



Programa de Pós-Graduação em

Computação Aplicada

Mestrado/Doutorado Acadêmico

Vinicius Ferreira Sales

Aperfeiçoamento do Treinamento de Redes de Super-Resolução
Deep Learning a partir de Imagens Hiperespectrais aprimoradas

São Leopoldo, 2023

Vinícius Ferreira Sales

**APERFEIÇOAMENTO DO TREINAMENTO DE REDES DE SUPER-RESOLUÇÃO
DEEP LEARNING A PARTIR DE IMAGENS HIPERESPECTRAIS APRIMORADAS**

Dissertação apresentada como requisito para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Computação Aplicada da Universidade do Vale do Rio dos Sinos - Unisinos.

Orientador:
Prof. Dr. Luiz Gonzaga da Silveira Junior

Coorientador:
Prof. Dr. Maurício Roberto Veronez

São Leopoldo
2023

S163a

Sales, Vinícius Ferreira.

Aperfeiçoamento do treinamento de redes de super-resolução *deep learning* a partir de imagens hiperespectrais aprimoradas / Victor Manoel Fernandes. – 2023.

103 f. : il. ; 30 cm.

Dissertação (mestrado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, 2023.

“Orientador: Prof. Dr. Luiz Gonzaga Junior

Coorientador: Prof. Dr. Maurício Roberto Veronez.”

1. Computação visual. 2. Deep Learning. 3. Sensoriamento remoto. 4. Super-resolução. I. Título.

CDU 004.4

Dados Internacionais de Catalogação na Publicação (CIP)
(Bibliotecária: Silvana Dornelles Studzinski – CRB 10/2524)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) - Código de Financiamento 001 /"This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

AGRADECIMENTOS

Primeiramente agradeço a Deus, pois se não fosse da vontade dele, hoje esse trabalho não estaria concluído.

Aos meus pais pelos inúmeros esforços para que eu pudesse alcançar os meus objetivos sempre me apoiando e servindo de inspiração.

À minha noiva, pela compressão e paciência ao longo desses anos, que não foram fáceis.

Minha eterna gratidão aos meus amigos Ademir Marques e Túlio Santana, por todo incentivo e apoio prestado a mim ao longo dessa jornada. À Marianne, Graciela, Diego, Marcela, Zanotta, Anderson, Tainá e Alysson, por todos os momentos valiosos que me proporcionaram e pelo apoio oferecido.

Aos Profs. Drs. Luiz Gonzaga Jr. e Mauricio Veronez por aceitarem me orientar no desenvolvimento da dissertação, na produção de artigos, e por ampliar os meus horizontes na minha vida profissional.

Ao Vizlab e à toda equipe do laboratório, por toda infraestrutura que me foi oferecida e pela oportunidade de conhecer e trabalhar com grandes pesquisadores. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES pelo financiamento.

À Petrobras pelo apoio financeiro no desenvolvimento deste trabalho, apoio esse que foi de fundamental importância.

À Universidade do Vale do Rio dos Sinos – Unisinos e ao Programa de Pós-Graduação em Computação Aplicada (PPGCA) por me fornecerem uma formação de qualidade.

RESUMO

Inserida no contexto de Computação Visual e Sensoriamento Remoto, a Super-resolução consiste no processo de restaurar informações de alta frequência em imagens de baixa resolução espacial. Tradicionalmente, esse tipo de técnica busca solucionar às limitações físicas de alguns sensores imageadores em identificar e analisar específicos alvos. Com aumento do uso dos métodos de *Deep Learning* abordagens mais robustas de Super-resolução vem ganhando cada vez mais espaço, como são os casos das redes de S-R baseadas em *Convolutional neural network*. Embora tal abordagem se mostrou superior as técnicas tradicionais de processamento digital de imagens, principalmente em cenas RGB, imagens multiespectrais e essencialmente hiperespectrais carecem de mais atenção pelo método, uma vez que, ao aprimorar as suas resoluções espaciais, a consistência espectral deve ser mantida, fato esse considerado um dos grandes desafios dentro da Super-resolução. Ainda nesse contexto, em consequência da dificuldade de se obter imagens hiperespectrais de baixa e alta resoluções espaciais devidamente registradas, as cenas de baixa resolução espacial são sintetizadas a partir do processo de degradação, reamostragem e ruído das suas correspondentes de alta resolução espacial. Apesar desse fluxo ser comumente adotado, ainda não foi avaliado qual é a real influência que as técnicas de reamostragem exercem sobre os métodos de Super-resolução inteligentes, já que não existe um consenso sobre qual a melhor técnica a ser empregada. Sendo assim, a hipótese é que a identificação da melhor função de reamostragem de HIs de LR possibilita que os modelos de Super-resolução em *Deep Learning* gerem HIs de HR de melhor qualidade. Portanto, esse trabalho tem como objetivo de avaliar funções de reamostragem e identificar a melhor função para ao aperfeiçoamento do treinamento de redes de super-resolução em *deep learning*. Como proposta para este trabalho, dois diferentes *datasets* de imagens hiperespectrais consagrados na literatura foram escolhidos e empregados no processo de sintetização de imagens de baixa resolução espacial. Posteriormente, com os dados gerados foram avaliados os seus comportamentos dentro Super-resoluçãoo melhor modelo de Super-resoluçãoo selecionado com base nos trabalhos relacionados. Essa avaliação foi realizada a partir de diferentes métricas de comparação de imagens hiperespectrais, com destaque para as métricas Peak Signal-to-Noise Ratio, *Spectral Angle Mapper* e *Structure Similarity Index Measurement*. Dos valores obtidos testes de hipótese como são o caso dos teste de Friedman e Nemeyi foram aplicados, de modo a identificar estatisticamente qual a melhor técnica aplicada. Por fim, os resultados obtidos foram confrontados e avaliados a partir de um novo conjunto de dados obtido de forma controlada, de modo que a consistência espectral pôde ser avaliada com base em imagens espectrais de alta resolução preditas e leituras pontuais com espectrômetro não imageador. Dos resultados obtidos, as reamostragem do tipo Lanczos e Cubic apresentaram os melhores resultados em relação as demais, comprovando assim, a hipótese avaliada.

Palavras-chave: Computação Visual. Sensoriamento Remoto. *Deep Learning*. Super-Resolução.

ABSTRACT

Inserted in the context of Visual Computing and Remote Sensing, Super-resolution consists of the process of restoring high frequency information in low spatial resolution images. Traditionally, this type of technique seeks to solve the physical limitations of some imaging sensors in identify and analyze specific targets. With the increasing use of Deep Learning methods more robust approaches of Super-resolution has been gaining more and more space, as are the cases of Super-resolution networks based on Convolutional neural network. Although such an approach proved to be superior to traditional digital image processing techniques, mainly in RGB scenes, multispectral and essentially hyperspectral images need more attention by the method, since, by improving their spatial resolutions, the spectral consistency must be maintained, a fact considered one of the great challenges within the Super-resolution. Still in this context, due to the difficulty of obtaining hyperspectral images of low and high spatial resolutions properly registered, the scenes of low spatial resolution are synthesized from the process of degradation, resampling and noise of their corresponding of high spatial resolution . Although this flow is commonly adopted, it has not yet been evaluated what is the real influence that resampling techniques have on intelligent Super-resolution methods, since there is no consensus on the best technique to be used. Thus, the hypothesis is that the identification of the best resampling function of HIs de LR enables the Super-resolution models in Deep Learning to generate HIs of HR of better quality. Therefore, this work aims to evaluate resampling functions and identify the best function for the improvement of training of super-resolution networks in deep learning. As a proposal for this work, two different datasets of hyperperctral images consecrated in the literature were chosen and used in the process of synthesizing low spatial resolution images. Subsequently, with the data generated, their behaviors were evaluated within the best Super-resolution model selected based on the related works. This evaluation was performed from different metrics of comparison of hyperspectral images, especially the metrics: Peak Signal-to-Noise Ratio , Spectral Angle Mapper e Structure Similarity Index Measurement. From the values obtained hypothesis tests such as the case of the Friedman and Nemeyi tests were applied, in order to identify statistically which technique was best applied. Finally, the results obtained were compared and evaluated from a new set of data obtained in a controlled way, so that the spectral consistency could be evaluated based on predicted high resolution spectral images and point readings with non-iImageer spectrometer. From the results obtained, the resampling type Lanczos and Cubic presented the best results in relation to the others, thus proving, the hypothesis evaluated.

Keywords: Imagens Espectral.

LISTA DE FIGURAS

1	Modelo Ondulatório da REM	21
2	Espectro Eletromagnético	22
3	Combinação entre bandas espectrais	23
4	Comparação entre imagens Multiespectrais e Hiperespectrais	24
5	Visualização e Comparação entre as Resoluções Espaciais de diferentes sensores espaciais	26
6	Relação entre inteligência artificial, machine learning e deep learning	27
7	Operação de Convolução	29
8	Exemplo de aplicação das funções <i>Max pooling</i> , <i>Min pooling</i> e <i>Average pooling</i>	30
9	Visualização e Comparação entre os tipos de convoluções	31
10	Exemplo de Bloco Residual.	33
11	Tipos de Mecanismos de Atenção.	34
12	Mecanismos de Atenção espacial e por canal	35
13	Fluxograma de Investigação dos Trabalhos Relacionados	40
14	Resultados das Etapas de Busca dos Trabalhos Relacionados	43
15	Divisão entre dados de treinamento e teste	52
16	Comparação Qualitativa	56
17	Fluxograma da metodologia empregada	59
18	Elaboração dos Dados de Treinamento	64
19	Estruturas da Rede ERCSR	65
20	Arquitetura completa do método de S-R ERCSR	66
21	Validação Cruzada do tipo K-fold	67
22	Resultado do Imageamento Hiperespectral	70
23	Aquisição das Assinaturas Espectrais	71
24	Avaliação Estatística dos Métodos de Reamostragem para dataset de Chikusei	74
25	Comparação do erro médio entre as imagens de HR preditas e sua verdade de campo para o dataset de Chikusei	75
26	Aplicação da Técnica de S-R para o dataset Chikusei	75
27	Avaliação Estatística dos Métodos de Reamostragem para dataset de Harvard	76
28	Comparação do erro médio entre as imagens de HR preditas e sua verdade de campo para o dataset de Harvard	77
29	Aplicação da Técnica de S-R para o dataset Harvard	77
30	Avaliação Estatística dos Métodos de Reamostragem para dataset Unisinos	79
31	Comparação entre a imagem original e as imagens preditas	80
32	Comparação qualitativa entre as assinaturas espectrais coletadas in-loco com as assinaturas espectrais preditas	81

LISTA DE TABELAS

1	Consulta realizadas nas bases de dados juntamente com seu período de filtragem	39
2	Consulta realizadas nas bases de dados juntamente com seu período de filtragem	41
3	Classificação das Questões de Avaliação em Relação a média	42
4	Síntese dos Algoritmos empregados nos trabalhos analisados	47
5	Lista dos datasets empregados e sua frequência	50
6	Lista dos diferentes tipos de métricas empregadas	53
7	Métodos de super-resolução de HIs	64
8	Comparação entre as técnicas de reamostragem para o dataset de Chikusei .	73
9	Comparação entre as técnicas de reamostragem para o dataset de Chikusei .	76
10	Comparação entre as técnicas de reamostragem para o dataset Unisinos . . .	79
11	Comparação quantitativa entre as assinaturas coletadas em campo e as assinaturas preditas	81
12	Produção científica autoral e participação em artigos.	103

LISTA DE SIGLAS

QA	Avaliação da qualidade do Estudo
FAPERGS	Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul
HI	Imagem Hiperespectral
MI	Imagem Multiespectral
DL	<i>Deep Learning</i>
QP	Questão de Pesquisa
S-R	Super-Resolução
R	Rótulo
μm	Micrômetro
nm	Nanômetro
E	Campo Elétrico
M	Campo Magnético
C	Velocidade da Luz
IFOV	Campo de Visada Instantâneo
DBN	<i>Autoencoders, Deep Belief Networks</i>
CBAM	<i>Generative Adversarial Networks</i>
IGARSS	<i>International Geoscience and Remote Sensing Symposium</i>
CV	Validação Cruzada
CD	Distância Crítica
RMS	<i>Root Mean Square</i>
PSNR	<i>PeakSignal – to – NoiseRatio</i>
SAM	<i>Spectral Angle Mapper</i>
SSIM	<i>Structure Similarity Index Measurement</i>
ERGAS	<i>Erreur Relative Globae Adimensionnelle de Synthèse</i>
CC	<i>Correlation Coefficient</i>

SUMÁRIO

1	INTRODUÇÃO	17
2	FUNDAMENTAÇÃO TEÓRICA	21
2.1	Imageamento Espectral	21
2.1.1	Resoluções das Imagens	23
2.2	Deep Learning	26
2.2.1	Redes Neurais Convolucionais - CNN	28
2.3	Super-Resolução em HI a partir de técnicas de Deep Learning	34
3	TRABALHOS RELACIONADOS	37
3.1	Questões de Pesquisa	37
3.2	Planejamento de Pesquisa	38
3.3	Seleção de Estudos	39
3.3.1	Critérios de Seleção e Exclusão	39
3.3.2	Avaliação da qualidade do Estudo	39
3.3.3	Resultados da Seleção dos Estudos	40
3.4	Extração das Informações	42
3.5	Síntese dos Trabalhos	44
3.5.1	Informações gerais sobre os artigos selecionados	44
3.5.2	Principais problemáticas abordadas e suas soluções	44
3.5.3	Tipos de algoritmos utilizados	46
3.5.4	Dataset empregados	49
3.5.5	Fluxo de treinamento e Validação	51
3.5.6	Métricas de validação	53
3.5.7	Avaliação da resposta espectral	55
3.6	Conclusão	56
4	MATERIAL E MÉTODOS	59
4.1	Formulação e Avaliação da melhor função de Reamostragem de HI de Baixa Resolução	60
4.1.1	Técnicas de Reamostragem	60
4.1.2	Aquisição de dados	62
4.1.3	Modelos de Super-Resolução por <i>Deep Learning</i>	63
4.1.4	Treinamento e Validação	66
4.2	Análise e Validação da consistência espectral	69
5	RESULTADOS E DISCUSSÕES	73
5.1	Avaliação das técnicas de Reamostragem	73
5.1.1	Dataset de Chikusei	73
5.1.2	Dataset de Harvard	75
5.2	Avaliação da Consistência Espectral das técnicas de Reamostragem Selecionadas	78
5.2.1	Dataset Unisinos	78
6	CONSIDERAÇÕES FINAIS	83
6.1	Limitações do Trabalho	84
6.2	Trabalhos Futuros	84
	REFERÊNCIAS	85

APÊNDICE A – MONOGRAFIA DE ASSINATURA ESPECTRAL	93
APÊNDICE B – RESULTADOS OBTIDOS NA VALIDAÇÃO CRUZADA	99
APÊNDICE C – PRODUÇÃO CIENTÍFICA	103

1 INTRODUÇÃO

Com a grande necessidade de conhecer, em detalhes, o espaço onde se habita, seja para fins de exploração ou ocupação, a Ciência Cartográfica surge como arte, ciência e tecnologia capaz de propor métodos e conceitos para que o homem consiga representar o meio em que vive, para fins de planejamento, ensino e localização (TAYLOR, 1991).

A Ciência Cartográfica dispõe de fluxos de trabalho que vão desde a aquisição da informação até a sua representação, seja por meios analógicos ou digitais, em mapas bidimensionais (2D) ou tridimensionais (3D), conhecidos como modelos realísticos. As formas de aquisição, processamento e representação das mais variadas informações disponíveis sobre a superfície terrestre incluem Geodésia, Topografia, Fotogrametria, Sensoriamento Remoto e Visão Computacional (KRAAK, 1998).

Dentre os métodos de mapeamento, o Sensoriamento Remoto desponta como a técnica mais bem sucedida na coleta automática de informações (SCHOWENGERDT, 2006; MENESES; ALMEIDA, 2012). A perfeita combinação entre os sensores imageadores acoplados em sistemas de monitoramento como satélites, aeronaves, e atualmente os veículos aéreos não tripulados (VANTs), têm proporcionado o mapeamento de vastas áreas de recursos naturais e artificiais (HAUT et al., 2018). Com este respectivo avanço, um aumento na disponibilidade de imagens com distintas características, com destaque às imagens hiperespectrais (*hyperspectral images* - HIs), vem sendo utilizado para a observação da Terra e seus recursos (ZHENG et al., 2019).

Ao contrário dos olhos humanos, que só podem ser expostos à luz visível, as HIs são empregadas na coleta e processamento de informações presentes em várias faixas contínuas do Espectro Eletromagnético (EE). A sua análise tornou-se uma área de pesquisa próspera e ativa em Sensoriamento Remoto e Visão Computacional com uma ampla gama de aplicações (HAUT et al., 2018; LI et al., 2017), que incluem, por exemplo, reconhecimento e classificação de objetos (FAUVEL et al., 2012; CHEN et al., 2016; GHAMISI et al., 2017), monitoramento urbano e ambiental (STUART; MCGONIGLE; WILLMOTT, 2019; SCAFUTTO; SOUZA FILHO; OLIVEIRA, 2017), detecção mineralógica (KUMAR; YARRAKULA, 2017; PEYGHAMBARI; ZHANG, 2021) entre outros. Comparado com imagens multiespectrais (*multi-spectral image* - MI) ou imagens convencionais em cores (RGB), a HI é definida como um cubo de dados 3D que consiste em um conjunto de imagens 2D, de modo que suas centenas de bandas, uma contígua às outras dentro do EE, forneçam detalhes mais precisos de assinatura espectral em relação a diferentes materiais (QU; QI; KWAN, 2018).

Embora imagens com variações espectrais muito altas possibilitem a discriminação precisa de diferentes objetos por suas assinaturas espectrais, entretanto devido à limitação das tecnologias dos sensores, somente uma pequena fração da energia radiante pode ser coletada para cada banda de espectro. Logo, para atingir uma relação sinal-ruído aceitável, a área do sensor por pixel deve ser grande, levando a uma resolução espacial¹ mais grosseira (WANG; SHI, 2013;

¹Índice que determina qual é o menor tamanho de objeto registrado em uma imagem.

LI; ZHANG; YOU, 2019; DOU et al., 2020). Consequentemente, as imagens hiperespectrais são frequentemente adquiridas sob uma resolução espacial relativamente baixa, degradando seu desempenho em aplicações práticas de identificação de objetos no solo com precisão aceitável, como é o caso de mineralogia, manufatura e vigilância (HE; LIU, 2018; LI; ZHANG; YOU, 2019; TSAGKATAKIS et al., 2019; DOU et al., 2020). Este cenário é o oposto das câmeras multiespectrais, onde grande parte da informação espectral é sacrificada para alcançar alta resolução espacial, integrando a radiância da cena em menos bandas espectrais muito mais amplas (HE; LIU, 2018). Como meio de contornar o problema de baixa resolução em imagens hiperespectrais, abordagens baseadas em pós-processamento são bastante convenientes, como é o caso da super-resolução (S-R) (MEI et al., 2017; LI et al., 2019; DOU et al., 2020; ZHAO et al., 2021; HU et al., 2021).

Super-resolução consiste em um conjunto de técnicas que buscam aprimorar os detalhes minuciosos da imagem, aumentando a resolução espacial. Esse método supera os problemas das técnicas convencionais (Fusão, Análise Subpixel, entre outras) de aprimoramento de resolução, como introdução de ruído, distorção espacial e falta de clareza nos detalhes da imagem (ROY et al., 2014; HAYAT, 2018; ZHENG et al., 2019; ROHITH; KUMAR, 2020; RAN et al., 2020). A S-R pode ser definida como um método baseado em Computação Visual que busca recuperar uma imagem de alta resolução espacial capturada com uma grande densidade de pixels (*high resolution* - HR de uma imagem ruidosa, borrada e de baixa resolução espacial (*low resolution* - LR) observada (DONG et al., 2014; HU et al., 2017; MEI et al., 2017).

As primeiras abordagens pertinentes aos métodos de S-R são oriundas da década de 70. Porém foi com o recente surgimento das Redes Neurais Convolucionais - (*Convolution neural network* - CNN), que os métodos de S-R realmente alcançaram bom desempenho, especialmente no contexto de imagens convencionais (Imagens RGB) (LIU; LEE, 2019; ROHITH; KUMAR, 2020). A teoria básica presente na definição e construção dos respectivos modelos de S-R pressupõe que imagens LR sejam usadas como entrada para os modelos inteligentes, e os fragmentos correspondentes de referência (imagens HR) são usados como saída para o treinamento (NIKONOROV et al., 2018). Desse modo, é possível aprender uma função de mapeamento entre imagens LR e HR de forma de semi-supervisionada (QU; QI; KWAN, 2018).

Contudo, em função da dificuldade em se obter um conjunto de HIs robusto, com imagens de LR e HR devidamente registradas, e por se tratarem de um processo moroso e oneroso, a maioria das abordagens de S-R propõe que as imagens de LR sejam sintetizadas, de modo que a partir de processos de degradação, reamostragem e a adição de ruídos é possível obter uma imagem LR sintética (QU; QI; KWAN, 2018; LIU; LEE, 2019; DOU et al., 2020).

Apesar dos trabalhos em sua totalidade, conforme observado ao longo da revisão bibliográfica, adotarem o fluxo mencionado anteriormente, de forma que consigam obter uma função de mapeamento não linear entre HIs de LR e HR, este processo não reflete a situação do mundo real. As imagens reais geralmente contém estruturas complexas que envolvem fatores aleatórios, como, vibração do sensor e ruídos que são desconhecidos. Portanto, ao reduzir as imagens

de HR para LR, sem o conhecimento prévio do sensor, parte dessas informações estão sendo negligenciadas do modelo. Logo, modelos preditivos que não consideram respectivas informações não conseguem generalizar adequadamente ou obter resultados agradáveis (ZHANG et al., 2020). Além disso, específicos processos de reamostragem são irreversíveis (HU; LI; XIE, 2017), de modo que métodos de S-R não obtenham resultados fidedignos.

Ainda dentro do contexto da problemática supracitada, a S-R em HI não deve considerar apenas as perdas espaciais, mas também preservar a consistência espectral, tida como um dos maiores gargalos dentro das técnicas de S-R em imagens multi e hiperespectrais (ZHENG et al., 2018; HANG; LIU; LI, 2021). Diante do exposto, este trabalho aborda a seguinte questão de pesquisa: como aprimorar as técnicas de S-R em imagens hiperespectrais baseadas em *Deep Learning*, de modo que a consistência espectral seja preservada.

Diante da questão de pesquisa apresentada, a motivação para este trabalho é analisar como diferentes métodos de reamostragem de HI de LR espacial influenciam os modelos de S-R baseados em *Deep Learning* ao estimar HI de HR de melhor qualidade e logo, poder indicar o melhor método dentro de um conjunto específico de abordagens. Nesse sentido, a hipótese de pesquisa é: a identificação da melhor função de reamostragem de HIs de LR possibilita que os modelos de S-R em *Deep Learning* gerem HIs de HR de melhor qualidade.

A partir da hipótese apresentada, o objetivo do trabalho é:

- Avaliar funções de reamostragem e identificar a melhor função para o aperfeiçoamento do treinamento de redes de super-resolução em *deep learning*.

Para a execução do objetivo geral, foram estabelecidos os seguintes objetivos específicos:

- Gerar diferentes conjuntos *datasets* de imagens hiperespectrais de alta e baixa resoluções, de modo que cada conjunto empregue uma função específica de reamostragem;
- Identificar o modelo de super-resolução em imagens hiperespectrais com maior destaque na literatura, com base na revisão sistemática desenvolvida;
- Analisar os diferentes resultados gerados a partir das métricas de validação de imagens e testes estatísticos de modo a identificar a melhor função de reamostragem;
- Criar um novo *dataset* de HIs, com assinaturas espectrais geolocalizadas na mesma área imageada;
- Avaliar a consistência espectral gerada pelas imagens hiperespectrais de alta resolução preditas pelo modelo selecionado, e com auxílio da melhor função de reamostragem, por meio das assinaturas espectrais geolocalizadas em solo.

Este trabalho se destaca pelas seguintes contribuições:

- Ser pioneiro em investigar as técnicas de reamostragem de imagens de LR no contexto de S-R, utilizando para isso imageamento hiperespectral associado às assinaturas espectrais coletadas diretamente no solo.
- Propor a construção de um novo *dataset* de HIs georreferenciadas obtido de forma controlada associado com assinaturas hiperespectrais geolocalizadas.
- Realizar uma extensa avaliação experimental para averiguar o desempenho da melhor abordagem de reamostragem, empregando para isso a utilização de diferentes índices comparativos de HIs associados com testes estatísticos;

Esta proposta apresenta: uma revisão de literatura com conceitos sobre imageamento espectral, conceitos de *Deep Learning*, mais precisamente sobre redes convolucionais e suas especificidades e conceitos de super-resolução a partir de técnicas de *Deep Learning* (Capítulo 2); uma revisão bibliográfica baseada em busca sistemática (Capítulo 3); a descrição dos métodos e materiais propostos (Capítulo 4); resultados preliminares e discussões do que foi desenvolvido para esta proposta (Capítulo 5); e cronograma até a conclusão do mestrado (Capítulo 6).

2 FUNDAMENTAÇÃO TEÓRICA

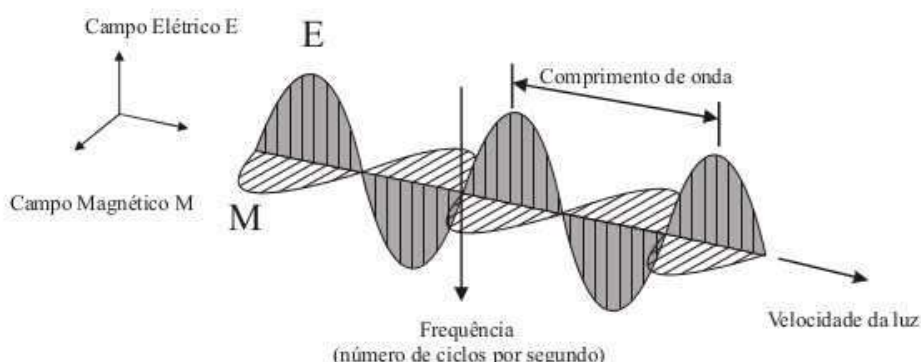
Este capítulo constitui o embasamento teórico que respalda a construção do presente trabalho. Na seção 2.1 serão destacados alguns tópicos pertinentes ao Sensoriamento Remoto, para posteriormente na seção 2.2 serem apresentados os conceitos de *Deep Learning* empregados nesse trabalho. Por fim, ambos os aspectos contribuíram para introdução dos conceitos de super-resolução em imagens Hiperespectrais tratadas na seção 2.3

2.1 Imageamento Espectral

Historicamente o termo Sensoriamento Remoto é estabelecido como a utilização de sensores com o intuito de estudar o ambiente terrestre por meio da interação entre radiação eletromagnética (REM) e as propriedades físicas ou químicas da superfície, de maneira que não ocorra o contato físico entre sensor e objeto (SCHOWENGERDT, 2006; MORAES NOVO, 2010; KHORRAM et al., 2012; MENESES; ALMEIDA, 2012).

O entendimento do processo interativo entre a REM e as propriedades físicas e químicas da superfície terrestre fundamenta-se, a princípio, no entendimento da natureza da REM. Formulada pelos modelos ondulatórios e corpuscular, a radiação eletromagnética como a luz solar propaga-se no espaço em forma de onda e energia ao mesmo tempo (MENESES; ALMEIDA, 2012). Enquanto o modelo corpuscular trabalha com os conceitos de pulsos de energia, o modelo ondulatório, com o respaldo na teoria das ondas, descreve a REM (Figura 1) como um campo elétrico (E) de várias magnitudes em direção perpendicular e transversal ao campo magnético (M) que propaga pelo espaço na velocidade da luz (c).

Figura 1: Modelo Ondulatório da REM



Fonte: Retirado de Meneses e Almeida (2012)

Essa dualidade do comportamento da REM se manifesta em função da interação da energia com a matéria. Caso as partículas se encontrem na escala microscópica, a energia incidente representada em forma de pulso, ao tocar sobre a superfície, promove que parte dessa energia incidida seja absorvida, enquanto o restante é refletida para atmosfera, de modo que modelo

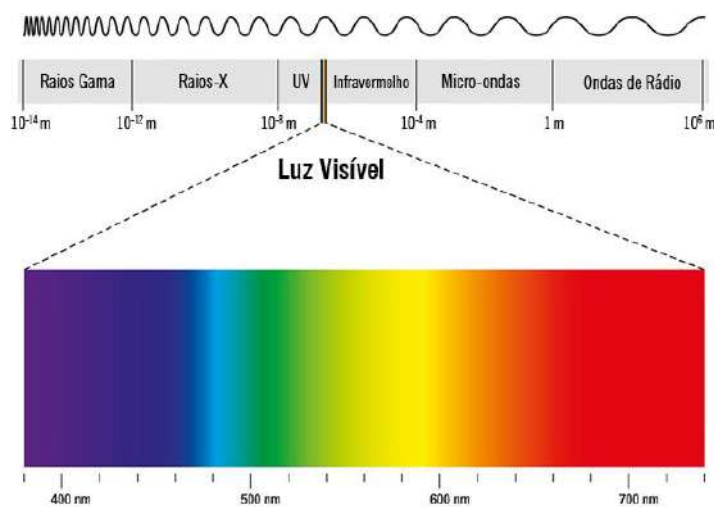
corpúscular seja representado nesse tipo de modalidade (CRACKNELL, 2007)

Em oposto a incidência da REM em escala microscópica, o modelo ondulatório está intrinsicamente ligado com a relação do comprimento de onda (λ) e o tamanho da partícula. Quando as partículas dos objetos detêm de tamanhos superiores a específicos λ incidida pela REM, tais comprimentos de onda tendem a serem refletidos de volta ao sensor, enquanto os outros são irradiados através das partículas.(CRACKNELL, 2007)

Na aquisição das informações pelos sensores, o processo de manifestação entre a matéria e a radiação eletromagnética é extremamente importante que se verificar a qualidade e intensidade irradiada pela fonte de REM (SCHOWENGERDT, 2006; KHORRAM et al., 2012; MENESSES; ALMEIDA, 2012). Todos os corpos no mundo acima de zero Kelvin ou $-273,15$ graus Celsius($^{\circ}\text{C}$) emitem radiação, porém no sistema planetário da Terra, o Sol é a principal fonte natural de radiação devido a sua alta temperatura de superfície, que se próxima da faixa dos 6000°C (MORAES NOVO, 2010).

Ao tocar na superfície terrestre, a REM emitida pelo Sol, proporciona que diferentes objetos na superfície e na atmosfera reflitam, absorvam ou transmitam Energia Eletromagnética (EE) em diferentes proporções, de modo que tais diferenças permitem que esses componentes sejam identificados (MATHER; KOCH, 2011). Logo, sensores acoplados em VANTs, aeronaves ou plataformas de satélites registram a magnitude do fluxo de energia refletido ou emitido por objetos. No Sensoriamento Remoto as faixas do EE mais empregadas são as faixas entre o intervalo do Visível e o Micro-Ondas (Figura 2).

Figura 2: Espectro Eletromagnético



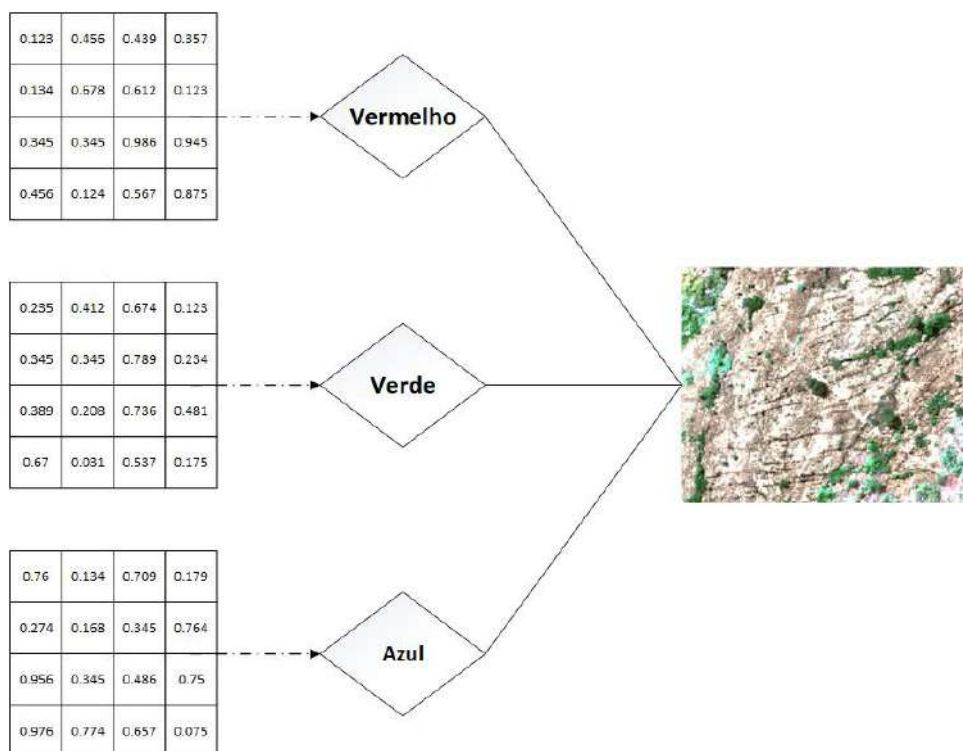
Fonte: Retirado de Moraes Novo (2010)

2.1.1 Resoluções das Imagens

Ao conhecer os principais comprimentos de ondas dentro do Sensoriamento Remoto aliado as suas distintas interações com o objeto, corrobora para que se possa definir as características de aquisição espectral do sensor (TEMPFLI et al., 2009). Essa aquisição se sucede de forma que cada Banda Espectral (ou canal da imagem), seja criada a partir da medição da intensidade da radiação que deixa a superfície dos objetos correspondente a um intervalo pré-definido de comprimentos de onda (KHORRAM et al., 2012). Cada intervalo, presente dentro do EE, auxilia para nomear as Bandas Espectrais, como é o caso das Bandas do Vermelho, Verde e Azul.

A imagem correspondente á Banda Espectral, por sua natureza digital, é composta por um arranjo de elementos em forma de *grid* regular retangular, de modo que cada célula do *grid(pixel)* é indexada por linha e coluna e corresponde a um valor de intensidade da radiação obtida da área imageada do objeto (CRACKNELL, 2007). Embora a Banda espectral seja formada em escala de cinza, a combinação de três bandas espectrais apresentadas nos componentes básicos de cor aditiva (vermelho, verde e azul), resulta na criação na imagem colorida (Figura 3), permitindo assim, a visualização de informações espectrais não verificadas na faixa do visível (SCHOWENGERDT, 2006).

Figura 3: Combinação entre bandas espectrais



Fonte: Adaptado de Mather e Koch (2011)

Além dos intervalos espectrais, características como o tamanho da área unitária imageada, os diferentes níveis de intensidade capturada dentro do intervalo espectral e a frequência temporal

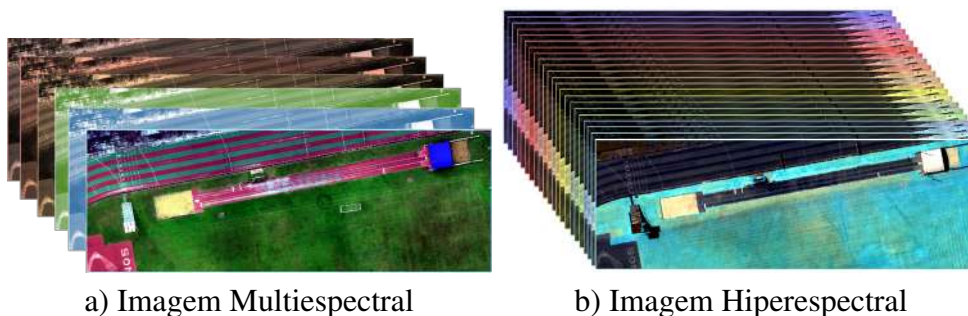
que essas informações são coletadas, são de suma importância para definir o tipo de aplicações no sensoriamento remoto (TEMPFLI et al., 2009). Essas quatro características são descritas em forma de resoluções respectivamente denominadas de resolução espectral, espacial, radiométrica e temporal (MORAES NOVO, 2010; MENESES; ALMEIDA, 2012).

