



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Modelos de Detecção de Nuvens Usando Redes Neurais Totalmente Convolucionais em Imagens Multiespectrais do Sentinel-2

Alan Breno Soares Corrêa

DM 06/2025

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil

2025

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Alan Breno Soares Corrêa

**Modelos de Detecção de Nuvens Usando Redes Neurais Totalmente
Convolucionais em Imagens Multiespectrais do Sentinel-2**

Dissertação submetida à Banca Examinadora
do Programa de Pós-Graduação em Engenharia
Elétrica da UFPA para obtenção do Grau de
Mestre em Engenharia Elétrica na Área de
Telecomunicações.

Orientador: Prof. Dr. Fabrício José Brito Barros

Coorientador: Prof. Dr. Marcos César da Rocha Seruffo

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil

2025

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

S676m Soares Corrêa, Alan Breno,
MODELOS DETECÇÃO DE NUVENS USANDO REDES
NEURAS TOTALMENTE CONVOLUCIONAIS EM IMAGENS
MULTIESPECTRAIS DO SENTINEL-2 / Alan Breno Soares
Corrêa. — 2025.
78 f. : il. color.

Orientador(a): Prof. Dr. Fabrício José Brito Barros
Coorientador(a): Prof. Dr. Marcos César da Rocha Seruffo
Dissertação (Mestrado) - Universidade Federal do Pará,
Instituto de Tecnologia, Programa de Pós-Graduação em
Engenharia Elétrica, Belém, 2025.

1. Detecção de Nuvens. 2. Segmentação de Imagens. 3.
Sentinel-2. 4. Redes Neurais Totalmente Convolucionais. 5.
Sensoriamento Remoto. I. Título.

CDD 621.307209811



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA


“MODELOS DETECÇÃO DE NUVENS USANDO REDES NEURASIS TOTALMENTE CONVOLUCIONAIS EM IMAGENS MULTIESPECTRAIS DO SENTINEL-2”

AUTOR: **ALAN BRENO SOARES CORREA**


DISSERTAÇÃO DE MESTRADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE TELECOMUNICAÇÕES.

APROVADA EM: 28/03/2025


BANCA EXAMINADORA:

Documento assinado digitalmente
 **FABRICIO JOSE BRITO BARROS**
Data: 02/04/2025 22:36:31-0300
Verifique em <https://validar.iti.gov.br>


Prof. Dr. Fabrício José Brito Barros
(Orientadora – PPGEE/ITEC/UFPA)

Documento assinado digitalmente
 **MARCOS CESAR DA ROCHA SERUFFO**
Data: 02/04/2025 22:14:37-0300
Verifique em <https://validar.iti.gov.br>


Prof. Dr. Marcos César da Rocha Seruffo
(Coorientador – PPGEE/ITEC/UFPA)

Documento assinado digitalmente
 **JASMINE PRISCYLA LEITE DE ARAUJO**
Data: 02/04/2025 22:23:45-0300
Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Jasmine Priscyla Leite de Araújo
(Avaliadora Interna – PPGEE/ITEC/UFPA)

Documento assinado digitalmente
 **GABRIEL BRITO COSTA**
Data: 02/04/2025 09:38:14-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Gabriel Brito Costa
(Avaliador Externo – UFOPA)

Documento assinado digitalmente
 **Diego Lisboa Cardoso**
Data: 10/04/2025 13:32:58-0300
Verifique em <https://validar.iti.gov.br>

VISTO:

Prof. Dr. Diego Lisboa Cardoso
(Coordenador do PPGEE/ITEC/UFPA)

*Dedico este trabalho ao meu filho, **Alan Miguel**, cuja existência me tornou uma pessoa melhor.*

Agradecimentos

A Deus, por me conceder força, sabedoria e perseverança para seguir adiante, mesmo diante dos desafios. Sem a Sua presença em minha vida, nada disso seria possível.

À minha esposa, Paula Cristina, e ao meu filho, Alan Miguel, que são a razão maior da minha luta e superação. O amor, o apoio e a paciência de vocês foram fundamentais em todos os momentos dessa caminhada.

Aos meus pais, Antonio Carlos e Adilce Meire, por todo amor, apoio e pelos valores que me ensinaram desde a infância. Foram exemplos de dedicação e coragem, e tudo o que conquistei até aqui carrega um pedaço do esforço de vocês.

Aos meus irmãos, Antonio, Alex, Eleticia e Lelma, por serem parte essencial da minha vida, mesmo quando distantes, sei que posso contar com vocês.

À minha tia Ana, que me acolheu no início da minha jornada acadêmica, permitindo que eu morasse com ela e pudesse dar os primeiros passos na universidade. Seu gesto de generosidade e carinho foi crucial para que eu pudesse chegar até aqui.

Aos amigos que fiz na UFPA - Lucian (meu amigo e coorientador de TCC), Mozart, Clara e Caio -, pelos momentos de parceria, aprendizado e amizade. A convivência com vocês tornou esta jornada acadêmica mais leve e enriquecedora.

Ao professor Dr. Fabrício Barros, por me acolher no Laboratório de Computação e Telecomunicações (LCT), me orientar desde o Trabalho de Conclusão de Curso (TCC) e, principalmente, por acreditar no meu potencial, incentivando-me a seguir no caminho da pesquisa e do mestrado. Sua orientação e apoio foram fundamentais para a realização deste trabalho.

Ao professor Dr. Marcos Seruffo, cuja orientação, apoio e incentivo foram essenciais para a continuidade deste sonho. Sua confiança e palavras de encorajamento me deram forças para seguir em frente.

À empresa *Carbonext - Nature and Future*, pela parceria no Programa de Mestrado e Doutorado Acadêmico para Inovação (MAI/DAI), que me possibilitou manter o foco e a dedicação durante o curso. A oportunidade de aliar pesquisa e prática profissional foi essencial para o desenvolvimento deste trabalho.

Aos novos colegas Jean, Leo, Pedro, Williane e Albert, cuja parceria e amizade ao longo dos últimos anos foram extremamente enriquecedoras, tanto para o crescimento pessoal quanto acadêmico.

A todos os meus professores, que, ao longo da minha formação, contribuíram com conhecimento, dedicação e inspiração, deixando marcas importantes na minha trajetória.

Por fim, a todos que, de alguma forma, fizeram parte dessa caminhada, o meu mais sincero e profundo agradecimento.

“O maior valor de uma imagem é quando ela nos força a notar o que nunca esperávamos ver.”

John W. Tukey

Resumo

A detecção precisa de nuvens em imagens de satélite é fundamental para diversas aplicações de sensoriamento remoto, como monitoramento ambiental e análise de mudanças na cobertura terrestre. Satélites como o Sentinel-2 desempenham um papel essencial nesse contexto, pois fornecem imagens de alta resolução em nível global, com curto período de revisita (5 dias). No entanto, a presença de nuvens e sombras de nuvens representa um grande desafio no pré-processamento dessas imagens, dificultando a extração precisa de informações. Diversas abordagens baseadas em limiares espectrais e aprendizado profundo foram desenvolvidas para mitigar esse problema, mas ainda há espaço para melhorias. Neste trabalho, propõe-se o uso de *Fully Convolutional Neural Networks* (FCNNs) para a segmentação de nuvens em imagens do Sentinel-2, explorando diferentes níveis de processamento (L1C e L2A) e combinações de bandas espectrais na entrada (todas as bandas e RGB+NIR). Modelos baseados na arquitetura UNet foram treinados utilizando os encoders EfficientNet-B1 e MobileNet-V2, visando comparar desempenho, eficiência na segmentação e impacto do número de bandas. O conjunto de dados CloudSen12, composto por 10.000 imagens de 512×512 pixels de diferentes regiões do mundo, foi utilizado para os experimentos, abrangendo diversas condições atmosféricas. A avaliação quantitativa incluiu métricas como Acurácia, *Intersection over Union* (IoU) e F1-Score, enquanto a análise qualitativa foi realizada por meio da inspeção visual das máscaras de segmentação. Os resultados demonstraram que o encoder EfficientNet-B1 apresentou o melhor desempenho, atingindo 95,21% de acurácia, 82,74% de IoU e 90,56% de F1-Score. Além disso, modelos treinados com apenas as bandas RGB+NIR apresentaram desempenho competitivo, com acurácia de 94,87%, IoU de 81,38% e F1-Score de 89,73%. A análise entre os níveis de processamento indicou que a remoção de efeitos atmosféricos no nível L2A teve pouca influência na segmentação em relação ao nível L1C. Por fim, os modelos propostos superaram abordagens tradicionais e outras arquiteturas da literatura, destacando o potencial das FCNNs para aprimorar a detecção de nuvens em aplicações de sensoriamento remoto.

Palavras-chave: Detecção de Nuvem, Segmentação de Imagens, Sentinel-2, Redes Neurais Totalmente Convolucionais, Sensoriamento Remoto.

Abstract

Accurate cloud detection in satellite imagery is essential for various remote sensing applications, such as environmental monitoring and land cover change analysis. Satellites like Sentinel-2 play a key role in this context, providing high-resolution imagery at a global scale with a short revisit period (5 days). However, the presence of clouds and cloud shadows poses a major challenge in the preprocessing of these images, hindering the precise extraction of information. Several approaches based on spectral thresholds and deep learning have been developed to mitigate this issue, but there is still room for improvement. This work proposes the use of *Fully Convolutional Neural Networks* (FCNNs) for cloud segmentation in Sentinel-2 images, exploring different processing levels (L1C and L2A) and combinations of input spectral bands (all bands and RGB+NIR). Models based on the UNet architecture were trained using EfficientNet-B1 and MobileNet-V2 encoders, aiming to compare performance, segmentation efficiency, and the impact of the number of bands. The CloudSen12 dataset, consisting of 10,000 images of 512×512 pixels from different regions around the world and covering diverse atmospheric conditions, was used for the experiments. Quantitative evaluation included metrics such as Accuracy, *Intersection over Union* (IoU), and F1-Score, while qualitative analysis was performed through visual inspection of the segmentation masks. The results showed that the EfficientNet-B1 encoder achieved the best performance, reaching 95.21% accuracy, 82.74% IoU, and 90.56% F1-Score. Additionally, models trained with only the RGB+NIR bands achieved competitive performance, with 94.87% accuracy, 81.38% IoU, and 89.73% F1-Score. The comparison between processing levels indicated that the removal of atmospheric effects in L2A had little influence on segmentation compared to L1C. Finally, the proposed models outperformed traditional approaches and other architectures from the literature, highlighting the potential of FCNNs to enhance cloud detection in remote sensing applications.

Key-words: Cloud Detection, Image Segmentation, Sentinel-2, Fully Convolutional Neural Networks, Remote Sensing.

Lista de figuras

Figura 1 – Comparação entre imagem L1C vs L2A.	13
Figura 2 – Localização das imagens de treino (cinza) e teste (preto).	15
Figura 3 – Mosaico de imagens e suas respectivas máscaras do CloudSEN12.	16
Figura 4 – Arquitetura básica da primeira FCNN proposta na literatura.	17
Figura 5 – Arquitetura da primeira U-Net proposta na literatura.	19
Figura 6 – Metodologia para treinamento e avaliação dos modelos.	24
Figura 7 – Exemplo de imagem com percentuais de cobertura de nuvem diferentes. . .	25
Figura 8 – Imagem com alto grau de dificuldade na segmentação. Da esquerda para a direita, na primeira linha as imagens L1C e L2A e suas respectivas máscaras. Na segunda linha as máscaras dos modelos que utilizam todas as bandas. Na terceira linha as máscaras dos modelos que utilizam apenas 4 bandas (RGB+NIR).	32
Figura 9 – Imagem com baixo grau de dificuldade na segmentação. Da esquerda para a direita, na primeira linha as imagens L1C e L2A e suas respectivas máscaras. Na segunda linha as máscaras dos modelos que utilizam todas as bandas. Na terceira linha as máscaras dos modelos que utilizam apenas 4 bandas (RGB+NIR).	33
Figura 10 – Gráfico da função de perda durante o treinamento dos modelos.	48
Figura 11 – Gráfico da função de perda durante a validação dos modelos.	48
Figura 12 – Gráfico da acurácia durante o treinamento dos modelos.	49
Figura 13 – Gráfico da acurácia durante a validação dos modelos.	49
Figura 14 – Gráfico do IoU durante o treinamento dos modelos.	50
Figura 15 – Gráfico do IoU durante a validação dos modelos.	50
Figura 16 – Gráfico do F1-Score durante o treinamento dos modelos.	51
Figura 17 – Gráfico do F1-Score durante a validação dos modelos.	51
Figura 18 – Máscaras dos modelos propostos em imagem com média cobertura de nuvens.	52
Figura 19 – Máscaras dos modelos propostos em imagem de difícil anotação de classes.	52
Figura 20 – Máscaras dos modelos propostos em imagem de alta cobertura de nuvens e sombras.	53
Figura 21 – Máscaras dos modelos propostos em imagem totalmente coberta por nuvens e sombras.	53
Figura 22 – Máscaras dos modelos propostos em imagem com baixa cobertura de nuvens e sombras.	54
Figura 23 – Máscaras dos modelos propostos em imagem sem nuvens e sombras.	54
Figura 24 – Máscaras dos modelos com classes iguais em imagem com média cobertura de nuvens.	55

Figura 25 – Máscaras dos modelos com classes iguais em imagem quase sem cobertura de nuvens.	55
Figura 26 – Máscaras dos modelos com classes iguais em imagem com alta cobertura de nuvens.	56
Figura 27 – Máscaras dos modelos com classes iguais em imagem com baixa cobertura de nuvens.	56
Figura 28 – Máscaras dos modelos em imagem sem nuvens, porém com superfície brilhante.	57
Figura 29 – Máscaras dos modelos em imagem com alta cobertura de nuvens.	57
Figura 30 – Máscaras dos modelos em imagem com média cobertura de nuvens.	58
Figura 31 – Máscaras dos modelos em imagem com alta cobertura de nuvens.	58
Figura 32 – Máscaras dos modelos em imagem com baixo percentual de pixels inválidos.	59
Figura 33 – Máscaras dos modelos em imagem com alto percentual de pixels inválidos. .	59
Figura 34 – Máscaras dos modelos em imagem com médio percentual de pixels inválidos.	60
Figura 35 – Máscaras dos modelos em imagem com alto percentual de pixels inválidos. .	60

Lista de tabelas

Tabela 1 – Resumo das Contribuições dos Trabalhos Correlatos.	9
Tabela 2 – Bandas Espectrais dos satélites Sentinel-2A e Sentinel-2B.	12
Tabela 3 – Comparação entre os produtos Sentinel-2 L1C e L2A	14
Tabela 4 – Classes semânticas das máscaras do CloudSen12.	16
Tabela 5 – Comparativo entre MobileNet-V2 e EfficientNet-B1	21
Tabela 6 – Comparação de métricas entre diferentes modelos para L1C e L2A	29
Tabela 7 – Comparação de Modelos com Classes Semelhantes.	35
Tabela 8 – Classes e Superclasses de pixels para avaliação dos modelos.	35
Tabela 9 – Resultados da comparação entre modelos para as classes Nuvem e Não Nuvem.	37
Tabela 10 – Resultados da comparação entre modelos para as classes válido e inválido.	38

Lista de abreviaturas e siglas

FCNN	<i>Fully Convolutional Neural Networks</i>
CNN	<i>Convolutional Neural Networks</i>
IoU	<i>Intersection Over Union</i>
NIR	<i>Near Infrared</i>
SWIR	<i>Short-Wave Infrared</i>
TOA	<i>Top-Of-Atmosphere</i>
BOA	<i>Bottom-Of-Atmosphere</i>
SCL	<i>Scene Classification Layer</i>
AOT	<i>Aerosol Optical Thickness</i>
WV	<i>Water Vapor</i>
Fmask	<i>Function of Mask</i>
XGBoost	<i>Extreme Gradient Boosting</i>
RF	<i>Random Forest</i>
SVM	<i>Support Vector Machine</i>
MCM	<i>Multitemporal Cloud Masking</i>
TSM	<i>Time Series Maximum and Minimum mask method</i>
NDTeI	<i>Normalized Difference Temporal Index</i>
SNIC	<i>Simple Non-Iterative Clustering</i>
CRF	<i>Conditional Random Fields</i>
GCP	<i>Global Convolutional Pooling</i>
LPP	<i>Local Pooling Pruning</i>
CAM	<i>Cloud Activation Map</i>
SSCA-net	<i>Spectral-Spatial-Context Attention Network</i>

ESA	<i>Europe Space Agency</i>
UE	União Européia
SMP	<i>Segmentation Models PyTorch</i>

Sumário

1	INTRODUÇÃO	1
1.1	Contexto	1
1.2	Motivação	1
1.3	Objetivos	2
1.3.1	Objetivo Geral	2
1.3.2	Objetivos Específicos	2
1.4	Estrutura do Trabalho	3
2	TRABALHOS CORRELATOS	4
2.1	Considerações Iniciais	4
2.2	Descrição dos Trabalhos Correlatos	4
2.3	Comparativo entre este Trabalho e os Correlatos	8
2.4	Síntese do Capítulo	10
3	REFERENCIAL TEÓRICO	11
3.1	Considerações Iniciais	11
3.2	Sobre o Sentinel-2	11
3.2.1	Características do Sentinel-2	11
3.2.2	Produtos do Sentinel-2	13
3.2.2.1	Produtos L1C	13
3.2.2.2	Produtos L2A	14
3.3	Dataset CloudSEN12	14
3.3.1	Processo de criação do Dataset	14
3.3.2	Divisão em Treinamento, Validação e Teste	15
3.3.3	Classes Semânticas e Cobertura de Nuvens	15
3.3.4	Relevância para Treinamento de Redes Neurais Totalmente Convolucionais	16
3.4	Redes Neurais Totalmente Convolucionais	17
3.4.1	Estrutura	17
3.4.2	Funcionamento	18
3.4.3	Modelo UNet	19
3.4.4	Encoders MobileNet-v2 e EfficientNet-B1	20
3.4.4.1	MobileNet-v2	20
3.4.4.2	EfficientNet-B1	20
3.5	Métricas	21
3.5.1	Acurácia	21
3.5.2	Intersection over Union (IoU)	22
3.5.3	F1-Score	22
3.6	Síntese do Capítulo	23

4	METODOLOGIA	24
4.1	Considerações Iniciais	24
4.2	Base de Dados	25
4.3	Modelos e Arquiteturas	26
4.4	Pré-Processamento e Treinamento	26
4.5	Avaliação dos Modelos	27
4.6	Síntese do Capítulo	28
5	RESULTADOS	29
5.1	Considerações Iniciais	29
5.2	Avaliação Quantitativa dos Modelos	29
5.2.1	Comparação entre Encoders	29
5.2.2	Impacto da Quantidade de Bandas Espectrais	30
5.2.3	Influência do Nível de Processamento (L1C vs. L2A)	30
5.2.4	Resumo da Avaliação Quantitativa	30
5.3	Avaliação Qualitativa dos Modelos	31
5.3.1	Imagem com Alto Grau de Dificuldade	31
5.3.2	Imagem com Baixo Grau de Dificuldade	32
5.3.3	Conclusões da Análise Qualitativa	33
5.4	Comparação com Outros Modelos	34
5.4.1	Comparação com Modelos com Classes Iguais	34
5.4.2	Comparação com Modelos com Classes Diferentes	35
5.4.2.1	Nuvem vs Não Nuvem	36
5.4.2.2	Válido vs Inválido	37
5.4.3	Conclusões das Comparações com Outros Modelos	38
5.5	Síntese do Capítulo	39
6	CONCLUSÃO	40
6.1	Trabalhos Futuros	41
6.2	Publicações	41
6.3	Agradecimentos à Carbonext	42
	Referências	44
	Apêndices	48
A	Apêndice A	48
B	Apêndice B	52
C	Apêndice C	55
D	Apêndice D	57

E	Apêndice E	59
---	----------------------	----

1 INTRODUÇÃO

1.1 Contexto

A detecção de nuvens em imagens de satélite constitui uma etapa crucial do pré-processamento em diversas aplicações de sensoriamento remoto, como monitoramento ambiental, previsão meteorológica, gestão de recursos hídricos e análise de uso e cobertura do solo (ZHU; WOODCOCK, 2014). Em muitas situações, as nuvens e suas sombras podem obstruir informações essenciais, prejudicando a qualidade das imagens e dificultando a extração de dados confiáveis. Técnicas tradicionais de correção atmosférica ou baseadas em limiares espectrais têm sido amplamente utilizadas, mas ainda enfrentam limitações na identificação de nuvens finas, sombras de nuvens ou áreas com superfícies altamente refletivas (TARRIO et al., 2020). Apesar de muitos desenvolvimentos, a detecção de nuvens continua sendo um problema não resolvido, com espaço para melhorias, especialmente em superfícies claras e nuvens finas (LÓPEZ-PUIGDOLLERS et al., 2021).

O satélite Sentinel-2, parte do programa Copernicus da Agência Espacial Europeia (ESA), destaca-se como uma das principais fontes de dados para este tipo de aplicação. Composto por uma constelação de dois satélites (Sentinel-2A e Sentinel-2B), o sistema fornece imagens multiespectrais com resolução espacial de 10m, 20m e 60m, abrangendo 13 bandas espectrais do visível ao infravermelho de onda curta (SWIR) (ESA, 2015). As bandas de maior resolução (10m) incluem o azul (B2: 490 nm), verde (B3: 560 nm), vermelho (B4: 665 nm) e infravermelho próximo (B8: 842 nm), enquanto bandas específicas como o SWIR (B11: 1610 nm) e o infravermelho de borda vermelha (B5: 705 nm) são particularmente úteis para a detecção de nuvens devido às suas características de absorção.

Nos últimos anos, a popularização de sensores multiespectrais e a crescente disponibilidade de imagens de alta resolução espacial impulsionaram a pesquisa em novas abordagens de detecção de nuvens. Ao mesmo tempo, métodos de aprendizado profundo vêm se mostrando cada vez mais promissores, especialmente quando comparados às abordagens convencionais. Nesse cenário, a adoção de redes neurais totalmente convolucionais destaca-se por sua capacidade de segmentar nuvens de forma precisa e automática, classificando cada pixel com base em informações espectrais e espaciais (AYBAR et al., 2022).

1.2 Motivação

O uso de *Fully Convolutional Neural Networks* (FCNNs) para a detecção de nuvens é motivado pela capacidade desses modelos de gerar máscaras de segmentação que preservam a

resolução original das imagens, possibilitando uma classificação detalhada em nível de pixel (LONG; SHELHAMER; DARRELL, 2014). A evolução de arquiteturas baseadas em codificadores (*encoders*) e decodificadores (*decoders*), como a UNet (RONNEBERGER; FISCHER; BROX, 2015), permitiu o desenvolvimento de estruturas mais leves e eficientes computacionalmente. Neste trabalho, são explorados dois codificadores modernos: MobileNet-V2 (SANDLER et al., 2018) e EfficientNet-B1 (TAN; LE, 2020).

O MobileNet-V2 utiliza convoluções separáveis em profundidade (*depthwise separable convolutions*) e blocos residuais invertidos para reduzir a complexidade computacional, sendo particularmente adequado para aplicações que exigem eficiência em recursos limitados. Já o EfficientNet-B1 faz parte da família de modelos escalonados de forma composta (*compound scaling*), que equilibram uniformemente profundidade, largura e resolução da rede para maximizar a precisão mantendo a eficiência computacional. A combinação desses codificadores com arquiteturas de decodificação adaptadas para imagens de satélite apresenta potencial para superar desafios específicos do domínio.

Entretanto, ainda há questões em aberto na literatura: o impacto do nível de processamento das imagens (*Level-1C* - L1C, com valores de reflectância no topo da atmosfera, e o *Level-2A* - L2A, com reflectância de superfície corrigida atmosféricamente) na qualidade da detecção, e a influência do número de bandas espectrais utilizadas. Sendo assim, este trabalho busca investigar como o uso de FCNNs pode superar as limitações das abordagens tradicionais na detecção de nuvens em imagens do Sentinel-2. Diferentemente de trabalhos anteriores, esta pesquisa avalia sistematicamente a influência do nível de processamento (L1C vs. L2A) e da quantidade de bandas espectrais na segmentação de nuvens, utilizando arquiteturas modernas de aprendizado profundo.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral deste trabalho é desenvolver e avaliar modelos de redes neurais totalmente convolucionais para a detecção de nuvens em imagens do satélite Sentinel-2, considerando diferentes níveis de processamento (L1C e L2A) e combinações espectrais (todas as bandas e RGB + NIR).

1.3.2 Objetivos Específicos

- Comparar o desempenho de diferentes de encoders (EfficientNet-B1 e MobileNet-V2) na tarefa de segmentação de nuvens, avaliando métricas como acurácia, IoU e F1-Score.
- Analisar o impacto da utilização de diferentes conjuntos de bandas (todas as bandas disponíveis ou apenas RGB + NIR) no desempenho dos modelos, verificando possíveis

ganhos ou perdas de acurácia e eficiência, buscando o desenvolvimento de modelos mais leves e eficientes.

- Investigar o efeito do nível de processamento das imagens (L1C vs. L2A) sobre a detecção de nuvens, avaliando se a correção atmosférica (L2A) traz benefícios significativos para a segmentação.
- Comparar os resultados obtidos com métodos tradicionais e outros modelos da literatura, de modo a posicionar a abordagem proposta em relação ao estado da arte na detecção de nuvens.

