



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE SOBRAL
CURSO DE GRADUAÇÃO EM ENGENHARIA DA COMPUTAÇÃO

SAMUEL GOMES RIBEIRO

**SEGMENTAÇÃO SEMÂNTICA DE INUNDAÇÕES EM IMAGENS DE SATÉLITE POR
MEIO DE REDES NEURAS TOTALMENTE CONVOLUCIONAIS**

SOBRAL

2023

SAMUEL GOMES RIBEIRO

SEGMENTAÇÃO SEMÂNTICA DE INUNDAÇÕES EM IMAGENS DE SATÉLITE POR
MEIO DE REDES NEURAIS TOTALMENTE CONVOLUCIONAIS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia da Computação do Campus de Sobral da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia da Computação.

Orientador: Prof. Dr. Marcelo Marques Simões de Souza.

SOBRAL

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R372s Ribeiro, Samuel Gomes.

Segmentação Semântica de Inundações em Imagens de Satélite por Meio de Redes Neurais Totalmente Convolucionais / Samuel Gomes Ribeiro. – 2023.
46 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Sobral, Curso de Engenharia da Computação, Sobral, 2023.

Orientação: Prof. Dr. Marcelo Marques Simões de Souza.

1. Redes Neurais. 2. Segmentação Semântica. 3. Radar de Abertura Sintética. 4. Inundação. 5. Funções de Perda. I. Título.

CDD 621.39

SAMUEL GOMES RIBEIRO

SEGMENTAÇÃO SEMÂNTICA DE INUNDAÇÕES EM IMAGENS DE SATÉLITE POR
MEIO DE REDES NEURAS TOTALMENTE CONVOLUCIONAIS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia da Computação do Campus de Sobral da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia da Computação.

Aprovada em: 06 DE DEZEMBRO DE 2023

BANCA EXAMINADORA

Prof. Dr. Marcelo Marques Simões de
Souza (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Iális Cavalcante de Paula Júnior
Universidade Federal do Ceará (UFC)

Alan Marques da Rocha
Programa de Pós Graduação em Engenharia Elétrica e
Computação (PPGEEC)

Dedico este trabalho à minha família, que sempre acreditou em mim e me apoiou incondicionalmente.

AGRADECIMENTOS

Agradeço ao meu pai, Francisco Antonio, à minha mãe, Daniele, e ao meu irmão, Cauan, pelo incentivo e apoio constante durante a minha graduação.

À minha namorada e melhor amiga Vitória, pelo companheirismo e compreensão durante os momentos em que precisei me ausentar para me dedicar à elaboração deste trabalho.

Ao Prof. Dr. Marcelo Marques Simões de Souza pelo tempo dedicado na excelente orientação em todas as etapas deste trabalho.

À banca examinadora, composta pelo Prof. Dr. Iális Cavalcante de Paula Júnior e pelo Me. Alan Marques da Rocha, expressei minha gratidão pelas preciosas colaborações e sugestões que enriqueceram este trabalho.

Aos amigos da graduação, Ailton Guarinho, Ananda Barros, Daniel Araújo, Jannielly Garcia, Francisco Paiva, Levi Frota, Klayver Ximenes, Lucas Freitas, Mateus Soares e Vitor Oliveira, cujas contribuições, reflexões, críticas e sugestões no âmbito pessoal e profissional foram inestimáveis.

"Se eu vi mais longe, foi por estar sobre ombros
de gigantes"

(Sir Isaac Newton)

RESUMO

Os desastres naturais representam um desafio global, acarretando danos significativos em diversos países. Especificamente, as inundações são eventos extremos que geram perdas de vidas, danos à infraestrutura e prejuízos econômicos substanciais. O sensoriamento remoto via radar oferece uma perspectiva valiosa para identificar áreas inundadas, devido à sua sensibilidade à presença de água. As Redes Neurais Artificiais (RNA) têm sido utilizadas com sucesso para detectar inundações em imagens de radar, graças à capacidade de aprender padrões complexos presentes nessas imagens. Este estudo propõe o desenvolvimento de uma rede neural utilizando as arquiteturas U-Net e U-Net++ para segmentar semanticamente áreas inundadas em imagens de radar. Foram exploradas diferentes funções de perda, como *Binary Cross-Entropy*, *Dice* e *Tversky* ($\alpha = \beta = 1$), visando otimizar a precisão da segmentação. Os experimentos indicaram que a U-Net++ com *Binary Cross-Entropy* se destacou ao analisar os resultados preditivos dos modelos.

Palavras-chave: Redes Neurais; Segmentação Semântica; Radar de Abertura Sintética; Inundação; Funções de Perda.