2.1.1.1 Resolução Espectral

É o tipo de resolução que possibilita distinguir diferentes objetos entre si. A sua análise leva em conta três parâmetros de medidas que são o número de bandas que o sensor possui, a largura em comprimento de onda das bandas e as posições que as bandas estão situadas no EE (MENESES; ALMEIDA, 2012). Ao considerar o número de bandas espectrais que o sensor possui, contribui para definir quanto ao tipo de sensor, seja ele Multiespectral ou Hiperespectral.

Os sensores Multiespectrais estão capacitados a registrar imagens relativamente com poucas bandas espectrais, com uma variação entre 3 a 20 bandas, de maneira que tais bandas podem ou não ser contíguas umas às outras enquanto os valores de intensidade nos comprimentos de onda são aferidos de forma discreta (CRACKNELL, 2007). Em oposição a essas características, os sensores Hiperespectrais se destacam por contemplarem um maior número de bandas de maneira contínua dentro do EE, oferecendo assim, para a medição do espectro contínuo ou assinatura espectral do objeto em análise (CRACKNELL, 2007)

Figura 4: Comparação entre imagens Multiespectrais e Hiperespectrais



Fonte: Elaborado pelo Autor

2.1.1.2 Resolução Espacial

Em complemento aos tipos de resolução, a resolução espacial é representada como o menor tamanho de objeto registrado em uma imagem, a sua aplicação em Sensoriamento Remoto está diretamente relacionada com o nível de detalhamento que se deseja obter. Dessa forma, um objeto só pode ser detectado, quando o seu tamanho é, no mínimo, igual ou maior do que o tamanho do elemento de resolução do terreno, ou seja, a resolução espacial (MENESES; ALMEIDA, 2012).

A estimativa da resolução espacial está relacionada diretamente ao campo de visada instantâneo do sensor (IFOV), que representa o tamanho da área, mensurado pelo ângulo β , a qual o sensor recebe a energia em um dado instante de tempo (MATHER; TSO, 2016). Com base na estimativa de abertura do ângulo β em miliradianos mais a distância em metros do sensor em relação ao objeto imageado, a resolução espacial poder ser obtida pela equação 2.1.

$$R.Espacial = H * \beta \quad (2.1)$$

Para aquisição de altos detalhes espaciais nas imagens é necessário um pequeno IFOV, pois garante assim, uma maior capacidade de se detectar pequenos elementos que integram a cena. Contudo, reduzir a resolução espacial em termos de estreitamento do IFOV implica em diminuir a quantidade de energia recebida por um sensor, uma vez que com a redução do ângulo β , acarreta uma restrição na sensibilidade do instrumento em mudanças nos níveis de energia. Portanto, sensores com pequenas variação de IFOV podem degradar a resolução espectral, por resultarem no aumento da relação sinal-ruído (MATHER; TSO, 2016).

Embora tal diminuição na relação sinal-ruído possa ser compensada com o aumento da largura do comprimento da banda, isso ocasionará na redução espectral. Consequentemente o aprimoramento da resolução espacial e espectral não podem ser alcançados em conjunto (MATHER; TSO, 2016).

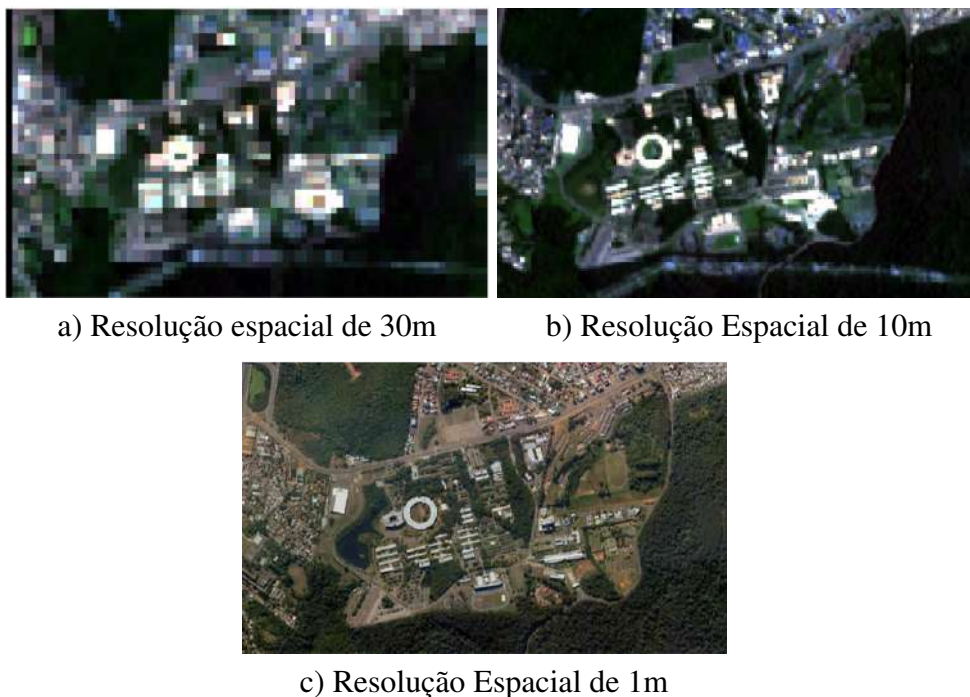
Para ilustrar as possíveis diferenças de resolução espacial, a Figura 5, mostra três imagens adquiridas por diferentes sensores na Universidade do Vale do Rio dos Sinos. Na imagem a) representa um exemplo do sensor OLI do Satélite 8, a cena b) corresponde ao imageamento feito pelo satélite Sentinel 2A e a imagem c) do Satélite Geoeye-1.

2.1.1.3 Resolução Radiométrica e Temporal

Em função da variabilidade dos níveis de intensidades existentes em um único intervalo de comprimento de onda, o termo resolução radiométrica surge como medida capaz de distinguir tais diferenças. Essa quantização ocorre de forma que a intensidade da radiação captada no sensor é convertida eletronicamente em um número digital discreto (DN), de modo que esse número, formulado por dígitos binários, represente um valor específico de intensidade mensurado em uma faixa de comprimento de onda (SCHOWENGERDT, 2006; KHORRAM et al., 2012). A resolução radiométrica mais empregada nos sensores espectrais é a resolução de 8 bits, que corresponde á variação de 256 diferentes tipos de intensidades (MENESES; ALMEIDA, 2012). Entretanto, com o avanço tecnológico no S-R valores superiores a 8 bits estão sendo propostos como é o caso do sensor OLI do Landsat 8 que trabalha com uma resolução radiométrica de 12 e 16 bits (ROY et al., 2014).

Diferente das outras resoluções, a resolução temporal está diretamente ligado ao espaço de tempo que o sensor demora para imagear a mesma área na superfície consecutiva. Esse tipo de resolução é amplamente empregada para estudos em séries temporais que resulta em acom-

Figura 5: Visualização e Comparação entre as Resoluções Espaciais de diferentes sensores espaciais



Fonte: Elaborado pelo Autor

panhar ou detectar a evolução de mudanças que ocorrem na Terra (MENESES; ALMEIDA, 2012)

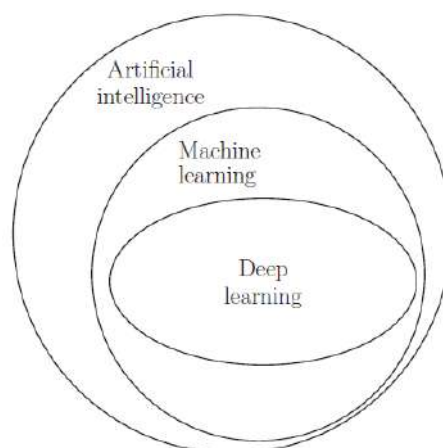
2.2 *Deep Learning*

Deep Learning é um subgrupo dentro do contexto de *Machine Learning* (Figura 6) que foca na criação de modelos robustos de redes neurais, que são capazes de identificar e extrair padrões de grandes conjuntos de dados, além de mapear com precisão entradas de informações complexas, para uma boa tomada de decisão (KELLEHER, 2019)

De modo geral, todos métodos de *Deep Learning* podem ser classificados em três categorias distintas, que são Redes Neurais Convolucionais (CNNs), Redes Não Supervisionadas Pré-treinadas (PUNs) e Redes Neurais Recorrentes/Recursivas (RNNs) (PEDRYCZ; CHEN, 2020). Contudo, antes de adentrar nos tipos de categoria é necessário conhecer os tipos de treinamento para modelos de *Deep Learning*.

Basicamente existem três tipos de treinamentos, que são o aprendizado supervisionado, o não supervisionado e o semi-supervisionado. No treinamento baseado no aprendizado supervisionado, o modelo é alimentado com um conjunto de dados de treinamento contendo tanto as observações quanto os seus resultados correspondentes, de modo que o modelo ao final, consiga inferir por mapeamento matemático, a classificação de futuros pontos de entrada no modelo (RASCHKA; MIRJALILI, 2019; PEDRYCZ; CHEN, 2020)

Figura 6: Relação entre inteligência artificial, machine learning e deep learning



Fonte: Retirado de Kelleher (2019)

Em oposto ao método supervisionado, no aprendizado não supervisionado o modelo é alimentado apenas com os dados de treinamento, que posteriormente são categorizados a partir de características correspondentes entre eles, de maneira que é possível classificar os dados de entrada com base na aparência dos recursos extraídos (KELLEHER, 2019; PEDRYCZ; CHEN, 2020)

Em consonância as duas abordagens supracitadas, o aprendizado por meio semi-supervisionado herda as propriedades de ambos os métodos. Um conjunto de dados semi-supervisionados contém principalmente dados de treinamento não classificados, simultaneamente com pequenas quantidades de dados de classificação. Essa junção corrobora para que duas vantagens importantes sejam destacada nesse tipo de aprendizado, sendo elas uma maior precisão em relação aos modelos não supervisionados com a adição de alguns pontos de dados classificados e significativamente menos trabalhosos e demorados em comparação com o aprendizado supervisionado (PEDRYCZ; CHEN, 2020).

Por meio das abordagens de aprendizagem mencionadas, é possível treinar e adaptar os modelos para que se possam solucionar problemas considerados complexos. Dentre, os tipos de modelos categorizados existentes em *Deep Learning*, as CNNs são definidas como um tipo de rede neural artificial, projetada para trabalhar com entradas estruturadas em forma de *grid*, que contêm fortes dependências espaciais entre específicas regiões do *grid* (AGGARWAL et al., 2018). Um exemplo claro para esse tipo de estrutura são as Imagens 2D, de modo que a sua dependência está ligada com a cor dos vizinhos adjacentes (GOODFELLOW; BENGIO; COURVILLE, 2016).

Outra categoria também muito importante no contexto de *Deep Learning*, são as PUNs, esse tipo de rede foi desenvolvido no intuito de solucionar a falta de dados no treinamento dos modelos. Essa técnica baseia-se no aprendizado das camadas ocultas de um rede neural a partir do método não supervisionado, no intuito de obter um ajuste mais preciso do conjunto de dados (PEDRYCZ; CHEN, 2020). Tipos de PUNs incluem *Autoencoders*, *Deep Belief Networks*

(DBN) e Generative Adversarial Networks (GAN).

Por último mas não menos importantes, as RNNs são redes adaptadas para o processamento de dados sequenciais. Estruturada com apenas uma única camada oculta e um *buffer* de memória, as RNNs é adaptada para que o *buffer* armazene os resultados da camada oculta para posteriormente alimentá-la de volta com os resultados armazenados juntamente com os novos dados de entrada (KELLEHER, 2019). Esse fluxo permite a rede manter uma memória do que viu anteriormente na sequência para ajudá-la o que fazer com a camada atual (AGGARWAL et al., 2018). Esse tipo de rede é usado para reconhecer padrões em sequência de dados, como texto, genomas, caligrafia, palavra falada ou dados de séries numéricas (CHAKRABARTI; ZICKLER, 2011)

Embora os métodos de *Deep Learning* mencionados anteriormente apresentem o mesmo nível de significância, está previsto para essa pesquisa apenas o uso das CNNs e arquiteturas do tipo Autoencoders, para mais informações sobre as demais abordagens sugere-se as seguintes bibliografias: Goodfellow, Bengio e Courville (2016),Skansi (2018),Aggarwal et al. (2018) e Pedrycz e Chen (2020).

2.2.1 Redes Neurais Convolucionais - CNN

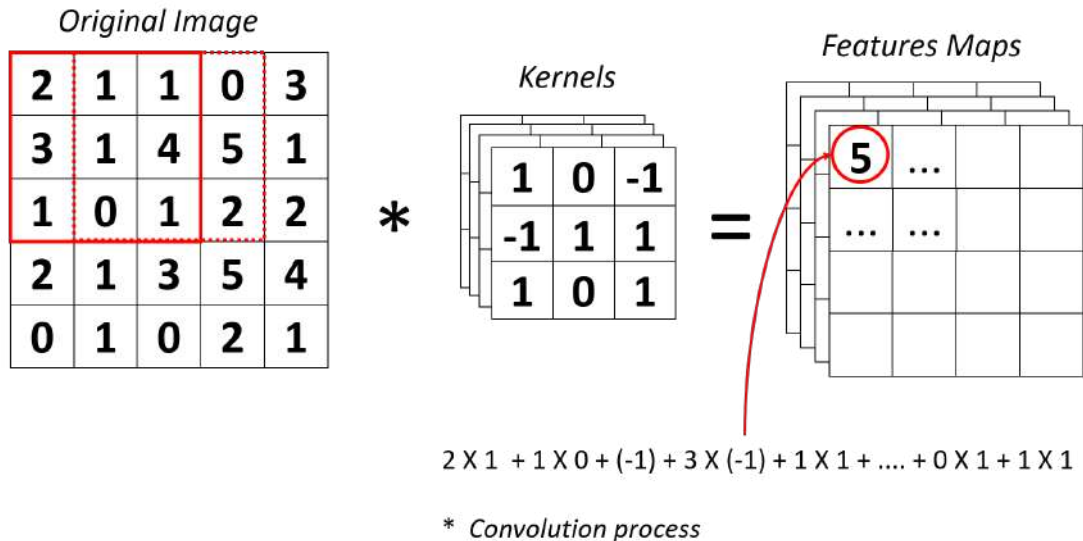
Com base nas inspirações biológicas entre as conexões presentes no córtex visual do cérebro, a CNN é um algoritmo capacitado em extrair de uma imagem de entrada, específicos atributos, e aprender por meio desses recursos, como identificar de forma robusta a presença, ou ausência, de características visuais de uma imagem (AGGARWAL et al., 2018; KELLEHER, 2019; PEDRYCZ; CHEN, 2020).Essa configuração busca reconhecer ou diferenciar específicos recursos, que ocorrem em função das estruturas das CNNs, que são implementadas com conexões esparsas e com alto nível de parâmetros compartilhados entre as camadas (PEDRYCZ; CHEN, 2020).

Cada camada da CNN é uma estrutura em forma de *grid* tridimensional, que possui altura, largura e a profundidade que é representado pelo número de canais/bandas.Por meio dessa estrutura, relações espaciais são herdadas de uma camada para a próxima, uma vez que, a operação de convolução e a transformação para a próxima camada são criticamente dependentes dessas relações (AGGARWAL et al., 2018).

A arquitetura de uma CNN consiste em reduzir a dimensão da imagem em cada camada, a medida que, as características não conhecidas das camadas anteriores são extraídas.Comumente, três tipos de camadas são responsáveis em realizar essa tarefa, que são: a camada de convolução, a camada de *pooling* e a *fully connected layer* (AGGARWAL et al., 2018).A camada de convolução tem como finalidade extrair, por meio de *kernels*/filtros de convolução, as principais propriedades das imagens, de modo que ao final, resulte em *feature maps* que são empilhados e podem servir como entrada para uma próxima camada convolucional ou saída da rede (PEDRYCZ; CHEN, 2020). Na Figura 7 é sintetizado o processo matemático que envolve os

filtros convolucionais. É válido ressaltar que cada *kernel* é responsável em identificar um tipo de feição na imagem, a maneira que ao final cada *kernel* irá produzir um *feature maps* distinto.

Figura 7: Operação de Convolução



Fonte: Elaborado pelo autor

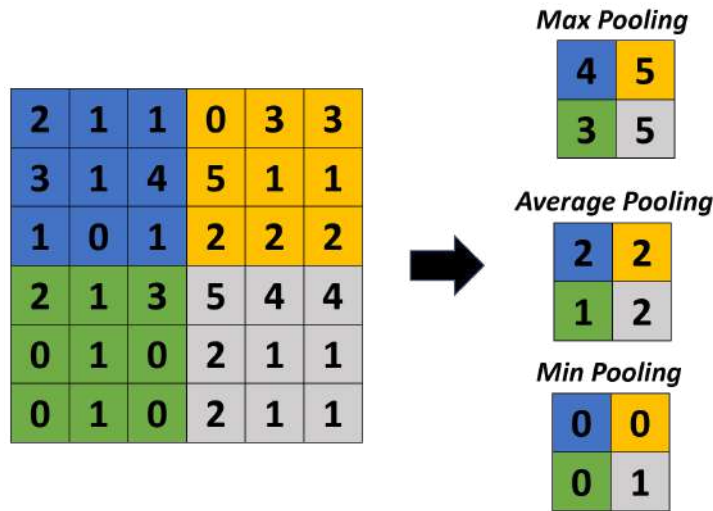
Ainda dentro do contexto dos *Kernels*, duas propriedades são extremamente importantes para que processo ocorra, que são *stride* e o *padding*. O *stride* representa a distância que um *kernels* irá se deslocar de uma posição para a próxima posição, seja ela vertical ou horizontal. Enquanto o *padding*, é empregado no intuito de solucionar a perda da informação nas bordas, a partir da adição de células, que geralmente são preenchidas com valor zero (AGGARWAL et al., 2018).

Após a operação de convolução, os valores obtidos, geralmente passam por uma função de ativação que desempenhará o papel de decidir quais pixels devem ser ativados em uma específica camada. Dentre as diversas funções de ativação, as mais empregadas são ReLU, Tanh e Sigmoid (ZAFAR et al., 2018; SINGHAL, 2017).

Em complemento aos principais tipos de camadas que estruturam uma CNN, a camada *pooling* é definida como o processo que busca reduzir a taxa amostral de um sinal, a fim de reduzir o custo computacional da rede, além de evitar problemas de *overfitting* (KELLEHER, 2019). Essa etapa, geralmente ocorre após a camada de convolução, na qual os *features maps* são reduzidos espacialmente, por meio de um filtro que percorre toda a imagem, de modo a conservar as informações de interesse e gerando novos *features maps*.

Entre os tipos de camada *pooling* os mais básicos são: *Max pooling*, *Max pooling* e *Average pooling* (KELLEHER, 2019). Na função *Max pooling* é armazenado o maior valor dentro do filtro, enquanto *Min pooling* armazena o menor valor e *Average pooling* o valor médio. Na Figura 8 é detalhado um exemplo representativo do funcionamento de cada *pooling* a partir de um filtro 3x3.

Figura 8: Exemplo de aplicação das funções *Max pooling*, *Min pooling* e *Average pooling*



Fonte: Elaborado pelo autor

Por último mas não menos importante, a *fully connected layer* é a última etapa de uma CNN, é o estágio em que cada neurônio na camada anterior está ligado a cada neurônio na próxima camada, seguida pela aplicação da função de ativação, que geralmente é do tipo *softmax* (KELLEHER, 2019). Nessa função, é calculada a distribuição probabilística do dado de entrada em pertencer a uma classe na qual foi treinada. Uma restrição que essa função determina é que os dados de saída sejam normalizados entre os valores 0 e 1 e a soma de cada resultante seja igual a 1 (SEWAK; KARIM; PUJARI, 2018).

É válido salientar que opcionalmente camadas de *Dropout* podem ser empregadas após a camada *pooling*. Essa camada tem como objetivo de reduzir a probabilidade de que problemas de *overfitting* ocorram. O seu funcionamento se baseia em desativar um conjunto de neurônios da *fully connected layer*, a fim de que a rede não fique dependente de um ou mais combinações de neurônios durante o treinamento (ZAFAR et al., 2018; SINGHAL, 2017).

2.2.1.1 CNN 2D e CNN 3D

Entre os tipos de redes neurais de convolução, geralmente a CNN 2D é o tipo de CNN mais usualmente empregada, embora os tipos de CNN 1D e 3D também apresentam bastante relevância em específicas áreas de aplicação (MEI et al., 2017). A principal diferença entre esses tipos de redes é em relação a dimensão do kernel de convolução utilizado. Nas CNNs 2D, previsões podem ser realizadas a partir de *kernels* convolucionais 2D que extraem informação em relação a cada pixel da imagem bruta de entrada, conforme é detalhado na Equação 2.2.

$$c_{xy} = f\left(\sum_{ij} w_{ij} a_{(x+i)(y+j)} + b\right) \quad (2.2)$$

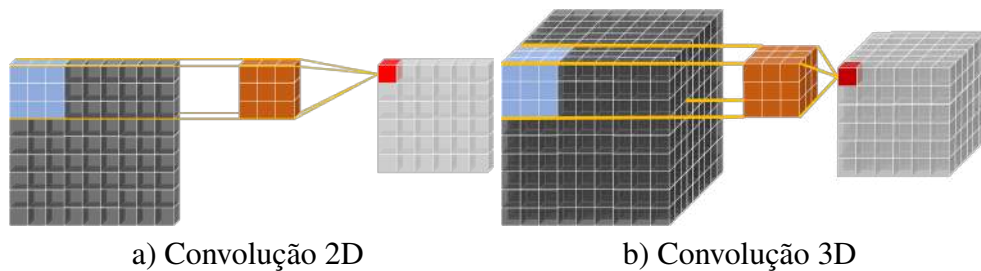
Onde c_{xy} representa resultado do processo pertinente a posição (x, y) , de modo que a entrada $a_{(x+i)(y+j)}$, com um deslocamento de (i, j) em relação a (x, y) , pesos w_{ij} para cada $a_{(x+i)(y+j)}$ localizado em um *kernels* de convolução 2D mais a adição de *bias* b , possam mensurar uma características para o pixel analisado a partir de uma função f de ativação.

No entanto, imagens constituídas com dezenas ou centenas de bandas/canais como são o caso de imagens multi e hiperespectral e dados de ressonância magnética requerem um grande número de *kernels*, que podem ser propensos a overfitting e com maiores custos computacionais (LI; ZHANG; SHEN, 2017). Nesse sentido, a CNN 3D surge como abordagem capaz de extrair informações tanto no contexto espacial quanto no contexto volumétrico/espectral, a partir de *kernels* convolucionais 3D. Esse tipo de convolução é usada principalmente para processamento de vídeo, quando a convolução é aplicada aos *frames* do vídeo (DOU et al., 2020). A convolução 3D é calculada como a soma ponderada de pixels em um cubo de dados 3D como mostra a Equação 2.3

$$c_{xyz} = f\left(\sum_{ijk} w_{ijk} a_{(x+i)(y+j)(z+k)} + b\right) \quad (2.3)$$

Onde c_{xyz} corresponde ao resultado do processo pertinente a posição (x, y, z) , de modo que a entrada $a_{(x+i)(y+j)(z+k)}$, com um deslocamento de (i, j, z) para a (x, y, z) , e pesos w_{ijk} para cada $a_{(x+i)(y+j)(z+k)}$ com um deslocamento de (i, j, z) no *kernel* de convolução 3D. Na figura a seguir são ilustrados os dois processo de convolução.

Figura 9: Visualização e Comparação entre os tipos de convoluções



Fonte: Elaborado pelo Autor

Assim como na CNN 2D, os *kernels* das CNNs 3D são repetidos em toda a imagem, e essas unidades repetidas compartilham as mesmas configurações de parâmetros, portanto é importante que vários tipos diferentes de *kernels* sejam empregados, de forma que se possam extrair distintos tipos de feições (DOU et al., 2020).

2.2.1.2 Autoencoders

Considerado como técnicas de aprendizado não supervisionado, as redes *Autoencoders* são estruturas capazes de obter uma camada de saída com a mesma dimensionalidade da camada

de entrada, de maneira que a entrada seja compactada em uma representação do espaço latente, para posteriormente essa representação seja decodificada para a saída (AGGARWAL et al., 2018). A sua composição é formada por duas partes: *encoder* e *decoder*. O *encoder* é a parte da rede responsável em compactar a entrada e em uma representação de espaço latente, que pode ser descrito por uma função de codificação (Equação 2.4). Enquanto o *decoder*, busca restaurar a entrada da representação do espaço latente (r) por uma função de decodificação (Equação 2.5).

$$r = f(e) = s(W^{(1)}e + b^{(1)}) \quad (2.4)$$

$$e = g(r) = g(W^{(2)}r + b^{(2)}) \quad (2.5)$$

Onde $W^{(1)}$ e $b^{(1)}$ representam os pesos e constantes de cada camada de entrada, f corresponde uma função de ativação que pode ser do tipo *sigmoid* ou *tanh* (LIANG; SHI; ZHANG, 2017). Enquanto $W^{(2)}$ e $b^{(2)}$ caracterizam os pesos e *bias* de cada camada e g como função de ativação da parte de decodificação.

Por fim, o erro de reconstrução para cada dado é definido pela Equação 2.6.

$$L = \| e - g(f(e)) \|^2 \quad (2.6)$$

Dentre as aplicações das redes *Autoencoders*, as tarefas que mais se destacam são a redução de dimensionalidade para visualização de dados e remoção de ruídos (AGGARWAL et al., 2018). No entanto, tarefas como reconhecimento de imagem com CNNs estão ganhando cada vez mais espaço (LIANG; SHI; ZHANG, 2017; SHAO et al., 2019)

2.2.1.3 Blocos Residuais

Semelhante a técnica de *skip connection*, as conexões residuais ou blocos residuais, nasceram no intuito de corrigir a perda do gradiente de treinamento em redes muito profundas (HE et al., 2016). O seu conceito baseia-se em adicionar atalhos que permitam que as informações sejam transmitidas mais facilmente para as próximas camadas da rede (Figura 10), ignorando assim o fluxo normal de uma CNN (WANG; PENG; SUN, 2019).

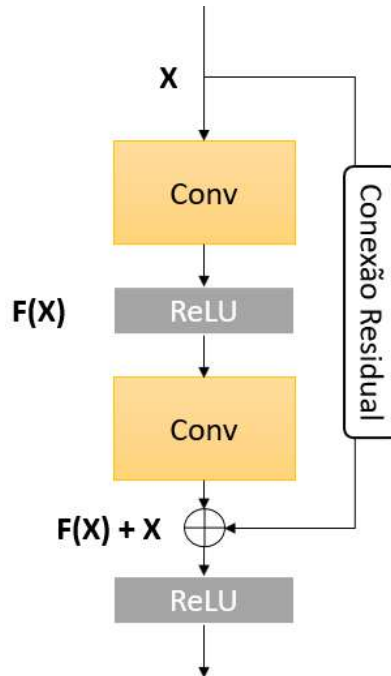
Proposto por He et al. (2016), os blocos residuais além de melhorar o fluxo de informações entre partes da rede, também contribuem em reduzir o ajuste excessivo de parâmetros (WANG; PENG; SUN, 2019). Os blocos residuais podem ser definidos pela seguinte equação 2.7:

$$X^l = H(X^{l-1}) = F(X^{l-1}) + X^{l-1} \quad (2.7)$$

Onde X^{l-1} e X^l representam a entrada e saída da l -ésima camada da rede e $H(\cdot)$ a função de mapeamento original. No qual o mapeamento $H(\cdot)$ pode ser recuperado pelo treinamento

da função $F(\cdot)$ indiretamente a partir de operações convencionais da CNN. Na Figura 10 é destacado um exemplo de bloco residual.

Figura 10: Exemplo de Bloco Residual.



Fonte: Adaptado de He et al. (2016)

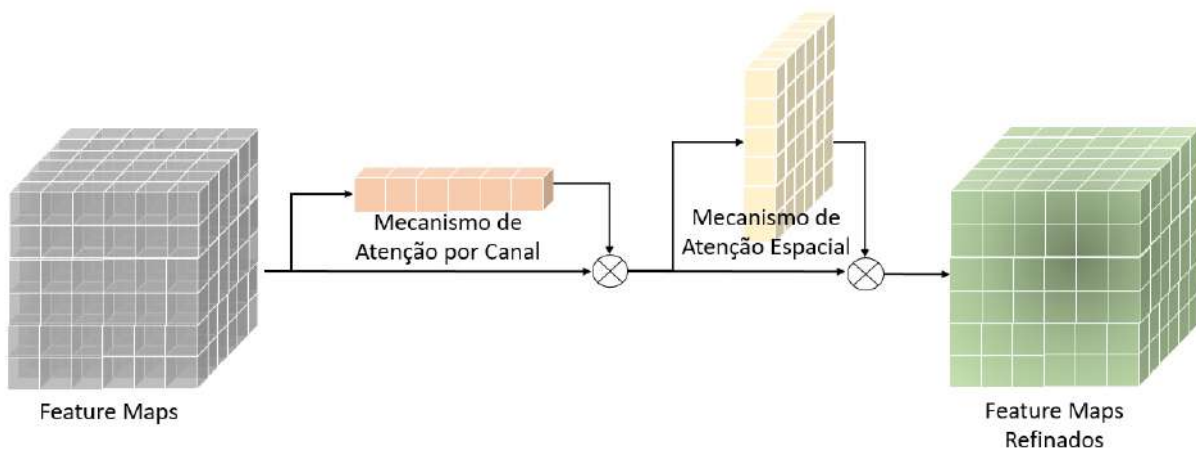
2.2.1.4 Mecanismo de Atenção

Os mecanismos de atenção são estruturas implementadas dentro das CNNs, a fim de que o modelo se concentre em informações mais relevantes, do que aprender informações básicas. Originalmente foram desenvolvidas com o intuito de aumentar o poder representativo dos modelos de *deep learning* (ZHU et al., 2021). A sua estrutura básica consiste em receber um *feature map* e fornecer como saída *attention map*, que posteriormente são multiplicados com o *feature map* com o objetivo de obter um resultado mais refinado (WOO et al., 2018).

De modo geral, os mecanismos de atenção são aplicados no contexto espaciais e de canal/bandas. No contexto espacial, conhecido como mecanismo de atenção do canal, é usado para explorar o relacionamento entre os canais para decidir quais canais são mais significativos, além de minerar a sua interdependência em relação ao modelo (WOO et al., 2018). Em oposição a categoria mencionada, o mecanismo de atenção espacial baseia-se em gerar recursos refinados para explorar a relação interespaical dos recursos originais (ZHU et al., 2021). Na Figura 11 é representado o respectivo processo.

Entre as diferentes abordagens implementadas de mecanismos de atenção, a abordagem proposta por Woo et al. (2018) conhecida como *Convolutional Block attention module* - CBAM, se destaca em comparação as outras abordagens em função de promover a junção entre os

Figura 11: Tipos de Mecanismos de Atenção.



Fonte: Adaptado de Woo et al. (2018)

mecanismos de atenção espacial e de canal. Essa junção garante que os valores mensurados enfatizem a importância dos recursos intermediários (WOO et al., 2018).

Para construção desse respectivo método, o autor a princípio, implementou o módulo de atenção do canal com duas funções do tipo *pooling max* e *avg*, para posteriormente serem transformados por uma rede neural de múltiplas camadas, dos quais os seus resultados foram mesclados. Por fim, uma função do tipo sigmoide é empregada para gerar um *attention map* com ênfase na resposta do canal.

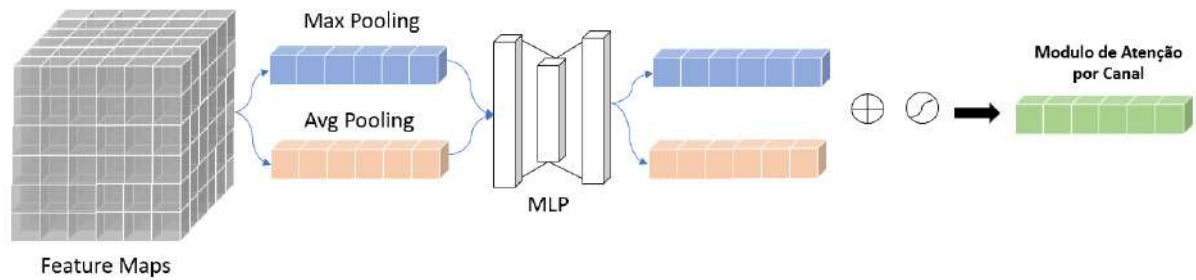
Na construção no módulo de atenção, semelhante a abordagem anteriormente mencionada, duas funções do tipo *pooling max* e *avg* são utilizadas ao longo do eixo do canal. Dos resultados obtidos, ambos são concatenados produzindo assim um novo *feature map*, no qual posteriormente é convertido para *attention map*, com ênfase na resposta espacial, a partir de uma operação de convolução e uma função de sigmoide. Na Figura 12 ilustra a descrição para ambos os módulos, sendo a) para o módulo de atenção do canal e b) para o módulo de atenção espacial.

2.3 Super-Resolução em HI a partir de técnicas de *Deep Learning*

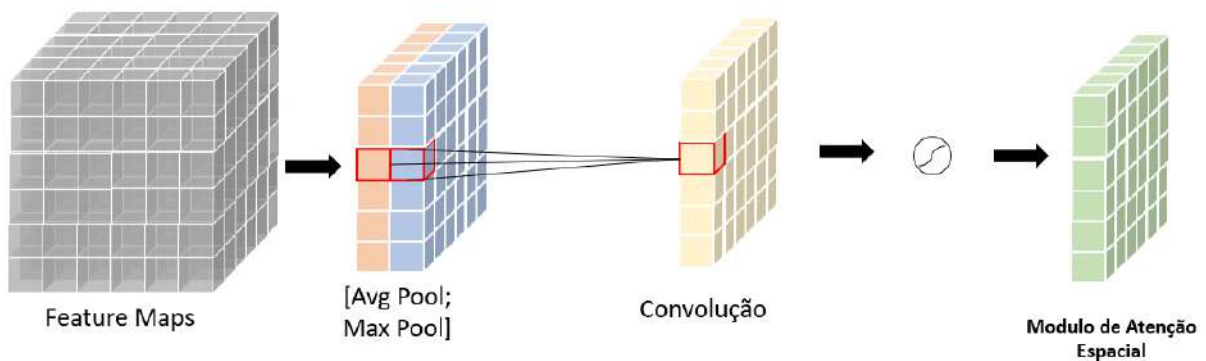
A super-resolução é um problema desafiador de Visão Computacional e Sensoriamento Remoto que visa aprimorar a resolução espacial de imagens de LR (ZHANG et al., 2021). O primeiro trabalho de S-R foi proposto no ano de 1974 por Gerchberg (1974), porém foi a partir do ano de 2016, baseado no poder revolucionário dos métodos de *Deep Learning* proporcionou um maior impacto na predição e retenção de informações de qualidade pelos métodos de S-R (ROHITH; KUMAR, 2020).