1.4 Estrutura do Trabalho

Este trabalho está organizado em cinco capítulos, conforme descrito a seguir:

- **Capítulo 2** — Trabalhos Correlatos: discute as principais pesquisas e soluções relacionadas ao tema, destacando as lacunas e oportunidades identificadas na literatura e comparando este trabalho com os correlatos.
- **Capítulo 3** — Referencial Teórico: descreve os detalhes acerca do Sentine-2 e suas imagens, além dos conceitos por trás das redes neurais totalmente convolucionais, segmentação de imagens e as métricas utilizadas para avaliação dos modelos.
- **Capítulo 4** — Metodologia: apresenta a estratégia adotada para o treinamento dos modelos, incluindo detalhes sobre o conjunto de dados utilizado, os modelos de redes neurais, as etapas de pré-processamento, treinamento e avaliação.
- **Capítulo 5** — Resultados: expõe as análises quantitativas e qualitativas, comparando o desempenho dos modelos propostos com outras abordagens da literatura, bem como os cenários de utilização e limitações identificadas.
- **Capítulo 6** — Conclusão: relembra os objetivos e discute os principais achados, além de apontar limitações e perspectivas para trabalhos futuros.

2 TRABALHOS CORRELATOS

2.1 Considerações Iniciais

Este capítulo apresenta os trabalhos correlatos ao tema desta dissertação, destacando estudos e abordagens relevantes para a compreensão do problema investigado. Inicialmente, são descritos os trabalhos correlatos, evidenciando suas contribuições, pontos positivos e limitações. O objetivo é situar esta pesquisa no contexto científico atual, identificando avanços, lacunas e desafios que justifique sua realização. Por fim, é realizada uma comparação entre este estudo e as abordagens existentes, destacando os diferenciais e contribuições esperadas.

2.2 Descrição dos Trabalhos Correlatos

A detecção de nuvens em imagens de satélite tem sido amplamente estudada, com uma variedade de métodos propostos para melhorar a precisão e eficiência. Em (SKAKUN et al., 2022) os autores realizaram um estudo abrangente de intercomparação de algoritmos de detecção de nuvens para imagens do Sentinel-2 e Landsat 8, avaliando 10 algoritmos diferentes em cinco conjuntos de dados de referência. O estudo destacou que algoritmos como Fmask, FORCE, MAJA e s2cloudless, apresentaram um bom equilíbrio entre erros de omissão (Falso Negativo) e comissão (Falso Positivo). No entanto, o estudo também identificou limitações, como a dependência dos resultados em relação aos conjuntos de dados de referência utilizados e a dificuldade na detecção de nuvens finas/semi-transparentes.

Complementando essa análise, (SINGH; BISWAS; PAL, 2023) realizou uma comparação abrangente de métodos de detecção de nuvens em imagens do Sentinel-2, incluindo XGBoost, Random Forest (RF), Support Vector Machine (SVM) e uma FCNN baseada em Resnet. O estudo demonstrou que XGBoost e RF apresentaram desempenho comparável, com RF alcançando a maior precisão (94,2%) ao utilizar uma combinação de características espectrais, texturais e morfológicas. Além disso, XGBoost mostrou melhor generalização em diferentes condições ambientais, superando métodos tradicionais baseados em limiares, como Fmask e Sen2cor. Entretanto, o estudo também identificou limitações, como a dificuldade na detecção de nuvens altas com visibilidade total do solo e o alto custo computacional associado às FCNNs.

Buscando aprimorar a detecção de nuvem, em (CANDRA; PHINN; SCARTH, 2020) os autores propuseram uma melhoria no algoritmo *Multitemporal Cloud Masking* (MCM), originalmente desenvolvido para o Landsat 8, para detectar nuvens e sombras de nuvens em imagens do Sentinel-2. O algoritmo utiliza uma abordagem multitemporal, comparando uma imagem alvo com uma imagem de referência sem nuvens, e demonstrou alta precisão na

detecção de nuvens e sombras em diferentes ambientes, incluindo áreas subtropicais e tropicais. Os autores também compararam o MCM com o algoritmo Fmask, mostrando que o MCM tem um desempenho superior, especialmente na detecção de sombras de nuvens. Todavia, o algoritmo depende de uma imagem de referência sem nuvens, o que pode ser um desafio em áreas com alta cobertura de nuvens. Além disso, o estudo não aborda a detecção de nuvens em áreas com neve, o que pode ser uma limitação em regiões frias.

Seguindo para métodos que exploram a variação temporal, em (LIANG et al., 2024) os autores propuseram um novo método para a detecção de nuvens e sombras de nuvens em imagens do Sentinel-2, denominado *Time Series Maximum and Minimum mask method* (TSMM). A abordagem se baseia na variação temporal das bandas azul e infravermelho próximo para segmentar nuvens e sombras de nuvens. Além disso, um processo de convolução espacial é aplicado para reduzir falsos positivos e melhorar a precisão nas bordas das nuvens. Os resultados demonstraram um desempenho superior do TSMM em relação a métodos tradicionais, como Fmask, Sen2Cor e S2cloudless, alcançando 93% de acurácia geral e 85% de F1-score. A abordagem se destacou especialmente na detecção de nuvens finas e sombras, que são frequentemente subestimadas por outros métodos. Entretanto, a disponibilidade de séries temporais de qualidade pode ser um problema para este método.

Em (AI et al., 2024) os autores desenvolveram um método para a detecção de nuvens denominado RF-NDTeI-SNIC, que combina aprendizado de máquina, índices espectro-temporais e segmentação de imagem. A abordagem utiliza um modelo *Random Forest* (RF) treinado com dados do Landsat-8 para uma primeira classificação de nuvens, seguido pela aplicação do *Normalized Difference Temporal Index* (NDTeI), um índice baseado na variação da banda azul ao longo do tempo, permitindo diferenciar nuvens de superfícies brilhantes, como areia, neve e áreas urbanas. Para refinar a segmentação e reduzir ruídos na classificação, foi utilizado o algoritmo de segmentação *Simple Non-Iterative Clustering* (SNIC), que agrupa pixels espectralmente homogêneos. Os experimentos conduzidos com nove cenas do Sentinel-2 demonstraram uma acurácia geral de 95,01%, superando algoritmos amplamente utilizados, como s2cloudless, Fmask e Tmask, especialmente na identificação de nuvens sobre alvos brilhantes. As principais limitações observadas foram: a dependência da disponibilidade de imagens livres de nuvens para a construção da referência temporal, o que pode ser um desafio em regiões persistentemente nubladas; mudanças no solo ao longo do tempo podem impactar a precisão do NDTeI, levando a classificações incorretas; e o modelo de RF pode exigir re-treinamento para diferentes regiões e condições atmosféricas.

Buscando superar o problema do desbalanceamento de classes e detecção de bordas de nuvens, em (CHEN; PAN; ZHANG, 2023) os autores desenvolveram um método para a detecção de nuvens combinando a extração de múltiplas características (espectrais, texturais e exponenciais) com um prior de segmentação para refinar a detecção de bordas. A abordagem baseia-se no treinamento de um modelo LightGBM ponderado, que atribui pesos adaptativos às

amostras para lidar com o problema de desbalanceamento dos dados. Além disso, o resultado da classificação inicial é refinado utilizando *Conditional Random Fields* (CRF) combinado com um algoritmo de segmentação, melhorando a separação entre nuvens e outros elementos da superfície terrestre. Os testes foram realizados em imagens Sentinel-2 com seis tipos de superfícies subjacentes (zonas urbanas, corpos d'água, áreas agrícolas, terrenos áridos, neve/gelo e florestas), obtendo acurácia superior a 93% para todos os cenários analisados. O método proposto foi comparado com outras técnicas tradicionais e baseadas em aprendizado profundo, como Fmask, Sen2Cor, CloudSegNet e Swin-Transformer, demonstrando maior estabilidade na detecção de nuvens em diferentes superfícies e menor necessidade de grandes conjuntos de treinamento. A abordagem apresenta bom desempenho mesmo com o limitado número de amostras, tornando-se uma alternativa promissora para aplicações onde o volume de dados rotulados é limitado. As principais limitações do trabalho estão ligadas à dependência da qualidade da segmentação inicial e da resolução espacial de 60 metros das imagens testadas, o que pode limitar sua aplicabilidade em imagens com resolução mais alta.

Avançando para métodos baseados em aprendizado profundo, em (LI et al., 2020) os autores propuseram um método de detecção de nuvens em imagens de sensoriamento remoto de alta resolução utilizando aprendizado profundo com supervisão fraca. O método é baseado em rótulos em nível de bloco, ou seja, a imagem é dividida em blocos e cada bloco é classificado se possui ou não pixels com nuvem, o que reduz significativamente o esforço de anotação em comparação com métodos que exigem anotações em nível de pixel. Os autores introduziram uma nova operação de pooling, chamada *Global Convolutional Pooling* (GCP), que melhora a capacidade da rede de capturar variações espaciais. Além disso, propuseram a estratégia de *Local Pooling Pruning* (LPP), que aumenta a resolução espacial do *Cloud Activation Map* (CAM), permitindo a detecção precisa de pequenas nuvens e bordas. O método alcançou uma precisão geral de 96,66% em um conjunto de dados global do satélite GaoFen-1 com resolução espacial de 8 metros. Ainda assim, o método não detecta sombras de nuvens, o que pode ser uma limitação em aplicações que exigem a remoção de sombras. Além disso, a complexidade computacional do método pode ser um desafio em aplicações que exigem processamento em tempo real.

Buscando uma abordagem mais inovadora, em (CHEN et al., 2020) os autores propuseram um método baseado em redes neurais convolucionais 3D multiescala para detecção de nuvens e sombras de nuvens em imagens multiespectrais de alta resolução. O método utiliza uma camada de convolução 3D para explorar tanto as características espaciais quanto espectrais das imagens, o que permite uma detecção mais precisa de nuvens e sombras de nuvens. Além disso, o método incorpora um módulo multiescala que processa a imagem em diferentes resoluções, permitindo a captura de informações contextuais em várias escalas. Os resultados experimentais demonstraram que o método alcançou alta precisão e recall, superando outras abordagens como MFC, Deeplab e PSPNet. Uma das principais vantagens do método é sua capacidade de distinguir nuvens, de neve e gelo, mesmo em imagens que não possuem bandas de infravermelho de ondas curtas, o que é particularmente útil em cenários onde a neve e o gelo são fontes de ruído. A

principal limitação está na complexidade computacional, devido as adições de convoluções 3D e módulos multiescala.

Em (HU; ZHANG; XIA, 2021) os autores propuseram uma variação da UNet, chamada CDUNet, para detecção de nuvens em imagens de sensoriamento remoto. A CDUNet introduz módulos inovadores, como o *High-frequency Feature Extractor* (HFE) e o *Multiscale Convolution* (MSC), para melhorar a extração de detalhes e a precisão nas bordas das nuvens. Além disso, o mecanismo de *Spatial Prior Self-Attention* (SPSA) foi utilizado para reconstruir informações espaciais, especialmente útil em cenários complexos com nuvens finas e fragmentadas. Os experimentos realizados em dois conjuntos de dados demonstraram que a CDUNet supera outros métodos de segmentação semântica, como UNet, PSPNet e HRNet, em termos de precisão de pixel (PA) e *Mean Intersection Over Union* (MIoU). A principal limitação da CDUNet é o custo computacional devido à introdução de múltiplos módulos (HFE, MSC e SPSA), o que pode limitar sua aplicação em cenários com restrições de hardware.

Em um trabalho que destaca a utilidade de máscaras de nuvens de alta resolução, (DOMNICH et al., 2021) propôs o KappaMask, um processador de máscaras de nuvens baseado em IA para imagens do Sentinel-2, utilizando a arquitetura U-Net. O trabalho destaca-se pela geração de máscaras de nuvens com resolução de 10 metros, o que permite uma delimitação mais fina das nuvens e sombras, aumentando a utilidade dos dados para aplicações de alto nível, como o monitoramento agrícola. O KappaMask foi treinado com um conjunto de dados anotado manualmente utilizando uma metodologia de aprendizado ativo, o que permitiu ao modelo lidar com cenários complexos de detecção de nuvens. Em comparação com métodos baseados em regras (Sen2Cor, Fmask, MAJA) e métodos baseados em aprendizado de máquina (S2cloudless), o KappaMask obteve um coeficiente de Dice superior, especialmente na detecção de sombras de nuvens e nuvens semi-transparentes. No entanto, o trabalho apresenta algumas limitações, como o escopo geográfico limitado ao Norte da Europa e a ausência de condições de inverno no conjunto de dados de treinamento. Além disso, o modelo ainda enfrenta desafios na detecção de nuvens semi-transparentes, o que pode levar a falsos negativos em algumas situações.

Finalmente, em (LÓPEZ-PUIGDOLLERS et al., 2021) os autores realizaram uma avaliação sistemática de modelos de deep learning para detecção de nuvens em imagens do Landsat-8 e Sentinel-2. Além disso, os autores utilizaram uma versão modificada da U-Net e compararam seu desempenho com métodos operacionais baseados em limiar, como o FMask e o Sen2Cor. O estudo demonstrou que modelos de deep learning têm um desempenho superior quando treinados e testados no mesmo conjunto de dados, mas o desempenho cai significativamente em validações *inter-dataset* (quando o modelo é treinado e testado em datasets diferentes) e *cross-sensor* (quando o modelo é treinado e testado com imagens de sensores diferentes). Apesar disso, a transferência de aprendizado entre Landsat-8 e Sentinel-2 mostrou-se viável, embora não tenha superado métodos específicos como o s2cloudless.

Em uma abordagem que busca reduzir a dependência de grandes conjuntos de dados

rotulados, (CHEN et al., 2023) propôs uma rede neural profunda chamada *Spectral-Spatial-Context Attention Network* (SSCA-net) para detecção automática de nuvens em imagens de satélite. A abordagem combina índices espectrais (Cloud Index) com módulos de atenção espectral-espacial-contextual, permitindo a detecção de nuvens sem a necessidade de coleta manual de amostras rotuladas. Os autores utilizaram um método de duplo limiar para gerar máscaras de nuvens automaticamente, o que é uma vantagem significativa em relação a métodos que dependem de grandes conjuntos de dados rotulados. A SSCA-net demonstrou alta precisão em cenas de alta luminosidade, superando métodos tradicionais como Fmask em termos de precisão geral e capacidade de distinguir nuvens de objetos brilhantes, como neve e rochas. As limitações do trabalho se encontram no curto conjunto de dados que pode ter gerado sobreajuste, além da arquitetura SSCA-net ter um alto custo computacional.

Por fim, em (LI et al., 2022) os autores propuseram um método leve baseado em aprendizado profundo, chamado CD-FM3SF, para detecção de nuvens em imagens do Sentinel-2A. O método utiliza uma arquitetura de codificador-decodificador com blocos residuais e opera em múltiplas escalas espaciais e espectrais, combinando características de todas as bandas do Sentinel-2A com as resoluções originais (10, 20 e 60 metros). O CD-FM3SF demonstrou ser eficiente em termos de parâmetros e computação, superando métodos tradicionais e outros métodos baseados em deep learning em termos de precisão e velocidade. Além disso, os autores introduziram um novo conjunto de dados, WHUS2-CD, com máscaras de nuvens rotuladas manualmente em alta resolução, o que permitiu uma avaliação mais precisa do método. No entanto, o método apresenta dificuldades na detecção de nuvens muito finas e não aborda a detecção de sombras de nuvens, o que pode ser uma limitação para aplicações práticas.

Os estudos mencionados demonstram que, nos últimos anos, diversas soluções têm sido propostas para a detecção de nuvens, embora ainda existam desafios em aberto. Os algoritmos baseados em regras espectrais são mais rápidos, mas enfrentam dificuldades na identificação de nuvens finas. Por outro lado, abordagens baseadas em deep learning tendem a superar essas limitações, porém apresentam um custo computacional mais elevado e dependem de grandes volumes de dados rotulados. A Tabela 1 resume as principais contribuições, vantagens e limitações dos trabalhos citados.

2.3 Comparativo entre este Trabalho e os Correlatos

Os estudos mostrados anteriormente sobre detecção e segmentação de nuvens em imagens de satélite adotaram diferentes abordagens, desde métodos baseados em limiares espectrais até técnicas mais avançadas de aprendizado profundo. Em particular, as redes neurais totalmente convolucionais (FCNNs) demonstraram um desempenho superior na segmentação de nuvens em comparação com métodos tradicionais. No entanto, seu alto custo computacional e a dependência de grandes conjuntos de dados rotulados manualmente ainda representam desafios significativos.

Tabela 1 – Resumo das Contribuições dos Trabalhos Correlatos.

Referência	Método	Contribuições	Vantagens	Limitações	Satélite
(SKAKUN et al., 2022)	Intercomparação de 10 algoritmos	Avaliação detalhada de métodos tradicionais	Bom equilíbrio entre erros de omissão e comissão	Dificuldade na detecção de nuvens finas	Sentinel-2, Landsat-8
(SINGH; BISWAS; PAL, 2023)	Comparação de XGBoost, RF, SVM e FCNN	RF com maior precisão (94,2%)	Generalização em diferentes condições ambientais	Alto custo computacional das FCNNs	Sentinel-2
(CANDRA; PHINN; SCARTH, 2020)	Melhoria do MCM	Deteção de sombras de nuvens	Alta precisão em áreas subtropicais e tropicais	Dependência de imagem de referência sem nuvens	Sentinel-2
(LIANG et al., 2024)	TSM	Deteção de nuvens finas e sombras	93% de acurácia geral	Dependência de séries temporais de qualidade	Sentinel-2
(AI et al., 2024)	RF-NDTeI-SNIC	Combinação de RF e NDTeI	95,01% de acurácia geral	Dependência de imagens livres de nuvens	Landsat-8, Sentinel-2
(CHEN; PAN; ZHANG, 2023)	LightGBM + CRF	Extração de múltiplas características	Estabilidade em diferentes superfícies	Dependência da qualidade da segmentação inicial	Sentinel-2
(LI et al., 2020)	Aprendizado profundo com supervisão fraca	GCP e LPP para deteção precisa	96,66% de precisão geral	Não detecta sombras de nuvens	GaoFen-1
(CHEN et al., 2020)	Redes neurais convolucionais 3D	Deteção de nuvens e sombras	Distinção entre nuvens e neve/gelo	Complexidade computacional	Gaofen-1, ZY-3
(HU; ZHANG; XIA, 2021)	CDUNet (variação da UNet)	Módulos HFE e MSC para precisão nas bordas	Superior a UNet, PSPNet e HRNet	Alto custo computacional	Landsat-8
(DOMNICH et al., 2021)	KappaMask (U-Net)	Máscaras de 10 metros	Deteção de sombras e nuvens semi-transparentes	Escopo geográfico limitado	Sentinel-2
(LÓPEZ-PUIGDOLLERS et al., 2021)	Avaliação de modelos de deep learning	Transferência de aprendizado entre Landsat-8 e Sentinel-2	Desempenho superior em datasets semelhantes	Desafios em validações inter-dataset	Sentinel-2, Landsat-8
(CHEN et al., 2023)	SSCA-net	Deteção automática sem amostras rotuladas	Alta precisão em cenas de alta luminosidade	Curto conjunto de dados e alto custo computacional	Não se aplica
(LI et al., 2022)	CD-FM3SF	Operação em múltiplas escalas	Eficiente em precisão e velocidade	Dificuldade na deteção de nuvens muito finas	Sentinel-2A

Logo, para garantir um bom desempenho, é essencial que as imagens de treinamento cubram a maior variedade possível de características atmosféricas e regionais, além da qualidade da segmentação de referência impactar diretamente no desempenho dos modelos treinados.

Apesar dos avanços alcançados com FCNNs na segmentação de nuvens, os correlatos mostram que ainda existem questões pouco exploradas quando aplicadas a imagens do Sentinel-2, tais como: Qual a influência do nível de processamento (L1C e L2A) na deteção de nuvens? É possível reduzir o número de bandas utilizadas e ainda assim obter um desempenho competitivo?

Neste trabalho, propõe-se o uso de FCNNs, especificamente a arquitetura UNet com os encoders EfficientNet-B1 e MobileNet-V2, para segmentação de nuvens em imagens Sentinel-2. Diferentemente de estudos anteriores, este estudo investiga o impacto da quantidade de bandas (todas as bandas e RGB+NIR) e a influência dos diferentes níveis de processamento (L1C e L2A) na qualidade da segmentação.

Outro diferencial desta pesquisa está na utilização do conjunto de dados CloudSen12, que contém 10.000 imagens de diferentes regiões do mundo, abrangendo uma ampla variedade

de condições atmosféricas. Enquanto alguns trabalhos correlatos utilizam conjuntos de dados menores e menos diversificados, a escolha do CloudSen12 possibilita uma avaliação mais abrangente da capacidade de generalização dos modelos treinados.

2.4 Síntese do Capítulo

A revisão dos trabalhos correlatos permitiu identificar diversas abordagens utilizadas na literatura para o problema em questão, evidenciando avanços e limitações nos estudos analisados. Observou-se que, apesar dos progressos alcançados, ainda existem desafios a serem superados, especialmente no que se refere a mascarar nuvens finas e semitransparentes, além do custo computacional para modelos baseados em aprendizado profundo. Dessa forma, a presente dissertação busca contribuir para essa área ao propor uma abordagem que busque melhorar a performance de detecção de nuvens e redução do custo computacional dos modelos através da utilização de poucas bandas espectrais na entrada.

3 REFERENCIAL TEÓRICO

3.1 Considerações Iniciais

O presente capítulo tem como objetivo apresentar o referencial teórico que fundamenta este trabalho, oferecendo uma visão abrangente sobre os principais conceitos e tecnologias envolvidos. Inicialmente, são descritas as características do satélite Sentinel-2, suas bandas espectrais e os produtos gerados, destacando sua relevância para o monitoramento ambiental. Em seguida, é apresentado o *dataset* CloudSen12, um conjunto de dados que combina imagens do Sentinel-2 com máscaras de nuvens e metadados essenciais para o treinamento e avaliação de modelos de segmentação.

Na sequência, são discutidas as Redes Neurais Totalmente Convolucionais (*Fully Convolutional Neural Networks* - FCNN), abordando sua origem, funcionamento e aplicabilidade em problemas de segmentação de imagens. Dentre esses modelos, a U-Net é destacada por sua estrutura eficiente para segmentação de imagens com alta precisão, sendo a arquitetura utilizada neste estudo. Por fim, são descritas as métricas empregadas para quantificar o desempenho do modelo de detecção de nuvens, fornecendo uma base sólida para a análise dos resultados obtidos.

3.2 Sobre o Sentinel-2

O Sentinel-2 é parte do Programa Copernicus, desenvolvido pela Agência Espacial Europeia (*Europe Space Agency* - ESA) e pela União Europeia (UE), com o objetivo de monitorar a superfície terrestre de forma eficaz. Os satélites Sentinel-2 são a segunda constelação das missões Sentinel da ESA e estão equipados com *scanners multiespectrais*. O principal objetivo da missão Sentinel-2 é fornecer dados de satélite de alta resolução para o monitoramento do uso/cobertura do solo, mudanças climáticas e desastres, além de complementar outras missões de satélite, como o Landsat (PHIRI et al., 2020).

3.2.1 Características do Sentinel-2

A missão Sentinel-2 consiste em dois satélites idênticos, Sentinel-2A e Sentinel-2B. Ambos operam em simultâneo, faseados a 180° entre si, numa órbita sincronizada com o Sol a uma altitude média de 786 km. O Sentinel-2 transporta uma carga útil de instrumento óptico que coleta amostras de 13 bandas espectrais, cobrindo uma larga faixa de espectro. A largura da faixa orbital é de 290 km. A frequência de revisita de cada satélite é de 10 dias e a revisita combinada da constelação é de 5 dias (EUROPEAN SPACE AGENCY, 2025).

A Tabela 2 mostra as características espectrais das bandas do Sentinel-2A e 2B, que cobrem do espectro visível até o infravermelho de ondas curtas (*Short-Wave Infrared* - SWIR). As bandas possuem diferentes resoluções espaciais: quatro bandas a 10 metros, seis bandas a 20 metros e três bandas a 60 metros. Destaque para as bandas 2 (Azul), 3 (Verde) e 4 (Vermelho) que compõem o espectro visível e são usadas para formar imagens com cores reais.

As bandas de 10 metros incluem as bandas Azul (2), Verde (3), Vermelha (4) e Infravermelho Próximo (*Near Infrared* - NIR)(8). Essas bandas são frequentemente usadas para análise detalhada da vegetação e classificação da cobertura do solo (WANG et al., 2016; LIU et al., 2023). As Bandas de 20 metros incluem as bandas *Red Edge* (6 e 7), NIR (8A) e SWIR, e são particularmente úteis para avaliar a saúde da vegetação e o conteúdo de clorofila (DELEGIDO et al., 2011; PERSSON; LINDBERG; REESE, 2018; REVILL et al., 2019). As bandas de 60 metros incluem bandas para correção atmosférica e triagem de nuvens, como as bandas de aerossol costeiro (1), vapor de água (9) e cirros (10) (LIU et al., 2023; SALGUEIRO; MARCELLO; VILAPLANA, 2021).

Tabela 2 – Bandas Espectrais dos satélites Sentinel-2A e Sentinel-2B.

