência de escuro (*.hdr* e *.RAW*) e outro com a referência de branco (*.hdr* e *.RAW*). As pastas também contém arquivos *.xls* e *.Xlm* com metadados das imagens produzidas. A Figura A.1 mostra os metadados de um arquivo *.xlm*. A base é composta por imagens nas quais se capturou, em alguns casos, a mesma cena, com taxa de quadros, velocidade de varredura e tempo de exposição diversos. Essa base de dados está disponível em <https://onedrive.live.com/?id=884F8921FAB5BBF4>

Lumo - Scanner measurement report

Sample name:

Date: Sunday, November 14, 2021

Time: 15:45:51

Error:

Description

fumacae fogo 11h38mi 14-11-2021 sol 30 fps

Camera

Attributes	Value
sensor type	FX10e
sensor id	123456789
frame rate	30.000007629395
integration time	7
samples	1024
active bands	448
spatial binning	1
spectral binning	1
frames recorded	1152
frames dropped	0
data format	BIL
calibration pack	

Metadata

Attributes	Value
operator	
description	fumacae fogo 11h38mi 14-11-2021 sol 30 fps
sensor temperature	52.109741210938
sensor prefix	

Figura A.1: Metadados gerados pelo software de controle da câmera FX-10

Apêndice B

Base de Dados de imagens hiperespectrais Specim IQ

A base de dados de imagens hiperespectrais Specim IQ foi construída com uso da câmera Specim modelo IQ, sendo composta por 280 imagens de motivos diversos como vegetação do cerrado, eucalipto, fogo, fumaça, frutas, vela e referência de branco. As imagens possuem formatos com 512 x 512 píxeis. É possível ajustar o tempo de integração, sendo a velocidade e ângulo de varredura fixos. A cada imagem capturada, corresponde um conjunto de arquivos armazenados em pastas específicas. Um arquivo .hdr contém a descrição dos parâmetros utilizados na captura da imagem, em formato ENVI. Esse arquivo é utilizado para leitura da imagem .RAW. Além de arquivos referentes à cena, há também arquivos com a referência de escuro (.hdr e .RAW) e outros com a referência de branco (.hdr e .RAW). As pastas também contém arquivos .Xlm com metadados das imagens produzidas, incluindo coordenadas geográficas (latitude e longitude da posição da câmera), bem como informações sobre reflectância da cena imageada. A base também é composta por 224 imagens capturadas com uso de uma câmera RGB Panasonic DMC-FZ5, correspondentes às mesmas cenas capturadas com a câmera Specim IQ. As câmeras foram montadas em uma estrutura mecânica para fixar a posição e orientação relativa das duas câmeras. Essa base de dados está disponível em <https://onedrive.live.com/?id=884F8921FAB5BBF4>

Apêndice C

Códigos fonte

Os códigos fonte, escritos em Python, utilizados na implementação dos algoritmos dessa tese estão disponíveis em <https://onedrive.live.com/?id=884F8921FAB5BBF4>