Definida como *Super Resolution Convolution Neural Network* - (SRCNN), método este proposto por Dong et al. (2014) e Dong et al. (2015), considerado como o pioneiro dos métodos

Figura 12: Mecanismos de Atenção espacial e por canal



a) Mecanismos de Atenção por canal



b) Mecanismos de Atenção Espacial

Fonte: Adaptado de Wang et al. (2021a)

de S-R baseado em *Deep Learning* em imagens RGB. Estruturado a partir de redes neurais convolucionais 2D, o respectivo método, fundamenta-se em uma imagem de baixa resolução \mathbf{X} pode ser aprimorada por uma função $\mathbf{F}(\cdot)$, de modo que o resultado encontrado possa ser o mais semelhante possível á imagem de alta resolução \mathbf{Y} . Porém para que tal processo ocorra, segundo o autor, os parâmetros da função θ devem ser estimados. Essa estimativa (Equação 2.8) é alcançada minimizando a diferença entre as imagens preditas $F(\mathbf{X};\theta)$ com as imagens de alta resolução Y_l , por meio da função do Erro Quadrático Médio (MSE).

$$L(\theta) = \frac{1}{n} \sum_{l=1}^n \|F(X_l; \theta) - Y_l\|^2, \quad (2.8)$$

Onde n representa o número total de amostras utilizadas.

Posteriormente, vários outros métodos S-R em imagens RGB foram desenvolvidos no intuito de aprimorar a técnica S-RCNN, dentre eles é válido citar as pesquisa de Kim, Lee e Lee (2016); Tai, Yang e Liu (2017) e Ledig et al. (2017). Embora esses modelos sejam capazes de gerar imagens S-R de alta qualidade, eles são projetados principalmente para imagens convencionais e não podem ser usados diretamente em HIs, uma vez que tais imagens contemplam de dezenas ou centenas de bandas espectrais altamente correlacionadas(LIU; LEE, 2019). Nesse sentido, adequações foram impostas a fim de que o modelo proposto por Dong et al. (2014) e

Dong et al. (2015) fossem adaptados para HIs.

De modo geral, os modelos de S-R em HIs partem do princípio que um conjunto de imagens hiperespectrais de HR e LR, representados por $D_n = \{X_n, Y_n\}$, de modo que uma HI de HR $Y \in R^{I,J,L}$ e uma HI de LR $X \in R^{i,j,L}$ são utilizados para formular o modelo de S-R, onde I, J, i e j ($I > i$ e $J > j$) representam a dimensão espacial e L o tamanho das dimensões espectral. Além disso, em função da grande dificuldade de se obter imagens de HR e LR devidamente alinhadas, é assumido que a imagem hiperespectral de LR é definida como a resultante do processos de degradação, reamostragem mais adição de ruídos da HI de HR (Equação 2.9). Logo, de posse de um D_n e uma função de mapeamento G_θ parametrizada por θ , é possível construir um modelo (Equação 2.10) que descreva a relação entre uma HI de HR desejado e o de HI de LR de entrada, a fim de que se possa obter um \hat{Y} o mais próximo possível da imagem verdadeira Y com a maior precisão (LI et al., 2017; WANG et al., 2021b).

$$X_l = (Y_l * D) \downarrow + r \quad (2.9)$$

Onde D corresponde ao filtro de degradação, $*$ a operação de convolução e \downarrow o processo de redução da amostragem de pixel por objeto (*down-sampling*) e r representa o ruído gerado na aquisição das imagens. A quantidade de bandas espectrais é representado por l , de maneira que $l = 1, 2, \dots, L$.

$$\hat{Y}_l = G_\theta(X_l; \theta) \quad (2.10)$$

3 TRABALHOS RELACIONADOS

Neste capítulo é descrito uma revisão da literatura pertinente aos métodos de Super-Resolução à partir de Deep Learning aplicados em imagens hiperespectrais. Para construção dessa revisão buscou-se levantar o real estado da arte desse assunto, no intuito de construir uma compressão abrangente sobre como os métodos de S-R baseados em DL estão sendo aplicados em imagens Hiper e Multiespectral, sem a perda de resposta espectral.

Para construção da revisão de literatura seguiu-se as abordagens propostas pelo relatório para revisões sistemáticas e meta-análises (PRISMA) (MOHER et al., 2010) e o protocolo de revisão proposto por Wen et al. (2012), para análise sintemática de modelos de *Machine Learning*. Na Seção 3.1 são apresentadas as principais questões de pesquisas a serem investigadas. Posteriormente na seção 3.2 é delimitado quais são os planejamentos de pesquisa como: as bases e termos de busca. Sucessivamente na Seção 3.3 é mostrado o processo de seleção de trabalho relacionados a partir dos critérios de exclusão, inclusão e pelo critério de avaliação da qualidade dos artigos. E por fim nas Seções 3.4 e 3.5 é feita a extração das informações e a síntese dos trabalhos que mais se aproximam da Hipótese dessa pesquisa respectivamente.

3.1 Questões de Pesquisa

Nesta seção é apresentado o problema de pesquisa concomitantemente com as questões de pesquisas investigadas com a revisão de literatura.

O *problema* a ser pesquisado é :

- Como aprimorar a resolução espacial em imagens hiperespectrais, de modo que ocorra o mínimo de perda de informação espectral ?

Com base na problemática supracitada e com o foco nas técnicas de inteligência artificial, é investigado como as abordagens baseados em *Deep Learning* são empregadas para solucionar tal problemática, a fim de destacar possíveis tendências, arquiteturas de Deep Learning empregadas, métricas, datasets e principalmente a consistência espectral gerada pelos resultados. A seguir são apontadas as Questões de Pesquisa formuladas (QP) a partir do problema mencionado.

QP1 : Os métodos de *Deep Learning* conseguem obter imagens de hiperespectrais de alta resolução de modo que não prejudique a resposta espectral?

A QP1 tem como objetivo identificar se os métodos baseado em *Deep Learning* são ferramentas aptas em aprimorar a resolução espacial de HI, uma vez que esse tipo de imagem apresenta uma grande quantidade de aplicações em diversas áreas do conhecimento, porém em função da sua baixa resolução espacial essas aplicações são limitadas (JIANG et al., 2020).

QP2 : Quais abordagens de *Deep Learning* estão sendo empregadas atualmente para o aprimoramentos de imagens hiperespectrais?

Essa pergunta busca investigar quais algoritmos, arquiteturas e métricas atuais mais utilizadas, a fim de catalogar tais abordagens, de maneira que se possa identificar pontos fortes e fracos, além de viabilizar possíveis comparações.

QP3 : Quais *datasets* mais utilizados para o treinamento dos modelos e como é feito esse treinamento?

Nessa questão é investigado se existe conjuntos de dataset de HI disponíveis e aptos para o treinamento dos modelos inteligente e como esses dataset estão sendo empregados para o treinamento e validação das imagens aprimoradas.

QP4 : Quais são as métricas ou métodos utilizados para avaliação da consistência espectral gerada pelas HI aprimoradas?

Considerado a principal questão a ser investigada, nessa pergunta busca analisar como os trabalhos estão avaliando a consistência espectral das HI, de modo identificar se existe métodos ou técnicas que proporcionam a respectiva avaliação.

3.2 Planejamento de Pesquisa

Nessa seção é descrito os passos empregados para a construção do planejamento de pesquisa na investigação dos artigos relacionados. A princípio, em função do problema e das questões de pesquisa apresentadas na Seção 3.1, palavras-chaves na língua inglesa foram extraídas, no intuito de construir *string* de busca a serem utilizadas nas fontes nas bases de dados. As palavras-chaves extraídas são as seguintes:

Super-Resolution OR Super Resolution OR SuperResolution
OR
Improvement Resolution OR Enhancement Resolution
AND
Hyperespectral Image
AND
Deep Learning

As fonte de busca utilizadas foram as seguintes bases:

- IEEE Xplore - ieeexplore.ieee.org
- Scopus - www.scopus.com
- ScienceDirect - www.sciencedirect.com
- MDPI - www.mdpi.com

As *strings* de busca empregadas para cada base estão descritas na Tabela 1:

Tabela 1: Consulta realizadas nas bases de dados juntamente com seu período de filtragem

Consulta	Período
(((((Super OR Improvement OR Enhancement) AND Resolution) OR super-resolution OR superresolution) AND (hyperspectral image) AND (deep learning))	2011 - 2021

Fonte: Elaborado pelo autor.

3.3 Seleção de Estudos

Nessa seção são descritos os passos utilizados no processo de filtragem dos trabalhos relacionados, de modo que ao final desse módulo é exposto de forma detalhada o número de trabalhos selecionados e os seus critérios de seleção.

3.3.1 Critérios de Seleção e Exclusão

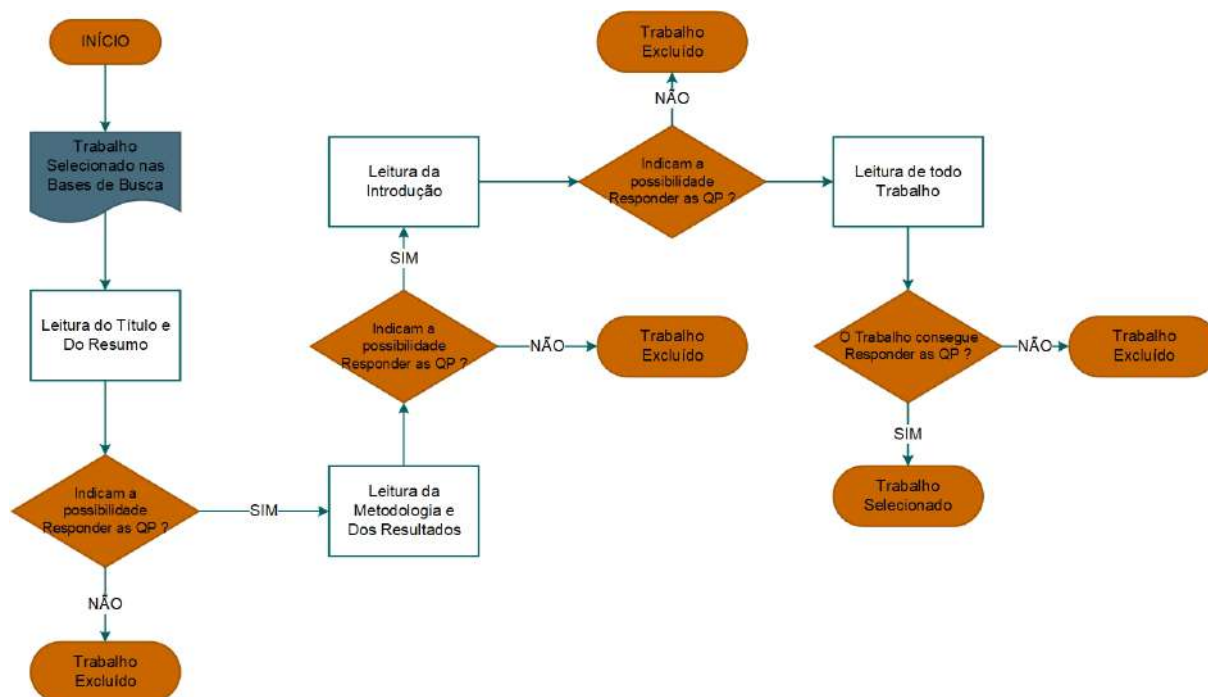
O objetivo de estabelecer critérios de seleção é o de viabilizar a identificação dos trabalhos que oferecem indícios para as respostas investigadas. Nesse sentido, o primeiro critérios de inclusão dos estudos relacionados foram aqueles publicados entre os anos de 2011 e 2021 que demonstravam contribuições nas abordagens de Super-Resolução em HI por meio dos métodos de *Deep Learning*. A escolha desse período se deve ao fato de proporcionar que as buscas efetuadas englobassem uma grande quantidade de trabalhos.

Para exclusão dos trabalhos adotou-se como critérios aqueles trabalhos duplicados, não escritos na língua inglesa, artigos de revisão e trabalhos que não utilizassem *Deep Learning* no aprimoramento de resolução espacial e espectral em HI. Além desses pontos referenciados, buscou-se investigar a congruência dos trabalhos de acordo com as questões de pesquisa (Seção 3.1). Essa investigação ocorreu por meio de 4 etapas, onde cada etapa corresponde a leitura de um tópico específico do texto. Caso um desses tópicos estivessem fora do contexto esperado, o trabalho não era selecionado. Na Figura 13 é ilustrado o fluxo de investigação juntamente com os tópicos de leitura.

3.3.2 Avaliação da qualidade do Estudo

Posteriormente abordagem descrita anteriormente, outro método de seleção de trabalho também foi aplicado, que é o método de Avaliação da qualidade do Estudo (QA). Proposto por Wen et al. (2012), o método QA consiste em elaborar uma série de questões de avaliação do rigor, da credibilidade e relevâncias dos estudos. Na Tabela 2, são destacadas as questões utilizadas.

Figura 13: Fluxograma de Investigação dos Trabalhos Relacionados



Fonte: Elaborado pela autor.

Cada pergunta contempla apenas três respostas: “Sim”, “Parcialmente” ou “Não”, de modo que as três respostas são pontuadas da seguinte forma: “Sim”=1, “Parcialmente”=0,5 e “Não”=0. Para cada determinado estudo é somado a sua pontuação de QA, dessa forma estudos cujo a soma seja inferior 6 foram excluídos.

3.3.3 Resultados da Seleção dos Estudos

Conforme mencionado nos tópicos anteriores, a seleção dos trabalhos relacionados foi subdividida em duas partes, a primeira parte correspondente aos critérios de inclusão e exclusão dos trabalhos que corroboram para integração daqueles estudos que indicassem possíveis respostas para as questões de pesquisa (Seção 3.1). Enquanto isso a segunda parte, correspondente aos critérios de qualidade que contribuíram na escolha das pesquisas mais relevantes. O número de cada trabalho identificado em cada estágio é possível ser observado no diagrama da Figura 14.

Dos resultados encontrados, apenas 28 trabalhos conseguiram atender aos critérios de inclusão e exclusão, porém esse número foi reduzido para 18, após passarem pela avaliação da qualidade do estudo, conforme apresentado na Subseção 3.3.2.

Com uma pontuação mínima de 6 pontos, os estudos selecionados obtiveram pontuação média de 7.8 com desvio padrão de 1.8, o que comprova que os trabalhos apresentam um QA semelhante. Em relação às médias por questões de avaliação, é possível classificá-las em três níveis que são: "Alto" para médias superiores a 0.8; "Médio" para valores de médias interme-

Tabela 2: Consulta realizadas nas bases de dados juntamente com seu período de filtragem

N°	Questão
QA1	Os objetivos da pesquisa estão claramente definidos?
QA2	O contexto do trabalho está descrito de forma adequada ?
QA3	Os métodos aplicados estão bem definidos ?
QA4	A avaliação do desempenho experimental é apropriado e justificável ?
QA5	O experimento é desenvolvido em conjuntos de datasets suficientes ? ('Sim':dois ou mais datasets, "Parcialmente" : apenas um dataset; "Não" : nenhum dataset.)
QA6	A precisão da estimativa é medida e relatada?
QA7	O método proposto é comparado com outros métodos?
QA8	As limitações do estudo são analisadas experimentalmente ?
QA9	As respostas espectrais oriundas dos métodos propostos são avaliadas de formar qualitativa(comparação gráfica) e quantitativa (métricas de comparação)? ("Sim": para as duas avaliações, "Parcialmente": para somente uma avaliação e "Não")
Q10	As métricas convencionais de comparação entre as imagens são utilizadas? ('Sim':três ou mais métricas, "Parcialmente" apenas duas métricas, "Não" : nenhuma métrica.)

Fonte: Adaptado de Wen et al. (2012).

Tabela 3: Classificação das Questões de Avaliação em Relação a média

Classificação	Questões de Avaliação
Alto	QA1, QA2, QA5, QA7 e QA10
Médio	QA3, QA4, QA6 e QA9
Baixo	QA8

Fonte: Elaborado pelo autor.

diárias entre 0.8 e 0.5 e 'Baixo' para médias inferiores a 0.5. Na Tabela 3, são destacados a classificação para cada questão.

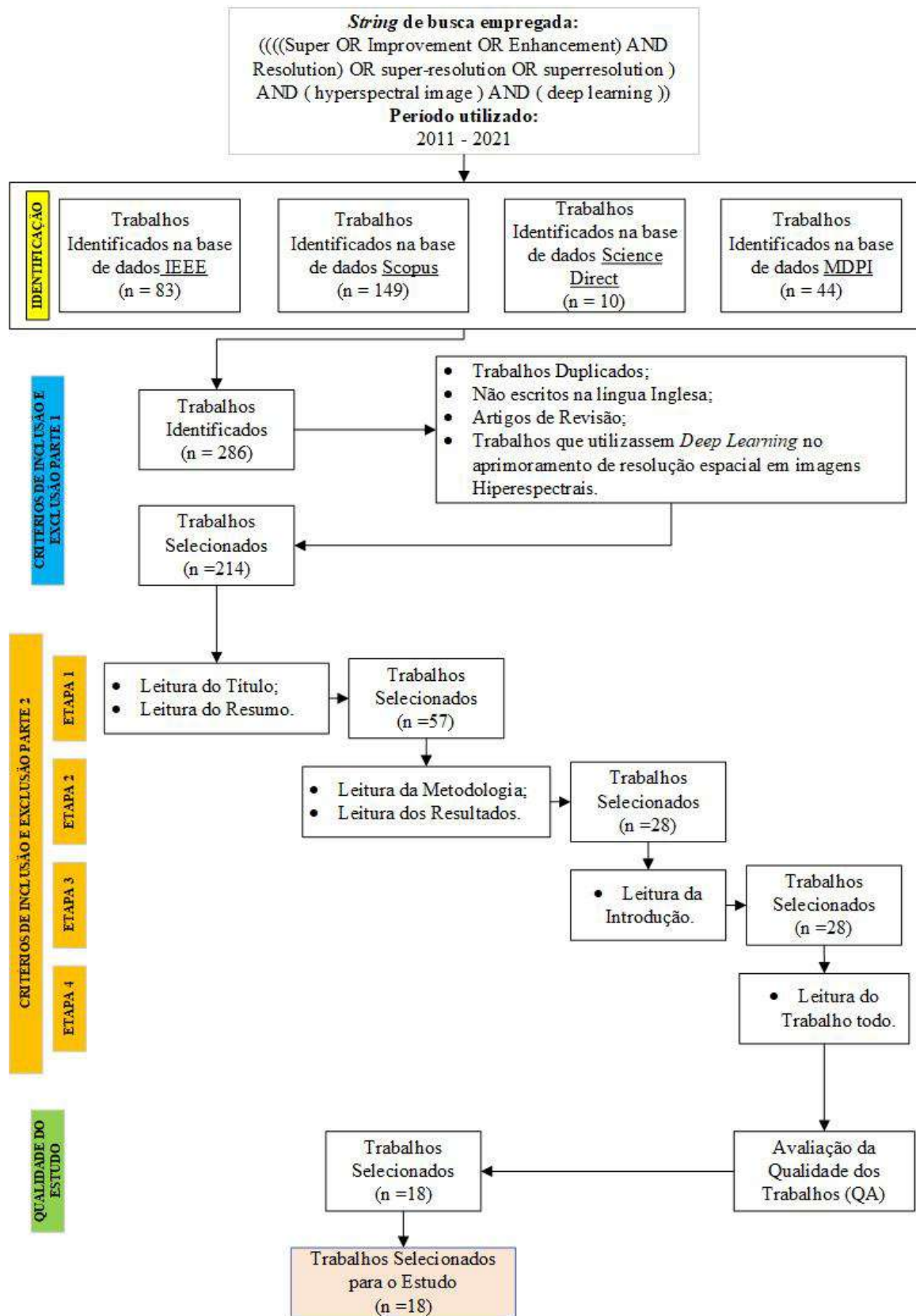
Embora apenas o critério QA8 seja classificado na categoria do tipo 'Baixo', o que indica que os estudos selecionados não se preocupam em analisar experimentalmente as suas limitações, de forma geral, os artigos eleitos apresentam fortes indícios que possam fornecer contribuições valiosas para este trabalho.

3.4 Extração das Informações

No intuito de sintetizar as principais informações encontradas nos respectivos trabalhos selecionados, que corroborassem em responder as questões de pesquisa levantadas (Seção 3.1), foi elaborado um formulário conforme apresentando em Wen et al. (2012), porém com adaptações para essa pesquisa. Nesse formulário, rótulos (R) foram propostos a fim de facilitar a extração e identificação dos principais pontos pertinentes a esse estudo. A seguir são detalhados os rótulos utilizados para extração e identificação das informações.

- R1 : Título do Artigo;
- R2 : Onde foi publicado e o Ano de Publicação;
- R3 : Problemática de Pesquisa;
- R4 : Hipótese Investigada;
- R5 : Qual método de *Deep Learning* foi empregado ;
- R6 : Quais são os dataset que foram utilizados;
- R7 : Como foi feito o treinamento e validação dos modelos;
- R8 : Quais métricas foram utilizadas para validação dos resultados;
- R9 : Como a resposta espectral é avaliada;
- R10 : Principal vantagem pontuada pelos autores em relação ao método proposto.

Figura 14: Resultados das Etapas de Busca dos Trabalhos Relacionados



Fonte: Elaborado pela autor.

3.5 Síntese dos Trabalhos

Mediante a extração dos potenciais pontos levantados pelos artigos selecionados para este trabalho, é extremamente importante sintetizar tais características, de modo que possibilite responder de forma clara e objetiva as questões de pesquisa abordadas. Os dados extraídos se resumem em dados quantitativos e qualitativos que foram extraídos por meio da abordagem de síntese narrativa e contagem de frequência.

3.5.1 Informações gerais sobre os artigos selecionados

Ao analisar os artigos revisados no ponto de vista temporal, percebe-se que em totalidade, todos os trabalhos foram publicados a partir do ano de 2017. Esse acontecimento se deve ao fato que os artigos pioneiros nos métodos de Super-Resolução por técnicas de Deep Learning em imagens RGB, foram publicados no ano de 2015 e posteriormente foram expandidos para HIs (LI; WANG; LI, 2021). Em relação aos anos de publicação, observou-se que os anos de 2019, 2020 e 2021 contemplaram com mais de 88% no número de publicações, o que mostra uma forte tendência de alta de estudos sobre o tema. Além disso é válido ressaltar que a busca dessa revisão foi finalizada em Setembro de 2021, logo acreditasse que mais trabalhos tenham sido publicados, aumento assim a quantidade de publicações para esses últimos 3 anos.

No que se refere ao tipo de publicação, 95% dos artigos selecionados foram publicados em periódicos, enquanto o restante foram submetidos para congressos ou conferências. Em relação as fontes de publicação os estudos foram submetidos para as seguintes fontes: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS)*, *IEEE Geoscience and Remote Sensing Letters*, *IEEE International Conference on Image Processing (ICIP)*, *IEEE Transactions on Computational Imaging*, *IEEE Transactions on Geoscience and Remote Sensing*, *Neurocomputing*, *International Journal of Machine Learning and Cybernetics*, *Signal Processing: Image Communication e Remote Sensing*. Por meio desses locais é possível inferir a forte preocupação tanto do Sensoriamento Remoto quanto na computação em aprimorar a resolução das imagens HI, uma vez que esse tipo de imagem é empregada em inúmeras aplicações.

3.5.2 Principais problemáticas abordadas e suas soluções

Nesse tópico é apresentado uma possível linha temporal dos métodos de Super-Resolução revisados, no intuito de encontrar tendências de pesquisa, de modo que se possa ao final traçar qual o direcionamento dos métodos de S-R atualmente.

Conforme já mencionado em tópicos anteriores, as primeiras linhas pesquisa de super-resolução baseadas em Deep Learning foram ancoradas por Dong et al. (2014, 2015). Em seu estudo, imagens naturais de alta e baixa resolução foram empregadas a fim de treinar uma rede

neural convolucional 2D, nomeada como S-RCNN. Ao final do treinamento a rede era capaz de aprimorar a resolução espacial das imagens RGB. Embora a metodologia desenvolvida pelos autores tenha se mostrado superior as abordagens tradicionais, imagens multi e hiperespectrais ainda eram consideradas como certos gargalos para metodologia proposta, uma vez que é extremamente complexo aprimorar a resolução espacial sem degradar a resposta espectral (LIU; LEE, 2019; YANG et al., 2020).

Nesse contexto, várias abordagens consideradas expansões da rede SRCNN foram proposta como é o caso dos trabalhos (LI et al., 2017; HU; LI; XIE, 2017). Em seus trabalhos, os autores além de empregarem CNN 2D, aplicaram modelos restritivos e adaptativos, de modo que limitações nos erros espaciais e espectrais fossem minimizados. Outra abordagem também considerada expansão do modelo SRCNN, são as publicações de Yuan, Zheng e Lu (2017); Li, Zhang e You (2019); Zou e Huang (2020). Baseado na teoria da transferência de aprendizado e assumindo que existe uma relação espacial entre as imagens naturais e Hiperespectrais, modelos de CNN 2D foram treinados com auxílio de *dataset* de imagens RGB e posteriormente transferido o conhecimento para HI. Após o respectivo processo, modelos de mistura espectral foram empregados com intenção de corrigir a resposta espectral.

Além das abordagens supracitadas, pesquisas como por exemplo de Zheng et al. (2019); Zhao et al. (2021); Hu et al. (2021) fundamentados na necessidade em separar e trabalhar de forma distinta a resposta espectral e espacial foram desenvolvidas. Zheng et al. (2019) em sua pesquisa propôs extrair os recursos espectrais e espaciais a partir de uma rede multi-caminhos por análise em banda individual, de modo que no final, pelo módulo de fusão, uma imagem de alta resolução espacial e espectral era restaurada. Semelhante ao estudo supracitado, porém fundamentado em módulos residuais guiados por gradientes espaciais, Zhao et al. (2021) sugeriu uma rede composta por módulos residuais para explorar totalmente os recursos hierárquicos aprendidos de todas as camadas convolucionais, que posteriormente eram utilizados para orientar por meio do gradiente o processo de S-R. Ainda apoiados dentro da conjuntura da técnica de *Deep Residual Networks*, Hu et al. (2021) estabeleceu um modelo residual baseado em conceitos de estrutura e textura, de forma que o cubo de estrutura e textura aprendidos eram integrados para obter um HI de alta resolução.

Apesar dos avanços dos métodos de CNN 2D para S-R em imagens multi e hiperespectrais, alguns trabalhos afirmam que os modelos baseados em CNN 2D infelizmente negligenciam o distúrbio espectral, uma vez que tal método não explora as informações de correlação espacial e espectral simultaneamente (MEI et al., 2017; LIU; LEE, 2019; DOU et al., 2020). No intuito de contornar tal impasse, Mei et al. (2017) construiu um método de S-R baseada em rede neural convolucional (CNN 3D), de modo que correlações espaciais e espectrais corroborassem na predição de HIs de alta resolução. Liu e Lee (2019), além de sugerir modelos de CNN3D implementou o aprendizado baseado em *Deep Residual Networks*, garantindo assim maior robustez nos resultados encontrados. Seguindo a mesma linha de módulos de CNN 3D, Li et al. (2019) implementou em sua pesquisa um rede neural convolucional 3D, de modo que

com auxílio de módulos de atenção permitiu empregar a estratégia de separação para extrair a informação espacial-espectral e, em seguida, fundi-los gradualmente.

Métodos mais avançados utilizando CNN 3D também foram empregados como são os casos de Dou et al. (2020) e Wang et al. (2021b). Em ambos os trabalhos Redes Adversarias Generativas - GANs foram implementadas com a intenção de aprimorar as HI não apenas ao explorar as características espaciais, mas também ao preservar as propriedades espectrais no processo. Ainda dentro do contexto de CNN3D, Yang et al. (2020) projetou uma rede com módulos de atenção local e não local, de forma que um modelo híbrido pudesse explorar as relações espectrais e espaciais, de maneira que imagens hiperespectrais de alta resolução espacial fossem preditas.

Embora as imagens geradas a partir de CNN 3D apresentaram resultados superiores aos modelos implementados com CNN 2D, limitações de GPU e memória acarretaram em longos tempos de treinamento, além de impossibilitar a construção de camadas mais profundas (JIANG et al., 2019). Nesse sentido, estudos cujo o foco principal é propor métodos inteligentes de SR em imagens hiperespectrais de forma otimizada passaram a ser a principal linha de pesquisa (LI et al., 2019; LI; WANG; LI, 2020, 2021). Li, Wang e Li (2020, 2021) sugeriram em seus trabalhos a construção de um módulo convolucional misto para extrair os recursos potenciais por convolução 2D / 3D em vez de uma convolução, o que permite que a rede extraia mais recursos espaciais de imagem hiperespectral reduzindo assim a complexidade do aprendizado.

Como resultado dessa busca sistemática dos artigos relacionados, percebe-se que os trabalhos atualmente buscam solucionar a falta de resolução espacial em HI partindo da necessidade de tratar de forma distinta as informações espaciais e espectrais para posteriormente serem integradas gradativamente. Concomitantemente a esse processo, a exigência de métodos otimizados se mostra como um ponto inevitável para esse processo.

3.5.3 Tipos de algoritmos utilizados

O objetivo desse item foi identificar quais técnicas, algoritmos ou arquiteturas baseadas em *Deep Learning* empregadas pelos artigos selecionados. A categorização dessas abordagens ocorreu após a extração dos dados no momento da leitura dos estudos. Por meio do respectivo levantamento, foi sintetizado todas as informações pertinentes aos algoritmos empregados. Tais informações podem ser visualizadas na Tabela 4.É válido salientar que muitos trabalhos não detalham de forma completa os procedimentos empregados, logo as informações levantadas foram sintetizadas na Tabela.

Tabela 4: Síntese dos Algoritmos empregados nos trabalhos analisados

ID	Autor	Métodos	Estratégia		Função de Perda	Otimizadores	Função de Ativação	Função de Sintetização
			Unidade 2D	Unidade 3D				
T1	Mei et al. (2017)	3D - FCNN	-	Convolução 3D	MSE	Adam	Relu	Kernel de Gauss
T2	Li et al. (2017)	SCT_SDCNN	3 Convoluções 2D + Restrição Espacial	-	MSE	x	Relu	Filtro Gaussiano + Ruído Gaussiano + Fator de Redução
T3	Yuan, Zheng e Lu (2017)	CNN_CNMF	3 Convoluções 2D + Unmixing	-	x	Adam	Relu	Filtro Gaussiano
T4	Hu, Li e Xie (2017)	SDCNN_SEC	3 Convoluções 2D + modelo de correção de erro espacial (SEC)	-	MSE	SGD	Relu	down-sampling
T5	Zheng et al. (2019)	SSIN	Convoluções 2D + módulo Residual	-	MSE	Adam	Relu	Bicúbica
T6	Li, Zhang e You (2019)	CSN	3 Convoluções 2D + Batch Normalization	-	x	x	Relu	Imagem foi desfocada e posteriormente reduzida
T7	Jiang et al. (2019)	2D-1D-HS-RGAN	Gerador: 5 Convoluções 2D + pixelshuffle + 3 Convoluções 1D; discriminador: 3 Convoluções 2D	-	MAE e SAM	Adam	LeakyReLU	Filtro de Passa-Baixa + Fator de Redução
T8	Liu e Lee (2019)	CNN3D Residual Learning	-	Convoluções 3D + Batch Normalization+ módulo residual	Charbonnier	Adam	Relu	Filtro gaussiano 3D + Fator de Redução
T9	Li et al. (2019)	1D-2D Attentional CNN	-	Convoluções 3D + módulo de atenção local e Residual	MAE e SAM	Adam	PReLU	Filtro Gaussiano + Fator de Redução
T10	Jiang et al. (2020)	SPPS-R	Convoluções 2D + módulo Residual	-	SSTV	Adam	Relu	Bicúbica
T11	Dou et al. (2020)	3DAS-RGAN	Discriminador: Convoluções 2D	Gerador: Convoluções 3D + módulo de atenção + pixelshuffle	MSE	Adam	LeakyReLU, Prelu e Sigmoid	down-sampling
T12	Li, Wang e Li (2020)	MCNet	2 convoluções 2D	2 convoluções 3D + módulo residual	MAE	Adam	Relu	Bicúbica
T13	Zou e Huang (2020)	DRCNN	Convolução 2D+ módulo Residual+ pixelshuffle	-	MAE	Adam	Relu	x
T14	Li, Wang e Li (2021)	ERCSR	Convoluções 2D+ módulo residual	Convoluções 3D+ módulo residual	MAE	Adam	Relu	Bicúbica
T15	Zhao et al. (2021)	ResNet_G-RDN	7 Convoluções 2D+ módulo residual	-	MAE	x	Relu	x
T16	Wang et al. (2021b)	HSS-RGAN	Gerador: Convoluções 2D + pixelshuffle + módulos residuais ; Discriminador: Convoluções 2D	-	MAE e SAM	Adam	Leaky e Relu	Filtro Gaussiano+ Fator de redução
T17	Yang et al. (2020)	LN-atten-CNN	-	Convoluções 3D + módulo residual+ atenção local e não local	MAE	Adam	Sigmoid e Relu	Filtro Gaussiano+ Fator de redução
T18	Hu et al. (2021)	RSTDN	duas redes paralelas de convoluções 2D(Estrutural e Textura) são concatenadas no final	-	x	Adam	Relu	Bicúbica

Ao analisar o tipo de treinamento implementado, 100% dos resultados se enquadram no tipo Semi-Supervisionado, o que demonstra possivelmente que os métodos não Supervisionados possam ser explorados futuramente. Em relação a etapa de rotulação entre os dados de entrada e saída dos modelos, as imagens de LR a princípio, são adquiridas por meio de técnicas de interpolações convencionais, como é o caso da interpolação bicúbica, que degrada a resolução espacial de uma imagem de HR, concebendo assim, uma imagem de LR. Desse modo, os modelos inteligentes geralmente empregam as imagens de LR com dados de entrada e como o resultado esperado as imagens de HR. Os tipos de técnicas empregadas em degradar as resoluções espaciais em cada trabalho são listadas na coluna Função de Sintetização.

Em relação ao tipo de estratégia utilizada, é possível notar que diferentes abordagens de arquiteturas foram testadas. Cada estrutura composta por características específicas, com destaque para as abordagens utilizando módulos residuais que corresponde a 67% dos trabalhos. Além de módulos residuais, estratégias como redes Adversárias Generativas, Transferência de Aprendizado e a união entre os modelos preditivos e técnicas consagradas na literatura, como é o caso do *Unmixing*, também foram experimentados.

Embora alguns trabalhos como são os caso de Mei et al. (2017); Liu e Lee (2019) e Dou et al. (2020), mencionem que a CNN 2D negligenciam o distúrbio espectral, 61% dos artigos selecionados são compostos por convoluções do tipo 2D, enquanto somente 22% aplicaram convoluções do tipo 3D e apenas 17% utilizaram de ambas as abordagens. Essa discrepância nas porcentagem pode ser justificada pelo fato que os modelos implementados por convoluções 3D requisitam esforço computacional superior as CNN 2D. Logo, abordagens que integram os dois recursos possam ser uma solução viável, conforme mencionado em Li, Wang e Li (2020, 2021).

Em comparação ao tipo de função de perda aplicadas pelos autores, dois tipos de função se destacaram em relação as demais, que são Erro Médio Absoluto (MAE) e Erro Médio Quadrático (MSE). Ambas os métodos foram utilizados em mais de 56% dos trabalhos revisados (10 Trabalhos). Em relação aos outros estudos, 28% (5 Trabalhos) adotaram funções customizadas e 16% (3 Trabalhos) não foram informados.