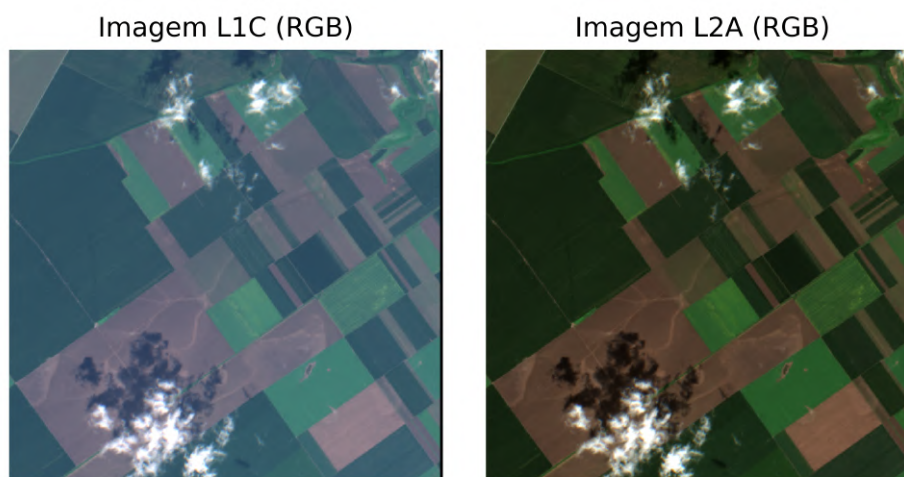
Resolução Espacial (m)	Banda	Comprimento de Onda Central (nm) S2A	Largura de Banda (nm) S2A	Comprimento de Onda Central (nm) S2B	Largura de Banda (nm) S2B
10	Banda 2 Azul	492.4	66	492.1	66
10	Banda 3 Verde	559.8	36	559	36
10	Banda 4 Vermelho	664.6	31	664.9	31
10	Banda 8 NIR	832.8	106	832.9	106
20	Banda 6 Red Edge	740.5	15	739.1	15
20	Banda 7 Red Edge	782.8	20	779.7	20
20	Banda 8A NIR Estreito	864.7	21	864	22
20	Banda 11 SWIR	1613.7	91	1610.4	94
20	Banda 12 SWIR	2202.4	175	2185.7	185
60	Banda 1 Aerossol Costeiro	442.7	21	442.2	21
60	Banda 9 Vapor D'água	945.1	20	943.2	21
60	Banda 10 SWIR Cirrus	1373.5	31	1376.9	30

Fonte: Adaptado de (PHIRI et al., 2020)

3.2.2 Produtos do Sentinel-2

O Sentinel-2 fornece uma variedade de produtos de dados que abrangem diferentes etapas do processamento, desde a captura bruta a bordo do satélite até os produtos prontos para análise. Entre esses, os mais amplamente utilizados e disponibilizados ao público são os produtos *Level-1C* (L1C) e *Level-2A* (L2A), que oferecem dados de observação da Terra com alta qualidade e resolução. O produto L1C fornece refletância no topo da atmosfera, com correção geométrica aplicada, enquanto o L2A inclui correção atmosférica, oferecendo refletância na superfície terrestre, além de camadas auxiliares. Esses produtos são amplamente empregados em aplicações como monitoramento de uso e cobertura da terra, agricultura de precisão e gestão ambiental. A Figura 1 apresenta um exemplo comparativo entre uma imagem sem correções atmosféricas (L1C) e outra com as correções aplicadas (L2A).

Figura 1 – Comparação entre imagem L1C vs L2A.



Fonte: O Autor

3.2.2.1 Produtos L1C

Os produtos de dados L1C do Sentinel-2 são imagens de reflectância no topo da atmosfera (*Top-Of-Atmosphere* - TOA) em geometria cartográfica. Estes produtos são distribuídos pela ESA e são a base para a geração de produtos de nível superior, como o L2A (SOLA et al., 2018). O L1C é essencialmente a imagem bruta capturada pelo sensor *Multi-Spectral Instrument* (MSI) e é processada para incluir correções geométricas (as imagens são alinhadas corretamente à superfície da terra) e radiométricas, mas ainda contém efeitos atmosféricos como espalhamento e absorção de luz que precisam ser corrigidos para aplicações mais precisas (USS et al., 2017). Os principais usos das imagens L1C são: monitoramento geral da terra e casos onde a correção atmosférica não é essencial.

3.2.2.2 Produtos L2A

Os produtos L2A são gerados a partir dos produtos L1C através de um processo de correção atmosférica, que remove os efeitos da atmosfera para fornecer dados de reflectância na superfície (Bottom-Of-Atmosphere, BOA) (MAIN-KNORN et al., 2017; VUOLO et al., 2016). O processador Sen2Cor é amplamente utilizado para essa conversão, corrigindo os produtos L1C para entregar produtos L2A que incluem ainda mapas de espessura óptica de aerossóis (*Aerosol Optical Thickness* - AOT), vapor d'água (*Water Vapor* - WV) e classificação de cena (*Scene Classification Layer* - SCL) (MAIN-KNORN et al., 2017; PFLUG et al., 2020). Esses produtos são fundamentais para aplicações que exigem alta precisão, como monitoramento de vegetação, análise de mudanças de uso do solo e estudos ambientais (VUOLO et al., 2016; SOLA et al., 2018). A Tabela 3 resume as principais diferenças entre os produtos L1C e L2A.

Tabela 3 – Comparação entre os produtos Sentinel-2 L1C e L2A

Característica	L1C	L2A
Correção Atmosférica	Não	Sim
Tipo de Reflectância	TOA (Top of Atmosphere)	BOA (Bottom of Atmosphere)
Classificação de cena	Não	Sim
Aplicações	Pré-processamento, análise bruta	Análises ambientais mais precisas
Tamanho dos arquivos	Menor	Maior (devido às camadas extras)

3.3 Dataset CloudSEN12

A detecção precisa de nuvens e suas sombras em imagens de satélite é um desafio recorrente no sensoriamento remoto, especialmente devido ao impacto significativo que essas contaminações têm na qualidade e utilidade dos dados. Nesse contexto, o dataset CloudSEN12 foi desenvolvido como uma solução abrangente para apoiar pesquisas em detecção semântica de nuvens e sombras de nuvens, com foco em imagens do Sentinel-2. O CloudSEN12 compreende 49.400 *patches* de imagem, distribuídos globalmente e organizados em três categorias principais de anotação: anotações de alta qualidade (10.000 *patches*), realizadas no nível de pixel; anotações rascunho (10.000 *patches*), que oferecem uma alternativa menos detalhada e mais econômica; e *patches* sem anotações (29.400 *patches*), que podem ser utilizados em abordagens de aprendizado semi-supervisionado. As imagens incluem dados multiespectrais do Sentinel-2 (níveis L1C e L2A), dados do Sentinel-1 e informações auxiliares, como elevação e mapas de cobertura do solo. Esta seção é baseada em (AYBAR et al., 2022; AYBAR et al., 2024).

3.3.1 Processo de criação do Dataset

O CloudSEN12 foi construído a partir de uma coleta cuidadosa de dados disponíveis publicamente, utilizando a plataforma *Google Earth Engine*. Foram selecionadas 20.000 regiões de interesse (ROIs) globais, posteriormente reduzidas a 9.880 após a aplicação de critérios

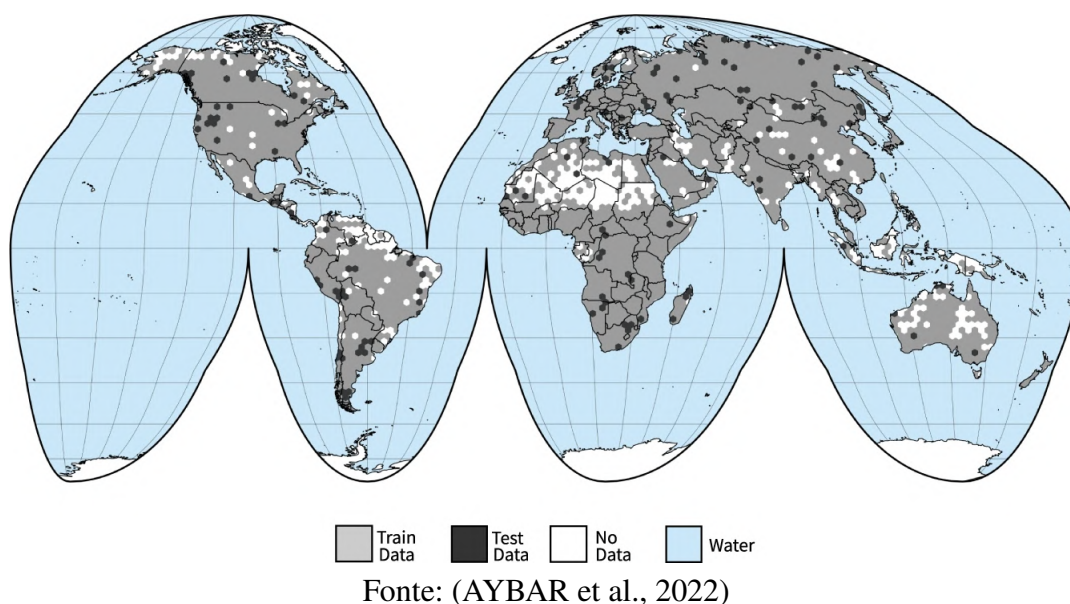
rigorosos, como a exclusão de áreas com saturação ou ausência de dados e a garantia de diversidade geográfica e temporal.

O processo de anotação contou com a participação de especialistas em detecção de nuvens, que utilizaram ferramentas de aprendizado ativo para otimizar o tempo e a qualidade das anotações. Um rigoroso protocolo de controle de qualidade foi implementado, garantindo que as anotações apresentassem precisão comparável ao nível humano (acurácia de 95,7% para todas as classes semânticas).

3.3.2 Divisão em Treinamento, Validação e Teste

Para evitar correlações espaciais que poderiam comprometer a generalização dos modelos, foi adotada uma estratégia de divisão espacialmente estratificada. O globo terrestre foi dividido em hexágonos de 50.000 km², e as regiões foram separadas em 90% para treinamento e 10% para teste, garantindo que os patches de imagem de treinamento e teste fossem espacialmente independentes. Dentro do conjunto de treinamento, 10% das imagens foram alocadas para validação, assegurando uma avaliação robusta do desempenho dos modelos. A Figura 2 mostra a distribuição das localizações de onde foram retiradas as imagens.

Figura 2 – Localização das imagens de treino (cinza) e teste (preto).



3.3.3 Classes Semânticas e Cobertura de Nuvens

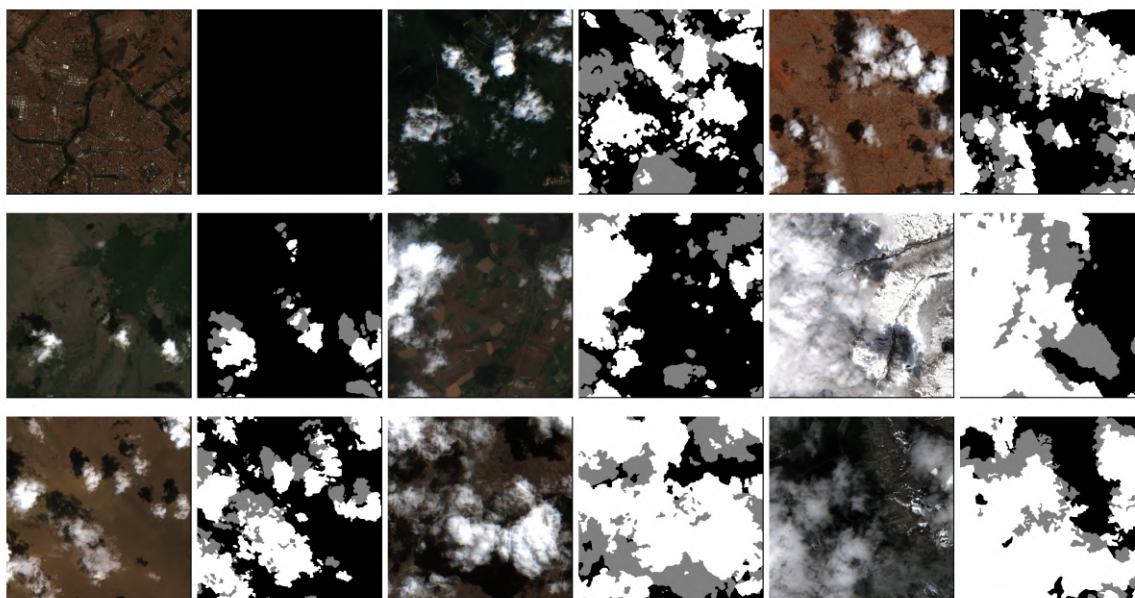
O dataset foi projetado para abordar diferentes cenários de cobertura de nuvens, onde cada região de interesse contém 5 imagens categorizadas de acordo com o percentual da cobertura de nuvem, sendo: céu limpo (0%), quase limpo (0–25%), pouco nublado (25–45%), moderadamente nublado (45–65%) e muito nublado (>65%). As classes semânticas incluem nuvens espessas, nuvens finas, sombras de nuvens e pixels claros, permitindo que modelos de aprendizado de

máquina tratem diferentes tipos de contaminação de forma diferenciada. A tabela 4 mostra as classes semânticas presentes no CloudSen12, enquanto a Figura 3 mostra um mosaico com imagens e suas respectivas máscaras de nuvem ao lado, exemplificando a variabilidade de terrenos e coberturas de nuvens presente no dataset.

Tabela 4 – Classes semânticas das máscaras do CloudSen12.

Código	Classe	Descrição
0	Claro	Pixels sem contaminação de nuvens e sombras de nuvens.
1	Nuvem Espessa	Nuvens opacas que bloqueiam toda a refletância da superfície da Terra.
2	Nuvem Fina	Nuvem semitransparente que altera o sinal espectral da superfície, mas ainda permite reconhecer o fundo.
3	Sombra de Nuvem	Pixels escuros emitidos por uma nuvem espessa ou fina.

Figura 3 – Mosaico de imagens e suas respectivas máscaras do CloudSEN12.



Fonte: O Autor

3.3.4 Relevância para Treinamento de Redes Neurais Totalmente Convolucionais

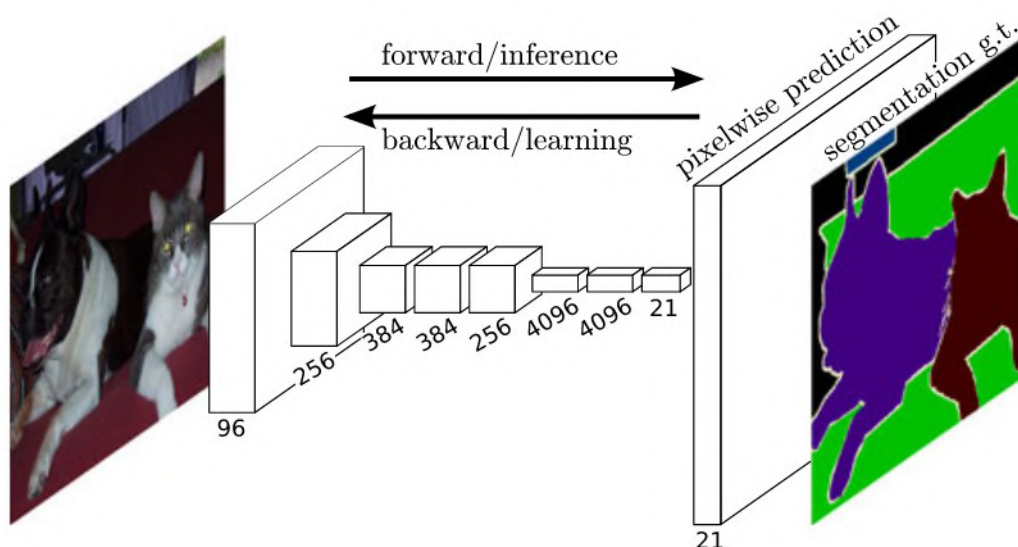
O CloudSEN12 é especialmente relevante para o treinamento de modelos de FCNN, devido à sua alta qualidade de anotação, diversidade de dados e organização voltada para aprendizado supervisionado. Datasets como o CloudSEN12 são fundamentais para o desenvolvimento de algoritmos robustos e generalizáveis, particularmente em tarefas de segmentação semântica, onde

a detecção precisa de nuvens e sombras é crítica para aplicações em monitoramento ambiental, agricultura e gestão de desastres.

3.4 Redes Neurais Totalmente Convolucionais

As FCNNs são uma variação das CNNs que eliminam as camadas totalmente conectadas, utilizando apenas camadas convolucionais para processar dados. O trabalho de (LONG; SHELHAMER; DARRELL, 2014) foi o primeiro a propor essa arquitetura, adaptando arquiteturas populares de CNN (AlexNet, VGG e GoogLeNet) para segmentação, transformando as camadas totalmente conectadas em convolucionais. Isso reduz o número de parâmetros e o risco de *overfitting*, além de melhorar a eficiência computacional (ZHU; ZU, 2022). As FCNNs são especialmente eficazes em tarefas de segmentação semântica, onde a saída precisa ter a mesma resolução espacial que a entrada. A Figura 4 mostra a arquitetura da primeira FCNN proposta, a ausência de camadas totalmente conectadas e a saída com a mesma dimensão da entrada são as principais novidades.

Figura 4 – Arquitetura básica da primeira FCNN proposta na literatura.



Fonte: (LONG; SHELHAMER; DARRELL, 2014)

3.4.1 Estrutura

A estrutura de uma FCNN consiste em múltiplas camadas convolucionais que processam a entrada de forma hierárquica, extraindo características em diferentes níveis de abstração. Ao contrário das CNNs tradicionais, as FCNNs não possuem camadas totalmente conectadas, o que permite que a rede aceite entradas de tamanho arbitrário e produza saídas correspondentes (LONG; SHELHAMER; DARRELL, 2014). A arquitetura pode incluir camadas de *pooling* para redução de dimensionalidade e camadas de ativação não lineares para introduzir complexidade

no modelo (DERRY; KRZYWINSKI; ALTMAN, 2023). Os principais componentes de uma Rede Neural Totalmente Convolutiva são apresentados a seguir:

- **Camadas Convolutivas:** FCNNs consistem inteiramente em camadas convolutivas, que são responsáveis pela extração de características. Essas camadas aplicam filtros aos dados de entrada para detectar padrões e características (ZHU; ZU, 2022; LONG; SHELHAMER; DARRELL, 2014).
- **Camadas de Pooling:** Essas camadas são usadas para reduzir as dimensões espaciais dos mapas de recursos, o que ajuda a gerenciar a carga computacional e controlar o overfitting. Os tipos comuns incluem *max pooling* e *mean pooling* (DERRY; KRZYWINSKI; ALTMAN, 2023).
- **Funções de ativação:** Funções de ativação não lineares, como *ReLU*, são aplicadas após operações convolutivas para introduzir não linearidade no modelo, o que é crucial para aprender padrões complexos (KUO, 2016).
- **Skip Connections:** FCNNs frequentemente incorporam *skip connections*, que combinam informações semânticas de camadas mais profundas com informações espaciais de camadas mais rasas. Isso ajuda a produzir saídas mais precisas e detalhadas, especialmente em tarefas de segmentação (LONG; SHELHAMER; DARRELL, 2014).
- **Funções de perda:** FCNNs normalmente usam funções de perda como Softmax Loss para tarefas de classificação. No entanto, funções de perda alternativas, como POD Loss, podem ser usadas para melhorar o desempenho e a robustez (ZHU; ZU, 2022).
- **Treinamento de ponta a ponta:** FCNNs são projetados para serem treinados de ponta a ponta, o que significa que eles recebem entradas de tamanho arbitrário e produzem saídas de tamanho correspondente, o que é particularmente útil para tarefas que exigem previsões espacialmente densas (LONG; SHELHAMER; DARRELL, 2014).

3.4.2 Funcionamento

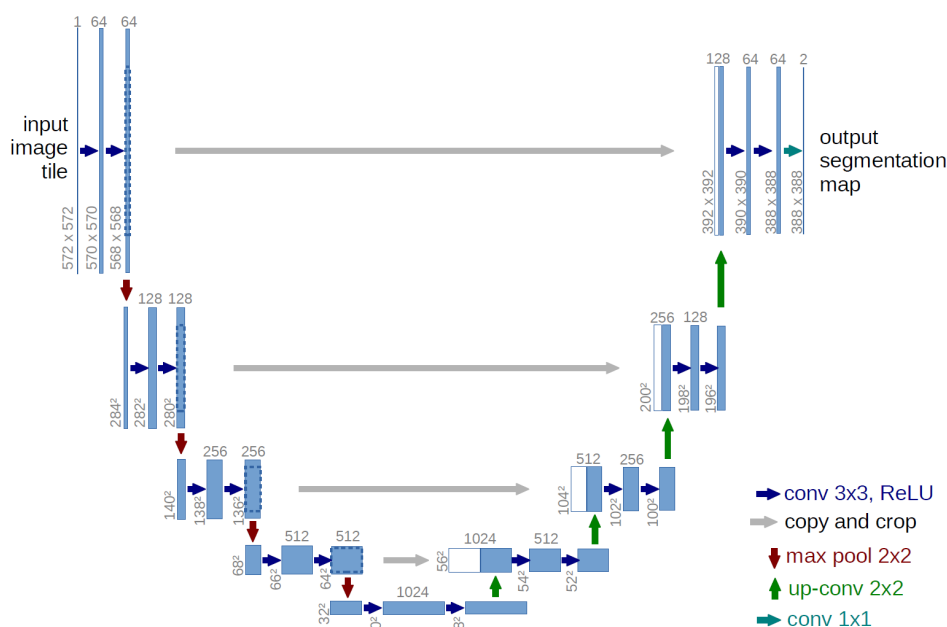
As FCNNs funcionam aplicando convoluções em toda a extensão da entrada, preservando a informação espacial. Isso é particularmente útil em tarefas que requerem previsões densas, como a segmentação de imagens, onde cada pixel da imagem de entrada é classificado (LONG; SHELHAMER; DARRELL, 2014). A eliminação das camadas totalmente conectadas não só reduz o número de parâmetros, mas também permite que a rede aprenda representações mais robustas e generalizáveis (LONG; SHELHAMER; DARRELL, 2014; ZHU; ZU, 2022).

3.4.3 Modelo UNet

Em (RONNEBERGER; FISCHER; BROX, 2015) é apresentada a arquitetura de rede neural totalmente convolucional chamada U-Net. Projetada inicialmente para segmentação de imagens biomédicas unidimensionais, a U-Net se destacou por ter excelentes resultados, mesmo com poucos dados anotados. A Figura 5 mostra a arquitetura simétrica da U-Net, que é composta por duas partes principais:

- **Caminho de Contração (Encoder):** Consiste em várias camadas de convolução 3x3 seguidas por uma função de ativação ReLU e uma operação de *max pooling* 2x2 para reduzir a resolução espacial. A cada etapa, o número de canais de características é dobrado, permitindo que a rede capture características cada vez mais complexas. O objetivo é extrair informações contextuais da imagem.
- **Caminho de Expansão (Decoder):** Utiliza *upsampling* e convolução transposta (*up-conv*) para aumentar a resolução espacial para reduzir o número de canais. As características de alta resolução do caminho de contração são concatenadas com as características do caminho de expansão, permitindo uma localização precisa. No final, uma convolução 1x1 é usada para mapear os vetores de características ao número de classes desejado.

Figura 5 – Arquitetura da primeira U-Net proposta na literatura.



Fonte: (RONNEBERGER; FISCHER; BROX, 2015)

Ao longo dos anos, diversas variantes e adaptações da U-Net foram desenvolvidas para otimizar seu desempenho e adequá-la a diferentes aplicações. Entre elas, destacam-se a 3D U-Net, a Attention U-Net, a Residual U-Net, a Recurrent U-Net, a Dense U-Net e a U-Net++,

cada uma introduzindo aprimoramentos específicos, como melhor captura de dependências espaciais, maior estabilidade no treinamento e refinamento na segmentação (SIDDIQUE et al., 2021). Com essas evoluções, a arquitetura U-Net passou a ser amplamente utilizada em diversas áreas, incluindo a classificação de uso e cobertura do solo, onde tem sido aplicada na análise de imagens multiespectrais captadas por satélites em órbita terrestre (SOLÓRZANO et al., 2021).

Além disso, o uso da U-Net foi facilitado pelo desenvolvimento de *frameworks* em Python, como o *Segmentation Models PyTorch* (SMP) (IAKUBOVSKII, 2019), que oferece uma implementação flexível e altamente personalizável. Esses frameworks permitem ajustes detalhados tanto nos hiperparâmetros quanto na estrutura da rede, possibilitando a escolha de diferentes encoders, variação na profundidade da arquitetura e outras configurações que impactam diretamente o desempenho da segmentação. Dessa forma, a combinação entre avanços na arquitetura da U-Net e ferramentas modernas de implementação tem impulsionado seu uso em um amplo espectro de aplicações científicas e industriais.

3.4.4 Encoders MobileNet-v2 e EfficientNet-B1

Em tarefas de visão computacional baseadas em redes neurais convolucionais, como classificação, detecção ou segmentação de imagens, os encoders desempenham o papel fundamental de extrair características relevantes das imagens de entrada. Neste trabalho são testados dois encoders: MobileNet-v2 e EfficientNet-B1.

3.4.4.1 MobileNet-v2

O MobileNet-V2, proposto por (SANDLER et al., 2018), foi desenvolvido com o objetivo de oferecer uma rede leve e eficiente para dispositivos com recursos computacionais limitados. Ele introduz duas principais inovações em relação ao seu antecessor: os blocos inversos residuais (*inverted residual blocks*) e o uso de camadas com conexões lineares no final de cada bloco.

Cada bloco inverso residual é composto por três etapas: uma expansão da dimensionalidade com convoluções 1×1 , seguida por uma convolução 3×3 (que atua separadamente por canal), e finalmente uma projeção linear com convolução 1×1 para reduzir a dimensionalidade. Esse design permite uma extração de características eficiente, mantendo baixo custo computacional e reduzindo a perda de informação.

3.4.4.2 EfficientNet-B1

O EfficientNet é uma família de redes proposta por (TAN; LE, 2020), baseada no conceito de *scaling* eficiente. Enquanto abordagens tradicionais aumentam a profundidade, largura ou resolução de forma isolada, o EfficientNet utiliza uma técnica de *compound scaling*, que expande essas três dimensões de maneira equilibrada e otimizada.

O modelo EfficientNet-B1 é uma das variações intermediárias da família, sendo maior e mais precisa que o EfficientNet-B0, mas ainda eficiente em termos de memória e velocidade. Sua arquitetura baseia-se em blocos *MBCConv* (Mobile Inverted Bottleneck Convolution), herdados da MobileNet-V2, com melhorias como o uso de *squeeze-and-excitation* (SE), que recalibra automaticamente a importância dos canais de características.

A escolha entre MobileNet-V2 e EfficientNet-B1 depende diretamente do equilíbrio entre custo computacional e desempenho desejado. O MobileNet-V2 é mais leve e adequado para aplicações em tempo real e em dispositivos com recursos limitados. Em contrapartida, o EfficientNet-B1 oferece maior acurácia com um aumento moderado de recursos, sendo mais indicado quando a qualidade da predição é prioritária. A Tabela 5 resume as principais características entre os encoders.

Tabela 5 – Comparativo entre MobileNet-V2 e EfficientNet-B1

Característica	MobileNet-V2	EfficientNet-B1
Tipo de bloco	<i>Inverted Residual</i>	<i>MBCConv + SE</i>
<i>Scaling</i>	Fixo	Composto
Parâmetros (milhões)	~3.4 M	~7.8 M
Precisão (ImageNet)	~72%	~79%
Squeeze-and-Excitation	Não	Sim
Desempenho	Leve	Eficiente

3.5 Métricas

Para avaliar o desempenho dos modelos, são utilizadas métricas amplamente adotadas na literatura, como Acurácia, Intersection over Union (IoU) e F1-Score. A seguir, as definições de cada métrica e suas respectivas fórmulas matemáticas.

3.5.1 Acurácia

A acurácia é uma métrica que mede a proporção de previsões corretas em relação ao total de amostras. É útil para avaliar o desempenho geral do modelo, mas pode ser enganosa em conjuntos de dados desbalanceados, pois pode superestimar o desempenho quando a classe majoritária é dominante. A acurácia é definida pela Equação 3.1:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.1)$$

onde:

- *TP* (True Positives) - Verdadeiros Positivos;

- TN (True Negatives) - Verdadeiros Negativos;
- FP (False Positives) - Falsos Positivos;
- FN (False Negatives) - Falsos Negativos.

3.5.2 Intersection over Union (IoU)

O *Intersection over Union* (IoU) é uma métrica amplamente utilizada na avaliação de modelos de segmentação de imagens. Mede a sobreposição entre a região prevista pelo modelo e a região real da classe de interesse, sendo calculada conforme a Equação 3.2:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (3.2)$$

onde:

- A representa a região prevista pelo modelo;
- B representa a região real;
- $|A \cap B|$ é a interseção entre as duas áreas;
- $|A \cup B|$ é a união entre as duas áreas.

No contexto da matriz de confusão, essa métrica também pode ser expressa em termos de verdadeiros positivos, falsos positivos e falsos negativos:

$$IoU = \frac{TP}{TP + FP + FN} \quad (3.3)$$

3.5.3 F1-Score

O *F1-Score* é uma métrica que equilibra precisão e revocação, sendo particularmente útil em conjuntos de dados desbalanceados. Ele é definido como a média harmônica entre precisão e revocação, conforme a Equação 3.4:

$$F1 = 2 \times \frac{\text{Precisão} \times \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (3.4)$$

onde:

- **Precisão** (*Precision*): mede a proporção de instâncias corretamente classificadas como positivas em relação ao total de instâncias previstas como positivas:

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (3.5)$$

- **Revocação** (*Recall* ou Sensibilidade): mede a proporção de instâncias positivas corretamente classificadas:

$$\text{Revocação} = \frac{TP}{TP + FN} \quad (3.6)$$

O *F1-Score* é especialmente útil quando há um desbalanceamento entre classes, pois garante que tanto a precisão quanto a revocação sejam levadas em consideração.

3.6 Síntese do Capítulo

Este capítulo apresentou os fundamentos teóricos essenciais para o desenvolvimento deste trabalho, abordando desde a aquisição dos dados até as metodologias empregadas para a análise das imagens. O conhecimento sobre o Sentinel-2 e o dataset CloudSen12 permitiu contextualizar as fontes de dados utilizadas, enquanto a revisão sobre redes neurais totalmente convolucionais forneceu embasamento para a escolha da U-Net como modelo principal de segmentação.

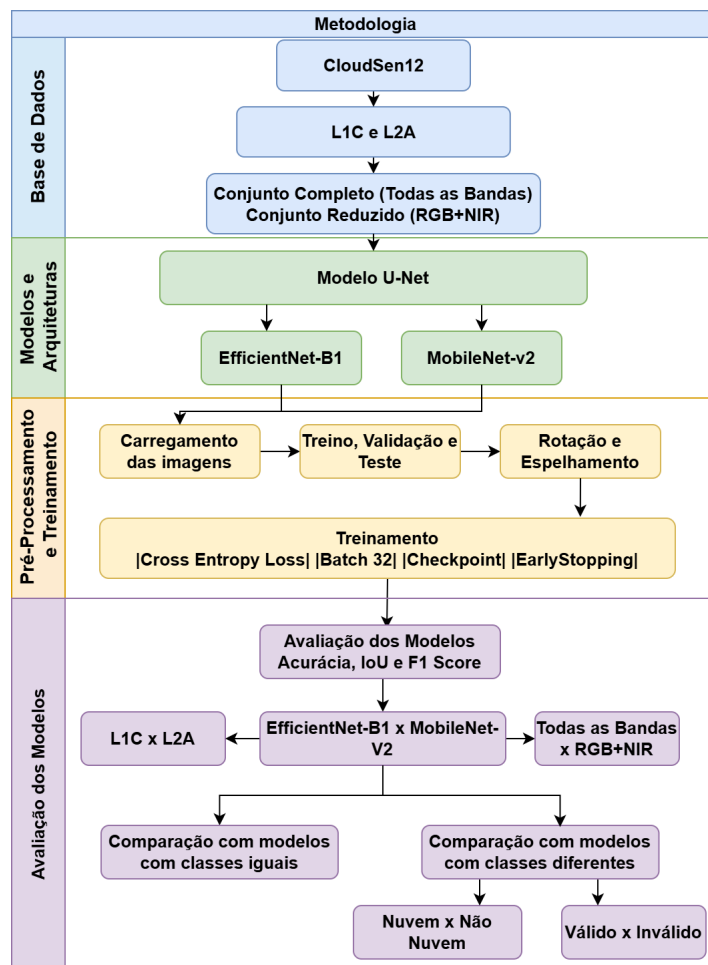
Além disso, a definição das métricas de avaliação foi crucial para garantir uma análise objetiva do desempenho do modelo de detecção de nuvens, permitindo a comparação quantitativa dos resultados. Dessa forma, o referencial teórico estabelecido neste capítulo sustenta as etapas metodológicas que serão detalhadas nos capítulos subsequentes, direcionando a implementação e a avaliação do modelo proposto.

4 METODOLOGIA

4.1 Considerações Iniciais

Neste capítulo, será apresentada a metodologia adotada para o desenvolvimento deste trabalho, conforme ilustrado na Figura 6. Inicialmente, descreve-se a base de dados utilizada na pesquisa, detalhando o dataset empregado e os conjuntos de entrada considerados nos treinamentos, levando em conta diferentes níveis de processamento e combinações espectrais. Em seguida, são apresentadas as arquiteturas dos modelos utilizados. Posteriormente, detalham-se os procedimentos de treinamento, incluindo a definição dos hiperparâmetros e a configuração dos *callbacks*. Por fim, são descritas as estratégias de avaliação dos modelos treinados, com a aplicação de métricas amplamente utilizadas na literatura.

Figura 6 – Metodologia para treinamento e avaliação dos modelos.



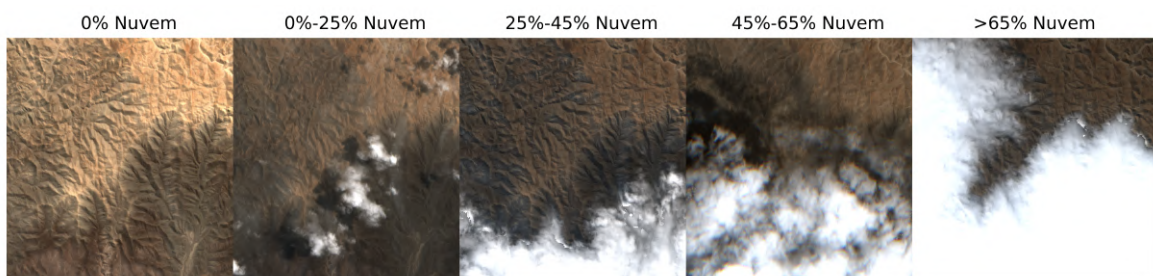
Fonte: O Autor

4.2 Base de Dados

Neste trabalho, foi utilizado o dataset CloudSen12, que contém 10.000 imagens adquiridas pelo satélite Sentinel-2 e rotuladas manualmente em alta qualidade, abrangendo diversas regiões do mundo. As imagens estão disponíveis nos níveis de processamento L1C, contendo todas as 13 bandas espectrais, e L2A, que disponibiliza 12 bandas, uma vez que a banda 10 (Cirrus) é utilizada exclusivamente para correções atmosféricas e, portanto, não está presente nesse nível. Cada imagem possui 512×512 pixels, e todas as bandas foram padronizadas para uma resolução espacial de 10 metros, utilizando interpolação por vizinho mais próximo para reamostragem (AYBAR et al., 2022).

O conjunto de dados compreende 2.000 cenas distintas, cada uma contendo cinco variações da mesma região com diferentes coberturas de nuvens. Essas variações são classificadas em cinco categorias, conforme a cobertura de nuvens: céu limpo (0%), quase limpo (0–25%), pouco nublado (25–45%), moderadamente nublado (45–65%) e muito nublado (>65%). A Figura 7 mostra um exemplo.

Figura 7 – Exemplo de imagem com percentuais de cobertura de nuvem diferentes.



Fonte: O Autor

Para o treinamento dos modelos, foram definidos dois conjuntos distintos de entrada:

- **Conjunto Completo:** Inclui todas as bandas disponíveis no respectivo nível de processamento (L1C ou L2A).
- **Conjunto Reduzido:** Contempla apenas as bandas RGB (Red, Green, Blue) e NIR (Near-Infrared), com o objetivo de reduzir a complexidade computacional do modelo e evitar a necessidade de reamostragem, visto que essas bandas já possuem resolução nativa de 10 metros.

As imagens foram normalizadas e os rótulos correspondentes (máscaras de nuvem) foram mantidos no formato original. A divisão dos dados seguiu o sistema proposto pelos autores, que, para evitar correlações espaciais e comprometer a generalização dos modelos, adotaram uma estratégia de divisão espacialmente estratificada. O globo terrestre foi dividido em hexágonos de

50.000 km², e as regiões foram separadas em 90% para treinamento e 10% para teste, garantindo que as de imagem de treinamento e teste fossem espacialmente independentes. Foram 8.490 imagens para o treinamento, 535 imagens para validação e 975 imagens para o teste.

4.3 Modelos e Arquiteturas

Foram utilizadas redes neurais totalmente convolucionais para a segmentação de nuvens, uma vez que esses modelos permitem a geração de máscaras com a mesma dimensão das imagens de entrada, realizando a classificação de forma pixel a pixel. A arquitetura base escolhida foi a UNet, amplamente reconhecida por sua capacidade de segmentação precisa, especialmente em aplicações médicas (AZAD et al., 2024) e em imagens multiespectrais (SOLÓRZANO et al., 2021). Dois encoders distintos foram avaliados:

- **EfficientNet-B1** – Encoder da família EfficientNet, projetado para otimizar o equilíbrio entre precisão e eficiência computacional, apresentando um desempenho robusto com menor consumo de recursos (TAN; LE, 2020).
- **MobileNet-V2** – Encoder otimizado para aplicações em dispositivos com restrições computacionais, oferecendo menor consumo de memória e tempo de inferência reduzido (SANDLER et al., 2018).

A implementação foi realizada utilizando as bibliotecas PyTorch (PASZKE et al., 2019), PyTorch Lightning (FALCON; The PyTorch Lightning team, 2019) e Segmentation Models PyTorch (IAKUBOVSKII, 2019), proporcionando modularidade e facilitando a experimentação com diferentes configurações de modelo.

4.4 Pré-Processamento e Treinamento

O fluxo de pré-processamento e treinamento seguiu as seguintes etapas:

1. **Carregamento das imagens** – Conversão para tensores e normalização das bandas espectrais.
2. **Divisão do dataset** – Separação dos dados em treino, validação e teste, seguindo a metodologia definida pelo dataset original.
3. **Aumento de dados** – Aplicação de transformações aleatórias com probabilidade de 50% para melhorar a generalização dos modelos, incluindo:
 - Espelhamento horizontal e vertical.
 - Rotação aleatória em ângulos múltiplos de 90° (90, 180 e 270).

- Outras transformações, como variações de zoom, brilho e contraste, não foram aplicadas, pois as imagens contêm bandas além do espectro visível, tornando tais ajustes inadequados.

4. **Configuração do treinamento** – Os modelos foram treinados considerando os seguintes parâmetros:

- Função de perda: Cross-Entropy Loss.
- Otimizador: Adam.
- Taxa de aprendizado inicial: 1×10^{-3} , com ajuste dinâmico via scheduler, reduzindo a taxa sempre que, após quatro épocas consecutivas, não houvesse melhora no desempenho da rede. O limite mínimo foi fixado em 1×10^{-6} .
- Tamanho do *batch*: 32.
- Número máximo de épocas: 100. Definido empiricamente. O ideal é que o treinamento nunca alcance o máximo de épocas, pois o modelo deve convergir antes e os callbacks encerrarão o treinamento.
- Callback de checkpoint: Monitoramento do erro de validação ao final de cada época, salvando o modelo sempre que um novo menor erro for alcançado.
- Callback de early stopping: Interrupção automática do treinamento caso não haja melhora no erro de validação após 12 épocas consecutivas.

4.5 Avaliação dos Modelos

Os modelos treinados foram avaliados com base em três métricas principais de segmentação:

- **Acurácia:** Mede a proporção de pixels corretamente classificados em relação ao total de pixels.
- **IoU:** Muito usada em segmentação de imagens, calcula a sobreposição entre a região prevista e a região real, dividindo a interseção pelo total combinado das duas áreas.
- **F1-Score:** Média harmônica entre precisão (quantos dos positivos previstos realmente são positivos) e revocação (quantos dos positivos reais foram corretamente previstos).

Após a avaliação individual, os resultados foram comparados entre os diferentes encoders (EfficientNet-B1 vs. MobileNet-V2) e entre os conjuntos espectrais (todas as bandas vs. RGB+NIR), analisando o impacto da escolha do encoder e das bandas na segmentação de nuvens.

Além da análise dos encoders e das bandas espectrais, também foi realizada uma comparação entre os modelos treinados com imagens L1C e L2A, a fim de avaliar qual nível de processamento proporciona melhores resultados na segmentação de nuvens. Para isso, foram comparadas as métricas obtidas em cada caso, buscando identificar se a remoção dos efeitos atmosféricos nas imagens L2A influencia positivamente na detecção de nuvens. Por fim, os melhores modelos treinados foram comparados com outros modelos da literatura, a fim de consolidar o desempenho frente a abordagens estatísticas e modelos também baseados em aprendizado profundo.

4.6 Síntese do Capítulo

No capítulo foi apresentada a metodologia que permitiu a implementação de um modelo eficaz para a segmentação de nuvens, utilizando abordagens modernas de aprendizado profundo. A definição dos conjuntos de entrada, a escolha dos encoders e os parâmetros de treinamento foram ajustados para otimizar a acurácia da segmentação, minimizando a complexidade computacional. A avaliação dos modelos possibilitou uma análise detalhada do impacto das diferentes configurações utilizadas, fornecendo subsídios para a interpretação dos resultados obtidos e para a escolha do melhor modelo para a tarefa proposta.

5 RESULTADOS

5.1 Considerações Iniciais

Nesta seção, serão apresentados os resultados obtidos nos modelos treinados para a segmentação de nuvens. Serão analisados o desempenho dos modelos treinados, considerando diferentes configurações, incluindo variações de encoders (EfficientNet-b1 e MobileNet-v2), bandas espectrais (todas as bandas e apenas 4 bandas) e níveis de processamento (L1C e L2A). No total, foram treinados 8 modelos. Além disso, será discutido o impacto dessas escolhas na qualidade da segmentação, utilizando métricas amplamente adotadas na literatura. No Apêndice A se encontra os gráficos das métricas durante o treinamento dos modelos.

5.2 Avaliação Quantitativa dos Modelos

A Tabela 6 mostra os resultados quantitativos obtidos pelos modelos treinados, considerando as métricas de Acurácia, IoU e F1-Score. A análise será realizada comparando o impacto da escolha do encoder, a quantidade de bandas espectrais e do nível de processamento das imagens (L1C e L2A).

Modelo	L1C			L2A		
	Acurácia	IoU	F1-Score	Acurácia	IoU	F1-Score
Todas as bandas - EfficientNet-B1	95.28	82.74	90.56	95.04	81.95	90.08
Todas as bandas - MobileNet-V2	94.93	81.60	89.87	94.59	80.46	89.17
RGB + NIR - EfficientNet-B1	94.86	81.36	89.72	94.87	81.38	89.73
RGB + NIR - MobileNet-V2	94.44	79.99	88.88	94.61	80.53	89.21

Tabela 6 – Comparação de métricas entre diferentes modelos para L1C e L2A

5.2.1 Comparação entre Encoders

Ao comparar os encoders, observa-se que o EfficientNet-B1 superou consistentemente o MobileNet-V2 em todas as configurações. Para as imagens L1C com todas as bandas, o EfficientNet-B1 atingiu uma acurácia de 95.28%, enquanto o MobileNet-V2 obteve 94.93%. A mesma tendência é observada para IoU (82.74% contra 81.60%) e F1-Score (90.56% contra 89.87%). Esse padrão de superioridade do EfficientNet-B1 também se repete nas imagens L2A, bem como na configuração reduzida de bandas (RGB + NIR), indicando que o encoder EfficientNet-B1 apresenta melhor capacidade de extração de características para a segmentação de nuvens.

5.2.2 Impacto da Quantidade de Bandas Espectrais

Ao comparar os modelos treinados com todas as bandas espectrais contra aqueles treinados apenas com RGB + NIR, nota-se que o uso de todas as bandas proporcionou um leve aumento nas métricas. Para L1C, a melhor configuração com todas as bandas (EfficientNet-B1) obteve um IoU de 82.74%, enquanto a configuração com RGB + NIR alcançou 81.36%, representando uma diferença de aproximadamente 1.38 pontos percentuais. Resultados similares foram observados para o MobileNet-V2. Essa diferença sugere que a inclusão de bandas adicionais melhora a segmentação, possivelmente devido ao aumento da informação espectral disponível para o modelo. Porém, essa diferença é pequena considerando o número de bandas não usadas nos modelos mais simples (9 bandas para L1C e 8 bandas para L2A). Além disso, o uso dos modelos RGB+NIR evita a necessidade de pré-processamento nas imagens, já que essas bandas possuem a resolução espacial de 10 metros, enquanto as demais bandas possuem resolução de 20 ou 60 metros, necessitando de reamostragem para aplicação nos modelos.

5.2.3 Influência do Nível de Processamento (L1C vs. L2A)

Ao analisar o impacto do nível de processamento das imagens, observa-se que, na configuração com todas as bandas, os modelos treinados com L1C apresentaram um desempenho ligeiramente superior aos treinados com L2A. No entanto, é importante destacar que as imagens L1C possuem 13 bandas, enquanto as imagens L2A possuem apenas 12, uma vez que uma das bandas é utilizada para correções atmosféricas. Dessa forma, para uma análise mais justa, considerando a mesma quantidade de bandas, observa-se que, na configuração RGB + NIR, os modelos treinados nas imagens L2A possuem ligeira vantagem, notada principalmente com o uso do encoder mobilenet-v2, com uma diferença máxima de 0,54% no IoU. O encoder EfficientNet-B1, que obteve o melhor desempenho, tem variações mínimas entre o modelo treinado com imagens L1C e L2A. Esses resultados sugerem que a remoção dos efeitos atmosféricos nas imagens L2A tem impacto mínimo quando a quantidade de bandas é igual e também indica que a diferença observada na configuração com todas as bandas pode estar associada mais à redução da informação espectral do que ao processamento atmosférico em si.

5.2.4 Resumo da Avaliação Quantitativa

De modo geral, os resultados indicam que:

- O encoder EfficientNet-B1 obteve desempenho superior ao MobileNet-V2 em todas as configurações testadas.
- O uso de todas as bandas espectrais resultou em métricas ligeiramente superiores em comparação com o uso apenas de RGB + NIR.

- Os modelos RGB + NIR mostram que o nível de processamento L2A foi levemente superior ao modelos treinados com imagens L1C, sugerindo que a remoção dos efeitos atmosféricos nas imagens L2A é benéfica, mas tem impacto mínimo no desempenho quando a quantidade de bandas é mantida igual.

5.3 Avaliação Qualitativa dos Modelos

A análise qualitativa dos resultados é essencial para complementar os achados quantitativos, pois permite uma inspeção visual da segmentação realizada pelos modelos e identificação de possíveis padrões de erro. Para essa análise, foram selecionadas duas imagens representativas, cada uma com diferentes níveis de complexidade na segmentação. Em ambas as imagens, são apresentadas as versões nos níveis de processamento L1C e L2A, suas respectivas máscaras de referência e as segmentações preditas pelos modelos treinados. O padrão de cores é: preto para classe limpo, branco para classe nuvem, amarelo para classe nuvem fina e cinza para classe sombra de nuvem.

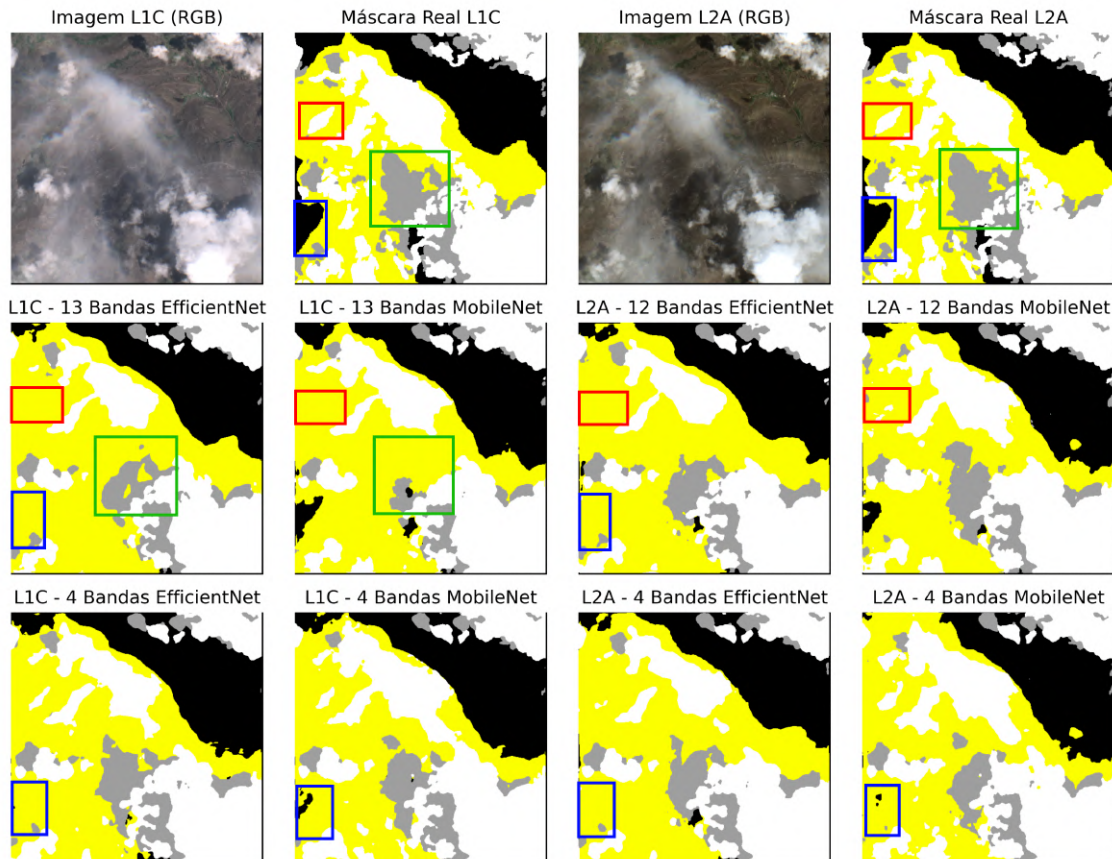
5.3.1 Imagem com Alto Grau de Dificuldade

A primeira imagem selecionada apresenta um cenário desafiador tanto para a anotação manual quanto para a predição dos modelos. Essa complexidade se deve à presença das quatro classes do problema (Limpo, Nuvem, Nuvem fina e Sombra de nuvem) distribuídas de forma heterogênea na cena. Além disso, há transições sutis entre as classes de interesse, o que aumenta a ambiguidade na segmentação.