Diferente das funções de perda tradicionais, as funções customizadas em métodos de super-resolução, partem do princípio que as abordagens tradicionais, como por exemplo MAE e MSE, conseguem preservar a informação espacial dos resultados, no entanto, informações espectrais podem ser distorcidas ao ignorar as correlações entre as características espectrais (JIANG et al., 2020). A fim de assegurar simultaneamente a credibilidade espacial e espectral dos resultados da reconstrução, funções de perda customizadas como é o caso da integração entre MAE e *Spectral Angle Mapper* - (SAM) podem ser considerada um opção em potencial, conforme apresentado nos trabalhos de Jiang et al. (2019); Liu e Lee (2019); Li et al. (2019); Jiang et al. (2020) e Wang et al. (2021b).

Ainda com base nas informações da Tabela 4, a função de ativação mais empregada entre as camadas convolucionais é a função *ReLU*, a qual corresponde a 72% (13 Trabalhos) dos

estudos. Em relação aos outros artigos, com exceção aos estudos de Jiang et al. (2019) e Li et al. (2019) que empregaram as funções LeakyReLU e PReLU respectivamente, os demais trabalhos construíram a arquitetura das suas redes aplicando o método *ReLU* acompanhado de outras funções como Sigmoid e LeakyReLU. Logo percebe-se a existência de uma hegemonia da função de ativação *ReLU* nas técnicas de S-R em imagens hiperespectrais.

Aliados às funções de perda, os otimizadores despontam como técnicas que corroboram em a acelerar os treinamentos das redes. Nesse sentido, o otimizador mais utilizado nos estudos foi *Adaptive Moment Estimation-Adam* com 78% (14 artigos) dos trabalhos, isso mostra que grande parte dos artigos não realizaram testes com distintos otimizadores para o treinamento das suas redes.

De modo geral, é possível notar que as técnicas e estruturas das redes convolucionais implementadas para os métodos de S-R em imagens hiperespectrais são bastante heterogêneas, não havendo assim, uma abordagem que se destaca em relação as demais. Logo, é plausível mencionar que existam inúmeras outras combinações de técnicas a serem exploradas, uma vez que uma variação muito pequena de estudos de S-R em HI a partir de técnicas de Deep Learning foram testadas.

3.5.4 Dataset empregados

Os datasets são de extrema importância para os métodos de Deep Learning, uma vez que são a partir deles que os modelos inteligentes conseguem modelar e validar possíveis problemas complexos. Para construção de um Deep Learning variáveis como dimensionalidade, escassez e resolução devem ser consideradas de forma que problemas como *Overfitting* e *Underfitting* não transcorram. Nesse sentido, e com base nos trabalhos selecionados foram levantados todos os principais datasets utilizados, a fim de analisar características como detalhes de aquisição, quantidade e disponibilidade. A Tabela 3.5.4 são destacados tais datasets e a frequência que eles aparecem.

É válido salientar que a maioria dos datasets são disponibilizados de forma online e de fácil acesso, com exceção do dataset de **Montana** (YUAN; ZHENG; LU, 2017) que não foi encontrado. Com base no valor de frequência mencionado na Tabela 3.5.4 percebe-se que os 5 primeiros dataset se encontram em mais de 75% dos trabalhos selecionados, o que indica possivelmente que esses conjuntos de dados foram os pioneiros nas pesquisas de S-R em imagens hiperespectrais ou são considerados os datasets de maior destaque para comunidade científica nessa área de pesquisa. Nesse contexto a seguir são detalhados os principais atributos de cada conjunto de dados que se encontram disponibilizados de forma online.

- **Pavia Centre e University:** Coletadas pelo sensor Rosis após as campanhas de mapeamento da cidade de Pavia no norte da Itália. Os dois dataset se tratam somente de duas imagens como dimensões de (1096 x 1096 x 103) e (610 x 610 x 102) respectivamente, sendo que 12 dessas bandas espectrais não contêm informações sendo necessário

Tabela 5: Lista dos datasets empregados e sua frequência

DATASETS	FREQUÊNCIA
Pavia Centre (YASUMA et al., 2008)	14
Cave (YASUMA et al., 2008)	11
Pavia University (YASUMA et al., 2008)	9
Washington DC Mall (YASUMA et al., 2008)	6
Harvard (CHAKRABARTI; ZICKLER, 2011)	5
Foster (NASCIMENTO; AMANO; FOSTER, 2016)	3
Cuprite (YASUMA et al., 2008)	2
Paris (SIMOES et al., 2014)	2
Urban (YASUMA et al., 2008)	2
Chikusei (YOKOYA; IWASAKI, 2016)	1
Indian Pine (YASUMA et al., 2008)	1
Natural (NASCIMENTO; FERREIRA; FOSTER, 2002)	1
Montana (YUAN; ZHENG; LU, 2017)	1

Fonte: Elaborado pelo autor.

descartá-las. Além disso as cenas desse sensor contemplam de resolução espacial de 1.3 metros e resolução espectral de 4nm coletadas nos comprimentos de ondas entre (430nm-800nm) no EE (GAMBA, 2015).

- **Cave:** Composto por 32 HI de objetos em geral, o banco de dados de Cave é construído a partir da captura das imagens por meio da câmera CCD. Esse tipo de sensor garante a aquisição de imagens com dimensões de (512 x 512) pixels e com 31 bandas espectrais compostas com a resolução espectral de 10nm adquiridos a partir dos comprimentos de ondas variando entre (400nm - 700nm) (YASUMA et al., 2008).
- **Washington DC Mall:** Representado somente por uma única imagem capturada por intermédio de um sensor hiperespectral aerotransportada sobre o *Washington DC Mall*, esse conjunto de dados contemplam uma imagem de aproximadamente 150 Megabytes com 210 bandas espectrais na região de 0,4 a 2,4 μm do espectro visível e infravermelho. É importante ressaltar que as bandas na região de 0,9 e 1,4 μm , onde a atmosfera é opaca, foram omitidas do conjunto de dados, deixando 191 bandas somente (FENG et al., 2016).
- **Harvard:** Representado por um banco de dados composto por 75 HI de diferentes locais com variação sob o tipo de iluminação, esse datasets foi construído a partir do sensor comercial *Nuance FX, CRI Inc*, permite adquirir imagens hiperespectrais com resolução espectral de aproximadamente 10nm na região entre 420 nm a 720 nm do EE (CHAKRABARTI; ZICKLER, 2011).
- **Foster:** Adquirido entre os anos de 2002 e 2003 em Portugal, o dataset Foster é composto de 30 imagens hiperespectrais de 33 bandas com o tamanho de 1204 x 1344 pixels. Cole-

tadas entre as faixas do espectro 400 nm até 700nm corrobora para que cada HI contemple uma resolução espectral de 10nm (NASCIMENTO; AMANO; FOSTER, 2016).

- **Cuprite** : Desenvolvido a partir do mapeamento da região de Cuprite em Las Vegas pelo sensor Aviris. Esse conjunto de dados é constituído somente por uma única imagem com 250x190 pixels e 224 bandas fracionadas entre as faixas 370 nm a 2480 nm do EE. Algumas bandas são consideradas ruidosas necessitando assim removê-las (M. GRAÑA MA. VEGANZONS, 2021).
- **Paris**: Criado somente por uma única imagem com dimensões de (72 x 72 x 128) capturada sobre a cidade de Paris pelo sensor Hyperion com resolução espacial de 30 metros (SIMOES et al., 2014).
- **Urban**: Corresponde ao mapeamento de uma área de $4m^2$, no qual foi gerado uma única imagem com dimensões de (307 x 307 x 210) na região entre os intervalos 400 nm a 2500 nm do EE. Semelhante ao conjunto de dados de Cuprite é necessário remover bandas ruidosas (HASANLOU, 2019).
- **Chikusei**: Conjuntos de dados hiperespectral aerotransportado obtidos pelo sensor *Headwall Hyperspec-VNIR-C* sobre áreas agrícolas e urbanas em Chikusei no Japão. A única Hi obtida é composta por 128 bandas espectrais situadas entre as faixas 363 nm a 1018 nm do EE. A cena consiste em 2517x2335 pixels e a distância de amostragem do solo foi de 2.5 m (YOKOYA; IWASAKI, 2016).
- **Indian Pine**: Cena coletada pelo sensor Aviris no noroeste do estado de Indiana com dimensões de 145 x 145 pixels e 224 bandas de refletância espectral na faixa de comprimento de onda 0.4–2.5 μm (BAUMGARDNER; BIEHL; LANDGREBE, 2015).
- **Natural**: *Dataset* composto por 8 imagens hiperespectrais coletadas de algumas áreas urbanas e rurais de Portugal. Cada cena é composto de 30 bandas espectrais na faixa de comprimento de onda 400 nm até 700nm. Diferente dos outros conjuntos de dados, as imagens apresentadas nesse dataset apresentam diferentes tipos de dimensões. (NASCIMENTO; FERREIRA; FOSTER, 2002).

3.5.5 Fluxo de treinamento e Validação

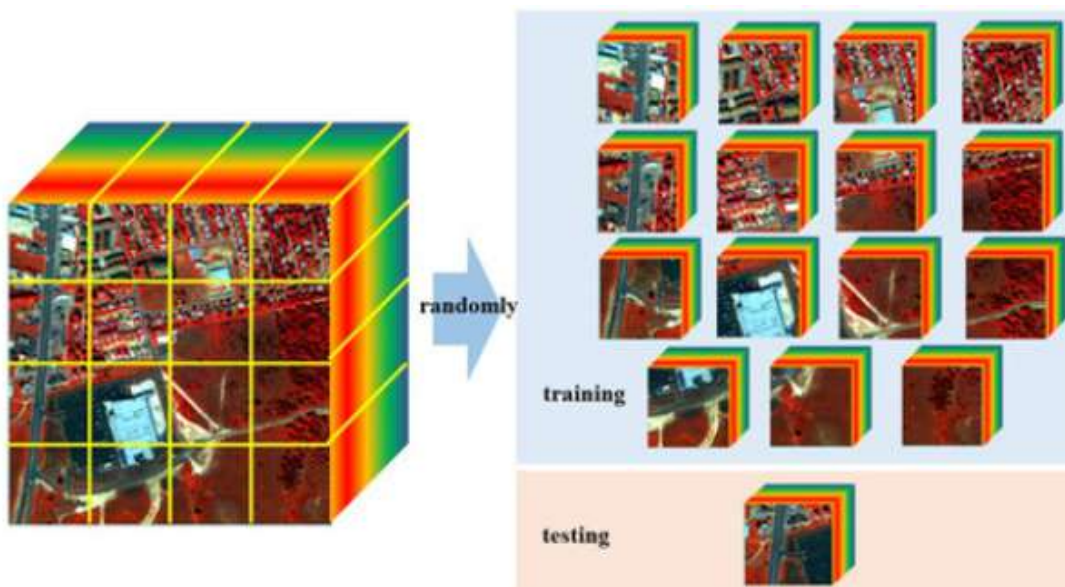
O objetivo desse tópico foi investigar e avaliar quais são os procedimentos descritos nos trabalhos relacionados pertinentes ao fluxo de treinamento e validação. Dessa forma, questões como: quais métodos foram utilizados, como foi feita a divisão entre os dados de treinamento e validação, e se abordagem como *data augmentation* foram efetuadas, possam ser respondidas.

Com base nas investigações percebe-se que diferentes abordagens para divisão dos conjuntos em dados de treinamento e validação foram aplicadas. Entretanto, antes de detalhar quais

são essas técnicas é importante acentuar que alguns *dataset*, conforme aprofundado na subseção 3.5.4, são constituídos por uma ou mais imagens de grandes dimensões, proporcionando assim a necessidade em fragmentar tais imagens em pequenos fragmentos, de maneira que se possa ampliar o conjunto de dados e reduzir o nível de esforço computacional da rede.

Dentre as abordagens de divisão entre dados de treinamento e validação, alguns trabalhos como por exemplo de Mei et al. (2017); Zheng et al. (2019); Jiang et al. (2019); Liu e Lee (2019); Li, Wang e Li (2021); Wang et al. (2021b); Yang et al. (2020) e Hu et al. (2021), adotaram a ideia de selecionar a imagem de teste, como sendo uma pequena fatia que contemplasse a maior variabilidade de detalhes, ao se comparar com toda a imagem, enquanto o restante era definida como dados de treinamento. Seguindo a mesma ideia de seleção, porém de forma randomizada, os trabalhos propostos por Yuan, Zheng e Lu (2017); Hu, Li e Xie (2017); Li et al. (2019); Jiang et al. (2020); Dou et al. (2020); Li, Wang e Li (2020); Zou e Huang (2020); Li, Wang e Li (2021) e Zhao et al. (2021), utilizaram da técnica de decompor a imagem em pequenas subimagens, de modo que ao final as imagens utilizadas para teste eram selecionadas de forma randômica, conforme mostra a Figura 15.

Figura 15: Divisão entre dados de treinamento e teste



Fonte: Retirado de (DOU et al., 2020)

Além das técnicas supracitadas, métodos mais robustos também foram empregados, como é caso da validação cruzada. Esse método fundamenta-se em treinar vários modelos de *deep learning* em diferentes subconjunto de dados de treinamento e validação, cujo a abordagem mais famosa é a validação cruzada *k-fold* (REFAEILZADEH; TANG; LIU, 2009). Dentre os trabalhos revisados, os artigos de Li et al. (2017) utilizaram de respectiva técnica.

Em complemento ao fluxo de treinamento revisados, proposta como *data augmentation* foram implementadas. Essa técnica concentra-se em aumentar a quantidade e diversidades de dados do *dataset* a partir de tratamentos na imagens como, Rotação, Translação, Ampliação, Mo-

Tabela 6: Lista dos diferentes tipos de métricas empregadas

MÉTRICAS DE AVALIAÇÃO	FREQUÊNCIA
PSNR	18
SAM	18
SSIM	18
ERGAS	10
RMSE	9
CC	5
UIQI	3

Fonte: Elaborado pelo autor.

dificação da perspectiva, Achatamento e Distorção (GOODFELLOW; BENGIO; COURVILLE, 2016). Dentre os trabalhos selecionados apenas 22% dos trabalhos utilizaram dessa técnica.

3.5.6 Métricas de validação

As métricas de validação são de extrema importância para qualquer trabalho que faça o uso de *Deep Learning*, pois a partir dessas métricas é mensurado a qualidade de predição do modelo proposto. Para avaliação dos modelos inteligentes é indispensável que mais de uma métrica seja empregada, visto que diferentes performance podem ser encontradas. Com base nesse contexto, foram apreciadas as principais métricas de comparação entre imagens reais e preditas referenciadas nos artigos revisados. Na Tabela 3.5.6 são registradas todas as métricas mencionadas e a frequência que elas foram aplicadas.

Dos tipos de métricas extraídas, conforme mostra a Na Tabela 3.5.6, todos os artigos selecionados utilizaram as métricas **PSNR**, **SAM** e **SSIM**, o que mostra uma forte preocupação dos artigos em analisar as variáveis como a relação do sinal ruído, a consistência espectral e a similaridade entre a imagem hiperespectral predita e sua verdade de campo. Mais detalhes sobre os métodos são mencionados a seguir:

- **Peak Signal-to-Noise Ratio - (PSNR)**: Calcula o índice entre a potência máxima possível de uma imagem e a potência do ruído que afeta a qualidade de sua representação. Sua equação 3.1 é estabelecida pela divisão entre o valor máximo do pixel na banda pela equação 3.2 que representa o Erro Quadrático Médio (MSE).

$$PSNR = 10 \log \left(\frac{\max(I_{HR})}{MSE} \right) \quad (3.1)$$

$$MSE = \frac{\sum_{I_{HR}=1}^X \sum_{\hat{I}_{HR}=1}^Y (I_{HR} - \hat{I}_{HR})^2}{XY} \quad (3.2)$$

Onde X e Y são o número de colunas e linhas da imagem.

- **Spectral Angle Mapper - (SAM):** Regularmente é usado para avaliar a preservação da resposta espectral. A sua equação 3.3 mensura a similaridade espectral a partir da variação angular entre o pixel de referência e o pixel predito.

$$SAM(I_{HR}, \hat{I}_{HR}) = \frac{1}{XY} \sum_{i=1}^{XY} \arccos\left(\frac{\langle I_{HR}, \hat{I}_{HR} \rangle}{\|I_{HR}\|_2 \|\hat{I}_{HR}\|}\right), \quad (3.3)$$

Onde \langle, \rangle representa a operação de produto escalar e $\|\cdot\|_2$ configura a normal euclidiana.

- **Structure Similarity Index Measurement - (SSIM):** Mede a consistência estrutural entre a imagem estimada e a imagem de referência. A sua formulação é dada pela equação 3.4, onde $\mu_{I_{HR\{i\}}}$ e $\mu_{\hat{I}_{HR\{i\}}}$ simbolizam o valor médio da banda espectral i ; $\sigma_{I_{HR\{i\}}}$, $\sigma_{\hat{I}_{HR\{i\}}}$ representam a variância da imagem real e predita e $\sigma_{I_{HR\{i\}}\hat{I}_{HR\{i\}}}$ corresponde ao valor da covariância.

$$SSIM_{(I_{HR}, \hat{I}_{HR})} = \frac{(2\mu_{I_{HR\{i\}}}\mu_{\hat{I}_{HR\{i\}}} + C_1)(2\sigma_{I_{HR\{i\}}\hat{I}_{HR\{i\}}} + C_2)}{(\mu_{I_{HR\{i\}}}^2 + \mu_{\hat{I}_{HR\{i\}}}^2 + C_1)(\sigma_{I_{HR\{i\}}}^2 + \sigma_{\hat{I}_{HR\{i\}}}^2 + C_2)} \quad (3.4)$$

As constantes C_1 e C_2 são adicionadas no intuito de estabilizar a função, evitando assim problemas de divisão por zero. Essas constantes são estabelecidas por $(K_1L)^2$ e $(K_2L)^2$ respectivamente, onde K_1 e K_2 são valores fixados e L é determinado pela dimensão do pixel.

- **Root Mean Square Error(RMSE):** Representa a raiz quadrada das diferenças entre a imagem real e a imagem predita, conforme ilustra a Equação 3.5. A variável I representa o número de bandas espectrais.

$$RMSE((I_{HR}, \hat{I}_{HR})) = \frac{1}{I} \sum_{i=1}^I \sqrt{\text{MSE}(I_{HR\{i\}}, \hat{I}_{HR\{i\}})} \quad (3.5)$$

- **Erreur Relative Globae Adimensionnelle de Synthèse - (ERGAS):** Em função do $RMSE$ (Equação 3.5) não ser sensível na detecção de diferenças espectrais e espaciais entre imagens, o índice ERGAS foi estabelecido no intuito de solucionar esse impasse (WALD, 2000). Esse índice (Equação 3.6) calcula RMSE normalizado por banda e então o divide pelo fator espacial entre as imagens.

$$ERGAS(I_{HR}, \hat{I}_{HR}) = 100 \frac{1}{r} \sqrt{\frac{1}{I} \sum_{i=1}^I \left(\frac{\text{RMSE}((I_{HR\{i\}}, \hat{I}_{HR\{i\}}))}{\mu(\hat{I}_{HR\{i\}})} \right)^2} \quad (3.6)$$

Onde r é o fator de ampliação entre as imagens de comparação.

- **Correlation Coefficient (CC):** Coeficiente que mede o grau de correlação entre a imagem real e predita, como mostra a equação 3.7.

$$CC(I_{HR}, \hat{I}_{HR}) = \frac{1}{I} \sum_{i=1}^I CCI(I_{HR\{i\}}, \hat{I}_{HR\{i\}}) \quad (3.7)$$

$$CCI(I_{HR}, \hat{I}_{HR}) = \frac{\sum_{j=1}^{XY} (I_{HR\{j\}} - \mu_{I_{HR}})(\hat{I}_{HR} - \mu_{\hat{I}_{HR}})}{\sqrt{\sum_{j=1}^{XY} (I_{HR\{j\}} - \mu_{I_{HR}})^2 \sum_{j=1}^{XY} (\hat{I}_{HR\{j\}} - \mu_{\hat{I}_{HR}})^2}} \quad (3.8)$$

- **Universal Quality Image Index (UIQI):** Essa métrica é a combinação de três componentes diferentes que são: a correlação linear, semelhança de contraste e equivalência de luminância (WANG; BOVIK, 2002). Esse índice é definido pela seguinte Equação

$$UIQI(I_{HR}, \hat{I}_{HR}) = \frac{1}{L} \sum_{i=1}^I \frac{\sigma_{I_{HR\{i\}} \hat{I}_{HR\{i\}}}}{\sigma_{I_{HR\{i\}}} \sigma_{\hat{I}_{HR\{i\}}}} \frac{2\mu_{I_{HR}} \mu_{\hat{I}_{HR}}}{\mu_{I_{HR}}^2 + \mu_{\hat{I}_{HR}}^2} \frac{2\sigma_{I_{HR\{i\}}} \sigma_{\hat{I}_{HR\{i\}}}}{\sigma_{I_{HR\{i\}}}^2 + \sigma_{\hat{I}_{HR\{i\}}}^2}. \quad (3.9)$$

Em resumo, quando os valores de PSNR, SSIM, CC e UIQI são maiores, o HI estimado é semelhante ao HI de referência. Além disso, quanto menores forem os valores de RMSE, ERGAS e SAM, melhor será a qualidade de reconstrução do HI estimado.

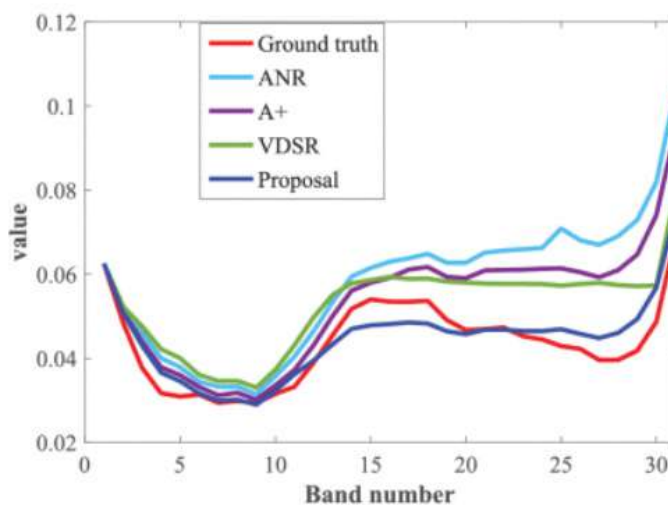
3.5.7 Avaliação da resposta espectral

Como já mencionado anteriormente, as imagens hiperespectrais são representadas por um cubo de dados tridimensionais (3D) de fusão espectral espacial que contém informações espaciais e espectrais. O espectrômetro desse tipo de imagem fornece dezenas a centenas de informações espectrais de banda estreita para cada pixel, produzindo uma curva espectral completa e contínua. Porém, esse ganho na resposta espectral acarreta na degradação da resposta espacial (MATHER; TSO, 2016). Os métodos de S-R baseados em *Deep Learning* vêm se mostrando como uma excelente abordagem como meio de contornar esse respectivo problema. Contudo, é válido ressaltar que a resposta espectral deve ser mantida ao longo desse processo. Nesse propósito essa questão de pesquisa foi elaborada no sentido de verificar como a resposta espectral das HI preditas são avaliadas.

Ao analisar os trabalhos, verificou-se que duas abordagens distintas foram utilizadas para avaliar a consistência espectral dos resultados encontrados. A primeira abordagem nomeada por alguns autores como análise quantitativa (MEI et al., 2017; ZHENG et al., 2019), baseiam-se somente na avaliação dos resultados encontrados pelas métricas SAM e ERGAS (Equações 3.3 e 3.6), de modo que se ambas as métricas apresentarem resultados próximos a 0 é considerado que a consistência espectral foi preservada.

Nomeada como abordagem Qualitativo, o segundo método de avaliação resumi-se em comparar visualmente a assinatura espectral. Essa comparação parte do princípio de selecionar feição ou pixel homólogo tanto na imagem real e predita, e comparar o comportamento da curva espectral de ambas a regiões. Se a curva espectral predita apresentar o comportamento geométrico semelhante a curva real, o resultado é considerado satisfatório. Na Figura 16 é ilustrado um exemplo tirado do trabalho de Hu, Li e Xie (2017).

Figura 16: Comparação Qualitativa



Fonte: Hu, Li e Xie (2017).

3.6 Conclusão

Fundamentando no princípio da possibilidade em aprender a função de mapeamento entre imagens de LR e HR por meio de supervisão, métodos baseados em *Deep Learning* vêm se destacando nos últimos anos, mostrando assim uma forte capacidade em gerenciar essa abordagem (MEI et al., 2017; DOU et al., 2020; LI; WANG; LI, 2021). Contudo, imagens multiespectral e especialmente hiperespectral requerem de procedimentos complexos, ao se compararem com as técnicas de S-R em imagens RGB, uma vez que a resposta espectral das bandas contemplam de características valiosas que podem ser facilmente degradadas. Nesse intuito, esse Capítulo foi criado com objetivo de averiguar as potenciais técnicas de *Deep Learning* nos métodos de super-resolução em HI, de modo que questões apresentadas no início desse capítulo pudessem ser averiguadas.

Em relação as questões de pesquisa QP1 e QP2, percebe-se que os métodos de *Deep Learning*, segundos os trabalhos revisados, se destacam em gerar resultados superiores aos processamentos de imagens tradicionais, já que em totalidade, os estudos fizeram comparações que comprovam essa respectiva afirmação. No entanto, ainda não exista uma abordagem, baseada em *Deep Learning* superior as demais, logo estudos que testam diferentes técnicas e arquiteturas

de redes possam agregar a essa linha de pesquisa.

Ainda em complemento as questões supracitadas, é possível notar que as redes convolucionais 2D e 3D, implementadas com as funções de perda e ativação MAE ou MSE e *ReLU* respectivamente e o otimizador do tipo Adam, são as abordagens com maior frequência de ocorrência dentre os artigos revisados. Para avaliar os respectivos modelos, métricas de comparação de imagens como PSNR, SAM, SSMI, ERGAS, UIQI e RSME também são usualmente empregadas.

Em contrapartida ao tipo de arquitetura, diferentes proposta foram testadas, porém algumas observações, como são descritas a seguir, foram mencionadas em mais de um estudo:

- Os modelos S-R adotam a rede convolucional sem camada de *pooling*, porque tal processo degrada a resposta espectral (LIU; LEE, 2019; WANG et al., 2021b).
- Em geral, modelos profundos têm melhor capacidade de aprendizagem do que uma pequena rede, que não pode aprender mapeamento entre as imagens devido à capacidade limitada, mas a rede mais complexa nem sempre garanti um melhor resultado (LI et al., 2019; LI; WANG; LI, 2020).
- Em S-R é crucial extrair mais informações de contexto em torno de cada pixel. Logo, kernel de maiores dimensões, por exemplo, 9×9 ou 7×7 , podem contribuir para melhorar o desempenho na predição da resposta espacial (JIANG et al., 2020; LI; WANG; LI, 2020; WANG et al., 2021b).
- O mapeamento residual é mais sensível às mudanças de saída, portanto, módulos residuais contribuem em aumentar a capacidade de reconhecimento de recursos e melhorar a eficiência do treinamento (ZHENG et al., 2019; LI; WANG; LI, 2020).
- CNN 3D geram melhores resultados no contexto espectral do que as abordagens empregando CNN 2D (MEI et al., 2017; LIU; LEE, 2019; LI; WANG; LI, 2021)

Em comparação aos tipos de *datasets* de imagens hiperespectrais, percebe-se que em função a não popularidade das imagens hiperespectrais e a dificuldade em obtê-las, o número de amostras de treinamento é extremamente pequena. Nesse sentido, a necessidade em se adquirir um *datasets* robusto que contemplem imagens áreas e terrestres com diferentes condições de iluminação, torna-se algo extremamente crucial para as técnicas de S-R.

Outro ponto pertinente ao *datasets* são os dados empregados como rótulos de treinamento. Para aquisição desses tipos de informações, as imagens consideradas de alta resolução são degradadas por funções de interpolação, concebendo assim, uma imagem de baixa resolução e utilizadas como rótulos de treinamento. Embora essa abordagem seja comumente aplicada dentro do contexto da S-R, conforme mostrado nos artigos referenciados, as imagens de LR interpoladas, não conseguem representar fielmente uma HI de LR real, pois essas imagens são formadas

por estruturas complexas, que contemplam inúmeros fatores, como vibração do sensor, corrupção de ruído, distorções das lentes, entre outros. Portanto, quando aplicado a cenários reais, a maioria das abordagens existentes não consegue generalizar de forma adequada as imagens (ZHANG et al., 2020).

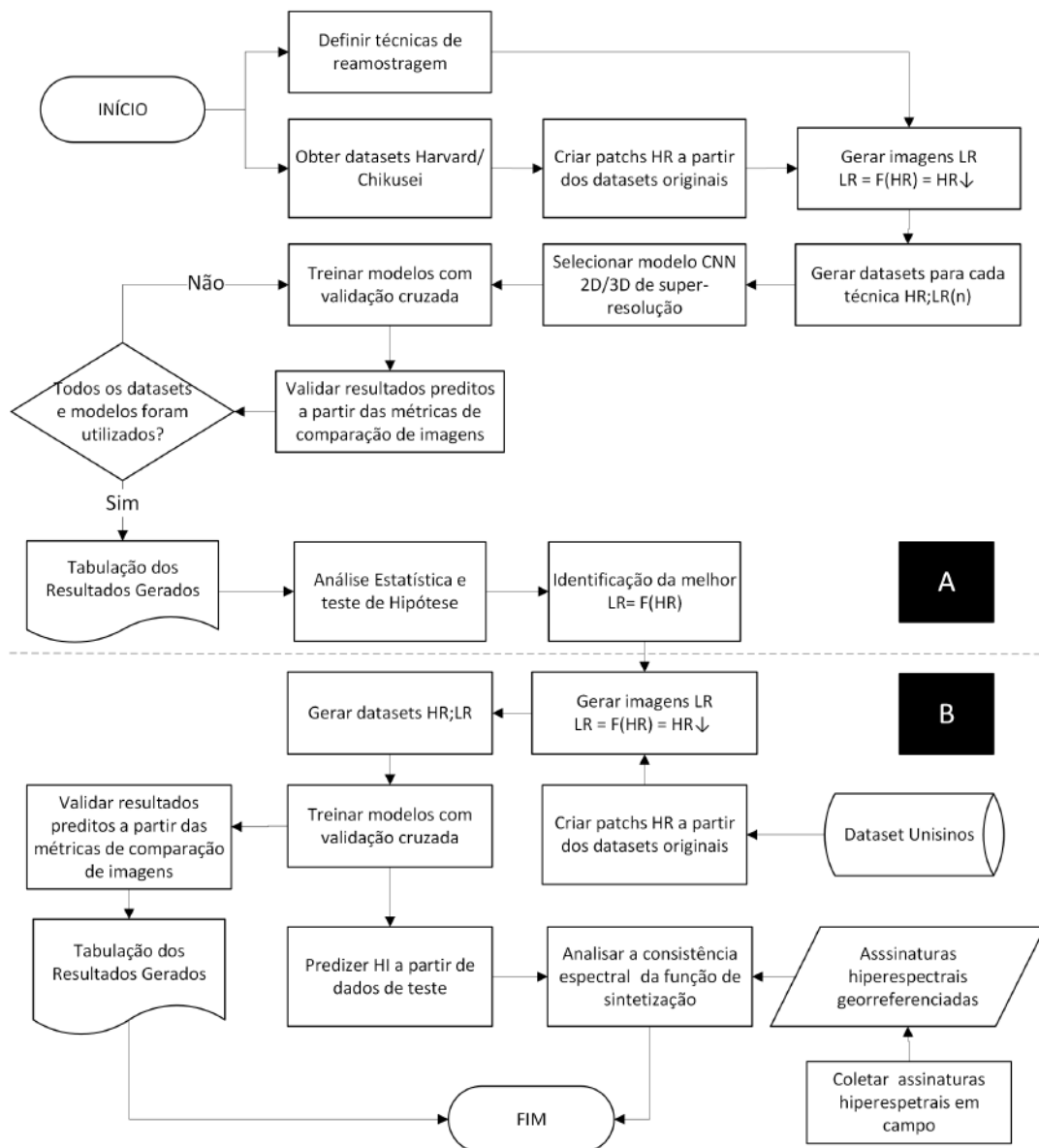
Para avaliar a consistência espectral das HI previstas pelos modelos, análise gráfica e métricas como SAM e ERGAS são bastante aplicadas e difundidas. Contudo, nota-se que outros tipos de abordagens comparativas ainda podem ser testadas, como por exemplo a comparação da curva espectral coletadas por sensores de alta resolução espectral, como é o caso do Espectrorradiômetro. Desse modo, comparações mais apuradas podem ser realizadas.

Em suma, pode se concluir que diferentes métodos baseados em Deep Learning, no intuito de aprimorar a resolução espacial sem degradar a resposta espectral em HI venham sendo testados. Entretanto, ainda há muitas combinações de técnicas a serem exploradas e aprimoradas, uma vez que essa metodologia se encontra em seu estágio inicial. Logo, estudos que busquem aprimorar e otimizar o processo garantindo resultados de ótima qualidade precisam ser explorados. Portanto essa pesquisa se enquadra nesse contexto, uma vez que foram avaliadas as técnicas de reamostragem em aprimorar os métodos de SR.

4 MATERIAL E MÉTODOS

Neste capítulo são descritos os aspectos metodológicos pertinentes para construção desse projeto de dissertação. Para descrição dos respectivos métodos, duas etapas foram criadas, de modo que cada etapa represente um subprocesso desenvolvido para criação do estudo. As duas etapas são referentes à: Formulação e Avaliação da melhor função de reamostragem de HI de Baixa Resolução Espacial e Análise e Validação da consistência espectral. Na Figura 17 é apresentando o Fluxograma resultante de todo processo, onde a Etapa 1 é representando pela letra **A** e a Etapa 2 pela letra **B**.

Figura 17: Fluxograma da metodologia empregada



Fonte: Elaborado pelo autor.

4.1 Formulação e Avaliação da melhor função de Reamostragem de HI de Baixa Resolução

Conforme referenciando na subseção 2.3, as abordagens de S-R no contexto de *deep learning* partem do princípio em aprender a função de mapeamento entre uma imagem de LR e HR, de modo que os dados de entrada na rede são as imagens de LR e como resultado esperado as imagens de HR. Contudo, em consequência das dificuldades em se obter HI ou MI de baixa e alta resolução espacial devidamente registradas, as cenas de LR são sintetizadas a partir das imagens de HR por meio da junção de processos, que envolvem técnicas de degradação, seguida posteriormente por uma reamostragem e finaliza com adição de ruídos (Equação 2.9). Dessa forma, é possível construir *datasets* que contemplem HIs ou MIs de baixa e alta resolução espacial. Para essa proposta de dissertação são avaliadas as influências das técnicas de reamostragem no contexto de S-R em HIs.

4.1.1 Técnicas de Reamostragem

O processo de reamostragem fundamenta-se em alterar o número de pixels na imagem, de maneira que possibilite aumentar ou diminuir a resolução espacial da imagem (SHLIEN, 1979). O resultado desse processo será uma imagem com valores reamostrados em novas posições do *grid*. A sua importância baseia-se em diversas operações de processamento digital de imagens, entre elas se destacam: registro entre as imagens de igual ou diferentes resoluções espaciais e a correção geométrica. Para o desenvolvimento dessa pesquisa, nove diferentes tipos de métodos de reamostragem foram referenciados, que são eles: *Average*, *Mode*, *Root Mean Square - RMS*, *Nearest Neighbor*, *Bilinear*, *Bicubic*, *CubicSpline*, *Gaussian*, e *Lanczos*. É válido salientar que a escolha dessas técnicas é fundamentado nas abordagens mais empregadas em Sensoriamento Remoto, com destaque para as abordagens *Bicubic* e *Gaussian*, métodos esses mais citados entre os trabalhos relacionados (Tabela 4). A descrição dos respectivos métodos são listados a seguir:

Average Nessa técnica o pixel reamostrado (x', y') é estimado com base no valor médio dos pixels presentes em um filtro de tamanho $(m \times n)$, onde m representa o número de linha e n o número de colunas.

$$f(x', y') = \frac{\sum_{i=1}^m \sum_{j=1}^n f(x_i, y_j)}{nm} \quad (4.1)$$

Mode : O pixel (x', y') é selecionado com base nos pixels com maior frequência de ocorrência em um filtro do tipo $(m \times n)$.

RMS O valor do pixel reamostrado é estimado a partir do RMS dos pixels, como mostra a Equação 4.2

$$f(x', y') = \sqrt{\frac{\sum_{i=1}^m \sum_{j=1}^n f(x_i, y_j)^2}{nm}} \quad (4.2)$$

Nearest Neighbor Essa abordagem apoia-se em assegurar que o valor do pixel (x', y') represente o valor mais próximo da posição ocupada pelo pixel (x, y) da imagem original (PATIL, 2018). A sua principal desvantagem consiste nas possíveis degradações em detalhes finos e o efeito de 'serrote' em bordas retas da imagem. A sua equação pode ser expressa como:

$$f(x', y) = \begin{cases} f(x, y), & dx < 0.5 \text{ and } dy < 0.5, \\ f(x + 1, y), & dx \geq 0.5 \text{ and } dy < 0.5, \\ f(x, y + 1), & dx < 0.5 \text{ and } dy \geq 0.5, \\ f(x + 1, y + 1), & dx \geq 0.5 \text{ and } dy \geq 0.5. \end{cases} \quad (4.3)$$

Onde os valores de dx e dy correspondem às distâncias euclidianas nas direções x e y respectivamente, entre os pontos (x', y') e (x, y) .

Bilinear Esse método calcula a média ponderada das distâncias dos quatro pixels mais próximos a fim de determinar a intensidade do pixel (x', y) na imagem reamostrada. Essa técnica resulta na suavização das bordas dos objetos (PEDRINI; SCHWARTZ, 2008), a sua equação é dada por:

$$f(x' y') = (1 - dx)(1 - dy)f(x, y) + dx(1 - dy)f(x + 1, y) + \Delta \quad (4.4)$$

$$\Delta = (1 - dx)dyf(x, y + 1) + dxdyf(x + 1, y + 1) \quad (4.5)$$

Bicubic Esse tipo de reamostragem é aplicado a fim de obter uma estimativa suavizada da imagem, de modo que os detalhes finos sejam preservados. A sua estimativa é obtida a partir da Equação 4.10, no qual utiliza 4 x 4 pixels ao redor do pixels em questão para calcular o valor de intensidade reamostrado (PEDRINI; SCHWARTZ, 2008).

$$f(x', y') = \sum_{m=0}^3 \sum_{n=0}^3 = f(x + m, y + n)R(m - dx)R(dy - n) \quad (4.6)$$

sendo

$$R(s) = \frac{1}{6}[P(s + 2)^3 - 4P(s + 1)^3 + 6P(s)^3 - 4P(s - 1)^3] \quad (4.7)$$

$$P(t) = \begin{cases} t, & t > 0, \\ 0, & t \leq 0. \end{cases} \quad (4.8)$$

CubicSpline Semelhante ao método de *Nearest Neighbor*, o algoritmos de reamostragem *Cubic Spline* se destaca por reamostrar os 16 pixels mais próximo do pixel em análise, no entanto, a partir do ajuste dos pontos sobre a curva garante maior suavidade nos resultados de saída, sendo essa característica com maior destaque para esse tipo de técnica (Equação 4.9). Para mais detalhe sobre o procedimento empregado nessa técnica sugeriu-se os seguintes trabalhos Keys (1981) e Chen, Li e Zhao (2016).

$$f(k) = \begin{cases} 3k^3 - 6k^2 + 4, & \text{for } 0 \leq \text{abs}(k) < 1 \\ -k^3 - 6k^2 + -12k + 8, & \text{for } 1 \leq \text{abs}(k) < 2 \\ 0 & \text{otherwise.} \end{cases} \quad (4.9)$$

Onde k representa o *kernel* analisado na imagem.

Gaussian Caracterizada por obter imagens suavizadas, a reamostragem do tipo gaussiana estima os novos valores de intensidade fundamentados na curva de Gauss, possibilitando assim uma melhor distribuição nos valores de intensidade. A sua estimativa é estabelecida pela seguinte função :

$$f(x', y') = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{(x-x_\mu)^2 + (y-y_\mu)^2}{2\sigma^2}} \quad (4.10)$$

Lanczos Essa técnica se destaca em preservar e suavizar as pequenas feições reamostradas da imagem original, de maneira a garantir maior nitidez e redução no efeito de serrilhamento. A sua função de ordem n é detalhada conforme mostra a Equação 4.11. Para mais detalhe sobre o método sugere-se a seguinte referência Dumitrescu e Boiangiu (2019).

$$f(k) = \begin{cases} \text{sinc}(k)\text{sinc}(\frac{k}{n}), & \text{for } \text{abs} \leq n \\ 0 & \text{otherwise.} \end{cases} \quad (4.11)$$

$$\text{sinc}(k) = \begin{cases} 1, & \text{for } k = 0 \\ \sin(\pi k)/(\pi k), & \text{otherwise.} \end{cases} \quad (4.12)$$

4.1.2 Aquisição de dados

Concomitantemente à escolha das técnicas reamostragem dois tipos de *datasets* de imagens hiperespectrais foram selecionados e aplicados nesse estudo. A escolha desses *datasets*

baseou-se naqueles conjuntos que apresentassem maior variabilidade de informações imageadas e disponibilizadas. Além disso, a definição de dois *datasets* é justificada no sentido de se obter tanto imagens aéreas como terrestres, uma vez que não foi encontrado na literatura um conjunto de dados que contemplassem ambas as características.

O primeiro *dataset* escolhido foi o *dataset* de *Harvard*. Conforme já relatado na subseção 3.5.4 esse conjunto de dados é composto por cinquenta imagens hiperespectrais de cenas internas e externas sob iluminação diurna e mais vinte e cinco imagens sob iluminação artificial. Ao todo, o arquivo contém 7.7 GB de imagens, sendo cada imagem formada por 31 bandas espectrais contínuas com uma largura espectral de 10nm e localizada entre a faixa de 420nm a 720nm do EE.

Diferente do *dataset* de *Harvard* que contempla somente imagens terrestre, o *dataset* de Chikusei é estruturado por dados hiperespectrais obtido pelo sensor Headwall Hyperspec-VNIR-C sobre áreas agrícolas e urbanas em Chikusei, Ibaraki no Japão. O conjunto de dados hiperespectral tem 128 bandas contínuas localizadas na faixa espectral de 363 nm a 1018 nm do EE. A cena consiste em 2517x2335 pixels e com uma largura espectral de 5.1nm. Ao todo o *dataset* contempla aproximadamente 1.5 GB de informações.

De posse dos respectivos *datasets*, as HIs foram fragmentadas em *patches*, de maneira que cada fragmento corresponda a uma parte específica da imagem. Esses *patches* representam o tamanho de informações que são inseridas na rede, quanto maior o seu tamanho maior é o esforço computacional exigido. Nesse sentido, para esse trabalho foram utilizados *patches* de 64x64 pixels. Cada *dataset* é responsável por gerar um conjunto específico, sendo os *datasets* de *Harvard* e Chikusei responsáveis por gerar 2587 e 1188 *patches* respectivamente.

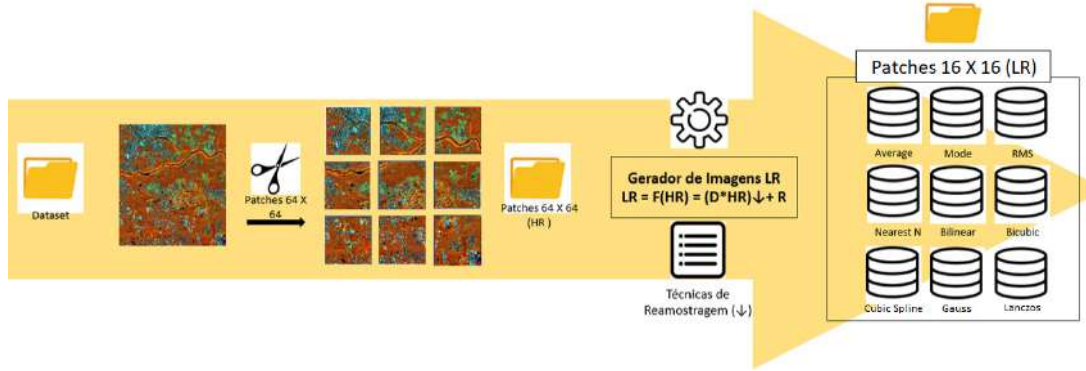
Com base no fluxo apresentado na Figura 17 o próximo estágio é definido pela sintetização das HIs de LR. Nesse processo, por meio das técnicas de reamostragem mais as imagens de HR adquiridas anteriormente, nove diferentes conjuntos compostos por imagens de 16 X 16 pixels foram gerados, de modo que cada conjunto seja representado por uma abordagem de reamostragem. Dessa forma cada *dataset* foi responsável por gerar 1 conjunto de HIs de HR e nove conjuntos de HIs de LR. Na Figura 18 é ilustrado todo o processo, no qual são utilizados os dados de *Chikusei* como exemplo de caso.

4.1.3 Modelos de Super-Resolução por *Deep Learning*

Tendo em vista os diferentes conjuntos de dados elaborados, a próxima fase proposta é a seleção do melhor método de super-resolução de HIs. Para seleção deste respectivo método, foram revisitados os 18 artigos resultantes da revisão da literatura (Capítulo 3) e escolhido os trabalhos que de alguma forma disponibilizam a implementação do processo. Dos trabalhos levantados apenas 4 dispuseram os códigos como mostra a Tabela 7.

Dentre os 4 métodos referenciados para essa proposta de dissertação foi escolhido o método

Figura 18: Elaboração dos Dados de Treinamento



Fonte: Elaborado pelo autor.

Tabela 7: Métodos de super-resolução de HIs

Autor	Nomeclatura	Links
Mei et al. (2017)	3D - FCNN	https://github.com/MeiShaohui/Hyperspectral-Image-Spatial-Super-Resolution-via-3D-Full-Convolutional-Neural-Network
Jiang et al. (2020)	SSPS-R	https://github.com/junjun-jiang/SSPS-R
Li, Wang e Li (2020)	MCNet	https://github.com/qianngli/MCNet
Li, Wang e Li (2021)	ERCSR	https://github.com/qianngli/ERCSR

Fonte: Elaborado pelo autor.

que apresentou os melhores resultados em comparação aos demais, no qual foram utilizados como critérios de escolha: o maior número de *datasets* empregados e avaliados, as melhores métricas de comparação de imagens (subseção 3.5.6) para um mesmo *dataset* e o índice QA (Capítulo 3). Das observações realizadas o algoritmo ERCSR demonstrou maior coerência entre os pontos analisados, sendo esse método empregado para essa proposta.

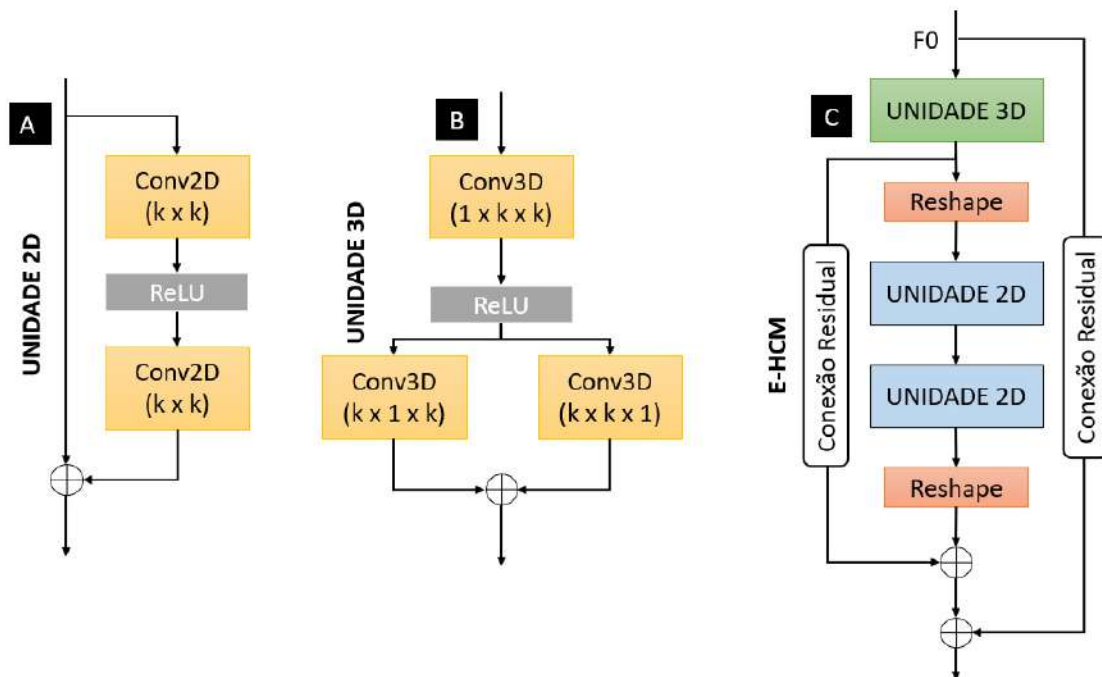
A arquitetura do método de S-R ERCSR é delimitada por três eixos principais sendo eles: extração de recursos, reconstrução da imagem e conexões residuais. Na primeira parte definida como extração de recursos, no intuito de enfatizar as informações espaciais e espectrais uma convolução 3D definida como *separable 3D convolution* ($f_{sconv3D}$) é empregada. Diferente da Convolução 3D tradicional, onde a entrada F_{d-1} é processada por uma camada de convolução com o *kernel* ($k \times k \times 1$), que é aplicado para explorar o conteúdo espacial, a *separable 3D convolution* estuda a relação da dimensão espacial e espectral separando o *kernel* em duas formas, ou seja, ($k \times 1 \times k$) e ($k \times k \times 1$) e unindo ao final por um processo de adição. Desse modo, a primeira função do método F_0 é definida de acordo com a seguinte equação:

$$F_0 = f_{sconv3D}(reshape(LR)) \quad (4.13)$$

Onde a função *reshape* redimensiona LR em quadro dimensões (1,L,i,j)

Os resultados encontrados pela Equação F_0 são inseridos no módulo E-HCM. Esse módulo é composto por uma Unidade 3D, e duas Unidades 2D e duas operações de *reshape*. Na unidade 3D é analisado as relações espectral no contexto horizontal e vertical do espaço, com intuito de aumentar a exploração espacial da imagem sob a condição espectral a ser obtida. Concluído esse processo, os *features maps* resultantes são remodelados em quatro dimensões para que as convolução 2D possam ser realizadas. Por meio de duas convoluções do tipo 2D, segundo os autores, proporcionam uma melhor otimização da rede além de viabilizar a capacidade do aprendizado espacial. Por fim, duas conexões residuais são empregadas com o intuito de fundir os resultados encontrados. A Figura 19 a seguir detalha os aspetos de cada parte onde a) representa a unidade 3D, b) a unidade 2D e c) a representação de todo o módulo E-HCM. É válido acentuar que o módulo E-HCM pode obter características de forma hierárquica, porém para esse trabalho só foram utilizados quatro módulos de E-HCM. Por fim, após as estimativas dos módulos, pela convolução 3D de *kernel* (1,1,1) seguida de uma $f_{sconv3D}$ a extração dos recursos é obtida após uma conexão residual.

Figura 19: Estruturas da Rede ERCSR

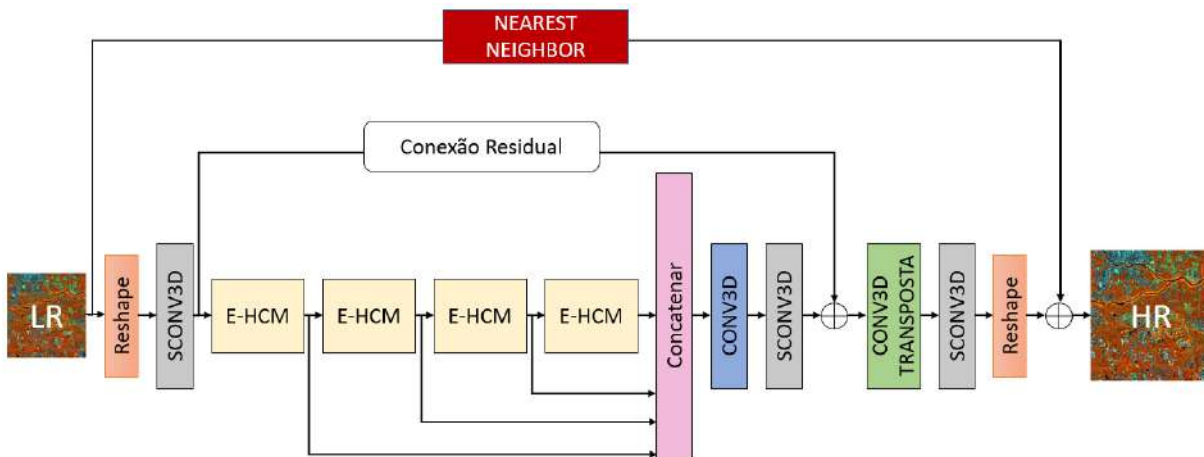


Fonte: Adaptado de Li, Wang e Li (2021).

Para a reconstrução da imagem, as informações obtidas pelas camadas de extração de recursos, são utilizadas em um processo de convolução 3D transposta seguido por uma $f_{sconv3D}$ para ao final serem estruturadas no formato (L,I,J) por uma função do tipo *reshape*. Concluído a estruturação, uma imagem de HR é estimada com base em uma conexão residual a partir da imagem estruturada e a imagem LR de entrada reamostrada a partir da técnica de *Nearest Neighbor*. Na Figura 20 é exemplificado todas as camadas que constituem o método de S-R ERCSR.

Para manter as mesmas condições propostas pelo método de S-R ERCSR é necessário que os parâmetros de treinamento da rede sejam iguais. Nesse sentido, foi adotado para o treinamento a função de otimização do tipo ADAM com uma taxa de aprendizado $lr = 0.0004$ e valores de betas iguais a $\beta_1 = 0.9$ e $\beta_2 = 0.999$. A função de perda empregada é do tipo L1 *loss*, função esta semelhante ao cálculo do erro absoluto. Ao todo foram utilizadas 200 épocas de treinamento com o tamanho de *Batch size*=12. Contudo, devido ao grande esforço computacional exigido, foi necessário a redução do *Batch size* para 6, sendo essa a única diferença pontuada em comparação a proposta original. Todo o algoritmo é desenvolvido na estrutura Pytorch dentro do ambiente Python. Para treinamento validação dos resultados, todo o processo foi executado na plataforma Colab PRO, onde são disponibilizadas tais GPUs: NVIDIA Tesla P100 e NVIDIA T4.

Figura 20: Arquitetura completa do método de S-R ERCSR



Fonte: Adaptado de Li, Wang e Li (2021).

4.1.4 Treinamento e Validação

O fluxo de treinamento e validação aplicados nesse estudo corresponde ao mesmo fluxo proposto no artigo *Analysis of machine learning techniques for carbonate outcrop image classification in Landsat 8 satellite data*. Artigo este elaborado pelo autor desse projeto de dissertação e aceito no *International Geoscience and Remote Sensing Symposium-(IGARSS)*, 2022. Em síntese o fluxo proposto é apresentado no Pseudocódigo 1. A sua estrutura é composta por quatro processos iterativos, com destaque para as iterações presentes nas linhas três e cinco que correspondem aos processos de Reprodutibilidade do Experimento e Validação Cruzada respectivamente. A Reprodutibilidade do Experimento consiste no processo em definir parâmetros aleatórios para um valor fixo (*seed*), de modo que ao importar o valor fixado para dentro do experimento os resultados encontrados podem ser replicados. Dessa forma é possível reestruturar os dados do experimento a fim de proporcionar maior variabilidade de possíveis combinações

nos conjuntos de treinamento e validação. Para este trabalho foram estabelecidos apenas dois valores de *seed*, sendo eles os números 1 e 2.

Algorithm 1 Treinamento e Validação das Funções de Sintetização de HI de LR

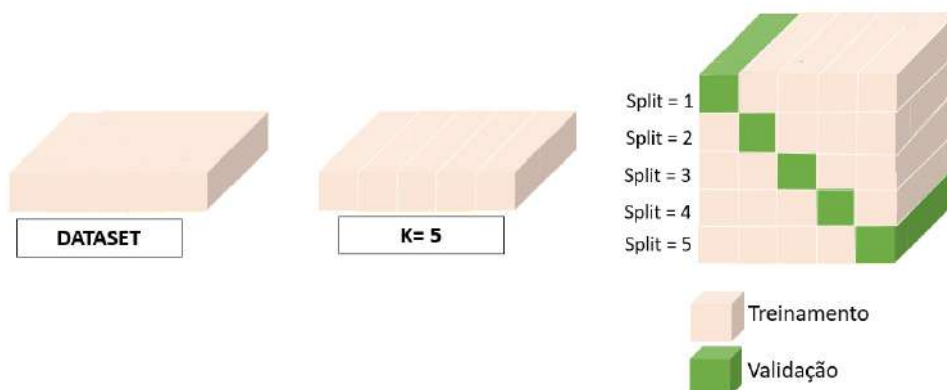
```

1: for iterações = 1,2...Nº de Funções do
2:   Criar Tabela
3:   for iterações = 1,2...seed do
4:     Definir os dados de treinamento e validação partir da Validação Cruzada
5:     for iterações = 1,2...k do
6:       Reinicializar modelo de super-resolução ERCSR
7:       for iterações = 1,2...Épocas do
8:         Treinar ERCSR
9:       end for
10:      Validar ERCSR treinado
11:      Salvar os resultados na Tabela
12:    end for
13:  end for
14:  Exportar Tabela
15: end for

```

O processo de Validação Cruzada (CV) é um método estatístico que busca comparar e avaliar os algoritmos de *Machine Learning* a partir da segmentação dos dados em dois grupos, sendo o primeiro empregado para treinar o modelo, enquanto o segundo, avalia as previsões (REFAEILZADEH; TANG; LIU, 2009). Para o trabalho, é adotado a CV do tipo k-fold com valor de $k=5$, conforme empregado em Li et al. (2017). Nesse tipo de validação, os dados são particionados em k segmentos de tamanho igual ou quase igual, para posteriormente dentro de um processo iterativo com k iterações, o treinamento e validação são realizados de maneira que cada iteração, com $(k-1)$ segmentos são usados para no aprendizado do modelo, enquanto o restante é empregado para validação. A Figura 21 demonstra um exemplo com $k = 5$, onde os *splits* correspondem a classificação dos conjunto em cada iteração.

Figura 21: Validação Cruzada do tipo K-fold



A fim de proporcionar redundância de resultados para posteriores avaliações, é estabelecido ao iniciar os *splits*, a reinicialização do modelo de S-R ERCSR. Dessa forma, é possível treinar e validar k modelos a partir do mesmo conjunto, variando apenas os segmentos de treinamento e validação. Essa variabilidade na quantidade de modelos corrobora para que métricas de avaliação de resultados possam ser analisadas de forma mais apurada, a partir de métodos estatísticos, como média e desvio padrão.

Para avaliação dos experimentos foram mensurados em cada *split* as comparações entre as imagens reais e preditas. Essas correlações foram realizadas com base nas métricas de comparação de HIs descritas na subseção 3.5.6 a partir dos segmentos de validação. Por meio de tais resultados, é possível analisar e identificar a melhor função de sintetização de imagens de HIs de LR utilizando para isso testes de hipóteses, como são os casos dos teste de Friedman e Nemenyi.

Aplicado para comparar dados amostrais vinculados, quando o mesmo indivíduo é avaliado mais de uma vez, o teste não-paramétrico de Friedman compara os ranks médios (R_j) de Q algoritmos em D conjuntos de dados, de modo que caso todos os R_j (Equação 4.14) sejam equivalentes, a hipótese nula H_0 é aceita (POHLERT, 2014). Nesse tipo de teste, os dados numéricos não são analisados diretamente, mas sim suas posições ocupadas após a ordenação de cada grupo separadamente. Por meio dessa ordenação, as hipóteses H_0 e alternativa H_1 pode ser analisadas das seguintes maneiras:

H_0 As pontuações da população em qualquer um dos grupos relacionados NÃO são sistematicamente superiores ou inferiores às pontuações da população em qualquer um dos outros grupos relacionados.

H_1 As pontuações da população em alguns dos grupos relacionados são sistematicamente mais altas ou mais baixas do que as pontuações da população em outros grupos relacionados

Para aferição estatística, o respectivo teste utilizada da distribuição χ_F^2 (Equação 4.15) com $Q - 1$ graus de liberdade, para posteriormente mensurar o valor de F_F com base na Equação 4.16. Por meio do valor de F_F mais a distribuição de F de de Fisher com $Q - 1$ e $(Q - 1)(D - 1)$ graus de liberdade ($F_{Q-1, (Q-1)(D-1)}$), a H_0 é testada de tal forma que se $F_F > F_{Q-1, (Q-1)(D-1)}$ a hipótese é rejeitada.

$$R_j = \frac{1}{D} \sum_i r_i^j \quad (4.14)$$

$$\chi_F^2 = \frac{12D}{Q(Q+1)} \sum_j R_j^2 - \frac{Q(Q+1)^2}{4} \quad (4.15)$$

$$F_F = \frac{(D-1)\chi_F^2}{D(Q-1) - \chi_F^2} \quad (4.16)$$

Onde j representa o algoritmo e i o i ésimo conjunto de dados.

Embora o teste de Friedman consiga mostrar a similaridade ou não entre os desempenhos dos modelos, quando a sua H_0 é rejeitada o método não consegue indicar entre quais modelos ocorrem essas diferenças, sendo necessário a utilização de testes do tipo *post-hoc*¹ que é o caso do teste de Nemenyi (DEMŠAR, 2006).

Semelhante ao teste de Tukey para ANOVA, o teste de Nemenyi consiste em fazer comparações em pares com o intuito de verificar qual dos fatores que diferem entre si. Essa diferenciação é estimada pela distância crítica (CD) entre os modelos (Equação 4.17).

$$CD = q_\alpha \sqrt{\frac{Q(Q+1)}{6D}} \quad (4.17)$$

Onde: q_α são os valores baseados na distribuição *Studentized range* dividido por $\sqrt{2}$.

4.2 Análise e Validação da consistência espectral

Definida como etapa B do Fluxograma (Figura 17), nessa seção são detalhados os processos usufruídos na validação da consistência espectral. Essa validação busca averiguar o comportamento espectral das imagens preditas pelo método de S-R, ao ser aprimorado pela melhor função de reamostragem. Para essa aferição um novo *dataset* foi desenvolvido de tal forma que seus parâmetros espaciais e espectrais pudessem ser confrontados com as verdades coletadas diretamente em solo.

A primeira etapa para construção do respectivo *dataset*, nomeado como *dataset* Unisinos, consiste no procedimento de levantamento e coleta dos dados. Para aquisição dos dados duas expedições distintas foram estabelecidas, que são a Expedição 1: para o Mapeamento Hiperespectral e a Expedição 2: para coleta das assinaturas espectrais. A área de para coleta se encontra na cidade de São Leopoldo, Rio Grande do Sul, dentro do campus da UNISINOS (Universidade do Vale do Rio dos Sinos), mais precisamente no complexo esportivo da Universidade.

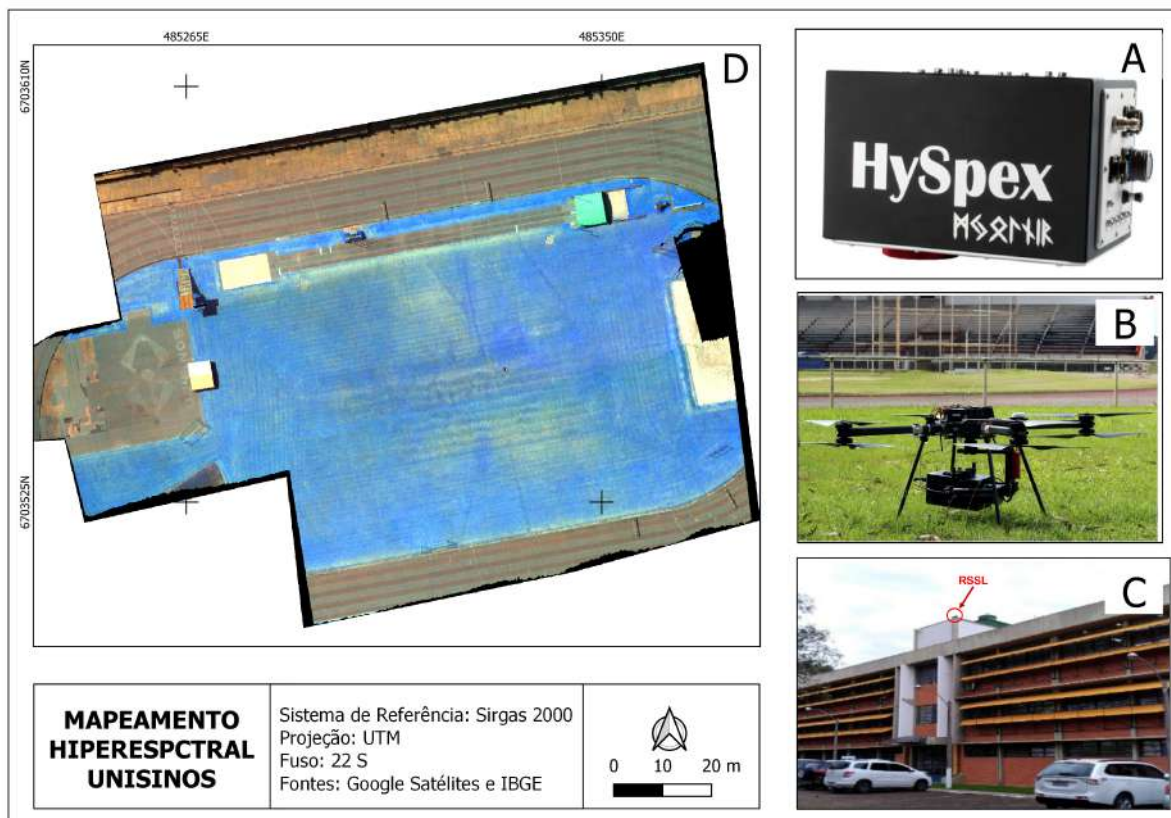
Ocorrida no dia 20/10/2021 a Expedição 1, foi realizada a partir do treinamento promovido pela empresa *HySpex* após a aquisição do sensor hiperespectral *HySpex Mjólnir S-620* pelo grupo VIZLAB. Com a possibilidade de imageamento tanto aéreo como terrestre, o sensor *Mjólnir S-620* contempla de uma resolução espectral entre as faixas 970-2500nm com intervalo de 5.1nm, compondo assim, uma imagem hiperespectral com 300 bandas. A sua resolução espacial é estimada a partir do IFOV equivalente a 0.54/0.54 mrad, sendo necessário uma distância mínima de 20 metros entre objeto e o sensor, de modo que problemas de foco sejam evitados. Por fim, uma resolução radiométrica de 16 bits permite registrar 65536 diferentes tipos de intensidades pelo sensor. Para aquisição da imagem, o respectivo sensor (Figura 22- A) foi embarcado em um Drone do tipo multirotor do modelo *Systems 1400-SE8* (22-B) da empresa BFD, o qual devidamente configurado, permitiu imagear uma área de aproximadamente 11 mil

¹Teste de comparação múltipla que é usado após a aplicação de teste não paramétricos com três ou mais algoritmos.

m^2 , a uma altitude de 120 metros, gerando assim, um mapeamento com uma resolução espacial de aproximadamente 7cm. Em complemento ao imageamento, um alvo de reflectância foi posicionado dentro da região mapeada, para ao final do processamento, correções e conversões pertinentes as respostas espectrais fossem executadas.

Em relação ao aliamento e a correção geométrica das imagens levantadas, foi empregado o *software* PosPac, o qual a partir das observações levantadas das Redes de Monitoramento Contínuo (RBMC), sendo elas as estações de POAL E SLRS (Figura 22-C), mais o posicionamento inercial registrado pelo drone, viabilizou que ajuste das imagens hiperespectrais fossem realizados, a fim de garantir qualidade posicional dos resultados. Por fim, o cubo hiperespectral (FIGURA22-D) foi gerado com suas devidas correções.

Figura 22: Resultado do Imageamento Hiperespectral



Fonte: Elaborado pelo autor.

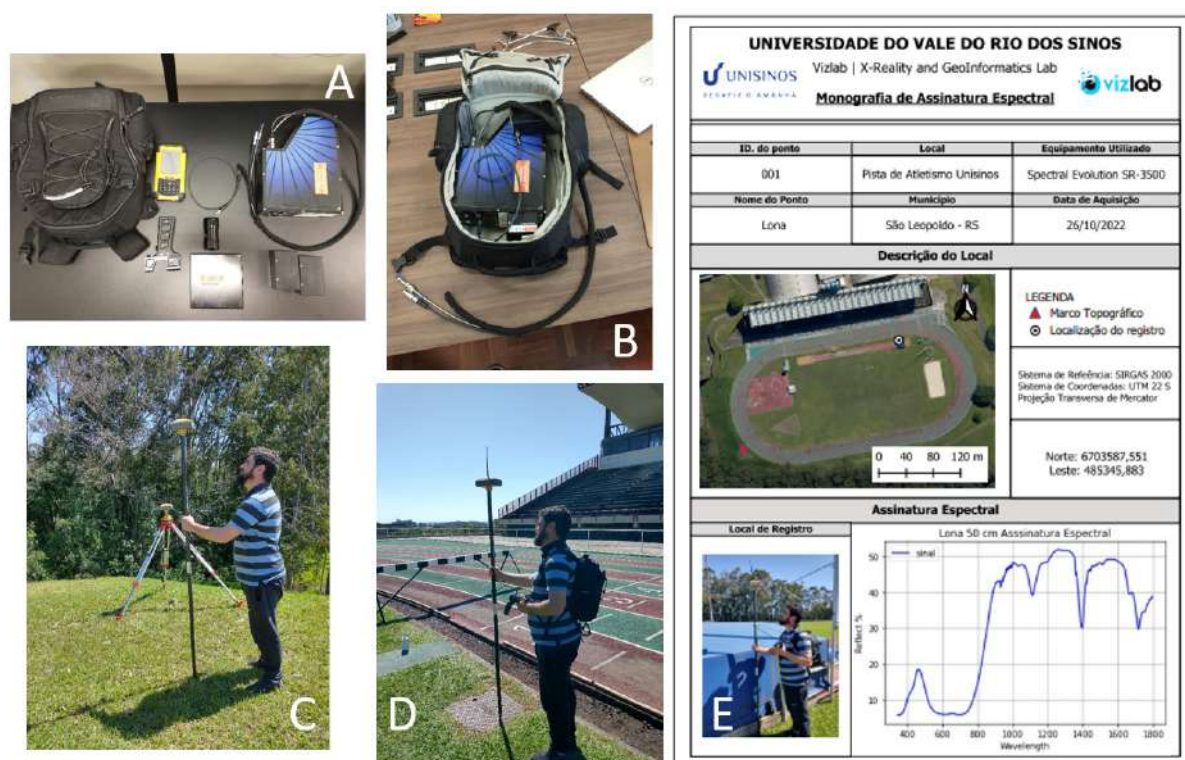
Em oposto a Expedição 1, o qual foi feito o mapeamento da área delimitada na Figura 22-D, na Expedição 2 foram adquiridas as leituras espectrais de forma pontual com auxílio do sensor do tipo não imageador. Essas leituras foram realizadas em feições foto identificáveis e geolocalizadas a partir do mapeamento executado na primeira expedição. Com base na seleção das respectivas feições, foram extraídas as suas coordenadas para aquisição das assinaturas espectrais *in-loco*.