A Figura 8 apresenta a comparação das previsões realizadas pelos oito modelos. Observa-se que todos os modelos apresentam dificuldades semelhantes, especialmente na distinção entre nuvem, nuvem fina e sombra de nuvem. Algumas regiões exibem inconsistências entre as previsões e a máscara de referência, o que pode ser reflexo tanto de limitações dos modelos quanto de incertezas no processo de anotação manual. Os boxes em vermelho mostram uma região da imagem classificada como nuvem e predita pela maior parte dos modelos como nuvem fina. Os boxes em azul mostram uma região classificada como limpo e predita pela maior parte dos modelos como nuvem fina. Os boxes em verde mostram uma região classificada como sombra de nuvem e predita com dificuldades pelos modelos parcialmente como nuvem fina.

Apesar dessas dificuldades, os modelos conseguiram identificar as principais regiões de nuvem e sombra, ainda que com variações na granularidade da segmentação. Além disso, a similaridade entre as previsões dos modelos complementa o que a Tabela 6 mostrou que, em geral, os modelos têm desempenhos parecidos e talvez as principais limitações ainda sejam os limiares humanos na anotação manual das classes presentes nas imagens.

Figura 8 – Imagem com alto grau de dificuldade na segmentação. Da esquerda para a direita, na primeira linha as imagens L1C e L2A e suas respectivas máscaras. Na segunda linha as máscaras dos modelos que utilizam todas as bandas. Na terceira linha as máscaras dos modelos que utilizam apenas 4 bandas (RGB+NIR).



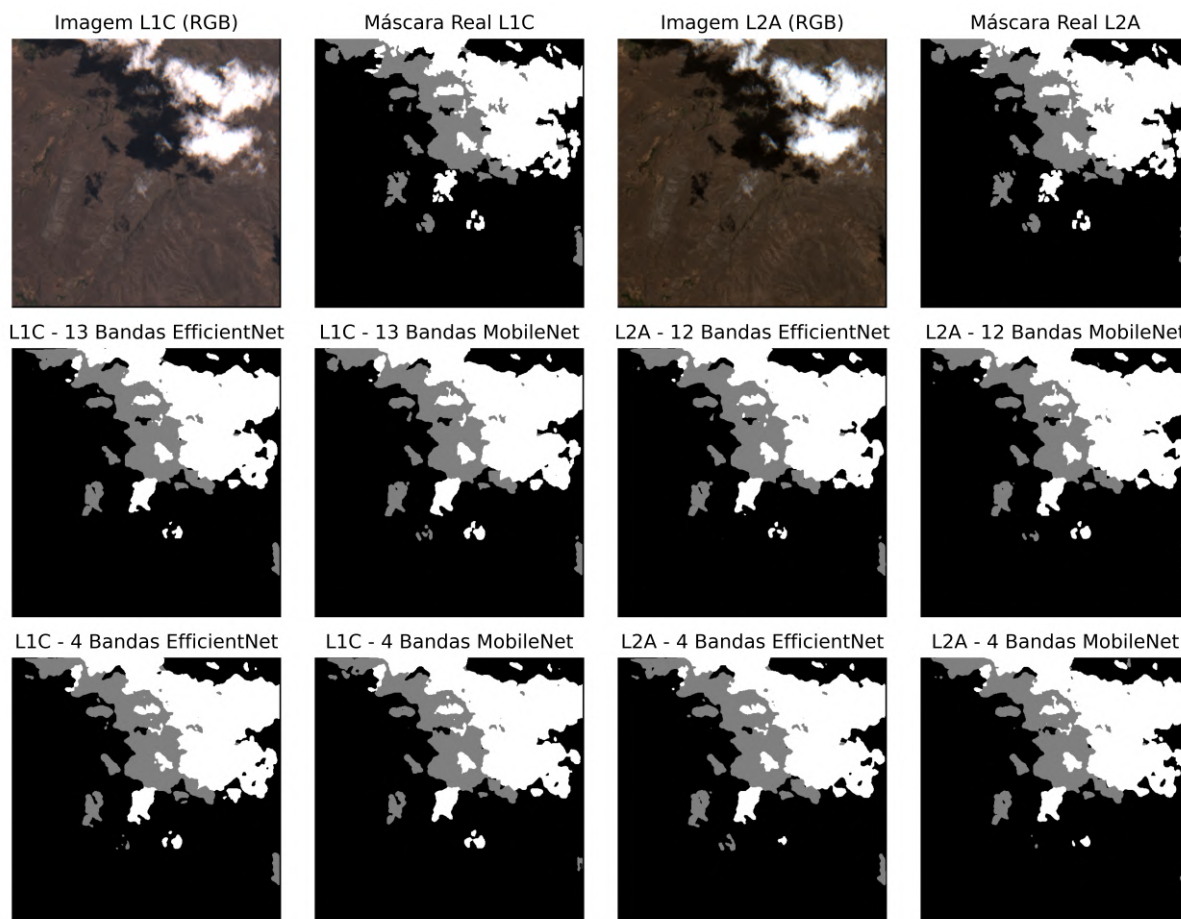
Fonte: O Autor

5.3.2 Imagem com Baixo Grau de Dificuldade

A segunda imagem apresenta um cenário de segmentação mais simples, com transições bem definidas entre as classes e menor ambiguidade entre as regiões. Como pode ser observado na Figura 9, os modelos apresentaram um desempenho notavelmente superior, conseguindo segmentar as regiões de nuvem e sombra com alto grau de precisão.

Outro ponto relevante é a semelhança entre as previsões dos modelos treinados com todas as bandas e aqueles treinados com apenas 4 bandas (RGB + NIR). Assim como indicado pela análise quantitativa, os modelos conseguem manter um alto desempenho mesmo com a redução do número de bandas, o que sugere que utilizar apenas essas quatro bandas pode ser um caminho viável para tornar os modelos mais leves e eficientes, além de dispensar a necessidade de reamostragem das imagens.

Figura 9 – Imagem com baixo grau de dificuldade na segmentação. Da esquerda para a direita, na primeira linha as imagens L1C e L2A e suas respectivas máscaras. Na segunda linha as máscaras dos modelos que utilizam todas as bandas. Na terceira linha as máscaras dos modelos que utilizam apenas 4 bandas (RGB+NIR).



Fonte: O Autor

5.3.3 Conclusões da Análise Qualitativa

Os resultados qualitativos corroboram os achados quantitativos, reforçando algumas tendências importantes:

- Os modelos apresentaram desempenho semelhante entre si, independentemente da arquitetura utilizada (EfficientNet-B1 ou MobileNet-V2).
- A segmentação se mostrou mais desafiadora em imagens complexas, especialmente devido à confusão entre as classes de nuvem, nuvem fina e sombra de nuvem.
- A redução do número de bandas (RGB + NIR) não compromete significativamente a qualidade da segmentação, sugerindo que essa abordagem pode ser adotada para tornar os modelos mais eficientes sem perda expressiva de desempenho.

- Os modelos treinados com L1C e L2A apresentaram previsões similares, indicando que a remoção dos efeitos atmosféricos no nível L2A não trouxe benefícios significativos para a segmentação em cenários com menos bandas.

Essa análise visual reforça que, apesar dos desafios impostos por imagens com alta complexidade, os modelos conseguem generalizar bem e manter um desempenho satisfatório, especialmente em imagens com transições mais definidas entre as classes. Outras imagens comparando as máscaras geradas pelos modelos propostos se encontram no Apêndice B.

5.4 Comparação com Outros Modelos

Nesta seção, são comparados os dois melhores modelos treinados neste trabalho para cada configuração de bandas - L1C 13 Bandas EfficientNet-B1 e L2A 4 Bandas EfficientNet-B1 - com outros modelos da literatura. O objetivo dessa análise é comparar o desempenho das abordagens propostas em relação a modelos previamente desenvolvidos.

A comparação será realizada em duas partes. Primeiro, serão analisados modelos treinados para a segmentação de classes semelhantes, permitindo uma avaliação direta do desempenho em condições equivalentes. Em seguida, será feita uma adaptação da comparação com modelos que foram treinados para diferentes categorias de segmentação, buscando identificar tendências gerais e potenciais desafios na aplicação das metodologias desenvolvidas.

Essa análise visa contextualizar os resultados obtidos e explorar a viabilidade do uso de abordagens mais leves, como o modelo L2A 4 Bandas EfficientNet-B1, em relação a modelos que utilizam um maior número de bandas espectrais.

5.4.1 Comparação com Modelos com Classes Iguais

Para validar o desempenho dos modelos treinados, comparamos os mais eficazes neste estudo com outros modelos encontrados na literatura que fazem a segmentação de nuvens em classes comparáveis (limpo, nuvem, nuvem fina e sombra de nuvem). A Tabela 6 apresenta os resultados das métricas de Acurácia, IoU e F1-Score no conjunto de teste para cada modelo analisado. Os modelos propostos obtiveram os melhores desempenhos em todas as métricas. O modelo L1C 13 Bandas EfficientNet-B1 apresentou a maior acurácia (95,28%), maior IoU (82,75%) e maior F1-Score (90,56%), seguido pelo modelo L2A 4 Bandas EfficientNet-B1, que obteve acurácia de 94,87%, IoU de 81,38% e F1-Score de 89,73%.

Em comparação com os modelos da literatura, o L2A 4 Bandas EfficientNet-B1 superou a UnetMob-V1, UnetMob-V2 e Sensei-V2 em todas as métricas, demonstrando a eficácia da arquitetura EfficientNet-B1 para o problema de segmentação de nuvens mesmo com uma quantidade reduzida de bandas. A UnetMob-V2 obteve resultados próximos aos do modelo L2A 4 Bandas EfficientNet-B1, com F1-Score de 89,67%, enquanto a UnetMob-V1 ficou levemente

abaixo, com 89,27%. O modelo Sensei-V2 apresentou os menores valores entre os modelos treinados, com F1-Score de 88,79%.

Os resultados reforçam a viabilidade do uso da arquitetura EfficientNet-B1 para a segmentação de nuvens, tanto com todas as bandas espectrais disponíveis quanto com uma configuração reduzida de bandas. A manutenção de um alto desempenho mesmo com apenas quatro bandas sugere que modelos mais leves e eficientes podem ser utilizados sem comprometer significativamente a qualidade da segmentação. Outras imagens comparando as máscaras geradas pelos modelos com classes iguais se encontram no Apêndice C.

Tabela 7 – Comparação de Modelos com Classes Semelhantes.

Modelos	Acurácia	IoU	F1-Score	Fonte
L1C 13 Bandas - EfficientNet-B1	95,28%	82,75%	90,56%	O Autor
L2A 4 Bandas - EfficientNet-B1	94,87%	81,38%	89,73%	O Autor
UnetMob-V1	94,63%	80,62%	89,27%	(AYBAR et al., 2022)
UnetMob-V2	94,83%	81,27%	89,67%	(AYBAR et al., 2024)
Sensei-V2	94,40%	79,84%	88,79%	(FRANCIS, 2024)

5.4.2 Comparação com Modelos com Classes Diferentes

Para ampliar a análise comparativa e avaliar o desempenho dos modelos propostos frente a abordagens já consolidadas na literatura, foi realizada uma adaptação das classes de segmentação para permitir uma comparação justa com outros métodos, seguindo a abordagem de superclasses proposta em (AYBAR et al., 2022). Como os modelos analisados diferem em suas definições de classes, foi adotada uma estratégia de simplificação baseada na reorganização das categorias originais em duas novas configurações: Nuvem vs. Não Nuvem e Válido vs. Inválido. A Tabela 8 apresenta a correspondência entre as classes originais e as superclasses utilizadas para essa avaliação.

Tabela 8 – Classes e Superclasses de pixels para avaliação dos modelos.

Código	Classe	Superclasse 1	Superclasse 2	Prioridade
0	Claro	Não-nuvem	Válido	4
1	Nuvem Espessa	Nuvem	Inválido	1
2	Nuvem Fina	Nuvem	Inválido	3
3	Sombra de Nuvem	Não-nuvem	Inválido	2

Na configuração Nuvem vs. Não Nuvem, todas as classes relacionadas à presença de nuvens (Nuvem Espessa e Nuvem Fina) foram agrupadas em uma única categoria (Nuvem), enquanto as demais (Claro e Sombra de Nuvem) foram consideradas como Não Nuvem. Já na configuração Válido vs. Inválido, além das classes de nuvem, também consideramos a Sombra de Nuvem como parte da classe Inválido, uma vez que sua presença pode comprometer a qualidade da segmentação e da análise subsequente. Essa reorganização das classes permite comparar

diretamente os modelos treinados neste trabalho com aqueles que não distinguem diferentes tipos de nuvem ou sombra, tornando a análise mais abrangente.

Avaliamos nossos modelos treinados, L1C 13 Bandas EfficientNet-B1 e L2A 4 Bandas EfficientNet-B1, em ambas as configurações e os comparamos com métodos amplamente utilizados para a detecção de nuvens, incluindo abordagens baseadas em regras e redes neurais. Os modelos considerados nesta análise são:

- **Sen2Cor:** Método tradicional de correção atmosférica do Sentinel-2, que inclui uma máscara de nuvens baseada em regras heurísticas, disponível no nível L2A (MAIN-KNORN et al., 2017).
- **S2Cloudless:** O s2cloudless é um algoritmo automatizado de detecção de nuvens para imagens do Sentinel-2 no nível L1C com base em um algoritmo de aumento de gradiente. Como entrada ele precisa de 10 bandas: B01, B02, B04, B05, B08, B8A, B09, B10, B11, B12 e a saída do algoritmo é um mapa de probabilidade de nuvem. Os usuários podem converter o mapa de probabilidade em uma máscara de nuvem utilizando um limiar variável. O valor recomendado para o limiar é 0,4 (SKAKUN et al., 2022).
- **Fmask:** Algoritmo baseado em regras espectrais para classificar nuvens, sombras de nuvens, água e neve (QIU; ZHU; HE, 2019). Os parâmetros de dilatação para nuvem, sombra de nuvem e neve foi definido em 3, 3 e 0 pixels, respectivamente. O raio de erosão (dilatação) é definido como 0 (90) metros, enquanto o limite de probabilidade de nuvem é fixado em 20%.
- **Kappamask_L1C e Kappamask_L2A:** Rede Neural U-Net com duas configurações distintas: todas as bandas Sentinel-2 L1C e todas as bandas Sentinel-2 L2A, exceto a banda Red Edge 3. Foi treinado usando as máscaras Sentinel-2 KappaZeta Cloud e Cloud Shadow e o catálogo Sentinel-2 Cloud Mask (DOMNICH et al., 2021).
- **CD_FCNN_RGBI e CD_FCNN_RGBISWIR:** Rede Neural U-Net com duas combinações diferentes de bandas do sentinel-2: RGBI (B2, B3, B4 e B8) e RGBISWIR (B2, B3, B4, B8, B11 e B12) treinados no conjunto de dados Landsat Biome-8 e adaptado para o sentinel-2 usando transferência de aprendizado do Landsat 8 para o Sentinel-2 (LÓPEZ-PUIGDOLLERS et al., 2021).

5.4.2.1 Nuvem vs Não Nuvem

A Tabela 9 apresenta os principais resultados da comparação entre os modelos treinados e os modelos da literatura, agora na configuração binária Nuvem vs. Não nuvem. As métricas utilizadas são: Acurácia, IoU, F1-Score.

Os resultados mostram que os modelos baseados na EfficientNet-B1 superam os demais métodos em todas as métricas. O modelo L1C 13 Bandas EfficientNet-B1 obteve os melhores

Tabela 9 – Resultados da comparação entre modelos para as classes Nuvem e Não Nuvem.

Modelo	Acurácia%	IoU%	F1-Score%
L1C 13 Bandas EfficientNet-B1	94.17	85.74	92.33
L2A 4 Bandas EfficientNet-B1	93.56	84.30	91.48
Sen2Cor	80.64	52.41	68.78
S2Cloudless	86.33	68.31	81.17
Fmask	87.26	71.66	83.49
KappaMask_L1C	85.69	69.38	81.92
KappaMask_L2A	80.00	63.08	77.36
CD_FCNN_RGBI	80.52	51.62	68.09
CD_FCNN_RGBISWIR	81.48	53.27	69.52

desempenhos, atingindo 94,17% de acurácia, 85,74% de IoU e 92,33% de F1-Score. O modelo L2A 4 Bandas EfficientNet-B1 apresentou um desempenho ligeiramente inferior, porém ainda superior aos modelos da literatura, com 93,56% de acurácia, 84,30% de IoU e 91,48% de F1-Score.

Dentre os modelos comparados, o Sen2Cor, amplamente utilizado para a geração de produtos Sentinel-2 L2A, obteve 80,64% de acurácia e um baixo IoU de 52,41%, evidenciando dificuldades na segmentação precisa das nuvens. O S2Cloudless e o Fmask tiveram melhores desempenhos em comparação ao Sen2Cor, alcançando 86,33% e 87,26% de acurácia, respectivamente, mas ainda ficaram abaixo dos modelos baseados na EfficientNet-B1.

Entre os modelos baseados em redes neurais (KappMask e CD_FCNN), o KappaMask_L1C apresentou o melhor desempenho, com 85,69% de acurácia e 81,92% de F1-Score, mas ainda bem inferior aos modelos propostos. Os modelos CD_FCNN_RGBI e RGBISWIR tiveram os piores desempenhos entre os modelos baseados em aprendizado profundo, com acurácias de 80,52% e 81,48%, respectivamente. A explicação se deve ao fato de ser um modelo treinado em imagens do Landsat-8 e usado no Sentinel-2 utilizando transferência de aprendizagem, que tende a obter bons resultados, porém ainda não alcança os modelos treinados diretamente em imagens do Sentinel-2. Algumas imagens comparando as máscaras geradas pelos modelos com a superclasse nuvem e não nuvem podem ser encontradas no Apêndice D.

5.4.2.2 Válido vs Inválido

A classificação válido/inválido é essencial para identificar áreas comprometidas por nuvens e sombras, garantindo que apenas regiões úteis sejam consideradas para aplicações subsequentes. A Tabela 10 apresenta os resultados da comparação entre os modelos propostos (L1C 13 Bandas EfficientNet-B1 e L2A 4 Bandas EfficientNet-B1) e outros métodos da literatura.

Os modelos baseados na EfficientNet-B1 demonstraram desempenho superior em todas as métricas. O modelo L1C 13 Bandas EfficientNet-B1 obteve os melhores resultados, atingindo 93,98% de acurácia, 88,03% de IoU e 93,64% de F1-Score, enquanto o modelo L2A 4 Bandas

EfficientNet-B1 apresentou desempenho ligeiramente inferior, com 93,32% de acurácia, 86,75% de IoU e 92,90% de F1-Score. Esses resultados indicam que ambos os modelos possuem alta capacidade de diferenciar corretamente regiões válidas e inválidas, com vantagem para a versão que utiliza todas as 13 bandas do Sentinel-2 L1C.

Os métodos tradicionais, como Sen2Cor, S2Cloudless e Fmask, apresentaram desempenho inferior. O Sen2Cor, obteve 73,36% de acurácia, 46,85% de IoU e 63,81% de F1-Score, indicando baixa precisão na segmentação. O S2Cloudless apresentou melhora em relação ao Sen2Cor, alcançando 80,26% de acurácia e 75,99% de F1-Score, mas ainda abaixo dos modelos baseados em EfficientNet-B1. O Fmask, por sua vez, apresentou os melhores resultados entre os métodos tradicionais, atingindo 86,21% de acurácia, 74,66% de IoU e 85,49% de F1-Score, mas ainda aquém dos modelos propostos.

Os modelos baseados em redes neurais também ficaram abaixo dos modelos EfficientNet-B1. O KappaMask_L1C obteve 85,23% de acurácia e 73,32% de IoU, enquanto sua versão KappaMask_L2A teve desempenho inferior, com 80,23% de acurácia e 67,96% de IoU, seguindo a tendência das análises anteriores. Já os modelos CD_FCNN_RGBI e CD_FCNN_RGBISWIR, apresentaram os piores desempenhos, com acurácias de 72,17% e 72,68%, respectivamente, além de baixas métricas de IoU (43,50% e 44,07%) e F1-Score (60,63% e 61,18%).

Os resultados reforçam a eficácia dos modelos EfficientNet-B1 na tarefa de segmentação válido/inválido, superando significativamente tanto os métodos tradicionais quanto as redes neurais da literatura. A inclusão de todas as 13 bandas do Sentinel-2 no modelo L1C resultou em um desempenho melhor, porém os modelos com as bandas RGB+NIR obtiveram resultados ainda satisfatórios. Algumas imagens comparando as máscaras geradas pelos modelos com a superclasse válido e inválido podem ser encontradas no Apêndice E.

Tabela 10 – Resultados da comparação entre modelos para as classes válido e inválido.

Modelo	Acurácia%	IoU%	F1%
L1C 13 Bandas EfficientNet-B1	93.98	88.03	93.64
L2A 4 Bandas EfficientNet-B1	93.32	86.75	92.90
Sen2Cor	73.36	46.85	63.81
S2Cloudless	80.26	61.28	75.99
Fmask	86.21	74.66	85.49
KappaMask_L1C	85.23	73.32	84.60
KappaMask_L2A	80.23	67.96	80.92
CD_FCNN_RGBI	72.17	43.50	60.63
CD_FCNN_RGBISWIR	72.68	44.07	61.18

5.4.3 Conclusões das Comparações com Outros Modelos

Os resultados das comparações com os modelos treinados e os modelos da literatura mostram as seguintes conclusões:

- Modelos baseados em redes neurais totalmente convolucionais, quando treinados em datasets com grande variabilidade de dados, tende a superar os modelos estatísticos baseados em regras espectrais.
- Os encoder EfficientNet se mostrou superior ao encoder MobileNet, tanto nos modelos treinados, quanto nas comparações com os modelos UnetMob_v1 e UnetMob_v2 que são baseados em mobilenet.
- Os modelos treinados com um número reduzido de bandas (RGB+NIR) ainda foram superior aos demais modelos, mostrando que é possível reduzir o número de bandas sem comprometer a eficiência.

5.5 Síntese do Capítulo

Neste capítulo, foi apresentada uma análise quantitativa, qualitativa e comparativa dos modelos treinados para a detecção de nuvens em imagens Sentinel-2. Os resultados quantitativos demonstraram que os modelos baseados em EfficientNet-B1, especialmente a versão com 13 bandas do Sentinel-2 L1C, obtiveram os melhores desempenhos nas métricas avaliadas. A análise qualitativa reforçou essa superioridade, porém mostrando que os modelos com quantidade reduzida de bandas também tiveram bons resultados. Na comparação com modelos da literatura, os modelos treinados se destacaram em ambas as configurações de avaliação: Nuvem vs. Não-Nuvem e Válido vs. Inválido. Esses resultados demonstram o potencial dos modelos propostos para aprimorar a detecção de nuvens em imagens de satélite, contribuindo para aplicações em monitoramento ambiental e sensoriamento remoto.

6 CONCLUSÃO

O objetivo principal desta dissertação foi desenvolver e avaliar modelos de redes neurais totalmente convolucionais para a detecção de nuvens em imagens do Sentinel-2, considerando diferentes níveis de processamento (L1C e L2A) e configurações de bandas espectrais (todas as bandas versus RGB + NIR). Para isso, utilizou-se o dataset CloudSen12, que provê um amplo conjunto de imagens provenientes de diversas regiões do mundo, possibilitando avaliar a robustez dos modelos em cenários variados. A metodologia incluiu o uso da Unet com duas variações de encoders (EfficientNet e MobileNet). Houve uma comparação quantitativa e qualitativa dos resultados, além de uma análise comparativa com modelos da literatura, a fim de posicionar as soluções propostas em relação ao estado da arte.

Os resultados obtidos indicam que a arquitetura baseada em EfficientNet-B1 apresentou desempenho superior em comparação ao MobileNet-v2, alcançando altas taxas de acurácia, IoU e F1-Score. Embora o uso de todas as bandas espectrais tenha se mostrado a melhor opção, a configuração reduzida (RGB + NIR) também apresentou resultados competitivos, sugerindo um caminho promissor para a adoção de modelos mais leves e eficientes. Na comparação com métodos tradicionais, como Sen2Cor e S2Cloudless, e com outras arquiteturas de redes neurais, como KappaMask e CD_FCNN, ficou evidente a vantagem dos modelos propostos, tanto em cenários com classes equivalentes quanto em configurações mais simplificadas (Nuvem vs. Não-Nuvem e Válido vs. Inválido). Ainda assim, observou-se dificuldades na distinção entre nuvens finas e sombras de nuvens.

Apesar dos avanços, algumas limitações foram identificadas. Uma delas diz respeito ao custo computacional associado ao treinamento de modelos com muitas bandas espectrais, especialmente em cenários de grande volume de dados. Além disso, a qualidade das anotações de nuvens, sobretudo na presença de nuvens finas ou sombras, afeta diretamente a precisão dos resultados, logo, um dataset amplo e bem anotado tem impacto direto na qualidade do modelo. Outra limitação é que o treinamento e teste foram realizados em um dataset com imagens de mesma dimensão e mesma estratégia de anotação, logo, uma avaliação em outros datasets pode ajudar a validar ainda mais os modelos propostos.

Em síntese, este trabalho representa uma contribuição relevante para o campo do sensoriamento remoto, ao evidenciar que modelos baseados em redes neurais totalmente convolucionais são capazes de alcançar desempenhos expressivos na detecção de nuvens, superando métodos tradicionais e abordagens existentes na literatura. A adoção de modelos mais leves, utilizando apenas bandas RGB+NIR, mostrou-se viável sem perdas significativas de desempenho, indicando uma rota promissora para aplicações operacionais que demandem menor custo computacional. Os achados aqui relatados podem servir de base para novos estudos e aplicações que envolvam

a identificação de nuvens em imagens de satélite, auxiliando em processos de monitoramento ambiental.

6.1 Trabalhos Futuros

- Testar os modelos propostos em outros conjuntos de dados abertos e compará-los com métodos da literatura, a fim de avaliar sua robustez e capacidade de generalização.
- Exploração de arquiteturas mais avançadas de redes neurais: Modelos como DeepLabv3+, SegFormer, ou redes baseadas em Transformers têm demonstrado excelente desempenho em tarefas de segmentação semântica e podem ser aplicados à detecção de nuvens com potencial de melhora nos resultados.
- Investigar soluções híbridas, combinando redes neurais com modelos estatísticos ou mecanismos de atenção, com o objetivo de aumentar a precisão e reduzir a ocorrência de falsos positivos na detecção de nuvens.
- Aplicar os modelos propostos em aplicações que dependem de máscaras de nuvem e sombra, como a detecção de desmatamento, para avaliar se as máscaras geradas contribuem para a melhoria da acurácia nos resultados dessas análises.
- Desenvolvimento de uma interface ou ferramenta prática: Criar uma aplicação com interface gráfica ou API para facilitar o uso dos modelos por pesquisadores e profissionais de sensoriamento remoto.
- Melhoria na segmentação com pós-processamento: Investigar o uso de técnicas como Conditional Random Fields (CRFs) ou filtros espaciais para refinar os resultados de segmentação obtidos pelos modelos.

6.2 Publicações

1. DIAS, JEAN A. C.; TAMASAUSKAS, LEONARDO DE O.; GUIMARÃES, PEDRO H. DO V.; CORRÊA, ALAN B. S.; D. NETO, JOÃO D. C.; SANTOS, ALBERT E. C. DOS; SOUZA, DANILO; PAIXÃO, ERMÍNIO R.; FERNANDES, JOSÉ G. DOS S.; COSTA, GABRIEL B.; SERUFFO, MARCOS C. DA R.. Inteligência Artificial Aplicada ao Aprimoramento de Parâmetros para o Monitoramento de Fluxo de Carbono por Satélite na Região Amazônica In: Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais, 2024, Brasil. Anais do XV Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais (WCAMA 2024). 2024, p.31
2. CORREA, A. B. S.; RIBEIRO, L. M.; FERREIRA, F. H. C. S.; CARDOSO, C. M. M.; RIBEIRO, F. F.; ARAUJO, J. P. L.; A. NETO, M. C.; BARROS, F. J. B.. Medidas e Modelo

- de Traçado de Raios para Ambientes Indoor em 3.5GHz In: Anais da XIII Conferência Nacional em Comunicações, Redes e Segurança da Informação, 2023, Belém. Anais da XIII Conferência Nacional em Comunicações, Redes e Segurança da Informação. 2023.
3. RIBEIRO, L. M.; CARDOSO, C. M. M.; CORREA, A. B. S.; MENEZES, D. F.; OLIVEIRA, K. S.; ARAUJO, J. P. L.; A. NETO, M. C.; BARROS, F. J. B.. Modelo de Perda de Propagação para LoRa em Terreno Misto:Terra-Rio In: XIII Conferência Nacional em Comunicações, Redes e Segurança da Informação, 2023, Belém. Anais da XIII Conferência Nacional em Comunicações, Redes e Segurança da Informação. 2023.
 4. TAMASAUSKAS, L. O.; DIAS, J. A. C.; GUIMARAES, P. H. V. ; CORREA, A. B. S.; SANTOS, A. E. C.; COSTA, G. B.; FERNANDES, J. G. S.; SOUZA, D.; DIAS NETO, J. D. C.; REIS, G.; MARTINS, R.; VALENTE, D.; SOUSA, A.; COSTA, A.; SERUFFO, M. C. R.. Enhancing Satellite-Based Estimations of Carbon Flux in Caxiuanã Forest Using Artificial Intelligence In: International Carbon Dioxide Conference - ICDC11, 2024, Manaus. PRELIMINARY PROGRAM ICDC11. 2024.
 5. DIAS, J. A. C.; TAMASAUSKAS, L. O.; GUIMARAES, P. H. V. ; CORREA, A. B. S.; DIAS NETO, J. D. C.; SANTOS, A. E. C.; COSTA, G. B.; FERNANDES, J. G. S.; VALENTE, D.; MARTINS, R.; SOUSA, A.; COSTA, A.; SERUFFO, M. C. R.. Improving Environmental Monitoring Data Collection in the Amazon by Remote Sensing: A Comparative Analysis of Cloud Detection Models for Sentinel-2. In: International Carbon Dioxide Conference - ICDC11, 2024, Manaus. PRELIMINARY PROGRAM ICDC11. 2024.
 6. GUIMARÃES, PEDRO H. DO V.; TAMASAUSKAS, L. O.; DIAS, J. A. C.; CORREA, A. B. S.; SERUFFO, M. C. R.; FERNANDES, J. G. S.; PIRES, Y. P.; COSTA, G. B.. Análise de Índices Espectrais e Temperatura da Superfície para Estudo do Impacto da Urbanização nas Ilhas de Calor em Castanhal-PA. Artigo Aceito no XXI Simpósio Brasileiro de Sensoriamento Remoto.. , 2025.
 7. DIAS, J. A. C.; GUIMARAES, P. H. V. ; TAMASAUSKAS, L. O.; CORREA, A. B. S.; SERUFFO, M. C. R.; COSTA, G. B.; FERNANDES, J. G. S.; SOUZA, D.. CarbonFluxAI: Previsão de Balanço de Carbono na Floresta Nacional do Tapajós Utilizando Inteligência Artificial. Artigo Aceito no XXI Simpósio Brasileiro de Sensoriamento Remoto. , 2025.

6.3 Agradecimentos à Carbonext

Agradeço à empresa **Carbonext - Nature and Future** (<<https://www.carbonext.com.br>>) pela parceria e apoio por meio do Programa de *Mestrado Acadêmico para Inovação* (MAI/DAI), que viabilizou a continuidade deste trabalho.

A Carbonext, referência no desenvolvimento de projetos voltados à preservação ambiental e à redução de emissões de carbono, atualmente responsável pela conservação de mais de 1,6 milhão de hectares de Selva Amazônica sob risco de desmatamento, proporcionou as condições necessárias para a execução desta pesquisa, unindo ciência, inovação e compromisso com a sustentabilidade.

Referências

- AI, X. et al. A novel spectro-temporal index NDTel combined with machine learning algorithm for sentinel-2 cloud detection. v. 45, n. 12, p. 4023–4045, 2024. ISSN 0143-1161, 1366-5901. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/01431161.2024.2357843>>. Citado 2 vezes nas páginas 5 e 9.
- AYBAR, C. et al. Cloudsen12+: The largest dataset of expert-labeled pixels for cloud and cloud shadow detection in sentinel-2. *Data in Brief*, Elsevier Inc., v. 56, 10 2024. ISSN 23523409. Citado 2 vezes nas páginas 14 e 35.
- AYBAR, C. et al. Cloudsen12, a global dataset for semantic understanding of cloud and cloud shadow in sentinel-2. *Scientific Data*, Nature Research, v. 9, 12 2022. ISSN 20524463. Citado 5 vezes nas páginas 1, 14, 15, 25 e 35.
- AZAD, R. et al. Medical image segmentation review: The success of u-net. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 46, n. 12, p. 10076–10095, 2024. Citado na página 26.
- CANDRA, D. S.; PHINN, S.; SCARTH, P. Cloud and cloud shadow masking for Sentinel-2 using multitemporal images in global area. *International Journal of Remote Sensing*, v. 41, n. 8, p. 2877–2904, abr. 2020. ISSN 0143-1161, 1366-5901. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/01431161.2019.1697006>>. Citado 2 vezes nas páginas 4 e 9.
- CHEN, L.; PAN, J.; ZHANG, Z. A novel cloud detection method based on segmentation prior and multiple features for sentinel-2 images. v. 44, n. 16, p. 5101–5120, 2023. ISSN 0143-1161, 1366-5901. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/01431161.2023.2243022>>. Citado 2 vezes nas páginas 5 e 9.
- CHEN, Y. et al. A novel spectral indices-driven spectral-spatial-context attention network for automatic cloud detection. v. 16, p. 3092–3103, 2023. ISSN 1939-1404, 2151-1535. Disponível em: <<https://ieeexplore.ieee.org/document/10078301/>>. Citado 2 vezes nas páginas 8 e 9.
- CHEN, Y. et al. Cloud and Cloud Shadow Detection Based on Multiscale 3D-CNN for High Resolution Multispectral Imagery. *IEEE Access*, v. 8, p. 16505–16516, 2020. ISSN 2169-3536. Disponível em: <<https://ieeexplore.ieee.org/document/8962046/>>. Citado 2 vezes nas páginas 6 e 9.
- DELEGIDO, J. et al. Evaluation of sentinel-2 red-edge bands for empirical estimation of green lai and chlorophyll content. *Sensors*, v. 11, p. 7063–7081, 7 2011. ISSN 14248220. Citado na página 12.
- DERRY, A.; KRZYWINSKI, M.; ALTMAN, N. *Convolutional neural networks*. [S.l.]: Nature Research, 2023. 1269-1270 p. Citado na página 18.
- DOMNICH, M. et al. KappaMask: AI-Based Cloudmask Processor for Sentinel-2. *Remote Sensing*, v. 13, n. 20, p. 4100, out. 2021. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/13/20/4100>>. Citado 3 vezes nas páginas 7, 9 e 36.

- ESA. *Sentinel-2 User Handbook*. 2015. Accessed: 2023-08-01. Disponível em: <https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook>. Citado na página 1.
- EUROPEAN SPACE AGENCY. *Sentinel-2 Mission Overview*. 2025. Acessado em: 2025-01-23. Disponível em: <<https://sentiwiki.copernicus.eu/web/s2-mission>>. Citado na página 11.
- FALCON, W.; The PyTorch Lightning team. *PyTorch Lightning*. 2019. Disponível em: <<https://github.com/Lightning-AI/lightning>>. Citado na página 26.
- FRANCIS, A. Sensor Independent Cloud and Shadow Masking With Partial Labels and Multimodal Inputs. *IEEE Transactions on Geoscience and Remote Sensing*, v. 62, p. 1–18, 2024. ISSN 0196-2892, 1558-0644. Disponível em: <<https://ieeexplore.ieee.org/document/10505181/>>. Citado na página 35.
- HU, K.; ZHANG, D.; XIA, M. CDUNet: Cloud Detection UNet for Remote Sensing Imagery. *Remote Sensing*, v. 13, n. 22, p. 4533, nov. 2021. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/13/22/4533>>. Citado 2 vezes nas páginas 7 e 9.
- IAKUBOVSKII, P. *Segmentation Models Pytorch*. [S.l.]: GitHub, 2019. <https://github.com/qubvel/segmentation_models_pytorch>. Citado 2 vezes nas páginas 20 e 26.
- KUO, C. C. Understanding convolutional neural networks with a mathematical model. *Journal of Visual Communication and Image Representation*, Academic Press Inc., v. 41, p. 406–413, 11 2016. ISSN 10959076. Citado na página 18.
- LI, Y. et al. Accurate cloud detection in high-resolution remote sensing imagery by weakly supervised deep learning. *Remote Sensing of Environment*, v. 250, p. 112045, dez. 2020. ISSN 00344257. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0034425720304156>>. Citado 2 vezes nas páginas 6 e 9.
- LI, Z. et al. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, Institute of Electrical and Electronics Engineers Inc., v. 33, p. 6999–7019, 12 2022. ISSN 21622388. Citado 2 vezes nas páginas 8 e 9.
- LIANG, K. et al. A novel method for cloud and cloud shadow detection based on the maximum and minimum values of sentinel-2 time series images. v. 16, n. 8, p. 1392, 2024. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/16/8/1392>>. Citado 2 vezes nas páginas 5 e 9.
- LIU, X. et al. Multistage progressive interactive fusion network for sentinel-2: High resolution for all bands. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 16, p. 10191–10202, 2023. ISSN 21511535. Citado na página 12.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. 11 2014. Disponível em: <<http://arxiv.org/abs/1411.4038>>. Citado 3 vezes nas páginas 2, 17 e 18.
- LÓPEZ-PUIGDOLLERS, D. et al. Benchmarking Deep Learning Models for Cloud Detection in Landsat-8 and Sentinel-2 Images. *Remote Sensing*, v. 13, n. 5, p. 992, mar. 2021. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/13/5/992>>. Citado 4 vezes nas páginas 1, 7, 9 e 36.

MAIN-KNORN, M. et al. Sen2cor for sentinel-2. In: . [S.l.]: SPIE-Intl Soc Optical Eng, 2017. p. 3. ISBN 9781510613188. ISSN 1996756X. Citado 2 vezes nas páginas 14 e 36.

PASZKE, A. et al. Pytorch: An imperative style, high-performance deep learning library. *CoRR*, abs/1912.01703, 2019. Disponível em: <<http://arxiv.org/abs/1912.01703>>. Citado na página 26.

PERSSON, M.; LINDBERG, E.; REESE, H. Tree species classification with multi-temporal sentinel-2 data. *Remote Sensing*, MDPI AG, v. 10, 11 2018. ISSN 20724292. Citado na página 12.

PFLUG, B. et al. Next updates of atmospheric correction processor sen2cor. In: . [S.l.]: SPIE-Intl Soc Optical Eng, 2020. p. 2. ISBN 9781510638792. ISSN 1996756X. Citado na página 14.

PHIRI, D. et al. Sentinel-2 data for land cover/use mapping: A review. *Remote. Sens.*, v. 12, p. 2291, 2020. Citado 2 vezes nas páginas 11 e 12.

QIU, S.; ZHU, Z.; HE, B. Fmask 4.0: Improved cloud and cloud shadow detection in landsats 4–8 and sentinel-2 imagery. *Remote Sensing of Environment*, v. 231, p. 111205, 2019. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425719302172>>. Citado na página 36.

REVILL, A. et al. The value of sentinel-2 spectral bands for the assessment of winter wheat growth and development. *Remote Sensing*, MDPI AG, v. 11, 2019. ISSN 20724292. Citado na página 12.

RONNEBERGER, O.; FISCHER, P.; BROX, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. arXiv, 2015. Disponível em: <<http://arxiv.org/abs/1505.04597>>. Citado 2 vezes nas páginas 2 e 19.

SALGUEIRO, L.; MARCELLO, J.; VILAPLANA, V. Single-image super-resolution of sentinel-2 low resolution bands with residual dense convolutional neural networks. *Remote Sensing*, MDPI, v. 13, 12 2021. ISSN 20724292. Citado na página 12.

SANDLER, M. et al. Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation. *CoRR*, abs/1801.04381, 2018. Disponível em: <<http://arxiv.org/abs/1801.04381>>. Citado 3 vezes nas páginas 2, 20 e 26.

SIDDIQUE, N. et al. U-net and its variants for medical image segmentation: A review of theory and applications. v. 9, p. 82031–82057, 2021. ISSN 2169-3536. Disponível em: <<https://ieeexplore.ieee.org/document/9446143/>>. Citado na página 20.

SINGH, R.; BISWAS, M.; PAL, M. Cloud detection using sentinel 2 imageries: a comparison of XGBoost, RF, SVM, and CNN algorithms. *Geocarto International*, v. 38, n. 1, p. 1–32, dez. 2023. ISSN 1010-6049, 1752-0762. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/10106049.2022.2146211>>. Citado 2 vezes nas páginas 4 e 9.

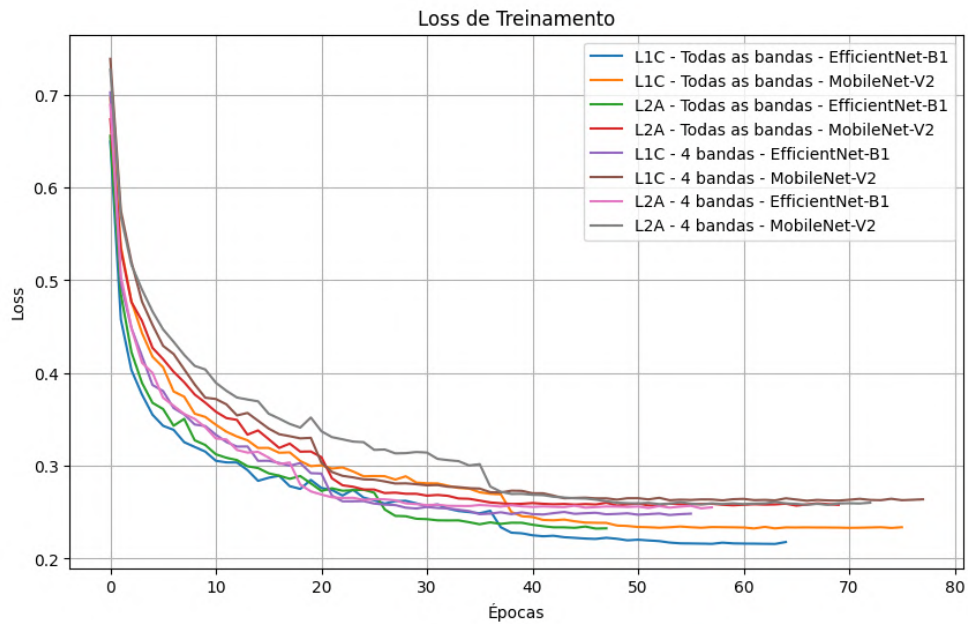
SKAKUN, S. et al. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. *Remote Sensing of Environment*, v. 274, p. 112990, jun. 2022. ISSN 00344257. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0034425722001043>>. Citado 3 vezes nas páginas 4, 9 e 36.

- SOLA, I. et al. Assessment of atmospheric correction methods for sentinel-2 images in mediterranean landscapes. *Int. J. Appl. Earth Obs. Geoinformation*, v. 73, p. 63–76, 2018. Citado 2 vezes nas páginas 13 e 14.
- SOLÓRZANO, J. V. et al. Land use land cover classification with u-net: Advantages of combining sentinel-1 and sentinel-2 imagery. v. 13, n. 18, p. 3600, 2021. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/13/18/3600>>. Citado 2 vezes nas páginas 20 e 26.
- TAN, M.; LE, Q. V. *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. arXiv, 2020. ArXiv:1905.11946 [cs]. Disponível em: <<http://arxiv.org/abs/1905.11946>>. Citado 3 vezes nas páginas 2, 20 e 26.
- TARRIO, K. et al. Comparison of cloud detection algorithms for Sentinel-2 imagery. *Science of Remote Sensing*, v. 2, p. 100010, dez. 2020. ISSN 26660172. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S2666017220300092>>. Citado na página 1.
- USS, M. et al. Analysis of signal-dependent sensor noise on jpeg 2000-compressed sentinel-2 multi-spectral images. *SPIE-Intl Soc Optical Eng*, v. 10427, 2017. Citado na página 13.
- VUOLO, F. et al. Data service platform for sentinel-2 surface reflectance and value-added products: System use and examples. *Remote Sensing*, MDPI AG, v. 8, 11 2016. ISSN 20724292. Citado na página 14.
- WANG, Q. et al. Fusion of sentinel-2 images. *Remote Sensing of Environment*, Elsevier Inc., v. 187, p. 241–252, 12 2016. ISSN 00344257. Citado na página 12.
- ZHU, Q.; ZU, X. Fully convolutional neural network structure and its loss function for image classification. *IEEE Access*, Institute of Electrical and Electronics Engineers Inc., v. 10, p. 35541–35549, 2022. ISSN 21693536. Citado 2 vezes nas páginas 17 e 18.
- ZHU, Z.; WOODCOCK, C. E. Continuous change detection and classification of land cover using all available landsat data. *Remote Sensing of Environment*, v. 144, p. 152–171, 2014. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425714000248>>. Citado na página 1.

Apêndices

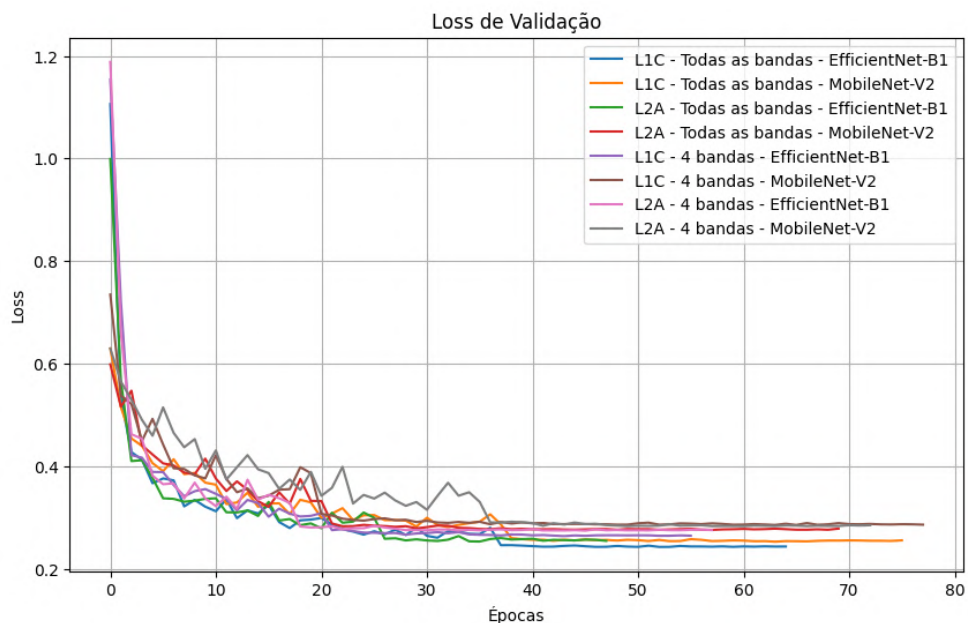
A Métricas de Treinamento dos Modelos Propostos

Figura 10 – Gráfico da função de perda durante o treinamento dos modelos.



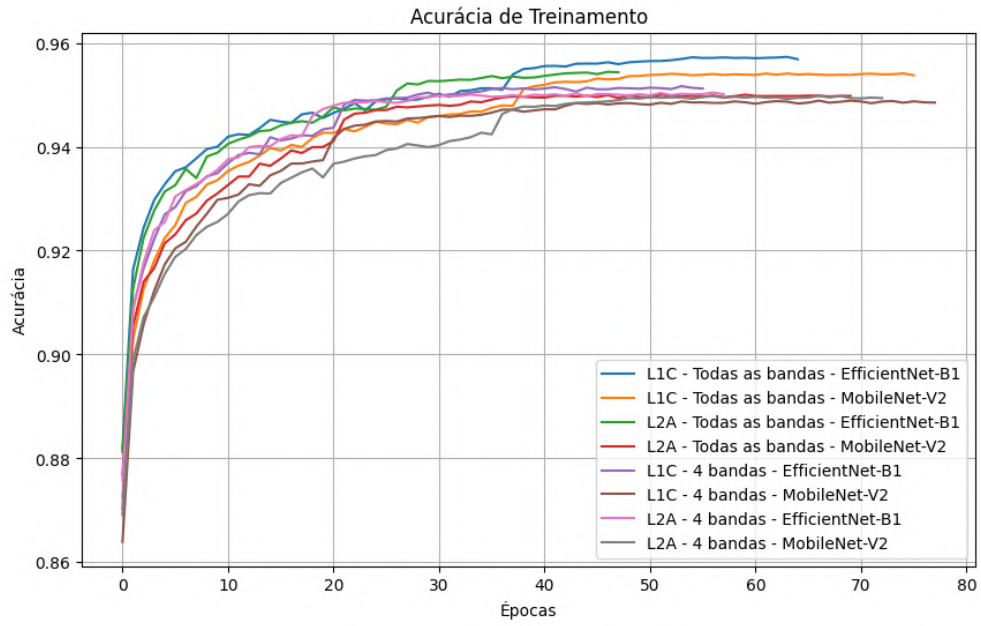
Fonte: O Autor

Figura 11 – Gráfico da função de perda durante a validação dos modelos.