Para a aquisição das assinaturas espectrais nos locais selecionados foi utilizado o sensor

espectral não imagiador (SR-3500 series Spectroradiometer) (FIGURA23-A). Por meio desse tipo de sensor é possível realizar a leitura espectral dos materiais tanto em laboratório como em campo. Para as leituras realizadas em campo é necessário transportar o sensor em uma mochila adaptada para esse tipo de trabalho, além de possuir baterias específicas para o funcionamento do instrumento (FIGURA23-B). É válido salientar que para esse experimento foi utilizado uma lente de 8° de FOV e uma placa para conversão de radiância em reflectância, placa esta utilizada como referência da radiação solar pelo sensor.

Para identificação em campo das feições selecionadas foi utilizado o receptor *Global Navigation Satellite System-GNSS RTK Hiper Plus++* (FIGURA23-C). Por meio desse receptor é possível localizar de forma precisa específicos pontos a partir das suas coordenadas ((FIGURA23-D). Ao todo foram localizados 6 pontos, no qual foram coletadas as suas respectivas assinaturas espectrais. Para mais detalhes sobre os pontos coletados foram geradas monografias de campo((FIGURA23-E)) para cada feição coletada que se encontram no APÊNDICE A.

Figura 23: Aquisição das Assinaturas Espectrais



Fonte: Elaborado pelo autor.

De posse das assinaturas levantadas é possível comparar as informações coletadas por ambos os sensores, proporcionando assim, a construção de um novo dataset de HIs contemplado com assinaturas espectrais coletadas na mesma área imageada. A construção do dataset, nomeado como dataset Unisinos, seguiu o mesmo fluxo proposto na Etapa A. Nesse fluxo as HIs de LR foram sintetizadas pela melhor técnica encontrada, para posteriormente serem inseridas no treinamento do método de S-R ERCSR.

Para o treinamento e validação do método de S-R, foi empregado da técnica de CV do tipo k-fold com $k = 5$. A partir dos resultados gerados foram avaliados as métricas de comparação de imagens entre os resultados preditos e a verdade de campo.

A análise da consistência espectral das HIs de HR preditas foi feita com base nas assinaturas espectrais coletadas na Expedição 2. Para essa avaliação foram feitas análises quantitativa e qualitativas. Na análise quantitativa foi utilizado da métrica SAM, enquanto na qualitativa foi avaliado visualmente o comportamento da assinatura espectral predita e assinatura espectral coletada *in-loco*.

5 RESULTADOS E DISCUSSÕES

Este Capítulo são relatados os resultados pertinentes aos experimentos realizados a fim de identificar a função de reamostragem que melhor corrobore no aprimoramento dos métodos inteligentes de S-R. Para primeira parte do experimento, representando pela parte A do Fluxograma (Figura 17), foram avaliados o comportamento dos métodos de reamostragem em dois dataset diferentes, de modo que ao final, a identificação das melhores abordagens pudessem ser realizadas. Posteriormente a respectiva identificação, por meio da análise espectral entre as assinaturas geradas pelo método S-R aprimorado pelas melhores funções encontradas e as assinaturas coletadas diretamente em campo (parte B do Fluxograma) corroboram em avaliar a consistência espectral do resultado gerado.

5.1 Avaliação das técnicas de Reamostragem

Para averiguar o comportamento das técnicas de reamostragem, os experimentos foram realizados a partir da rede de SR ERCSR, com valor de *seed* igual a 2, k valendo 5 e 200 épocas de treinamento. Dessa forma, para cada técnica selecionada, 10 modelos foram gerados, de maneira que foi possível analisar valores como média e desvio padrão. Como já relatado anteriormente as técnicas de reamostragem utilizadas foram *Average*, *Gaussian*, *Bilinear Cubic*, *Cubic-Spline*, *Nearest*, *Lanczos*, *Mode* e *RMS*, de modo que cada abordagem foi responsável por gerar um conjunto de imagens de LR para o aprendizado.

5.1.1 Dataset de Chikusei

Dos resultados obtidos para o dataset de Chikusei, conforme mostra a Tabela 10 a técnica de reamostragem *Lanczos* se mostrou matematicamente superior as demais abordagens avaliadas. Contudo, dentre as métricas examinadas, somente a métrica ERGAS apresentou grandes variações nos resultados encontrados. Essa variação pode ser justificada pelo grande número de bandas espectrais que compõem o respectivo dataset, que pode ter influenciado no resultado obtido pela métrica.

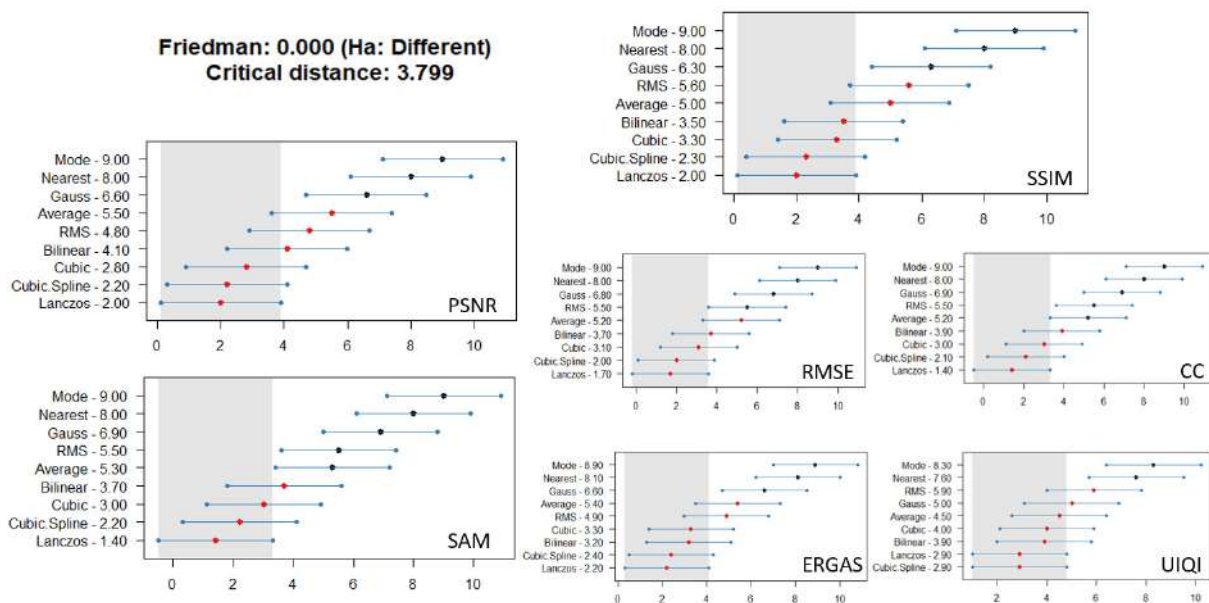
No entanto, ao averiguar os testes de hipótese não paramétrico de Friedman e Nemenyi (Fi-

Tabela 8: Comparação entre as técnicas de reamostragem para o dataset de Chikusei

Reamostragem	PSNR		SAM		SSIM		ERGAS		RMSE		CC		UIQI	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
<i>Average</i>	27.6937	0.7474	0.1156	0.0139	0.9954	0.0010	22.3149	2.6891	0.0019	0.0002	0.9778	0.0054	0.9771	0.0093
<i>Gaussian</i>	27.5348	0.7002	0.1171	0.0144	0.9953	0.0009	22.5548	2.6208	0.0020	0.0002	0.9769	0.0057	0.9785	0.0039
<i>Bilinear</i>	27.9123	0.7257	0.1129	0.0133	0.9955	0.0010	21.4652	2.2591	0.0019	0.0002	0.9786	0.0051	0.9761	0.0098
<i>Cubic</i>	28.0439	0.6666	0.1114	0.0136	0.9958	0.0007	21.3304	2.1054	0.0019	0.0002	0.9795	0.0053	0.9766	0.0084
<i>Cubic-Spline</i>	28.1439	0.7201	0.1115	0.0133	0.9959	0.0008	21.2257	2.4298	0.0018	0.0002	0.9795	0.0050	0.9803	0.0050
<i>Nearest</i>	26.0142	0.6269	0.1380	0.0157	0.9927	0.0012	27.0811	2.9400	0.0024	0.0002	0.9680	0.0073	0.9671	0.0054
<i>Lanczos</i>	28.1628	0.6385	0.1105	0.0134	0.9960	0.0006	21.1260	2.1338	0.0018	0.0002	0.9800	0.0051	0.9783	0.0076
<i>Mode</i>	25.5548	0.5361	0.1453	0.0153	0.9919	0.0013	28.7234	2.5850	0.0025	0.0002	0.9653	0.0078	0.9576	0.0286
<i>RMS</i>	27.7724	0.6691	0.1154	0.0140	0.9954	0.0009	22.1745	2.5493	0.0019	0.0002	0.9781	0.0053	0.9742	0.0093

gura 24), com um nível de confiança de 95%, o primeiro teste mostrou que existe diferença significativa entre as técnicas. Para avaliação dessa diferença, o segundo teste, com o mesmo nível de significância aplicado anteriormente, mostrou que as técnicas de reamostragem *Lanczos*, *Cubic-Spline*, *Cubic* e *Bilinear* são estatisticamente iguais, ao auxiliar no processo de S-R, visto que as diferenças das CDs de tais abordagens são inferiores a CD estimada (3.799). Contudo, ao comparar as funções selecionadas com as demais técnicas, fica comprovado pelo teste estatístico de Friedman e Nemenyi, que as funções de *Lanczos*, *Cubic-Spline*, *Cubic* e *Bilinear* apresentam maior influência no método de S-R em obter HIs de HR de melhor qualidade, ao analisar de forma conjunta todas as métricas de comparação.

Figura 24: Avaliação Estatística dos Métodos de Reamostragem para dataset de Chikusei



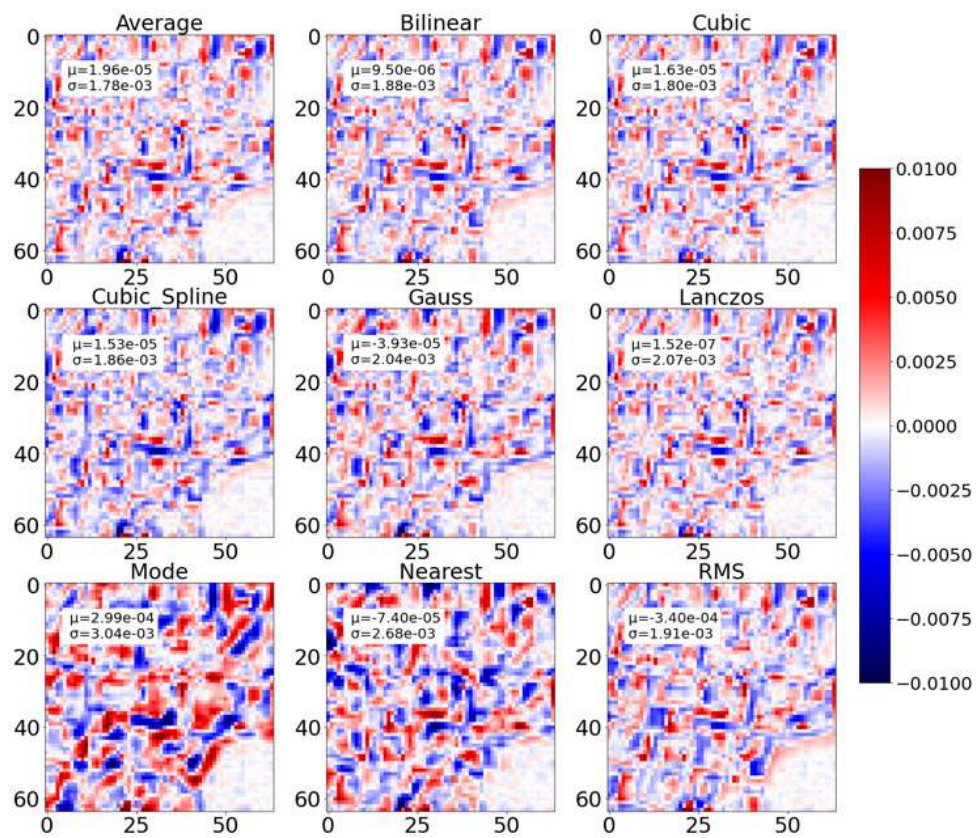
Fonte: Elaborado pelo autor.

Para uma comparação visual foi mensurado o Erro Médio entre as imagens de HR preditas com sua verdade de campo, conforme mostra a Figura 25. Por meio da respectiva figura gerada é possível observar que as técnicas de reamostragem *Mode*, *Nearest* e *RMS* apresentaram os piores resultados, uma vez que as maiores diferenças se encontram nessas técnicas. Ao analisar as demais técnicas de reamostragem, não foi possível observar grandes diferenças.

Contudo, com base nos resultados obtidos ficou comprovado que para o dataset de Chikusei, as técnicas de reamostragem do tipo *Lanczos*, *Cubic-Spline*, *Cubic* e *Bilinear* melhor corroboram em auxiliar que os métodos de SR auxiliem gerar HI de HR, uma vez que os valores obtidos para as métricas se mostram superiores as demais técnicas.

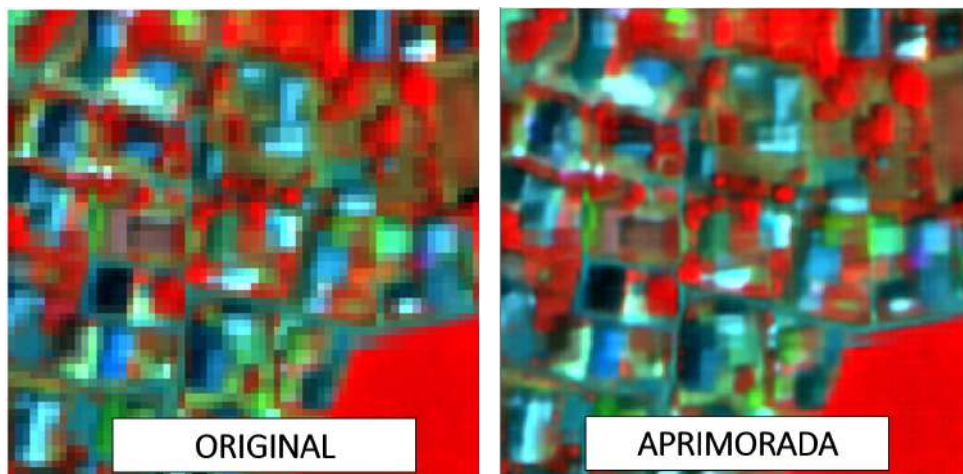
Na Figura 26 é ilustrado a potencialidade do método de SR empregado. É possível observar que a partir da imagem aprimorada é possível delimitar melhor os objetos do que a imagem original.

Figura 25: Comparação do erro médio entre as imagens de HR preditas e sua verdade de campo para o dataset de Chikusei



Fonte: Elaborado pelo autor.

Figura 26: Aplicação da Técnica de S-R para o dataset Chikusei



Fonte: Elaborado pelo autor.

5.1.2 Dataset de Harvard

Em relação aos resultados obtidos para o dataset de Harvard, conforme mostra a Tabela 10, mais uma vez a técnica de reamostragem *Lanczos* se mostrou matematicamente superior as

Tabela 9: Comparação entre as técnicas de reamostragem para o dataset de Chikusei

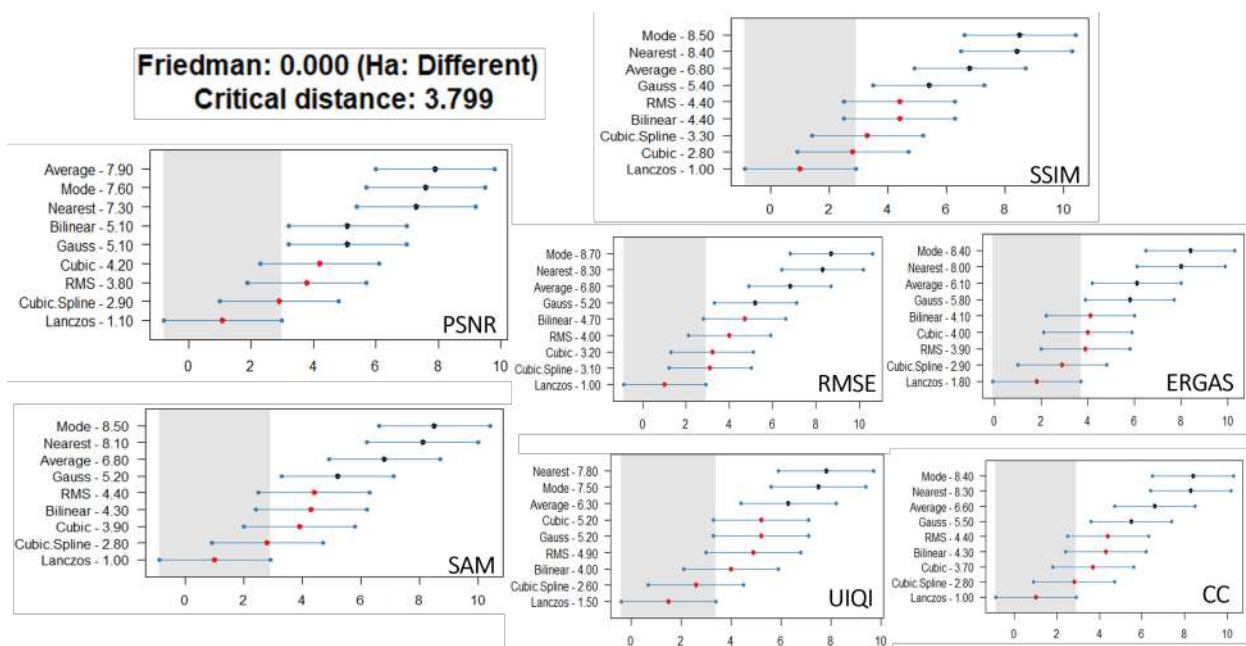
Reamostragem	PSNR		SAM		SSIM		ERGAS		RMSE		CC		UIQI	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
<i>Average</i>	29.8546	2.5733	0.0919	0.0247	0.9990	0.0007	19.0443	4.9942	0.0006	0.0002	0.9832	0.0097	0.9877	0.0055
<i>Gaussian</i>	31.1234	2.1183	0.0877	0.0243	0.9993	0.0005	19.9098	7.2276	0.0005	0.0002	0.9834	0.0103	0.9828	0.00166
<i>Bilinear</i>	31.1206	2.7757	0.0839	0.0237	0.9993	0.0006	17.2703	5.1708	0.0005	0.0002	0.9864	0.0085	0.9898	0.0064
<i>Cubic</i>	31.4452	1.7626	0.0831	0.0200	0.9994	0.0004	17.6516	4.8253	0.0005	0.0002	0.9865	0.0068	0.9869	0.0083
<i>Cubic-Spline</i>	29.8546	2.5733	0.0919	0.0247	0.9990	0.00007	19.0443	4.9942	0.0006	0.0002	0.9832	0.0097	0.9877	0.0055
<i>Nearest</i>	29.9449	2.0680	0.0992	0.0279	0.9988	0.0008	22.4700	7.8567	0.0006	0.0002	0.9783	0.0127	0.9788	0.0163
<i>Lanczos</i>	32.2227	2.2303	0.0791	0.0199	0.9986	0.0010	23.0441	9.8089	0.0006	0.0003	0.9777	0.0153	0.9804	0.0209
<i>Mode</i>	29.8660	2.5145	0.1016	0.0311	0.9986	0.0010	23.0441	9.8089	0.0006	0.0003	0.9777	0.0153	0.9804	0.0209
<i>RMS</i>	31.3749	2.5423	0.0855	0.0253	0.9993	0.0005	18.5142	6.9412	0.0005	0.0002	0.9845	0.0104	0.9840	0.0151

demais abordagens estudadas. Dos valores encontrados somente a métrica ERGAS apresentou resultados diferentes dos resultados esperados.

Em uma análise mais aprofundada das métricas avaliadas, os resultados obtidos se mostram superiores aos valores obtidos para o dataset de Chikusei. Essa variação pode ser justificada pelo grande número de bandas espectrais que compõem os datasets, uma vez que o dataset de Chikusei contém mais bandas espectrais, proporcionando assim, que o modelo de SR tenha mais dificuldade em prever HIs de HR.

Em relação os testes de hipótese não paramétrico de Friedman e Nemenyi (Figura ??), com um nível de confiança de 95% , o primeiro teste mostrou que existe diferença significativa entre as técnicas. Para avaliação dessa diferença, o segundo teste, com o mesmo nível de significância aplicado anteriormente, mostrou que as técnicas de reamostragem *Lanczos*, *Cubic-Spline*, *Cubic* e *RMS* são estatisticamente iguais, ao auxiliar no processo de S-R, visto que as diferenças das CDs de tais abordagens são inferiores a CD estimada (3.799).

Figura 27: Avaliação Estatística dos Métodos de Reamostragem para dataset de Harvard

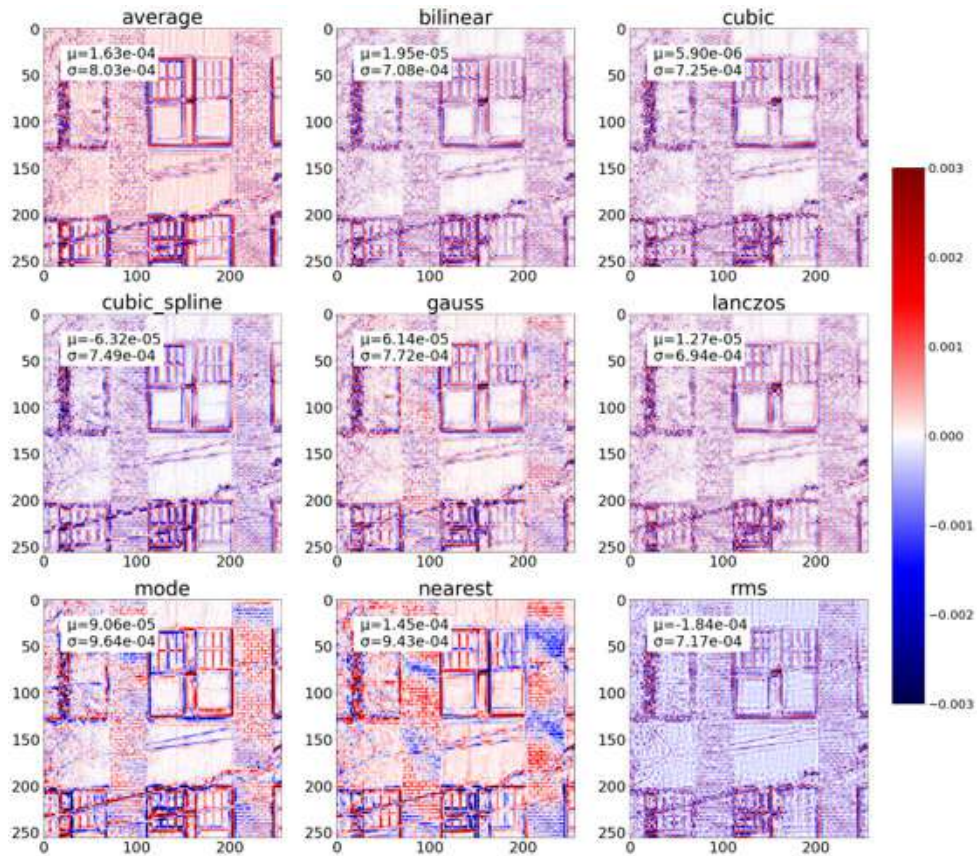


Fonte: Elaborado pelo autor.

Em uma comparação visual, o Erro Médio (Figura 28) mostra que as maiores diferenças

são observadas nas técnicas *Nearest*, *Mode* e *Average*. As demais técnicas não apresentaram grandes variações.

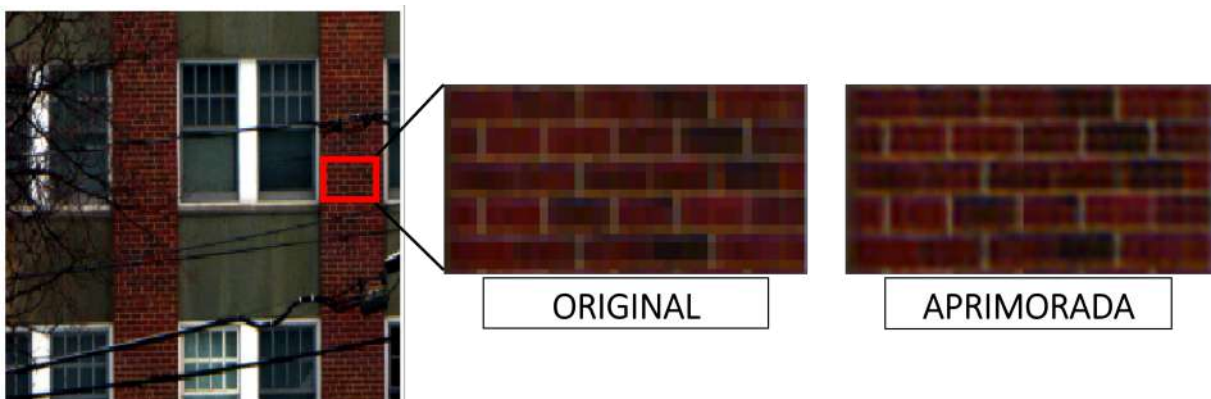
Figura 28: Comparação do erro médio entre as imagens de HR preditas e sua verdade de campo para o dataset de Harvard



Fonte: Elaborado pelo autor.

A fim de destacar a potencialidade do método de S-R ERCSR, na Figura 29 mostra a diferença entre a imagem original e a imagem aprimorada.

Figura 29: Aplicação da Técnica de S-R para o dataset Harvard



Fonte: Elaborado pelo autor.

De posse dos resultados obtidos, tanto para o dataset de Chikusei como para o dataset de Harvard, é possível notar que as técnicas de reamostragem *Lanczos*, *CubicSpline* e *Cubic* apresentaram os melhores resultados para ambos os datasets. Com base nessa semelhança percebe-se a existência de uma tendência, das respectivas técnicas, em proporcionar que os métodos de S-R gerem melhores resultados, em relação as demais abordagens avaliadas. Além disso é válido salientar que a reamostragem do tipo *Lanczos* se mostrou como ponto diferencial da pesquisa desenvolvida, uma vez que nenhum dos trabalhos relacionados avaliados (Tabela 4) utilizou de tal técnica ao longo dos seus experimentos.

5.2 Avaliação da Consistência Espectral das técnicas de Reamostragem Selecionadas

Finalizada a etapa A do fluxograma proposto (Figura 17) e de posse das melhores técnicas de reamostragem avaliadas, a segunda parte (Etapa B do fluxograma) do trabalho consiste em avaliar a consistência espectral dos resultados obtidos. Para essa avaliação foi criado de forma controlada o dataset Unisinos. Esse dataset é composto aproximadamente por 500 patches com dimensões de 64 x 64 pixels. Diferente da imagem original, composta de 300 bandas espectrais, foi construído cada *paches* com apenas 50 bandas espectrais contínuas, com intuito de otimizar a avaliação das técnicas. A faixa selecionada para essa avaliação corresponde a faixa espectral entre os intervalos de 1480 até 1735 nm.

Para averiguar o comportamento das técnicas de reamostragem identificadas, os experimentos foram realizados semelhantes aos passos executados para os datasets anteriores, que são : a rede de SR ERCSR, com valor de *seed* igual a 2, k valendo 5, 200 épocas de treinamento e por fim, com os testes de hipótese de Friedman e Nemenyi.

Contemplada a análise a estatística foi avaliado a consistência espectral para cada resultado gerado.

5.2.1 Dataset Unisinos

Dos resultados obtidos para o dataset Unisinos, conforme mostra a Tabela 10, a técnica de reamostragem *Lanczos* se mostrou matematicamente superior as demais abordagens estudadas. Dos valores encontrados, as métricas SAM e ERGAS apresentaram melhores valores em relação aos resultados encontrados para as mesmas métricas avaliadas para os datasets de Chikusei e Harvard. Essa variação pode ser justificada pelo fato que o dataset Unisinos não apresenta uma grande variabilidade de alvos imageados, interferindo assim diretamente no resultado. Contudo, ainda é necessário que mais análises sejam feitas. Em relação as demais métricas averiguadas, os resultados se mantiveram próximos aos valores obtidos anteriormente.

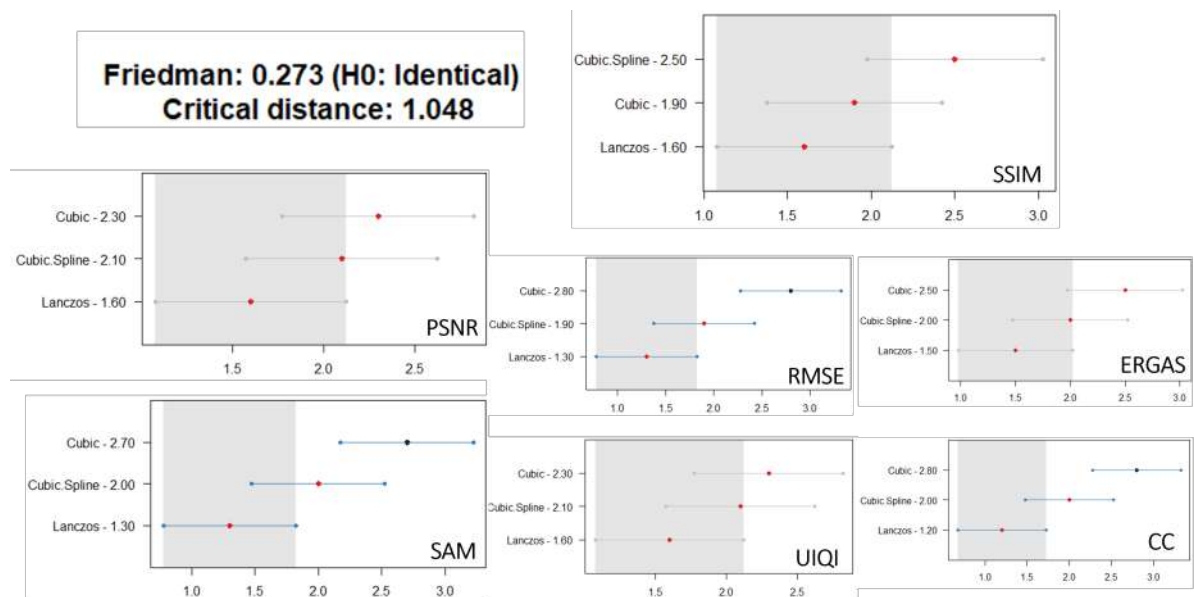
Os testes de hipótese não paramétrico de Friedman e Nemenyi (Figura 30), com um nível de confiança de 95% mostram que para o primeiro teste a maioria das técnicas avaliadas (PSNR, SSIM, UIQI e CC) não apresentaram diferenças significativas entre os dados avaliados,

Tabela 10: Comparação entre as técnicas de reamostragem para o dataset Unisinos

Reamostragem	PSNR		SAM		SSIM		ERGAS		RMSE		CC		UIQI	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
<i>Cubic</i>	30.9751	0.7863	0.0439	0.0071	0.9254	0.0134	6.7797	1.2732	1.6635	0.2840	0.9976	0.0013	0.9976	0.0012
<i>Cubic-Spline</i>	31.1176	0.7803	0.0431	0.0069	0.9234	0.0113	6.6401	1.2117	1.6368	0.2659	0.9977	0.0013	0.9977	0.00511
<i>Lanczos</i>	31.1265	0.8519	0.0428	0.0071	0.9264	0.0130	6.6227	1.2679	1.6248	0.2762	0.9978	0.0013	0.9977	0.0012

somente as métricas SAM, RMSE e CC obtiveram tal diferença, com uma CD estimada de 0.273. Embora respectiva diferença tenha ocorrido, mostrando que a técnica de reamostragem do tipo Cubic corrobore a gerar resultados inferiores as reamostragens do tipo Lanczos e Cubic-Spline, a maioria das métricas investigadas apresentaram diferenças insignificativas. Portanto, foi considerado que as técnicas de Cubic, Lanczos e Cubic-Spline apresentam a mesma influência no método de S-R em obter HIs de HR de melhor qualidade, uma vez que as três foram consideradas iguais para os datasets avaliados.

Figura 30: Avaliação Estatística dos Métodos de Reamostragem para dataset Unisinos



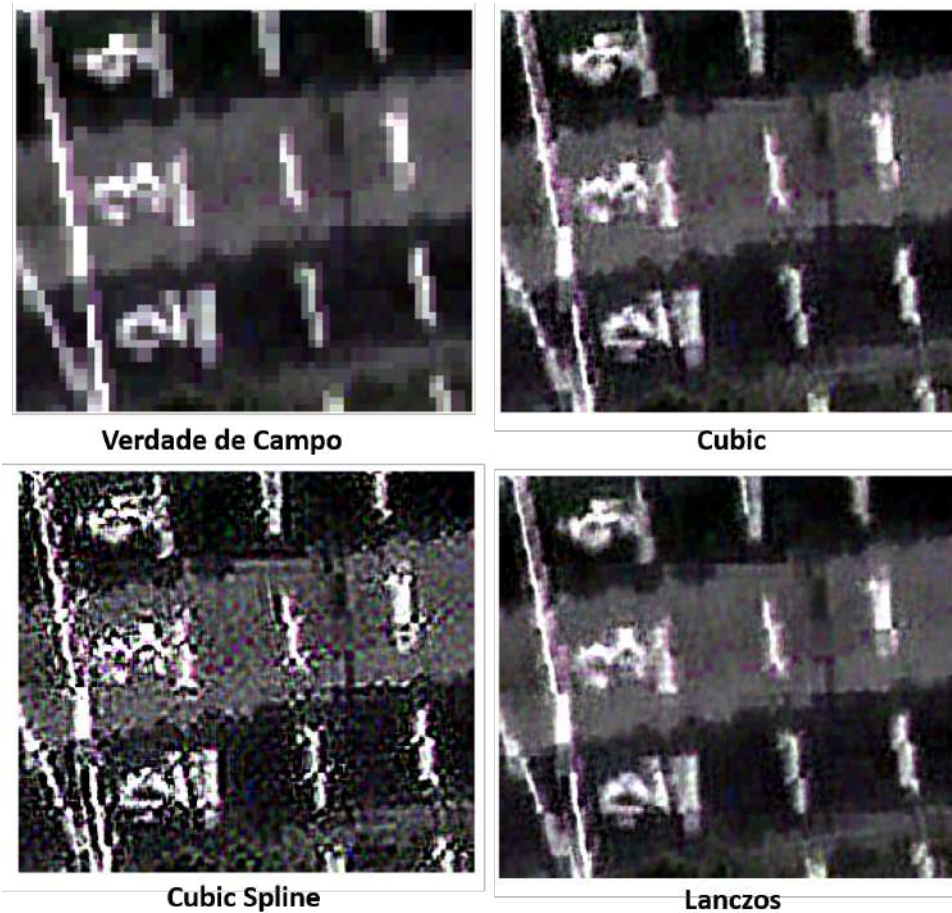
Fonte: Elaborado pelo autor.