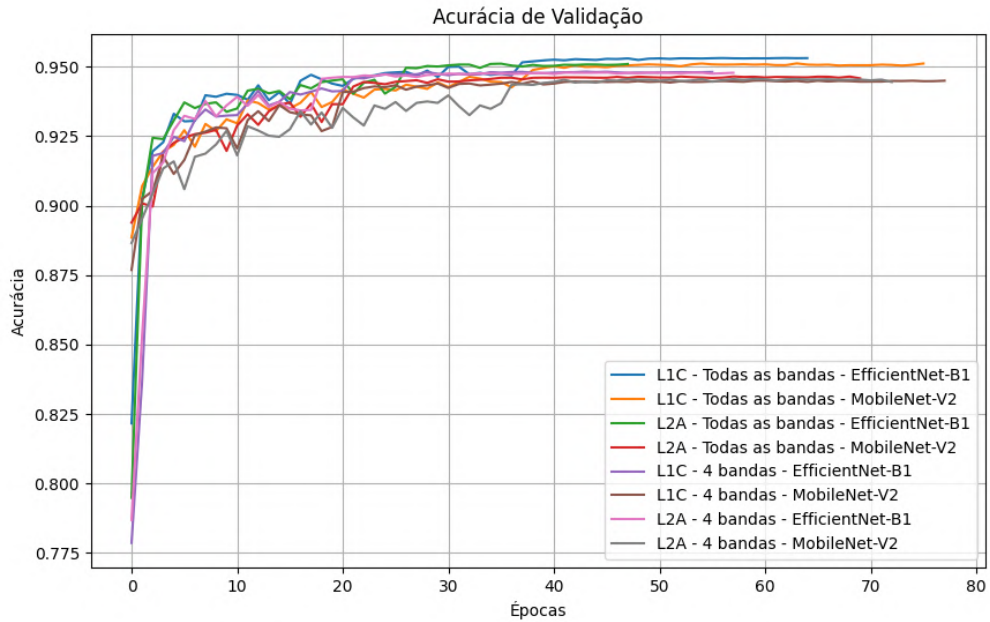
Fonte: O Autor

Figura 12 – Gráfico da acurácia durante o treinamento dos modelos.



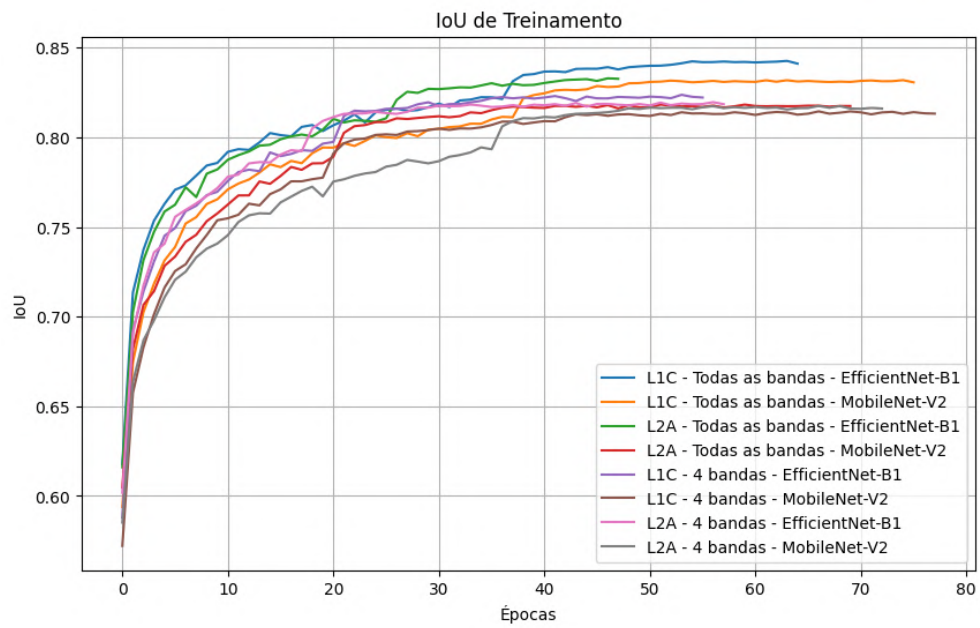
Fonte: O Autor

Figura 13 – Gráfico da acurácia durante a validação dos modelos.



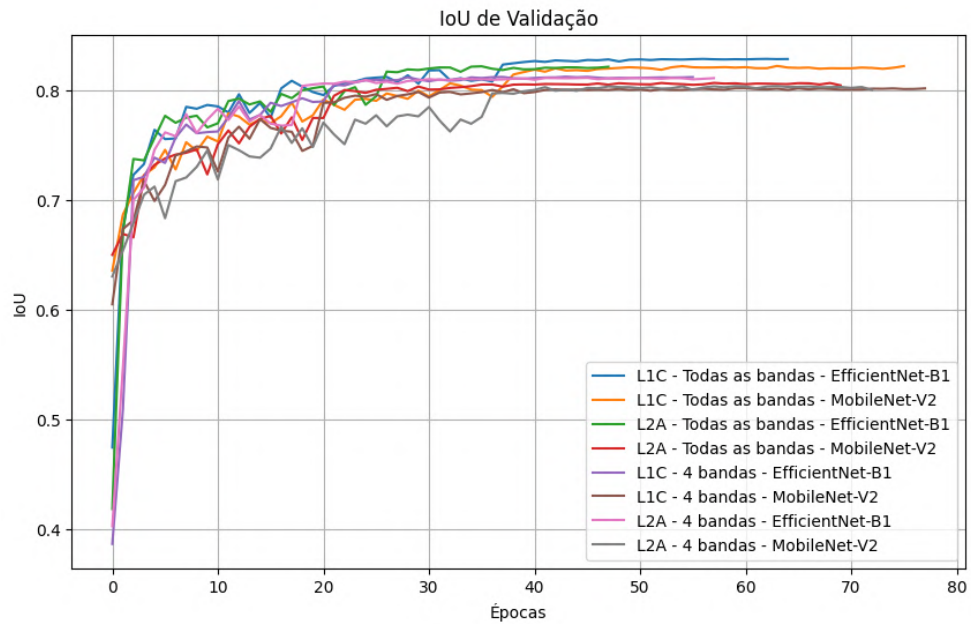
Fonte: O Autor

Figura 14 – Gráfico do IoU durante o treinamento dos modelos.



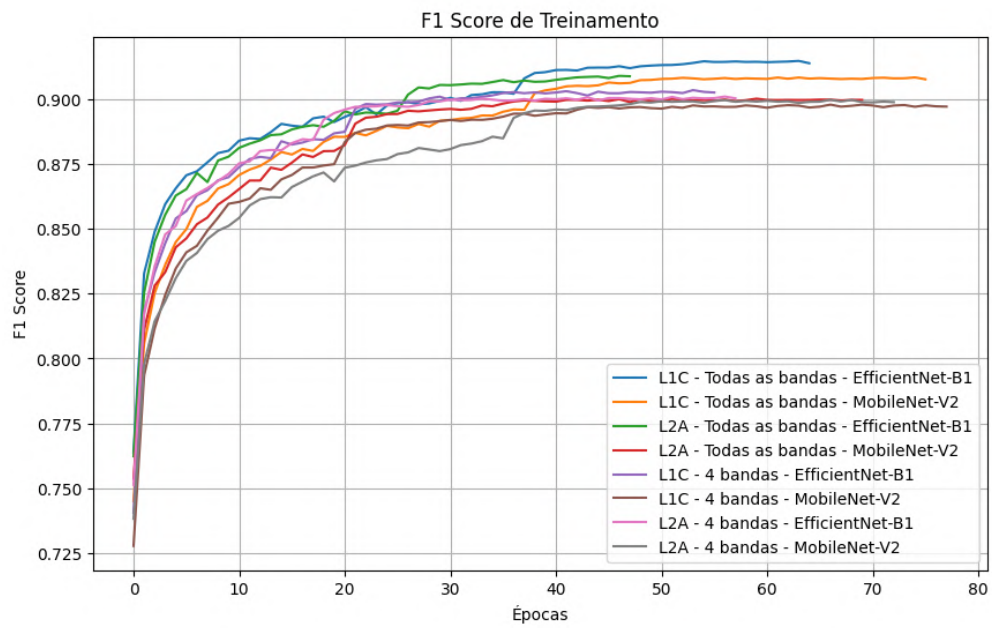
Fonte: O Autor

Figura 15 – Gráfico do IoU durante a validação dos modelos.



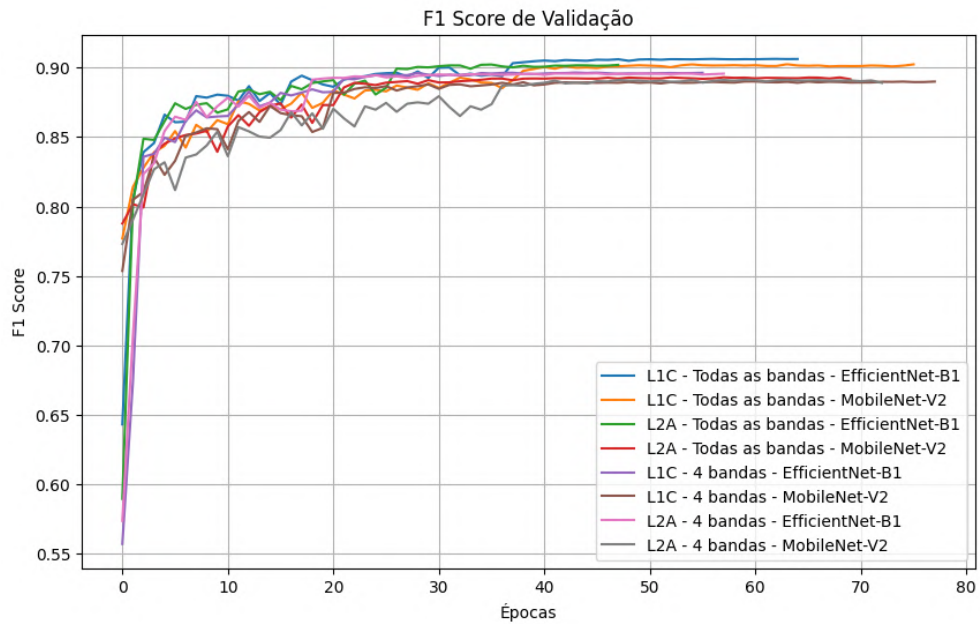
Fonte: O Autor

Figura 16 – Gráfico do F1-Score durante o treinamento dos modelos.



Fonte: O Autor

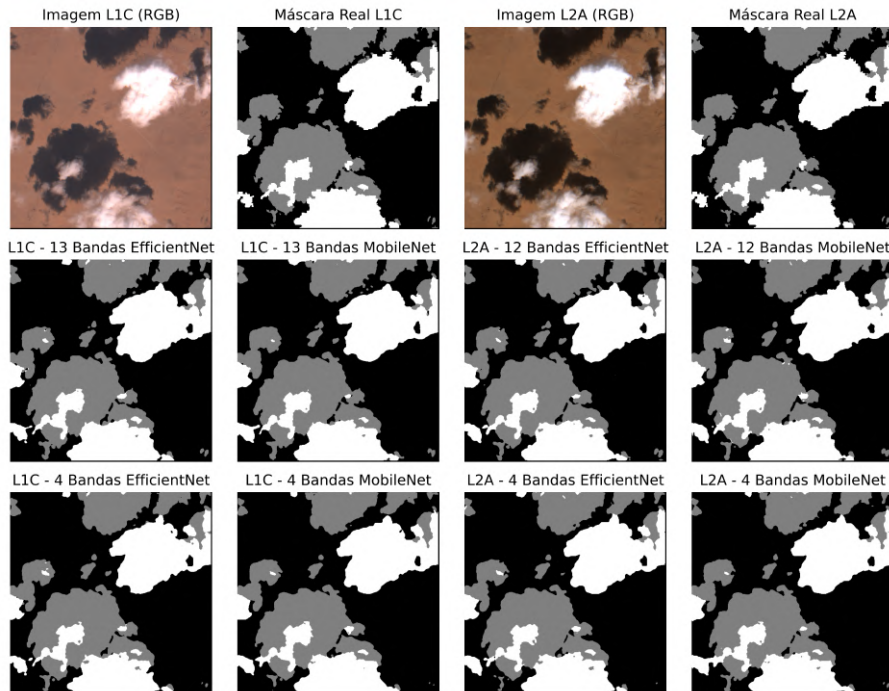
Figura 17 – Gráfico do F1-Score durante a validação dos modelos.



Fonte: O Autor

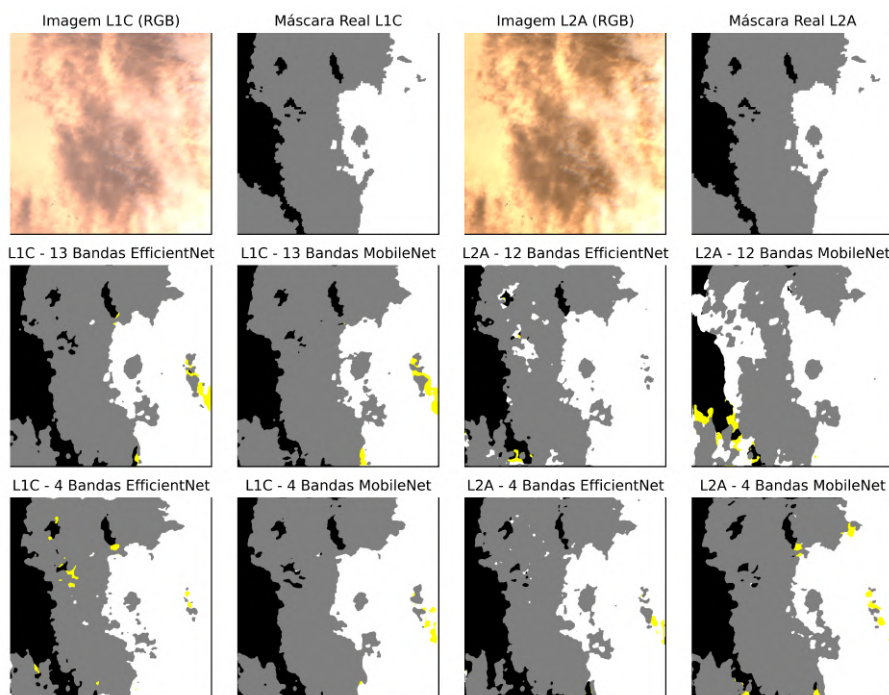
B Comparação dos Modelos Propostos

Figura 18 – Máscaras dos modelos propostos em imagem com média cobertura de nuvens.



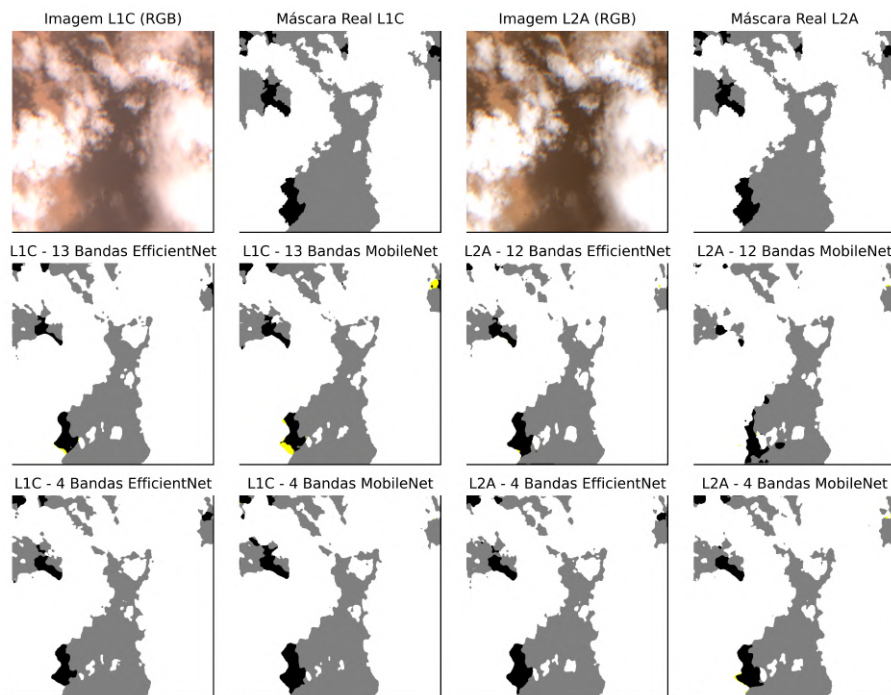
Fonte: O Autor

Figura 19 – Máscaras dos modelos propostos em imagem de difícil anotação de classes.



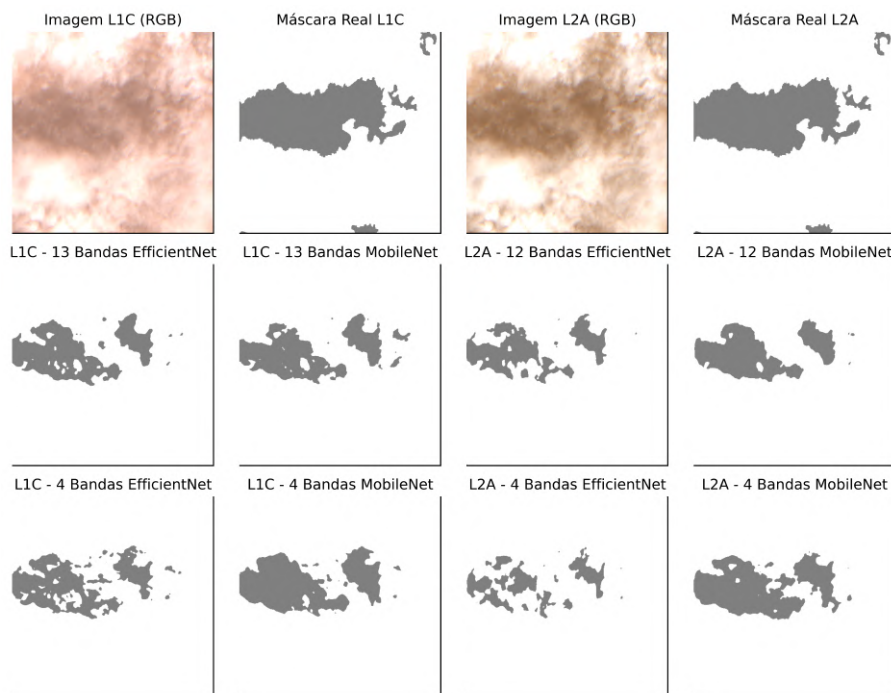
Fonte: O Autor

Figura 20 – Máscaras dos modelos propostos em imagem de alta cobertura de nuvens e sombras.



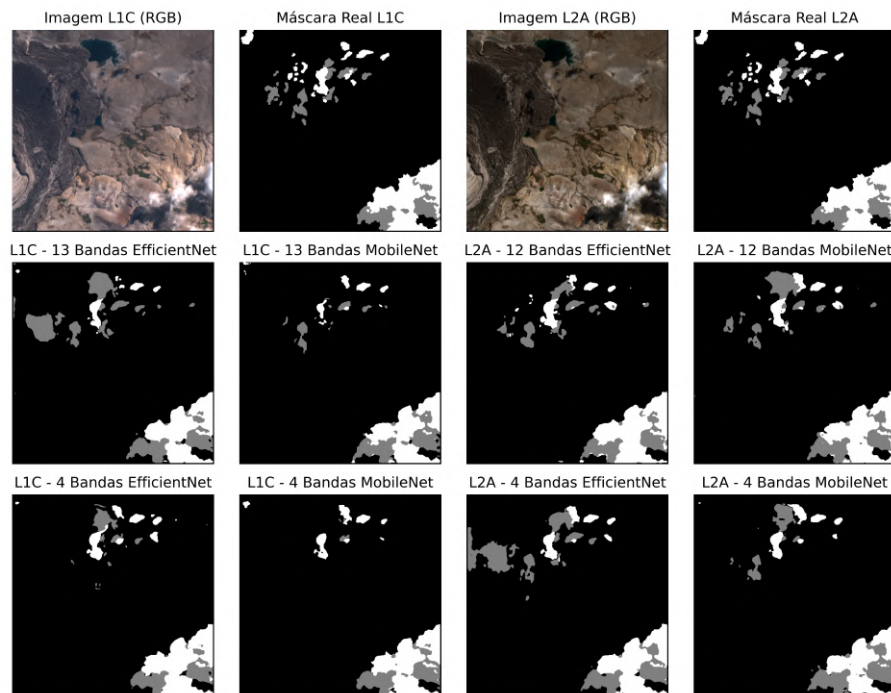
Fonte: O Autor

Figura 21 – Máscaras dos modelos propostos em imagem totalmente coberta por nuvens e sombras.



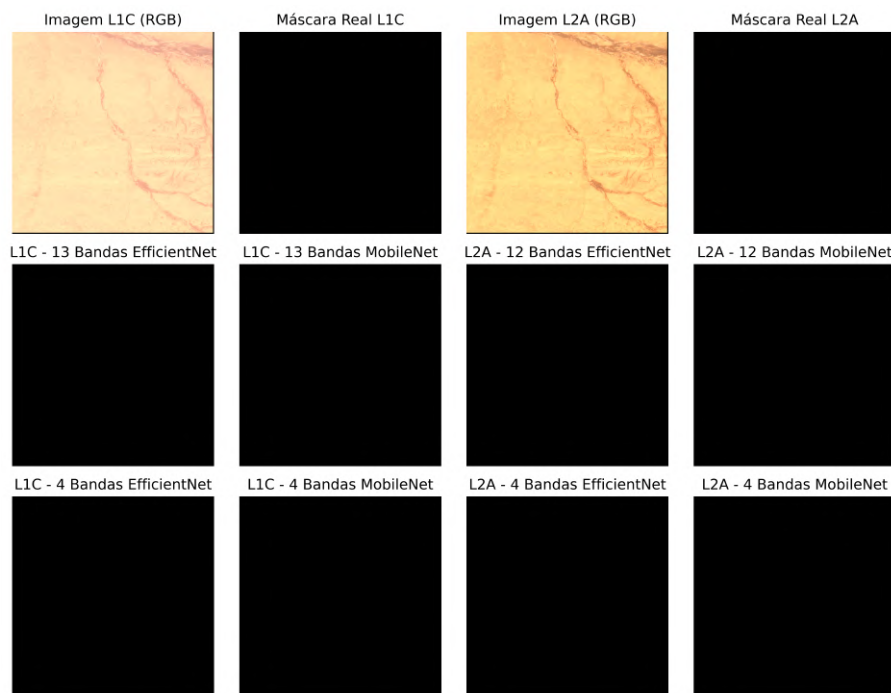
Fonte: O Autor

Figura 22 – Máscaras dos modelos propostos em imagem com baixa cobertura de nuvens e sombras.



Fonte: O Autor

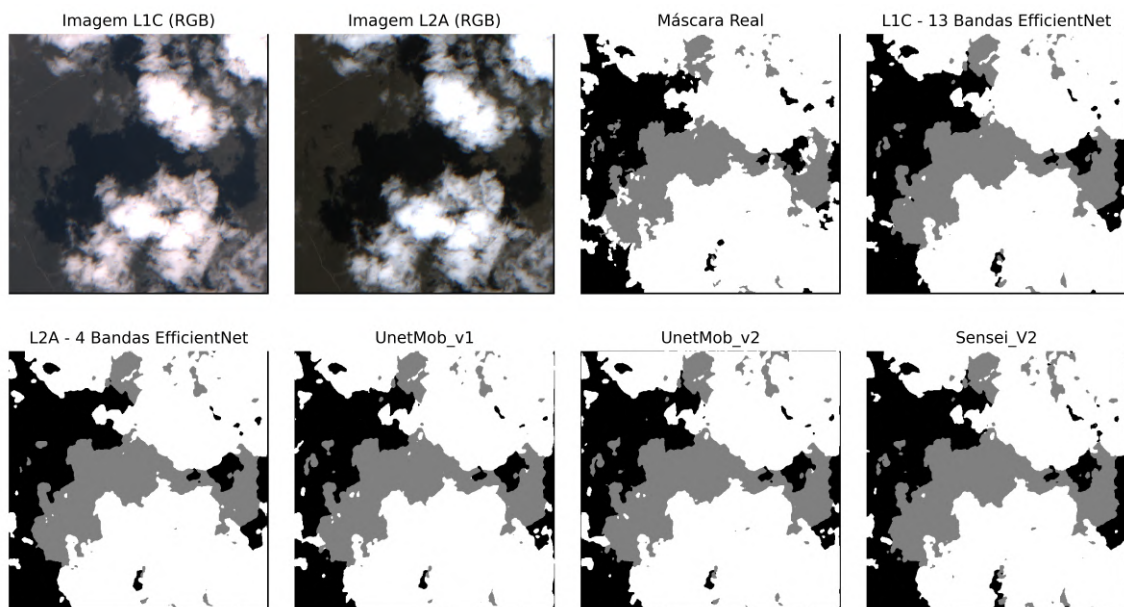
Figura 23 – Máscaras dos modelos propostos em imagem sem nuvens e sombras.