Para uma comparação visual, a Figura 31 destaca os resultados obtidos na predição do método de S-R ERCSR para cada técnica selecionada. É possível observar que o respectivo método consegue aprimorar as bordas dos objetivos fornecendo maior facilidade em delimitar um objeto de outro. Contudo, entre as imagens obtidas, a imagem predita pela reamostragem Cubic-Spline apresentou variações espectrais que degradaram a qualidade final do produto, comprovando assim a importância em avaliar a resposta espectral, já que essas variações não foram observadas nas métricas avaliadas.

Em relação as imagens preditas para as técnicas de reamostragem do tipo Lanczos e Cubic, ambas as técnicas conseguiram gerar produtos de melhor qualidade, uma vez que não foram observadas anomalias ao longo das cenas preditas.

Para avaliação espectral dos resultados obtidos, foram coletados *in-loco* com espectrorra-

Figura 31: Comparação entre a imagem original e as imagens previstas



Fonte: Elaborado pelo autor.

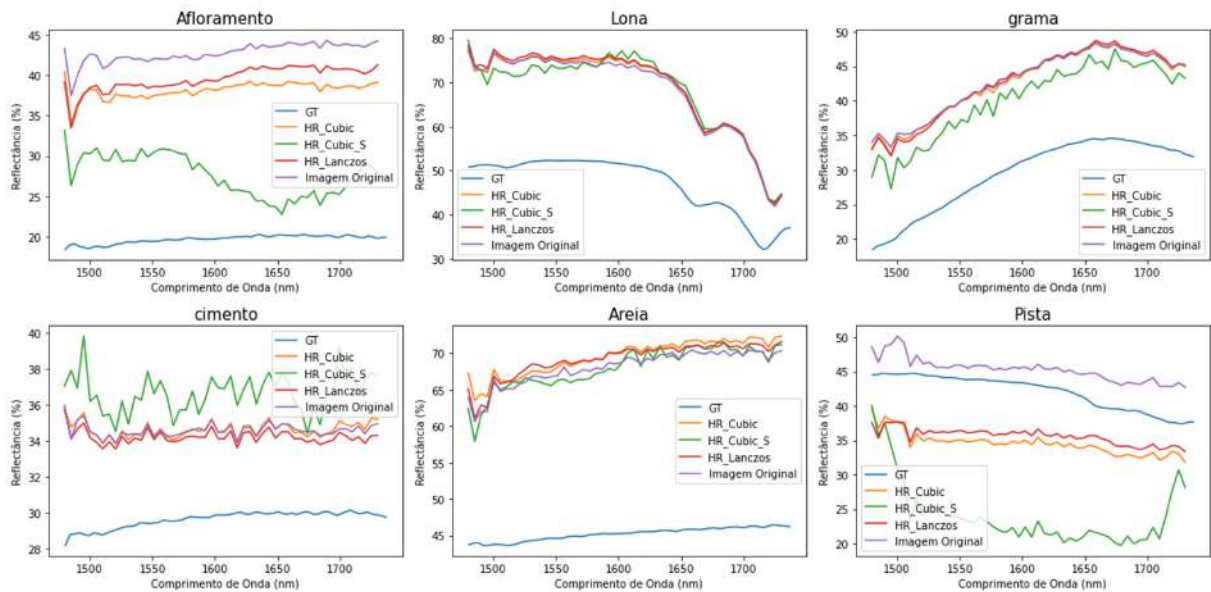
diômetro, seis pontos localizados no mesmo local onde foi adquirida a imagem para o dataset Unisinos. Os respectivos pontos coletados correspondem as seguintes feições: Afloramento, Lona, Grama, Cimento, Areia e Pista de Corrida. De posse de tais dados duas análises foram propostas, sendo a primeira uma análise quantitativa e a segunda qualitativa.

Em comparação as assinaturas espectrais coletadas, é possível observar pela Figura 32 que as assinaturas espectrais das imagens aprimoradas pelas técnicas de reamostragem do tipo Lanczos e Cubic (HR_{Cubic} e $HR_{Lanczos}$), apresentaram o mesmo padrão das assinaturas observadas nos alvos coletados (GT). Além disso é válido ressaltar que as pequenas variações nas assinaturas previstas podem ser oriundas das Imagem Original, uma vez que o comportamento das assinaturas analisadas são semelhantes.

Ainda dentro da análise qualitativa é válido destacar que o método de reamostragem Cubic-Spline apresentou resultados diferentes das métricas avaliadas na Etapa A do fluxograma proposto. Logo, é importante ressaltar a necessidade dos trabalhos de S-R em avaliar a consistência espectral, uma vez que respectivas variações podem ser observadas.

Para análise quantitativa dos resultados, foi utilizado da métrica SAM a fim de avaliar a diferença angular entre as assinaturas coletadas in-loco com as assinaturas previstas. Dos resultados

Figura 32: Comparação qualitativa entre as assinaturas espectrais coletadas in-loco com as assinaturas espectrais previstas



Fonte: Elaborado pelo autor.

obtidos, a técnica de reamostragem Cubic-Spline apresentou os piores resultados em relação as demais técnicas avaliadas, conforme mostra a Tabela 11.

Tabela 11: Comparação quantitativa entre as assinaturas coletadas em campo e as assinaturas previstas

Reamostragem	Afloramento	Lona	Grama	Cimento	Areia	Pista	μ	σ
<i>Cubic</i>	0,0206	0,0506	0,0403	0,0159	0,0182	0,0346	0,0300	0,0140
<i>Cubic-Spline</i>	0,1042	0,0658	0,0354	0,0310	0,0284	0,1717	0,0727	0,0565
<i>Lanczos</i>	0,0252	0,0483	0,0379	0,0162	0,0207	0,0381	0,0311	0,0123

6 CONSIDERAÇÕES FINAIS

As imagens hiperespectrais têm mostrado ao longo dos anos as suas potencialidades em diferentes aplicações. A sua alta variabilidade espectral tem fornecido detalhes mais precisos, de modo que diferentes matérias possam ser estudados e avaliados, utilizando apenas as suas assinaturas. Embora imagens com variações espectrais muito altas possibilitem a discriminação precisa de diferentes objetos, devido à limitação das tecnologias dos sensores, somente uma pequena fração da energia radiante pode ser coletada para cada banda do espectro. Logo, para atingir uma relação sinal-ruído aceitável, a área do sensor por pixel deve ser grande, levando a uma resolução espacial mais grosseira. A fim de contornar esse problema, técnicas pós-processamento foram criadas como é o caso da Super-Resolução a partir de métodos de *Deep Learning*. Contudo, em função da dificuldade em se obter um conjunto de HIs robusto, com imagens de LR e HR devidamente registradas, as abordagens de S-R propõe que as imagens de LR sejam sintetizadas, de modo que a partir de processos de degradação, reamostragem e a adição de ruídos é possível obter uma imagem LR sintética. Apesar dos trabalhos em sua totalidade, conforme observado ao longo da revisão bibliográfica, não foi observado um padrão de qual a melhor abordagem a ser utilizada. Nesse sentido, com intuito de compreender melhor a respectiva abordagem, esse trabalho teve como objetivo responder a seguinte hipótese: a identificação da melhor função de reamostragem de HIs de LR possibilita que os modelos de S-R em *Deep Learning* gerem HIs de HR de melhor qualidade.

Diante da hipótese apresentada, é avaliado as técnicas de reamostragem, comumente aplicadas na área do Sensoriamento Remoto, no aprimoramento dos métodos inteligentes de S-R. Em nossa primeira avaliação, as técnicas de reamostragem *Cubic*, *Cubic-Spline* e *Lanczos* apresentaram os melhores desempenhos perante as diferentes métricas de avaliação. Complementarmente, esses resultados foram comprovados pelos testes de hipótese de Friedman e Nemenyi, com um nível de significância de 95%, após uma extensa análise estatística, reforçando assim, a capacidade das técnicas de reamostragem em aprimorar os métodos de S-R.

Em relação a consistência espectral examinada, notou-se a importância de avaliar a resposta espectral predita e real. A criação do *dataset* proposto, associado com assinaturas espectrais coletadas *in-loco* possibilitou o refinamento do processo, fato esse, considerado um diferencial do trabalho proposto. Dos resultados obtidos, as reamostragem do tipo *Cubic* e *Lanczos* apresentaram as melhores consistências, com destaque para técnica de *Lanczos*, técnica essa não observada em nenhum dos trabalhos revisados.

Ao final, de posse das análises estatística e espectral encontradas, pode-se afirmar que a hipótese de pesquisa foi corroborada, uma vez que as técnicas de reamostragem do tipo *Cubic* e *Lanczos* apresentaram os melhores resultados ao aprimorarem os métodos de S-R.

6.1 Limitações do Trabalho

Durante o processo de desenvolvimento desse trabalho foram observadas as seguintes limitações.

- A dificuldade em encontrar datasets de imagens Hiperespectrais com alta variabilidade de feições mapeadas.
- Limitação de hardware, GPU e memória RAM, para o treinamento das redes, o que acarretou na necessidade em encontrar diferentes ambientes para a realização da pesquisa, como é o caso da versão PRO do Google Colab;
- Tempo de processamento para avaliação dos resultados. Em função da alta complexidade das HIs acarretou em um maior tempo de treinamento das redes, ocasionando assim, na necessidade em reduzir o escopo da pesquisa.

6.2 Trabalhos Futuros

Como trabalhos futuros são propostos os seguintes itens:

- Investigar o comportamento das funções de Degradação e Ruído, de modo a identificar qual a melhor ou as melhores funções que corroborem em aprimorar os métodos de S-R.
- Construir um novo dataset que contemple maior variabilidade de objetos mapeados, de maneira que tais objetos sejam conhecidas as suas assinaturas espectrais.
- Avaliar os métodos de Super-Resolução na identificação mineralógica de rochas.

REFERÊNCIAS

- AGGARWAL, C. C. et al. Neural networks and deep learning. **Springer**, [S.l.], v. 10, p. 978–3, 2018.
- BAUMGARDNER, M. F.; BIEHL, L. L.; LANDGREBE, D. A. **220 band aviris hyperspectral image data set**: june 12, 1992 indian pine test site 3. 2015.
- CHAKRABARTI, A.; ZICKLER, T. **Statistics of Real-World Hyperspectral Images**. 2011. 193–200 p.
- CHEN, Y. et al. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. **IEEE Transactions on Geoscience and Remote Sensing**, [S.l.], v. 54, n. 10, p. 6232–6251, 2016.
- CHEN, Y.; LI, Y.; ZHAO, Y. Sub-pixel detection algorithm based on cubic b-spline curve and multi-scale adaptive wavelet transform. **Optik**, [S.l.], v. 127, n. 1, p. 11–14, 2016.
- CRACKNELL, A. P. **Introduction to remote sensing**. [S.l.]: CRC press, 2007.
- DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. **The Journal of Machine Learning Research**, [S.l.], v. 7, p. 1–30, 2006.
- DONG, C. et al. Learning a deep convolutional network for image super-resolution. In: EUROPEAN CONFERENCE ON COMPUTER VISION, 2014. **Anais...** [S.l.: s.n.], 2014. p. 184–199.
- DONG, C. et al. Image super-resolution using deep convolutional networks. **IEEE transactions on pattern analysis and machine intelligence**, [S.l.], v. 38, n. 2, p. 295–307, 2015.
- DOU, X. et al. Super-resolution for hyperspectral remote sensing images based on the 3d attention-srgan network. **Remote Sensing**, [S.l.], v. 12, n. 7, p. 1204, 2020.
- DUMITRESCU, D.; BOIANGIU, C.-A. A study of image upsampling and downsampling filters. **Computers**, [S.l.], v. 8, n. 2, p. 30, 2019.
- FAUVEL, M. et al. Advances in spectral-spatial classification of hyperspectral images. **Proceedings of the IEEE**, [S.l.], v. 101, n. 3, p. 652–675, 2012.
- FENG, R. et al. Nonlocal total variation subpixel mapping for hyperspectral remote sensing imagery. **Remote Sensing**, [S.l.], v. 8, n. 3, p. 250, 2016.
- GAMBA, P. **Hyperspectral data over pavia city**. [S.l.: s.n.], 2015.
- GERCHBERG, R. Super-resolution through error energy reduction. **Optica Acta: International Journal of Optics**, [S.l.], v. 21, n. 9, p. 709–720, 1974.
- GHAMISI, P. et al. Advanced spectral classifiers for hyperspectral images: a review. **IEEE Geoscience and Remote Sensing Magazine**, [S.l.], v. 5, n. 1, p. 8–32, 2017.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

HANG, R.; LIU, Q.; LI, Z. Spectral super-resolution network guided by intrinsic properties of hyperspectral imagery. **IEEE Transactions on Image Processing**, [S.l.], v. 30, p. 7256–7265, 2021.

HASANLOU, M. **Remote sensing laboratory**. [S.l.: s.n.], 2019.

HAUT, J. M. et al. A new deep generative network for unsupervised remote sensing single-image super-resolution. **IEEE Transactions on Geoscience and Remote sensing**, [S.l.], v. 56, n. 11, p. 6792–6810, 2018.

HAYAT, K. Multimedia super-resolution via deep learning: a survey. **Digital Signal Processing**, [S.l.], v. 81, p. 198–217, 2018.

HE, K. et al. Deep residual learning for image recognition. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2016. **Proceedings...** [S.l.: s.n.], 2016. p. 770–778.

HE, Z.; LIU, L. Hyperspectral image super-resolution inspired by deep laplacian pyramid network. **Remote Sensing**, [S.l.], v. 10, n. 12, p. 1939, 2018.

HU, J. et al. A spatial constraint and deep learning based hyperspectral image super-resolution method. In: IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS), 2017., 2017. **Anais...** [S.l.: s.n.], 2017. p. 5129–5132.

HU, J. et al. Hyperspectral image superresolution via deep structure and texture interfusion. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [S.l.], v. 14, p. 8665–8678, 2021.

HU, J.; LI, Y.; XIE, W. Hyperspectral image super-resolution by spectral difference learning and spatial error correction. **IEEE Geoscience and Remote Sensing Letters**, [S.l.], v. 14, n. 10, p. 1825–1829, 2017.

JIANG, J. et al. Learning spatial-spectral prior for super-resolution of hyperspectral imagery. **IEEE Transactions on Computational Imaging**, [S.l.], v. 6, p. 1082–1096, 2020.

JIANG, R. et al. Learning spatial and spectral features via 2d-1d generative adversarial network for hyperspectral image super-resolution. In: IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP), 2019., 2019. **Anais...** [S.l.: s.n.], 2019. p. 2149–2153.

KELLEHER, J. D. **Deep learning**. [S.l.]: MIT press, 2019.

KEYS, R. Cubic convolution interpolation for digital image processing. **IEEE transactions on acoustics, speech, and signal processing**, [S.l.], v. 29, n. 6, p. 1153–1160, 1981.

KHORRAM, S. et al. **Remote sensing**. [S.l.]: Springer Science & Business Media, 2012.

KIM, J.; LEE, J. K.; LEE, K. M. Accurate image super-resolution using very deep convolutional networks. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2016. **Proceedings...** [S.l.: s.n.], 2016. p. 1646–1654.

KRAAK, M.-J. The cartographic visualization process: from presentation to exploration. **The Cartographic Journal**, [S.l.], v. 35, n. 1, p. 11–15, 1998.

- KUMAR, M. V.; YARRAKULA, K. Comparison of efficient techniques of hyper-spectral image preprocessing for mineralogy and vegetation studies. , [S.l.], 2017.
- LEDIG, C. et al. Photo-realistic single image super-resolution using a generative adversarial network. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2017. **Proceedings...** [S.l.: s.n.], 2017. p. 4681–4690.
- LI, J. et al. Hyperspectral image super-resolution with 1d–2d attentional convolutional neural network. **Remote Sensing**, [S.l.], v. 11, n. 23, p. 2859, 2019.
- LI, Q.; WANG, Q.; LI, X. Mixed 2d/3d convolutional network for hyperspectral image super-resolution. **Remote Sensing**, [S.l.], v. 12, n. 10, p. 1660, 2020.
- LI, Q.; WANG, Q.; LI, X. Exploring the relationship between 2d/3d convolution for hyperspectral image super-resolution. **IEEE Transactions on Geoscience and Remote Sensing**, [S.l.], 2021.
- LI, X.; ZHANG, L.; YOU, J. Domain transfer learning for hyperspectral image super-resolution. **Remote Sensing**, [S.l.], v. 11, n. 6, p. 694, 2019.
- LI, Y. et al. Hyperspectral image super-resolution using deep convolutional neural network. **Neurocomputing**, [S.l.], v. 266, p. 29–41, 2017.
- LI, Y.; ZHANG, H.; SHEN, Q. Spectral–spatial classification of hyperspectral imagery with 3d convolutional neural network. **Remote Sensing**, [S.l.], v. 9, n. 1, p. 67, 2017.
- LIANG, P.; SHI, W.; ZHANG, X. Remote sensing image classification based on stacked denoising autoencoder. **Remote Sensing**, [S.l.], v. 10, n. 1, p. 16, 2017.
- LIU, W.; LEE, J. An efficient residual learning neural network for hyperspectral image superresolution. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [S.l.], v. 12, n. 4, p. 1240–1253, 2019.
- M. GRAÑA MA. VEGANZONS, B. A. **Hyperspectral remote sensing scenes**. [S.l.: s.n.], 2021.
- MATHER, P. M.; KOCH, M. **Computer processing of remotely-sensed images: an introduction**. [S.l.]: John Wiley & Sons, 2011.
- MATHER, P.; TSO, B. **Classification methods for remotely sensed data**. [S.l.]: CRC press, 2016.
- MEI, S. et al. Hyperspectral image spatial super-resolution via 3d full convolutional neural network. **Remote Sensing**, [S.l.], v. 9, n. 11, p. 1139, 2017.
- MENESES, P. R.; ALMEIDA, T. d. Introdução ao processamento de imagens de sensoriamento remoto. **Universidade de Brasília, Brasília**, [S.l.], 2012.
- MOHER, D. et al. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. **Int J Surg**, [S.l.], v. 8, n. 5, p. 336–341, 2010.
- MORAES NOVO, E. M. de. **Sensoriamento remoto: princípios e aplicações**. [S.l.]: Editora Blucher, 2010.



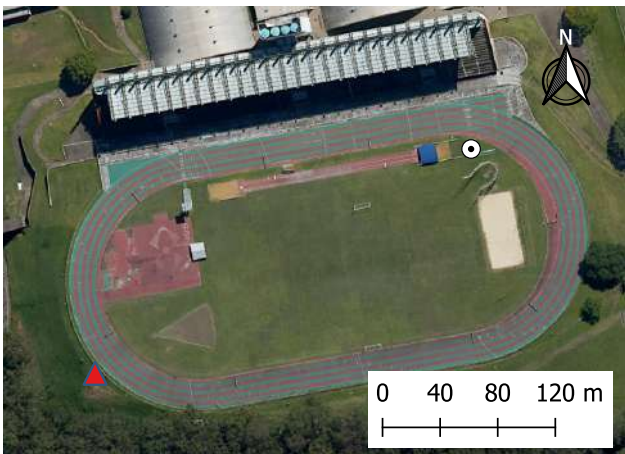

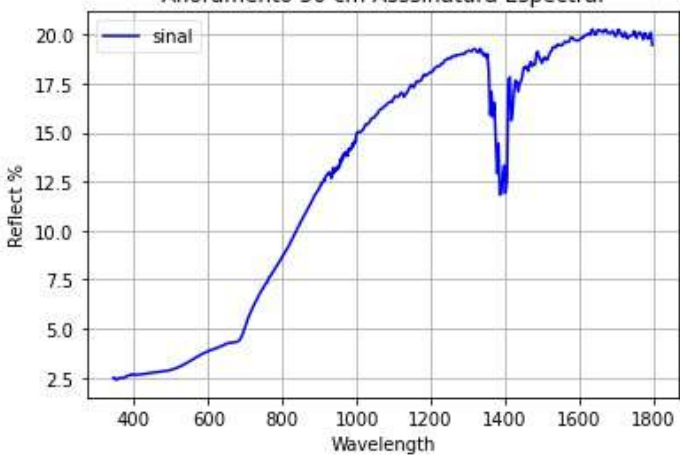
- NASCIMENTO, S. M.; AMANO, K.; FOSTER, D. H. Spatial distributions of local illumination color in natural scenes. **Vision research**, [S.l.], v. 120, p. 39–44, 2016.
- NASCIMENTO, S. M.; FERREIRA, F. P.; FOSTER, D. H. Statistics of spatial cone-excitation ratios in natural scenes. **JOSA A**, [S.l.], v. 19, n. 8, p. 1484–1490, 2002.
- NIKONOROV, A. et al. Deep learning-based enhancement of hyperspectral images using simulated ground truth. In: IAPR WORKSHOP ON PATTERN RECOGNITION IN REMOTE SENSING (PRRS), 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p. 1–9.
- PATIL, M. Interpolation techniques in image resampling. **Int. J. Eng. Technol**, [S.l.], v. 7, n. 3.34, p. 567–570, 2018.
- PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais: princípios, algoritmos e aplicações**. [S.l.]: Cengage Learning, 2008.
- PEDRYCZ, W.; CHEN, S.-M. **Deep learning: concepts and architectures**. [S.l.]: Springer, 2020.
- PEYGHAMBARI, S.; ZHANG, Y. Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: an updated review. **Journal of Applied Remote Sensing**, [S.l.], v. 15, n. 3, p. 031501, 2021.
- POHLERT, T. The pairwise multiple comparison of mean ranks package (pncmr). **R package**, [S.l.], v. 27, n. 2019, p. 9, 2014.
- QU, Y.; QI, H.; KWAN, C. Unsupervised sparse dirichlet-net for hyperspectral image super-resolution. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2018. **Proceedings...** [S.l.: s.n.], 2018. p. 2511–2520.
- RAN, Q. et al. Remote sensing images super-resolution with deep convolution networks. **Multimedia Tools and Applications**, [S.l.], v. 79, n. 13, p. 8985–9001, 2020.
- RASCHKA, S.; MIRJALILI, V. **Python machine learning: machine learning and deep learning with python, scikit-learn, and tensorflow 2**. [S.l.]: Packt Publishing Ltd, 2019.
- REFAEILZADEH, P.; TANG, L.; LIU, H. Cross-validation. **Encyclopedia of database systems**, [S.l.], v. 5, p. 532–538, 2009.
- ROHITH, G.; KUMAR, L. S. Paradigm shifts in super-resolution techniques for remote sensing applications. **The Visual Computer**, [S.l.], p. 1–44, 2020.
- ROY, D. P. et al. Landsat-8: science and product vision for terrestrial global change research. **Remote sensing of Environment**, [S.l.], v. 145, p. 154–172, 2014.
- SCAFUTTO, R. D. M.; SOUZA FILHO, C. R. de; OLIVEIRA, W. J. de. Hyperspectral remote sensing detection of petroleum hydrocarbons in mixtures with mineral substrates: implications for onshore exploration and monitoring. **ISPRS Journal of Photogrammetry and Remote Sensing**, [S.l.], v. 128, p. 146–157, 2017.
- SCHOWENGERDT, R. A. **Remote sensing: models and methods for image processing**. [S.l.]: Elsevier, 2006.

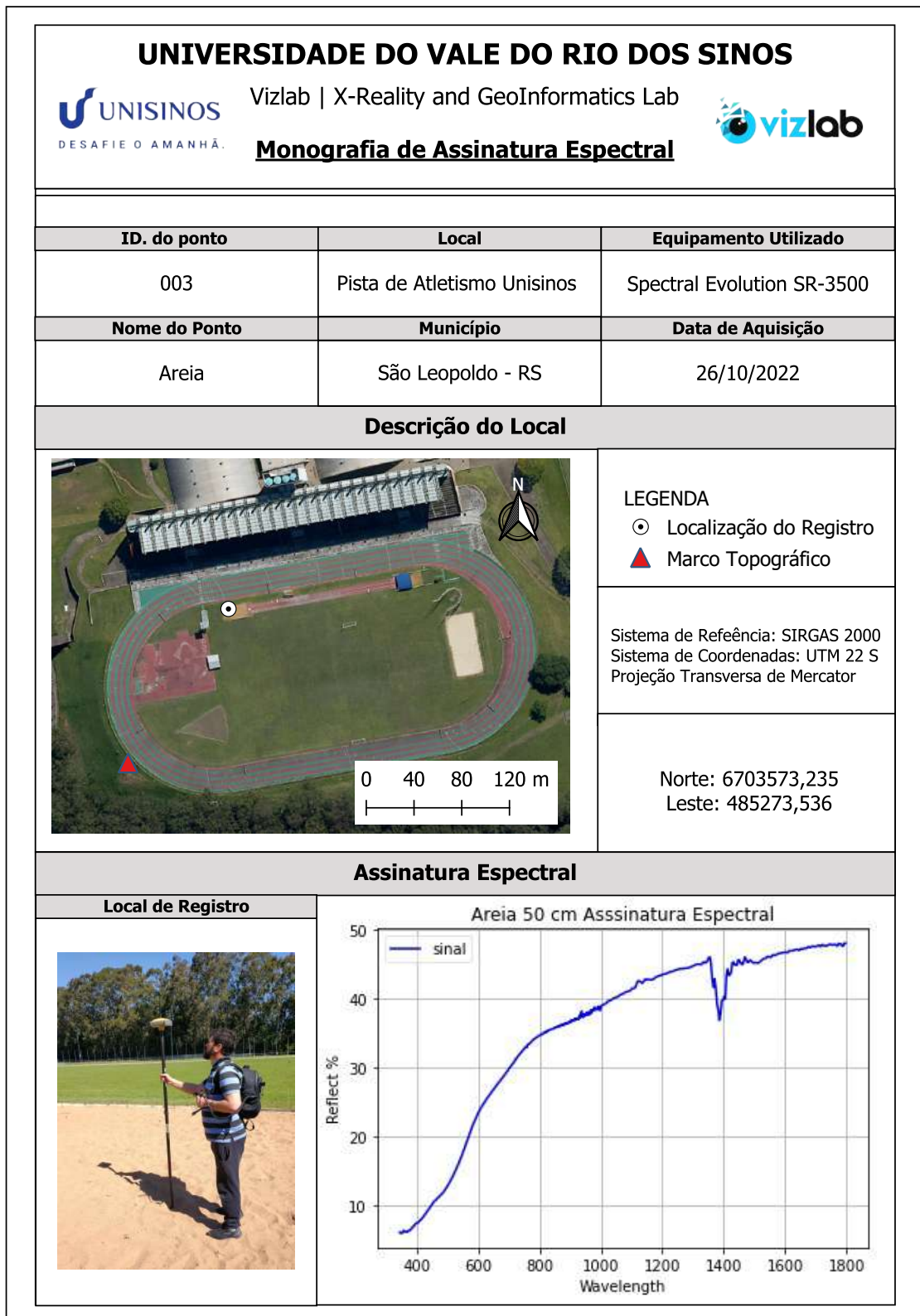
- SEWAK, M.; KARIM, M. R.; PUJARI, P. **Practical convolutional neural networks: implement advanced deep learning models using python**. [S.l.]: Packt Publishing Ltd, 2018.
- SHAO, Z. et al. Remote sensing image super-resolution using sparse representation and coupled sparse autoencoder. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [S.l.], v. 12, n. 8, p. 2663–2674, 2019.
- SHLIEN, S. Geometric correction, registration, and resampling of landsat imagery. **Canadian journal of remote Sensing**, [S.l.], v. 5, n. 1, p. 74–89, 1979.
- SIMOES, M. et al. A convex formulation for hyperspectral image superresolution via subspace-based regularization. **IEEE Transactions on Geoscience and Remote Sensing**, [S.l.], v. 53, n. 6, p. 3373–3388, 2014.
- SINGHAL, H. Convolutional neural network with tensorflow implementation. **Data Science Group, IITR**, [S.l.], 2017.
- SKANSI, S. **Introduction to deep learning: from logical calculus to artificial intelligence**. [S.l.]: Springer, 2018.
- STUART, M. B.; MCGONIGLE, A. J.; WILLMOTT, J. R. Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems. **Sensors**, [S.l.], v. 19, n. 14, p. 3071, 2019.
- TAI, Y.; YANG, J.; LIU, X. Image super-resolution via deep recursive residual network. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2017. **Proceedings...** [S.l.: s.n.], 2017. p. 3147–3155.
- TAYLOR, D. F. Geographic information systems: the microcomputer and modern cartography. In: **Modern cartography series**. [S.l.]: Elsevier, 1991. v. 1, p. 1–20.
- TEMPFLI, K. et al. **Principles of remote sensing: an introductory textbook**. [S.l.]: International Institute for Geo-Information Science and Earth Observation, 2009.
- TSAGKATAKIS, G. et al. Survey of deep-learning approaches for remote sensing observation enhancement. **Sensors**, [S.l.], v. 19, n. 18, p. 3929, 2019.
- WALD, L. Quality of high resolution synthesised images: is there a simple criterion? In: THIRD CONFERENCE "FUSION OF EARTH DATA: MERGING POINT MEASUREMENTS, RASTER MAPS AND REMOTELY SENSED IMAGES", 2000. **Anais...** [S.l.: s.n.], 2000. p. 99–103.
- WANG, B. et al. Hyperspectral imagery spatial super-resolution using generative adversarial network. **IEEE Transactions on Computational Imaging**, [S.l.], v. 7, p. 948–960, 2021.
- WANG, L.; PENG, J.; SUN, W. Spatial–spectral squeeze-and-excitation residual network for hyperspectral image classification. **Remote Sensing**, [S.l.], v. 11, n. 7, p. 884, 2019.
- WANG, Q.; SHI, W. Utilizing multiple subpixel shifted images in subpixel mapping with image interpolation. **IEEE Geoscience and Remote Sensing Letters**, [S.l.], v. 11, n. 4, p. 798–802, 2013.

- WANG, Y. et al. A new attention-based cnn approach for crop mapping using time series sentinel-2 images. **Computers and Electronics in Agriculture**, [S.l.], v. 184, p. 106090, 2021.
- WANG, Z.; BOVIK, A. C. A universal image quality index. **IEEE signal processing letters**, [S.l.], v. 9, n. 3, p. 81–84, 2002.
- WEN, J. et al. Systematic literature review of machine learning based software development effort estimation models. **Information and Software Technology**, [S.l.], v. 54, n. 1, p. 41–59, 2012.
- WOO, S. et al. Cbam: convolutional block attention module. In: EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV), 2018. **Proceedings...** [S.l.: s.n.], 2018. p. 3–19.
- YANG, J. et al. Hybrid local and nonlocal 3-d attentive cnn for hyperspectral image super-resolution. **IEEE Geoscience and Remote Sensing Letters**, [S.l.], 2020.
- YASUMA, F. et al. **Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum**. [S.l.: s.n.], 2008.
- YOKOYA, N.; IWASAKI, A. **Airborne hyperspectral data over chikusei**. Japan: Space Application Laboratory, University of Tokyo, 2016. (SAL-2016-05-27).
- YUAN, Y.; ZHENG, X.; LU, X. Hyperspectral image superresolution by transfer learning. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [S.l.], v. 10, n. 5, p. 1963–1974, 2017.
- ZAFAR, I. et al. **Hands-on convolutional neural networks with tensorflow: solve computer vision problems with modeling in tensorflow and python**. [S.l.]: Packt Publishing Ltd, 2018.
- ZHANG, L. et al. Unsupervised adaptation learning for hyperspectral imagery super-resolution. In: IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2020. **Proceedings...** [S.l.: s.n.], 2020. p. 3073–3082.
- ZHANG, S. et al. Spectral recovery-guided hyperspectral super-resolution using transfer learning: srl for hsr. **IET Image Processing**, [S.l.], v. 15, n. 11, p. 2656–2665, 2021.
- ZHAO, M. et al. Hyperspectral image super-resolution under the guidance of deep gradient information. **Remote Sensing**, [S.l.], v. 13, n. 12, p. 2382, 2021.
- ZHENG, K. et al. Multi-losses function based convolution neural network for single hyperspectral image super-resolution. In: FIFTH INTERNATIONAL WORKSHOP ON EARTH OBSERVATION AND REMOTE SENSING APPLICATIONS (EORSA), 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p. 1–4.
- ZHENG, K. et al. Separable-spectral convolution and inception network for hyperspectral image super-resolution. **International Journal of Machine Learning and Cybernetics**, [S.l.], v. 10, n. 10, p. 2593–2607, 2019.
- ZHU, H. et al. Attention mechanisms in cnn-based single image super-resolution: a brief review and a new perspective. **Electronics**, [S.l.], v. 10, n. 10, p. 1187, 2021.

ZOU, C.; HUANG, X. Hyperspectral image super-resolution combining with deep learning and spectral unmixing. **Signal Processing: Image Communication**, [S.l.], v. 84, p. 115833, 2020.