Fonte: O Autor

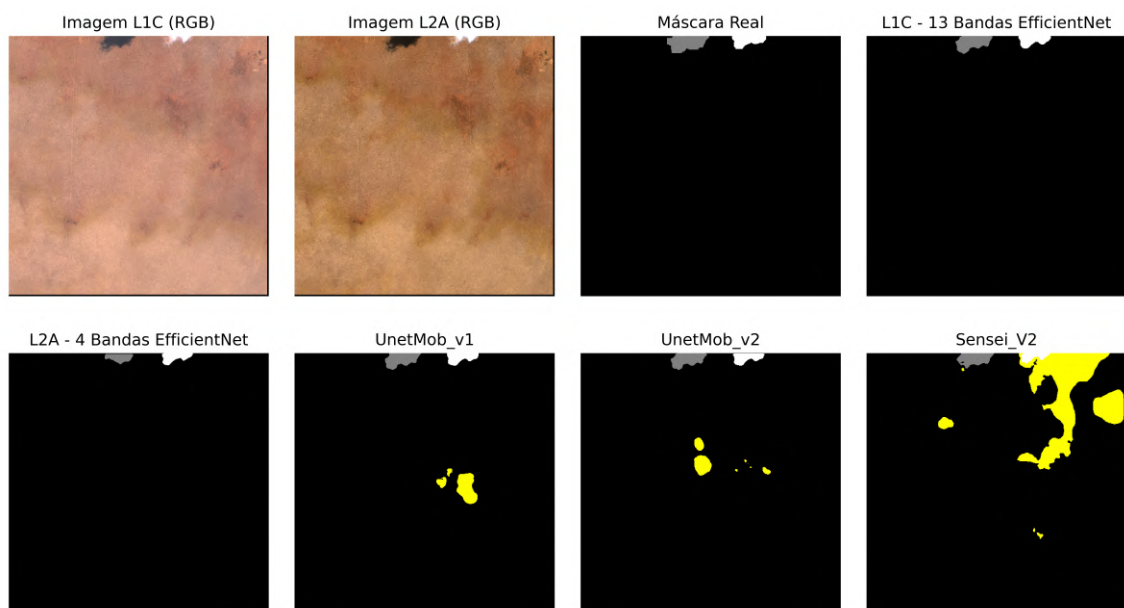
C Comparação dos Modelos Treinados com Classes Iguais

Figura 24 – Máscaras dos modelos com classes iguais em imagem com média cobertura de nuvens.



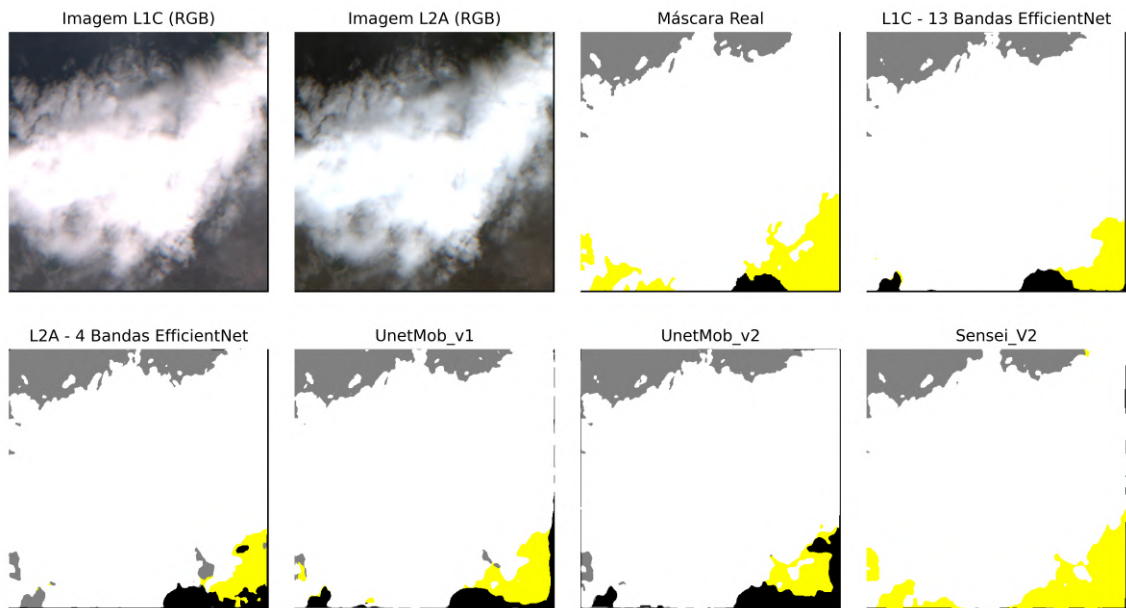
Fonte: O Autor

Figura 25 – Máscaras dos modelos com classes iguais em imagem quase sem cobertura de nuvens.



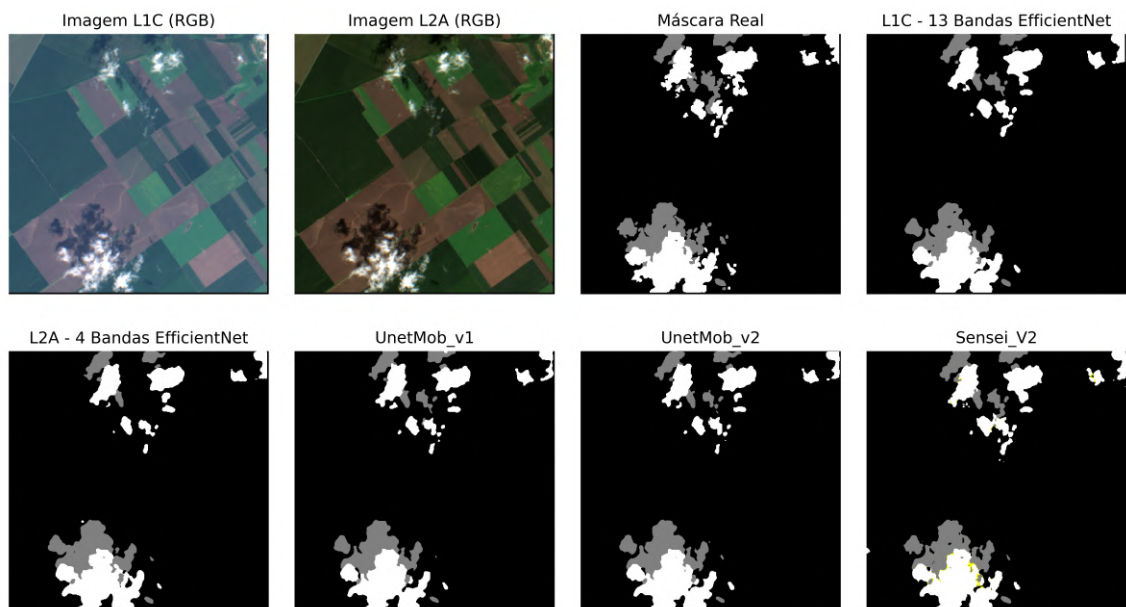
Fonte: O Autor

Figura 26 – Máscaras dos modelos com classes iguais em imagem com alta cobertura de nuvens.



Fonte: O Autor

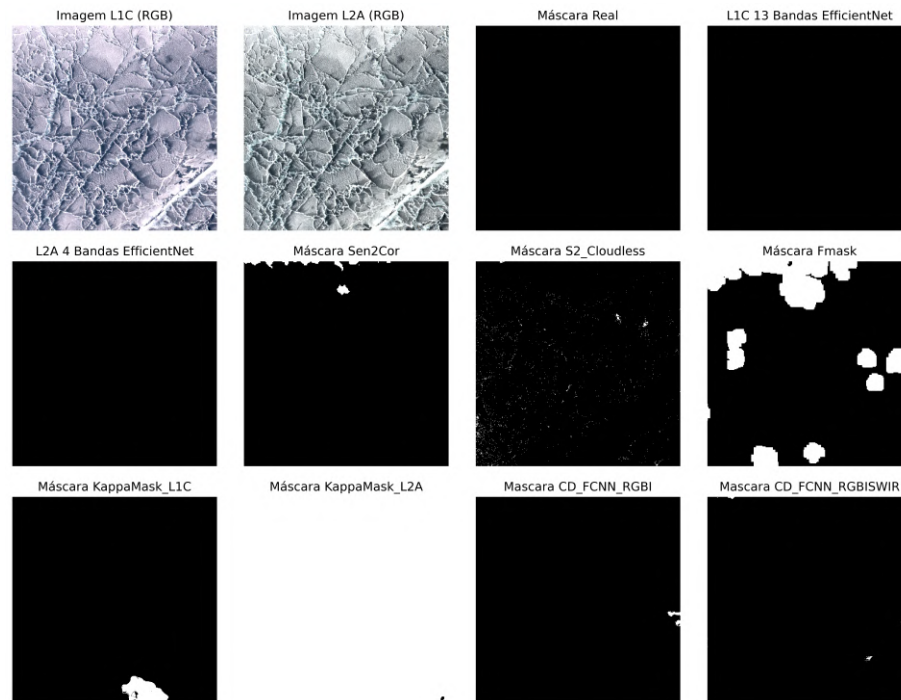
Figura 27 – Máscaras dos modelos com classes iguais em imagem com baixa cobertura de nuvens.



Fonte: O Autor

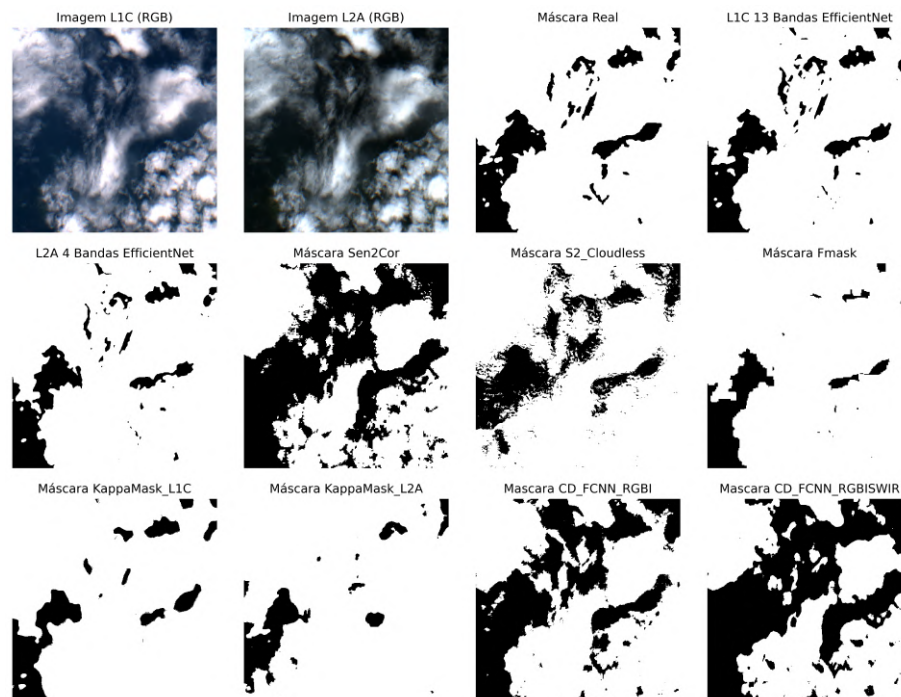
D Comparação entre Modelos com a Superclasse Nuvem vs Não Nuvem

Figura 28 – Máscaras dos modelos em imagem sem nuvens, porém com superfície brilhante.



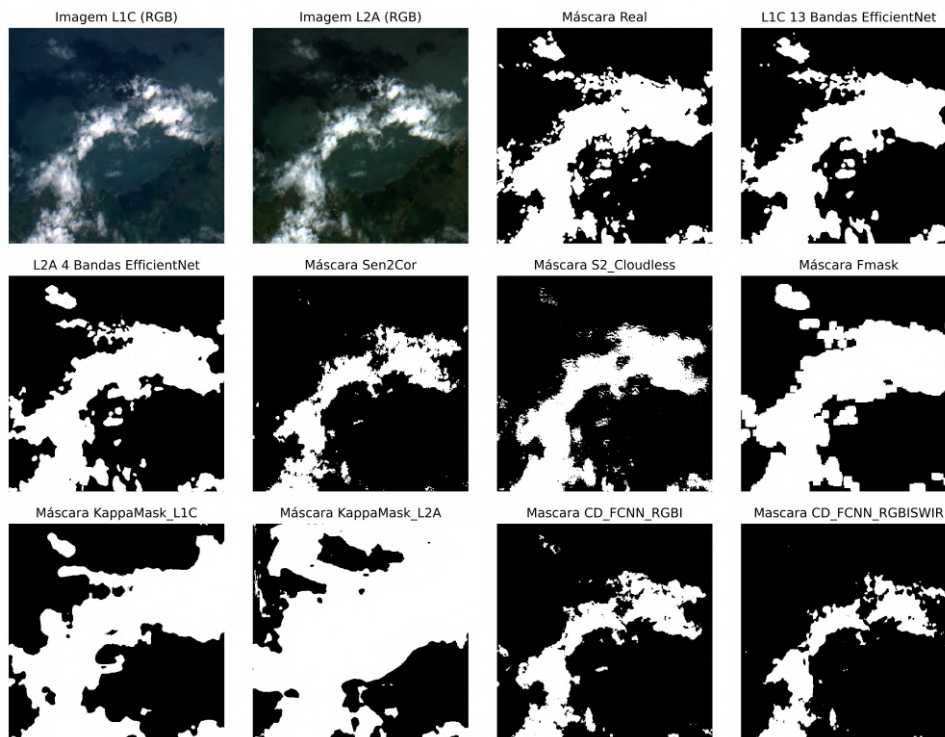
Fonte: O Autor

Figura 29 – Máscaras dos modelos em imagem com alta cobertura de nuvens.



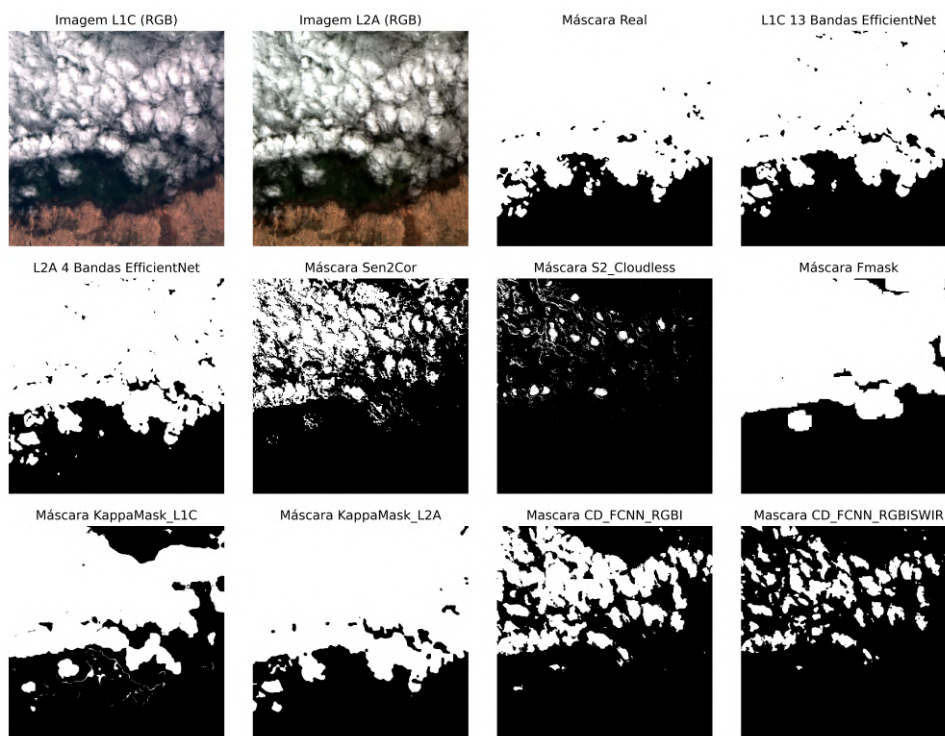
Fonte: O Autor

Figura 30 – Máscaras dos modelos em imagem com média cobertura de nuvens.



Fonte: O Autor

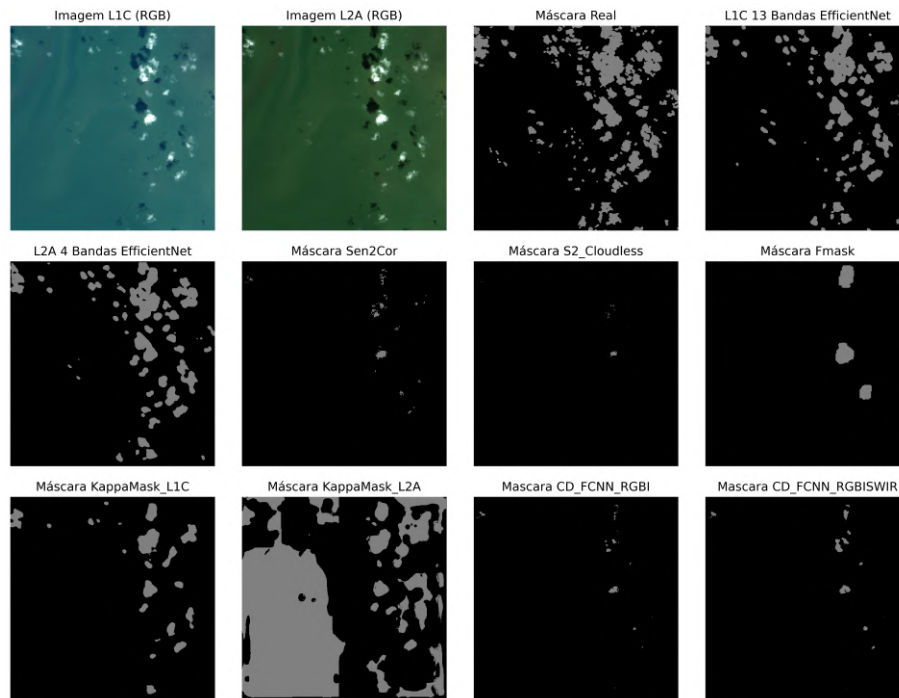
Figura 31 – Máscaras dos modelos em imagem com alta cobertura de nuvens.



Fonte: O Autor

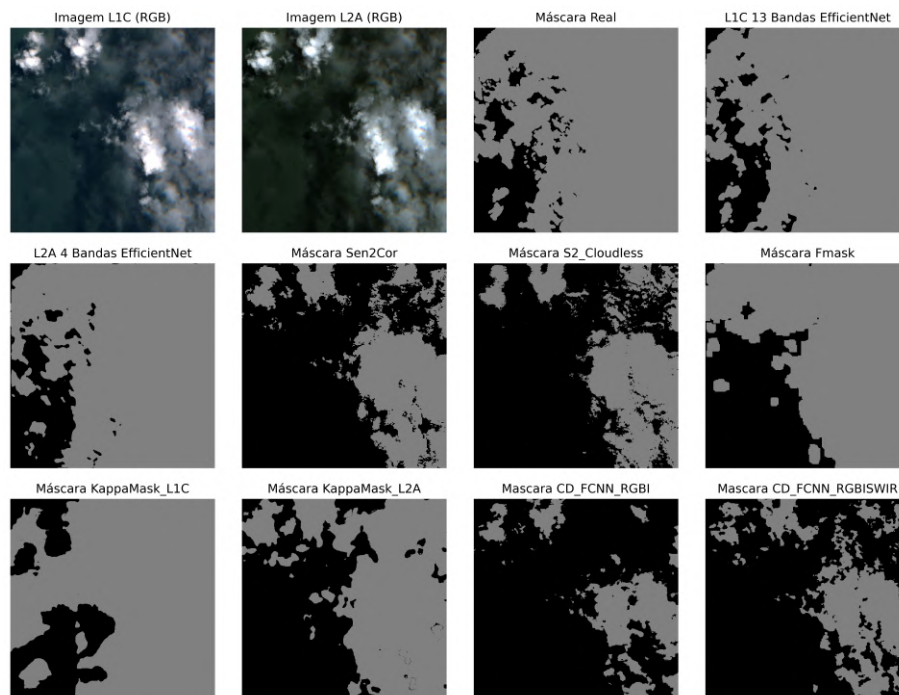
E Comparação entre Modelos com a Superclasse Válido vs Inválido

Figura 32 – Máscaras dos modelos em imagem com baixo percentual de pixels inválidos.



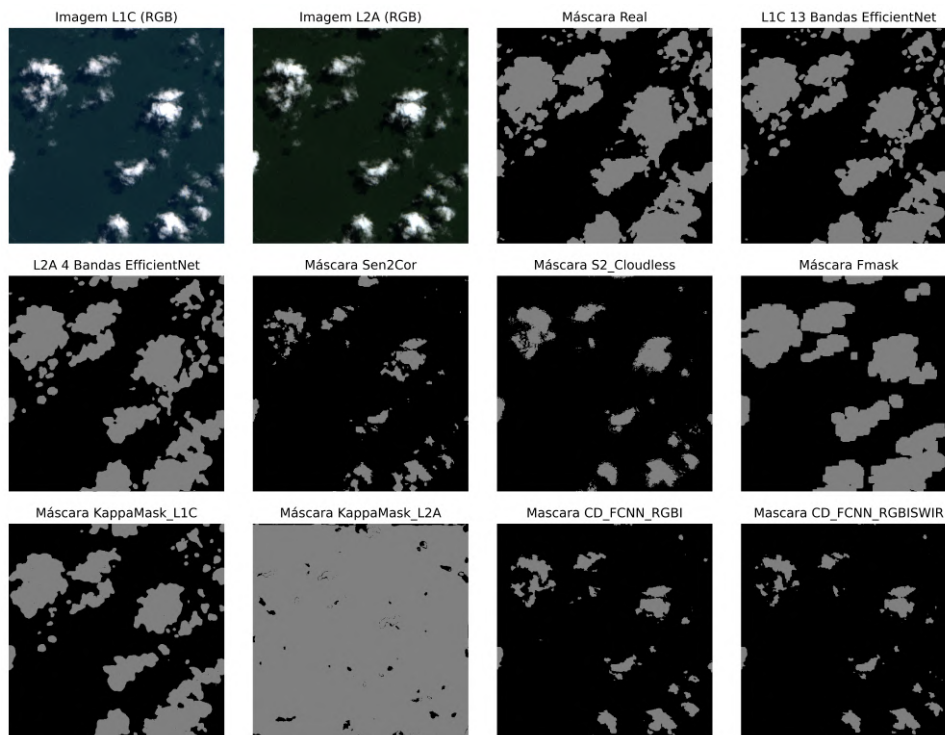
Fonte: O Autor

Figura 33 – Máscaras dos modelos em imagem com alto percentual de pixels inválidos.



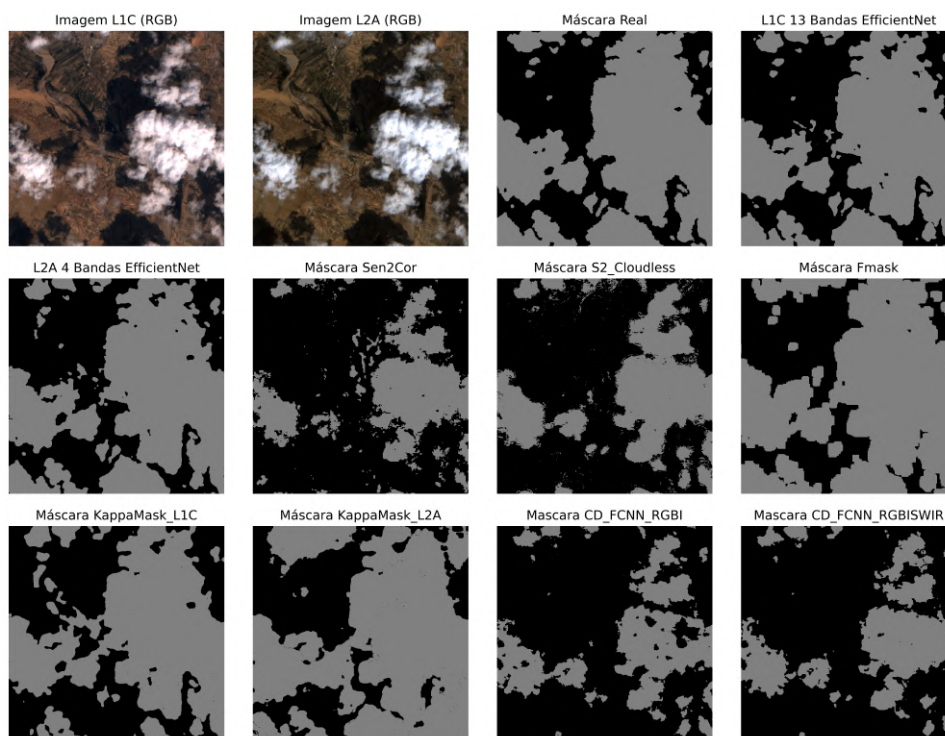
Fonte: O Autor

Figura 34 – Máscaras dos modelos em imagem com médio percentual de pixels inválidos.



Fonte: O Autor

Figura 35 – Máscaras dos modelos em imagem com alto percentual de pixels inválidos.



Fonte: O Autor