APÊNDICE A – MONOGRAFIA DE ASSINATURA ESPECTRAL

UNIVERSIDADE DO VALE DO RIO DOS SINOS		
		Vizlab X-Reality and GeoInformatics Lab 
Monografia de Assinatura Espectral		
<hr/>		
ID. do ponto	Local	Equipamento Utilizado
004	Pista de Atletismo Unisinos	Spectral Evolution SR-3500
Nome do Ponto	Município	Data de Aquisição
Afloramento	São Leopoldo - RS	26/10/2022
Descrição do Local		
		LEGENDA ○ Afloramento ▲ Marco Topográfico
		Sistema de Referência: SIRGAS 2000 Sistema de Coordenadas: UTM 22 S Projeção Transversa de Mercator
		Norte: 6703586,566 Leste: 485362,018
Assinatura Espectral		
Local de Registro	Afloramento 50 cm Assinatura Espectral	
		



UNIVERSIDADE DO VALE DO RIO DOS SINOS



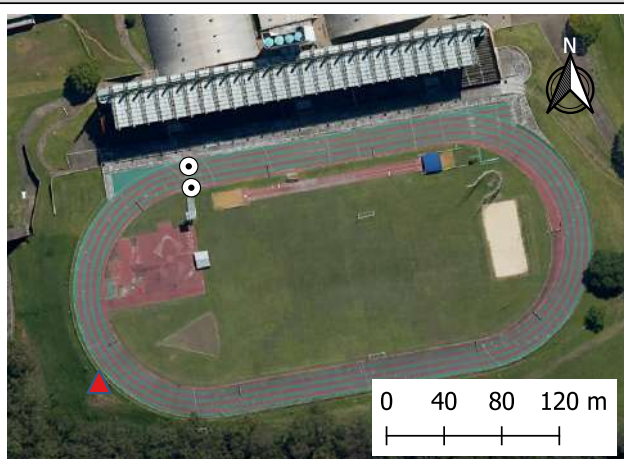
Vizlab | X-Reality and GeoInformatics Lab



Monografia de Assinatura Espectral

ID. do ponto	Local	Equipamento Utilizado
006	Pista de Atletismo Unisinos	Spectral Evolution SR-3500
Nome do Ponto	Município	Data de Aquisição
Cimento	São Leopoldo - RS	26/10/2022

Descrição do Local



LEGENDA

- ▲ Marco Topográfico
- ⊙ Pista de Atletismo

Sistema de Referência: SIRGAS 2000
Sistema de Coordenadas: UTM 22 S
Projeção Transversa de Mercator

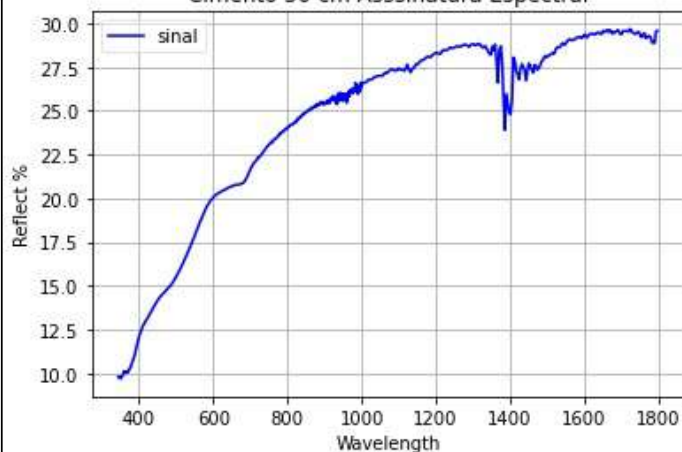
Norte: 6703594,114
Leste: 485296,179



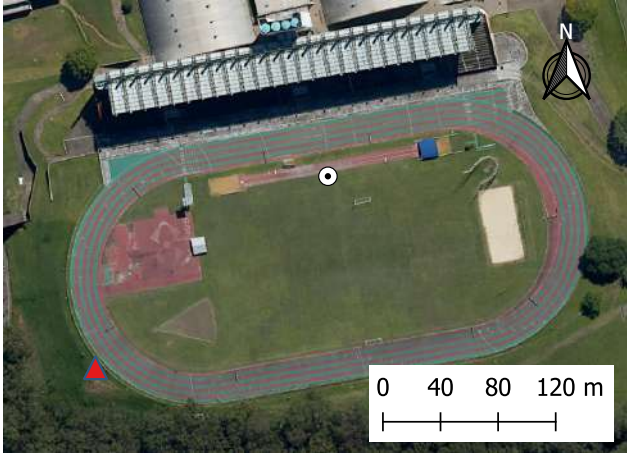



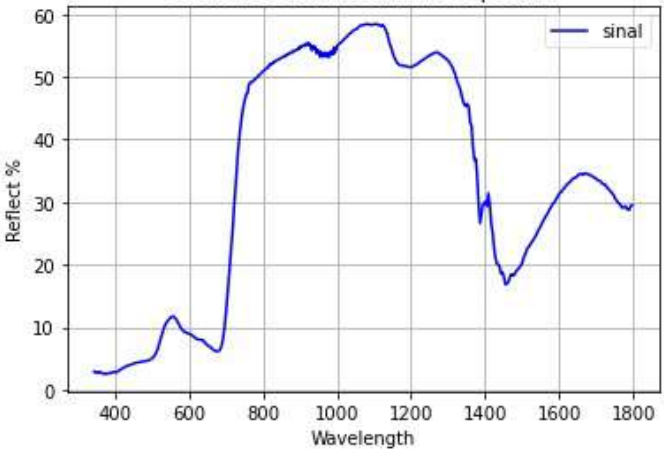
Assinatura Espectral

Local de Registro



Cimento 50 cm Assinatura Espectral



UNIVERSIDADE DO VALE DO RIO DOS SINOS		
		Vizlab X-Reality and GeoInformatics Lab 
Monografia de Assinatura Espectral		
<hr/>		
ID. do ponto	Local	Equipamento Utilizado
002	Pista de Atletismo Unisinos	Spectral Evolution SR-3500
Nome do Ponto	Município	Data de Aquisição
Grama	São Leopoldo - RS	26/10/2022
Descrição do Local		
		LEGENDA  Marco Topográfico  grama
		Sistema de Referência: SIRGAS 2000 Sistema de Coordenadas: UTM 22 S Projeção Transversa de Mercator
		Norte: 6703575,155 Leste: 485312,167
Assinatura Espectral		
Local de Registro	Grama 50 cm Assinatura Espectral	
		

UNIVERSIDADE DO VALE DO RIO DOS SINOS



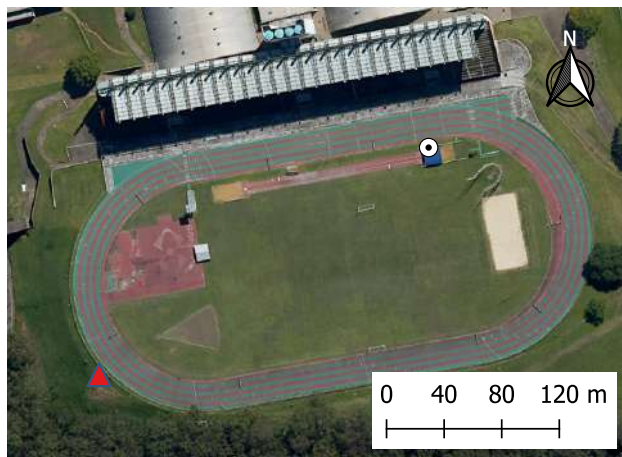
Vizlab | X-Reality and GeoInformatics Lab



Monografia de Assinatura Espectral

ID. do ponto	Local	Equipamento Utilizado
001	Pista de Atletismo Unisinos	Spectral Evolution SR-3500
Nome do Ponto	Município	Data de Aquisição
Lona	São Leopoldo - RS	26/10/2022

Descrição do Local



LEGENDA

- ▲ Marco Topográfico
- ⊙ Localização do registro

Sistema de Referência: SIRGAS 2000
Sistema de Coordenadas: UTM 22 S
Projeção Transversa de Mercator

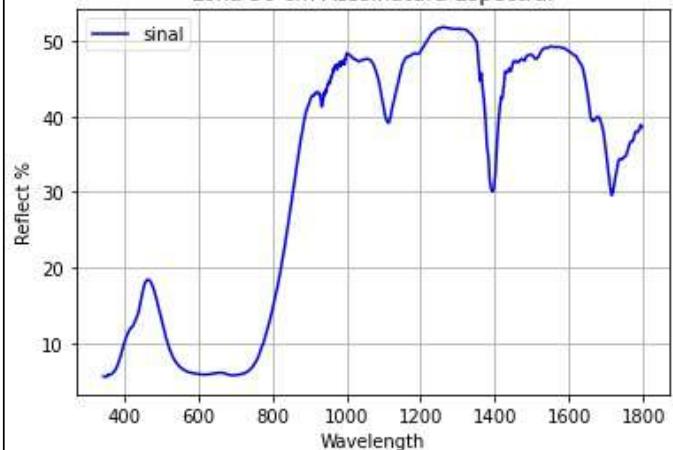
Norte: 6703587,551
Leste: 485345,883

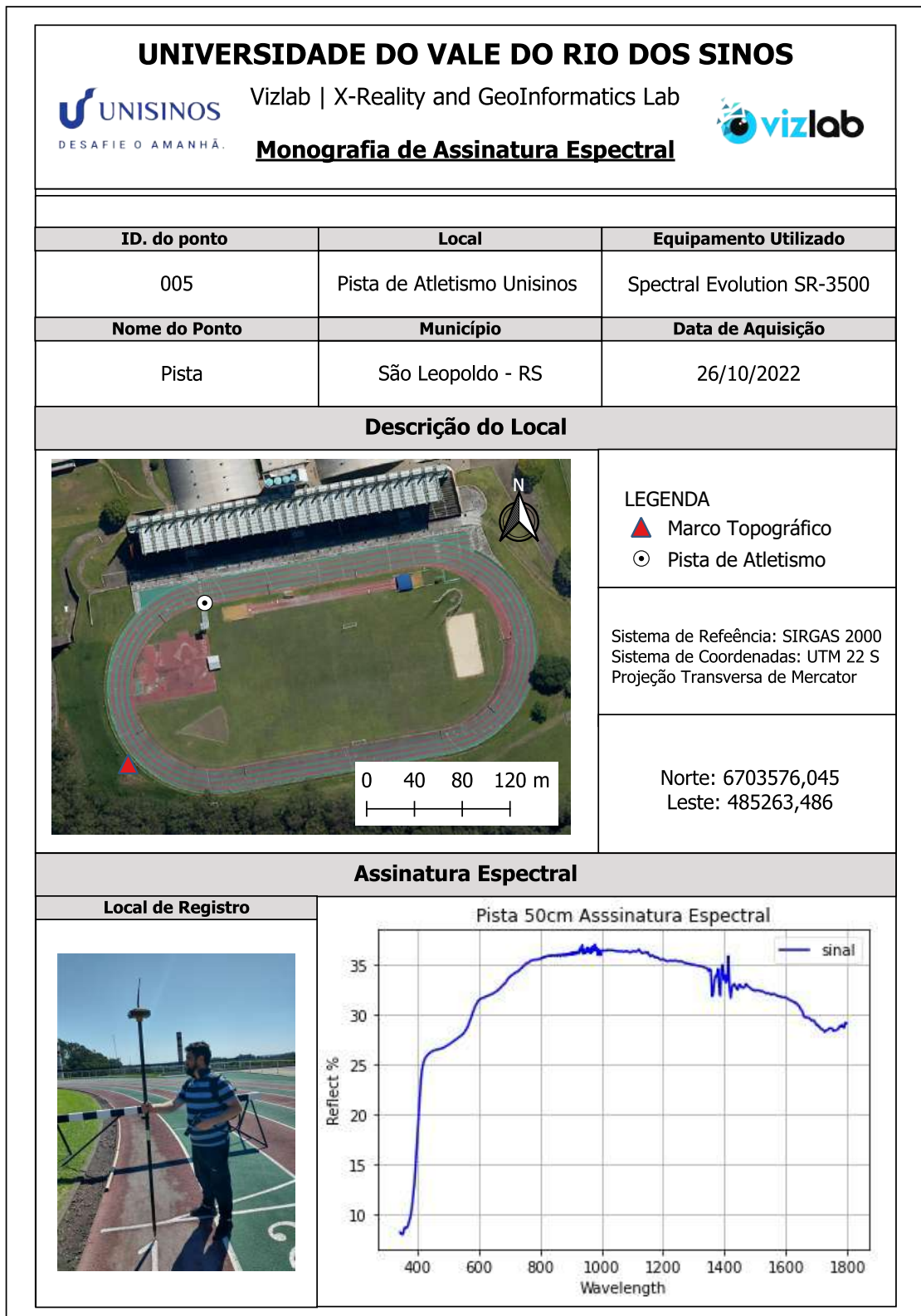
Assinatura Espectral

Local de Registro



Lona 50 cm Assinatura Espectral





APÊNDICE B – RESULTADOS OBTIDOS NA VALIDAÇÃO CRUZADA

Dataset Chikusei													
Resampling	Métricas	SEED 1					SEED 2					Média	Desvio Padrão
		Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5		
AVERAGE	PSNR	27,4319	26,9400	27,5023	28,1450	28,6585	27,2781	26,5079	27,4577	28,1466	28,8686	27,6937	0,7474
	SAM	0,1221	0,1336	0,1183	0,1030	0,0999	0,1233	0,1363	0,1177	0,1043	0,0970	0,1156	0,0139
	SSIM	0,9952	0,9945	0,9955	0,9962	0,9964	0,9949	0,9935	0,9956	0,9961	0,9965	0,9954	0,0010
	ERGAS	23,3881	25,2304	22,4341	20,0127	19,4208	24,0302	27,0876	22,3658	20,3503	18,8285	22,3149	2,6891
	RMSE	0,0020	0,0021	0,0020	0,0018	0,0017	0,0020	0,0022	0,0019	0,0018	0,0017	0,0019	0,0002
	CC	0,9740	0,9707	0,9774	0,9833	0,9832	0,9740	0,9704	0,9776	0,9832	0,9837	0,9778	0,0054
	UIQI	0,9787	0,9763	0,9804	0,9834	0,9836	0,9740	0,9520	0,9805	0,9796	0,9829	0,9771	0,0093
GAUSS	PSNR	27,0908	26,6511	27,2520	28,0451	28,3528	27,0488	26,6835	27,2994	27,9950	28,4857	27,5348	0,7002
	SAM	0,1254	0,1367	0,1213	0,1050	0,1014	0,1252	0,1370	0,1209	0,1057	0,1009	0,1171	0,0144
	SSIM	0,9948	0,9940	0,9952	0,9961	0,9962	0,9947	0,9941	0,9951	0,9960	0,9962	0,9953	0,0009
	ERGAS	24,4014	26,3448	22,9789	20,4716	19,7548	24,5414	25,9893	22,7378	20,4543	19,7200	22,5548	2,6208
	RMSE	0,0021	0,0022	0,0020	0,0018	0,0018	0,0021	0,0022	0,0020	0,0018	0,0017	0,0020	0,0002
	CC	0,9726	0,9694	0,9759	0,9825	0,9823	0,9728	0,9690	0,9757	0,9823	0,9824	0,9769	0,0057
	UIQI	0,9773	0,9728	0,9788	0,9831	0,9809	0,9754	0,9751	0,9755	0,9828	0,9823	0,9785	0,0039
Bilinear	PSNR	27,6852	27,2302	27,6663	27,5573	29,0738	27,5387	27,0827	27,6953	28,4451	29,1486	27,9123	0,7257
	SAM	0,1190	0,1304	0,1154	0,1060	0,0948	0,1199	0,1318	0,1152	0,1010	0,0951	0,1129	0,0133
	SSIM	0,9955	0,9949	0,9959	0,9941	0,9968	0,9953	0,9941	0,9957	0,9965	0,9967	0,9955	0,0010
	ERGAS	22,6948	24,6130	21,9188	20,1690	18,4744	22,5403	24,1973	22,3347	19,4376	18,2724	21,4652	2,2591
	RMSE	0,0019	0,0021	0,0019	0,0019	0,0016	0,0020	0,0021	0,0019	0,0017	0,0016	0,0019	0,0002
	CC	0,9755	0,9724	0,9787	0,9810	0,9848	0,9747	0,9706	0,9790	0,9841	0,9845	0,9786	0,0051
	UIQI	0,9800	0,9773	0,9814	0,9532	0,9851	0,9756	0,9657	0,9758	0,9847	0,9818	0,0098	
Cubic	PSNR	27,7504	27,2350	27,7758	27,9899	29,0935	27,9179	27,2251	27,7541	28,4919	28,9119	28,0439	0,6666
	SAM	0,1187	0,1298	0,1145	0,1019	0,0957	0,1187	0,1306	0,1148	0,1010	0,0954	0,1114	0,0136
	SSIM	0,9954	0,9949	0,9960	0,9956	0,9968	0,9956	0,9949	0,9957	0,9965	0,9964	0,9958	0,0007
	ERGAS	23,0122	24,1925	21,7052	20,7210	18,6523	22,4897	24,4029	21,3451	19,2494	19,2159	21,3304	2,1054
	RMSE	0,0019	0,0021	0,0019	0,0018	0,0016	0,0019	0,0021	0,0019	0,0017	0,0016	0,0019	0,0002
	CC	0,9759	0,9724	0,9791	0,9842	0,9848	0,9758	0,9721	0,9783	0,9840	0,9850	0,9795	0,0053
	UIQI	0,9741	0,9773	0,9817	0,9574	0,9848	0,9799	0,9766	0,9774	0,9836	0,9704	0,9766	0,0084
Cubic Spline	PSNR	27,9352	27,1298	27,9752	28,5491	29,2409	27,8468	27,1875	27,9390	28,4673	29,1684	28,1439	0,7201
	SAM	0,1178	0,1296	0,1141	0,1000	0,0952	0,1182	0,1309	0,1138	0,1012	0,0942	0,1115	0,0133
	SSIM	0,9956	0,9947	0,9960	0,9966	0,9968	0,9956	0,9948	0,9960	0,9965	0,9969	0,9959	0,0008
	ERGAS	22,1580	25,0238	21,4049	19,0819	18,2719	22,7352	24,3449	21,5824	19,4755	18,1781	21,2257	2,4298
	RMSE	0,0019	0,0021	0,0019	0,0017	0,0016	0,0019	0,0021	0,0019	0,0017	0,0016	0,0018	0,0002
	CC	0,9760	0,9733	0,9793	0,9844	0,9848	0,9764	0,9721	0,9795	0,9843	0,9850	0,9795	0,0050
	UIQI	0,9787	0,9689	0,9819	0,9850	0,9836	0,9786	0,9757	0,9813	0,9841	0,9851	0,9803	0,0050
Nearest	PSNR	25,5828	25,4385	25,7927	26,3470	26,8920	25,3655	25,2948	25,9493	26,52352	26,9562	26,0142	0,6269
	SAM	0,1467	0,1579	0,1418	0,1243	0,1189	0,1495	0,1593	0,1395	0,12454	0,1176	0,1380	0,0157
	SSIM	0,9916	0,9912	0,9927	0,9937	0,9944	0,9917	0,9909	0,9930	0,993901	0,9937	0,9927	0,0012
	ERGAS	28,2084	30,6011	27,7959	25,1462	23,6995	29,5860	31,3669	27,0414	24,27288	23,0924	27,0811	2,9400
	RMSE	0,0025	0,0026	0,0024	0,0022	0,0021	0,0025	0,0026	0,0023	0,00221	0,0021	0,0024	0,0002
	CC	0,9619	0,9588	0,9677	0,9757	0,9760	0,9620	0,9586	0,9682	0,975217	0,9755	0,9680	0,0073
	UIQI	0,9581	0,9661	0,9694	0,9663	0,9763	0,9663	0,9625	0,9718	0,972219	0,9622	0,9671	0,0054
Lanczos	PSNR	27,9871	27,3234	27,8834	28,5501	29,2050	28,0582	27,3018	27,9574	28,3956	28,7906	28,1628	0,6385
	SAM	0,1174	0,1289	0,1136	0,1000	0,0950	0,1177	0,1295	0,1133	0,1010	0,0952	0,1105	0,0134
	SSIM	0,9957	0,9951	0,9960	0,9964	0,9969	0,9957	0,9950	0,9961	0,9964	0,9961	0,9960	0,0006
	ERGAS	22,3100	24,1802	21,4155	18,9421	18,4396	22,3236	24,1237	21,4484	19,7530	19,5080	21,1260	2,1338
	RMSE	0,0019	0,0021	0,0019	0,0017	0,0016	0,0019	0,0021	0,0019	0,0017	0,0017	0,0018	0,0002
	CC	0,9763	0,9731	0,9794	0,9841	0,9851	0,9763	0,9726	0,9796	0,9844	0,9851	0,9800	0,0051
	UIQI	0,9802	0,9777	0,9823	0,9818	0,9851	0,9804	0,9766	0,9820	0,9794	0,9592	0,9783	0,0076
Mode	PSNR	25,2724	25,0287	25,4721	26,0143	26,7097	25,1972	24,8800	25,4752	25,8290	25,6699	25,5548	0,5361
	SAM	0,1521	0,1673	0,1479	0,1316	0,1235	0,1530	0,1668	0,1477	0,1327	0,1301	0,1453	0,0153
	SSIM	0,9915	0,9904	0,9922	0,9933	0,9940	0,9911	0,9904	0,9924	0,9930	0,9904	0,9919	0,0013
	ERGAS	30,0776	31,9995	28,9140	25,8591	24,4250	29,4804	32,6755	28,6435	26,5694	28,5903	28,7234	2,5850
	RMSE	0,0026	0,0028	0,0025	0,0023	0,0022	0,0026	0,0027	0,0025	0,0023	0,0023	0,0025	0,0002
	CC	0,9600	0,9545	0,9655	0,9732	0,9746	0,9588	0,9550	0,9654	0,9731	0,9725	0,9653	0,0078
	UIQI	0,9663	0,9609	0,9669	0,9724	0,9745	0,9587	0,9619	0,9684	0,9689	0,8775	0,9576	0,0286
RMS	PSNR	27,5895	27,1438	27,4076	28,2497	28,3358	27,5459	26,7097	27,6979	28,2055	28,8391	27,7724	0,6691
	SAM	0,1230	0,1335	0,1196	0,1036	0,1009	0,1230	0,1358	0,1187	0,1042	0,0988	0,1154	0,0140
	SSIM	0,9950	0,9944	0,9953	0,9962	0,9959	0,9949	0,9936	0,9955	0,9961	0,9964	0,9954	0,0009
	ERGAS	23,3462	25,0202	22,7480	19,8732	20,3629	23,7410	26,4591	22,2292	20,1129	19,0243	22,1745	2,5493
	RMSE	0,0020	0,0021	0,0020	0,0018	0,0017	0,0020	0,0022	0,0020	0,0018	0,0017	0,0019	0,0002
	CC	0,9738	0,9709	0,9773	0,9831	0,9832	0,9742	0,9707	0,9773	0,9831	0,9833	0,9781	0,0053
	UIQI	0,9777	0,9756	0,9765	0,9830	0,9631	0,9736	0,9552	0,9795	0,9805	0,9811	0,9742	0,0093

Dataset de Harvard													
Resampling	Métricas	SEED 1					SEED 2					Média	Desvio Padrão
		Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5		
AVERAGE	PSNR	26,7569	33,9583	28,6848	30,9551	29,0109	26,7215	33,9314	28,6035	30,9223	29,0011	29,8546	2,5733
	SAM	0,1293	0,0799	0,1086	0,0761	0,0657	0,1291	0,0800	0,1083	0,0762	0,0657	0,0919	0,0247
	SSIM	0,9978	0,9998	0,9992	0,9988	0,9993	0,9979	0,9998	0,9992	0,9988	0,9993	0,9990	0,0007
	ERGAS	24,9349	19,5008	22,5049	16,7557	11,3335	24,9645	19,5870	22,6944	16,8250	11,3428	19,0443	4,9942
	RMSE	0,0009	0,0003	0,0005	0,0006	0,0006	0,0009	0,0003	0,0005	0,0006	0,0006	0,0006	0,0002
	CC	0,9667	0,9883	0,9800	0,9882	0,9928	0,9669	0,9884	0,9802	0,9881	0,9928	0,9832	0,0097
GAUSS	UIQI	0,9791	0,9883	0,9854	0,9908	0,9948	0,9793	0,9884	0,9853	0,9908	0,9947	0,9877	0,0055
	PSNR	28,7958	34,7606	28,6050	33,0098	30,4173	28,8075	33,1040	31,0278	32,3378	30,3685	31,1234	2,1183
	SAM	0,1197	0,0755	0,1196	0,0662	0,0608	0,1155	0,0941	0,0955	0,0694	0,0606	0,0877	0,0243
	SSIM	0,9984	0,9999	0,9993	0,9994	0,9994	0,9985	0,9998	0,9995	0,9991	0,9994	0,9993	0,0005
	ERGAS	23,4365	19,1194	34,0364	14,3095	10,4941	22,3819	26,3387	20,2096	18,1896	10,5820	19,9098	7,2276
	RMSE	0,0008	0,0002	0,0005	0,0005	0,0005	0,0008	0,0003	0,0004	0,0005	0,0005	0,0005	0,0002
Bilinear	CC	0,9733	0,9894	0,9639	0,9916	0,9940	0,9746	0,9788	0,9846	0,9899	0,9941	0,9834	0,0103
	UIQI	0,9819	0,9888	0,9441	0,9939	0,9957	0,9838	0,9632	0,9886	0,9919	0,9956	0,9828	0,0166
	PSNR	26,3175	35,4012	30,5528	32,4361	30,0880	28,4816	34,9264	30,4204	32,4807	30,1011	31,1206	2,7757
	SAM	0,1288	0,0745	0,0954	0,0681	0,0603	0,1150	0,0738	0,0955	0,0674	0,0604	0,0839	0,0237
	SSIM	0,9979	0,9999	0,9995	0,9992	0,9995	0,9986	0,9999	0,9995	0,9993	0,9995	0,9993	0,0006
	ERGAS	27,2132	17,6401	19,3441	14,7602	10,2679	21,9764	17,4548	19,2028	14,5508	10,2922	17,2703	5,1708
Cubic	RMSE	0,0009	0,0002	0,0004	0,0005	0,0005	0,0008	0,0002	0,0004	0,0005	0,0005	0,0005	0,0002
	CC	0,9687	0,9900	0,9849	0,9908	0,9942	0,9747	0,9902	0,9847	0,9912	0,9942	0,9864	0,0085
	UIQI	0,9745	0,9911	0,9895	0,9930	0,9959	0,9844	0,9912	0,9892	0,9935	0,9960	0,9898	0,0064
	PSNR	29,4759	34,9346	31,4229	32,2491	30,5918	29,0324	33,4082	31,1006	31,6306	30,6058	31,4452	1,7626
	SAM	0,1131	0,0757	0,0901	0,0681	0,0601	0,1124	0,0880	0,0956	0,0694	0,0584	0,0831	0,0200
	SSIM	0,9987	0,9999	0,9997	0,9994	0,9995	0,9987	0,9998	0,9996	0,9993	0,9996	0,9994	0,0004
Cubic Spline	ERGAS	21,2971	18,4509	17,8678	14,8000	10,2434	21,1368	25,0396	19,9012	17,9194	9,8603	17,6516	4,8253
	RMSE	0,0008	0,0002	0,0004	0,0005	0,0005	0,0008	0,0003	0,0004	0,0005	0,0005	0,0005	0,0002
	CC	0,9764	0,9895	0,9864	0,9907	0,9943	0,9761	0,9809	0,9850	0,9905	0,9946	0,9865	0,0068
	UIQI	0,9848	0,9887	0,9899	0,9899	0,9959	0,9836	0,966571	0,9870	0,9866	0,9962	0,9869	0,0083
	PSNR	27,0397	35,8962	31,3409	33,0987	30,6365	29,2724	35,4675	31,1826	33,1185	30,6202	29,8546	2,5733
	SAM	0,1231	0,0726	0,0919	0,0665	0,0591	0,1126	0,0719	0,0921	0,0659	0,0595	0,0919	0,0247
Nearest	SSIM	0,9982	0,9999	0,9996	0,9992	0,9992	0,9987	0,9999	0,9996	0,9993	0,9995	0,9990	0,0007
	ERGAS	26,0220	17,1330	18,6673	14,4924	10,0508	21,5593	16,9097	18,6135	14,2910	10,2112	19,0443	4,9942
	RMSE	0,0009	0,0002	0,0004	0,0005	0,0005	0,0008	0,0002	0,0004	0,0005	0,0005	0,0006	0,0002
	CC	0,9718	0,9905	0,9861	0,9912	0,9945	0,9765	0,9907	0,9858	0,9915	0,9944	0,9832	0,0097
	UIQI	0,9758	0,9917	0,9901	0,9931	0,9961	0,9850	0,9919	0,9899	0,9936	0,9960	0,9877	0,0055
	PSNR	27,0664	33,1429	27,8762	31,8327	29,6219	27,6002	31,6041	29,3771	31,7640	29,5635	29,9449	2,0680
Lanczos	SAM	0,1389	0,0895	0,1303	0,0754	0,0687	0,1342	0,0995	0,1114	0,0753	0,0688	0,0992	0,0279
	SSIM	0,9972	0,9997	0,9990	0,9988	0,9991	0,9974	0,9996	0,9990	0,9987	0,9992	0,9988	0,0008
	ERGAS	29,4764	23,9861	37,2393	17,2765	12,3412	27,0763	24,1100	23,4340	17,7441	12,0156	22,4700	7,8567
	RMSE	0,0010	0,0003	0,0005	0,0006	0,0006	0,001	0,0003	0,0005	0,0006	0,0006	0,0006	0,0002
	CC	0,9633	0,9846	0,9580	0,9881	0,9922	0,9654	0,9743	0,9767	0,9882	0,9923	0,9783	0,0127
	UIQI	0,9701	0,9816	0,9431	0,9915	0,9939	0,9765	0,9635	0,9821	0,9915	0,9944	0,9788	0,0163
Mode	PSNR	29,6526	35,7024	31,6412	33,4312	30,7745	29,7152	35,6874	31,5342	33,3347	30,7540	32,2227	2,2303
	SAM	0,1109	0,0708	0,0899	0,0652	0,0584	0,1092	0,0711	0,0914	0,0659	0,0583	0,0791	0,0199
	SSIM	0,9988	0,9999	0,9997	0,9995	0,9996	0,9988	0,9999	0,9996	0,9994	0,9996	0,9995	0,0004
	ERGAS	20,9179	16,3265	17,9269	15,2821	9,8581	20,6173	16,4699	18,4516	14,9341	9,8449	16,0629	3,8409
	RMSE	0,0008	0,0002	0,0004	0,0005	0,0005	0,0008	0,0002	0,0004	0,0005	0,0005	0,0005	0,0002
	CC	0,9774	0,9911	0,9870	0,9919	0,9948	0,9780	0,9910	0,9866	0,9916	0,9948	0,9884	0,0063
RMS	UIQI	0,9855	0,9921	0,9905	0,9936	0,9963	0,9861	0,9920	0,9899	0,9934	0,9963	0,9916	0,0037
	PSNR	27,2396	33,7279	26,6409	30,5075	29,2729	27,2967	33,7055	29,3110	31,3516	29,6067	29,8660	2,5145
	SAM	0,1419	0,0852	0,1459	0,0826	0,0708	0,1424	0,0857	0,1118	0,0811	0,0684	0,1016	0,0311
	SSIM	0,9970	0,9998	0,9988	0,9982	0,9990	0,9970	0,9998	0,9991	0,9982	0,9992	0,9986	0,0010
	ERGAS	28,3013	20,6030	46,0805	19,1215	12,4924	28,1836	21,3707	23,9060	18,4798	11,9020	23,0441	9,8089
	RMSE	0,0010	0,0003	0,0006	0,0007	0,0006	0,0010	0,0003	0,0005	0,0007	0,0006	0,0006	0,0003
RMS	CC	0,9617	0,9863	0,9475	0,9850	0,9917	0,9617	0,9866	0,9786	0,9859	0,9923	0,9777	0,0153
	UIQI	0,9757	0,9890	0,9238	0,9866	0,9937	0,9761	0,9882	0,9863	0,9901	0,9946	0,9804	0,0209
	PSNR	29,3425	34,5406	28,8807	32,3103	30,4542	27,4952	35,5000	31,2697	33,2560	30,7000	31,3749	2,5423
	SAM	0,1135	0,0770	0,1178	0,0693	0,0606	0,1265	0,0722	0,0924	0,0663	0,0593	0,0855	0,0253
	SSIM	0,9986	0,9998	0,9994	0,9992	0,9995	0,9982	0,9999	0,9996	0,9993	0,9995	0,9993	0,0005
	ERGAS	21,1040	17,8630	32,4168	16,5165	10,4765	27,0082	16,8079	18,6562	14,2010	10,0923	18,5142	6,9412
RMS	RMSE	0,0008	0,0002	0,0005	0,0005	0,0005	0,0009	0,0002	0,0004	0,0005	0,0005	0,0005	0,0002
	CC	0,9758	0,9871	0,9651	0,9903	0,9942	0,9701	0,9904	0,9859	0,9913	0,9945	0,9845	0,0104
	UIQI	0,9842	0,9847	0,9480	0,9899	0,9957	0,9674	0,9912	0,9899	0,9933	0,9960	0,9840	0,0151

Dataset Unisinos													
Resampling	Métricas	SEED 1					SEED 2					Média	Desvio Padrão
		Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5		
Cubic	PSNR	31,9177	31,1153	31,7274	30,9359	29,7152	31,9375	31,1355	31,6108	30,8000	29,7985	30,9751	0,7863
	SAM	0,0373	0,0416	0,0378	0,0437	0,0558	0,0369	0,0414	0,0385	0,0440	0,0557	0,0439	0,0071
	SSIM	0,9124	0,9342	0,9474	0,9159	0,9154	0,9163	0,9350	0,9383	0,9086	0,9178	0,9254	0,0134
	ERGAS	5,7860	6,2592	5,6852	6,7104	8,9228	5,7440	6,2251	5,7910	6,7681	8,911313	6,7797	1,2732
	RMSE	1,4780	1,5359	1,4492	1,6182	2,1432	1,4432	1,5277	1,4780	1,6153	2,160931	1,6635	0,2840
	UIQI	0,9987	0,9981	0,9987	0,9978	0,9953	0,9987	0,9981	0,9987	0,9978	0,9953	0,9976	0,0013
Cubic Spline	PSNR	32,0696	31,0625	31,7250	30,9448	29,8712	32,0327	31,1346	31,6042	30,8197	29,9113	31,1176	0,7803
	SAM	0,0367	0,0419	0,0377	0,0436	0,0550	0,0367	0,0413	0,0386	0,0440	0,0552	0,0431	0,0069
	SSIM	0,9191	0,9234	0,9441	0,9145	0,9220	0,9138	0,9330	0,9358	0,9065	0,9222	0,9234	0,0113
	ERGAS	5,6820	6,3004	5,6721	6,6926	8,7839	5,6535	6,2147	5,8070	6,7683	8,8267	6,6401	1,2117
	RMSE	1,4508	1,5450	1,4459	1,6140	2,1239	1,4365	1,5256	1,4821	1,6148	2,1291	1,6368	0,2659
	UIQI	0,9987	0,9980	0,9987	0,9978	0,9954	0,9987	0,9981	0,9987	0,9978	0,9954	0,9977	0,0013
Lanczos	PSNR	32,1463	31,2082	31,7673	30,9818	29,6839	31,8788	31,2198	31,7533	30,8963	29,7295	31,1265	0,8519
	SAM	0,0365	0,0415	0,0375	0,0432	0,0549	0,0369	0,0411	0,0376	0,0432	0,0559	0,0428	0,0071
	SSIM	0,9233	0,9346	0,9460	0,9151	0,9196	0,9180	0,9332	0,9472	0,9142	0,9129	0,9264	0,0130
	ERGAS	5,6382	6,2334	5,6373	6,6494	8,9053	5,7744	6,1765	5,6494	6,6387	8,9245	6,6227	1,2679
	RMSE	1,4401	1,5295	1,4370	1,5911	2,1304	1,4384	1,5157	1,4411	1,5814	2,1433	1,6248	0,2762
	UIQI	0,9987	0,9981	0,9987	0,9978	0,9955	0,9987	0,9981	0,9987	0,9979	0,9954	0,9978	0,0013

APÊNDICE C – PRODUÇÃO CIENTÍFICA

Na Tabela a seguir estão listadas as contribuições científicas publicadas em periódicos ou anais de eventos que foram geradas direta ou indiretamente pela doutoranda durante o período do doutorado, sendo destacadas em negrito aquelas cujo primeiro autor seja a autora deste projeto de tese.

Tabela 12: Produção científica autoral e participação em artigos.

Status	Tipo	Qualis	H5	Publicação
Aceito	Evento	A3	21	MÜLLER, Mariane et al. A quantitative analysis on different carbonate indicators based on spaceborne data in controlled karst area. In: 2020 International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2020.
Aceito	Evento	A3	21	ZANOTTA, Daniel et al. A multi-looking approach for spatial super-resolution on laboratory-based hyperspectral images. In: 2021 International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2021.
Aceito	Evento	A3	21	ZANOTTA, Daniel et al. A geometric approach for hyperspectral and 3D point cloud integratio in natural scenes . In: 2022 International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2022.
Aceito	Evento	A3	21	MARQUES, Ademir et al. DC-GAN for fracture data generation based on segmented outcrop images acquired from UAV. In: 2022 International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2022.
Aceito	Evento	A3	21	MARQUES, Ademir et al. Low-cost bathymetry prototype and adapted sonar calibration for water reservoir measurement. In: 2022 International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2022.
Aceito	Evento	A3	21	SALES, Vinicius et al. Analysis of machine learning techniques for carbonate outcrop image classification in Landsat 8 satellite data. In: 2022 International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2022.