ABSTRACT

Natural disasters pose a global challenge, causing significant damage across various countries. Floods are extreme events leading to loss of lives, infrastructure damage, and substantial economic losses. Remote sensing through radar offers valuable insight into identifying flooded areas due to its sensitivity to water presence. Artificial neural networks have successfully been used to detect floods in radar images, thanks to their ability to learn complex patterns within these images. This study proposes the development of a neural network utilizing U-Net and U-Net++ architectures to semantically segment flooded areas in radar images. To optimize segmentation accuracy, various loss functions such as Binary Cross-Entropy, Dice, and Tversky ($\alpha = \beta = 1$) were explored. The experiments indicated that U-Net++ with Binary Cross-Entropy stood out when analyzing the predictive results of the models.

Palavras-chave: Neural Networks; Semantic Segmentation; Synthetic Aperture Radar; Flood; Loss Functions.

LISTA DE FIGURAS

Figura 1 – Distribuição de ocorrências de inundações no Brasil de 2000 a 2023	19
Figura 2 – Configuração das antenas do Seasat-1	20
Figura 3 – Espectro eletromagnético	21
Figura 4 – Retroespalhamento em função do comprimento de onda	23
Figura 5 – Sistema de polarização única	24
Figura 6 – Arquitetura típica de uma rede convolucional	25
Figura 7 – Operação de convolução	26
Figura 8 – Operação de <i>max-pooling</i>	27
Figura 9 – Imagens do satélite <i>Sentinel-1</i>	29
Figura 10 – Imagens do satélite <i>Sentinel-2</i>	30
Figura 11 – Pré-processamento das imagens do <i>Sentinel-1</i>	31
Figura 12 – Imagens de radar recortadas do conjunto <i>C2S-MS Floods</i>	32
Figura 13 – Representação visual da segmentação semântica e de instâncias	33
Figura 14 – Arquitetura das redes U-Net e U-Net++	34
Figura 15 – Arquitetura da rede ResNet-18	35
Figura 16 – Evolução das perdas de treinamento e validação do melhor modelo	41
Figura 17 – Evolução das métricas de desempenho do melhor modelo	42

LISTA DE TABELAS

Tabela 1 – Classificação de desastres naturais	17
Tabela 2 – Bandas de microondas típicas no sensoriamento remoto	23
Tabela 3 – Desempenhos de validação e teste utilizando a função de perda <i>Binary Cross-Entropy</i>	39
Tabela 4 – Desempenhos de validação e teste utilizando a função de perda <i>Dice</i>	39
Tabela 5 – Desempenhos de validação e teste utilizando a função de perda <i>Tversky</i>	40
Tabela 6 – Predições dos modelos U-Net e U-Net++ para cada função de perda	41

LISTA DE ABREVIATURAS E SIGLAS

ALT	<i>Radar Altimeter</i>
CD	<i>Change Detection</i>
CEMADEN	Centro Nacional de Monitoramento e Alertas de Desastres Naturais
CEMS	<i>Copernicus Emergency Management Service</i>
CNN	<i>Convolutional Neural Network</i>
ESA	<i>European Space Agency</i>
FCL	<i>Fully Connected Layers</i>
FCNN	<i>Fully Convolutional Neural Network</i>
FN	Falsos Negativos
FP	Falsos Positivos
GPU	<i>Graphics Processing Units</i>
GRU	<i>Gated Recurrent Unit</i>
IA	Inteligência Artificial
IoU	<i>Intersection over Union</i>
MDE	Modelos Digitais de Elevação
NASA	<i>National Aeronautics and Space Administration</i>
ReLU	<i>Rectified Linear Unit</i>
RGB	<i>Red Green Blue</i>
RNA	Redes Neurais Artificiais
S1TBX	<i>Sentinel-1 Toolbox</i>
SAR	<i>Synthetic Aperture Radar</i>
SASS	<i>Seasat-A Scatterometer System</i>
SLAR	<i>Side Looking Airborne Radar</i>
SMMR	<i>Scanning Multichannel Microwave Radiometer</i>
STAC	<i>SpatioTemporal Asset Catalog</i>
SWIR	<i>Short-Wave InfraRed</i>
SWIRP	<i>Submm-Wave and long wave InfraRed Polarimeter</i>
VIRR	<i>Visible and Infrared Radiometer</i>
VP	Verdadeiros Positivos

LISTA DE SÍMBOLOS

y	Máscara verdadeira
\hat{y}	Máscara predita
z	Imagem normalizada
x	Imagem original
μ	Média
σ	Desvio padrão

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Objetivos	14
1.2	Organização do Trabalho	14
1.3	Trabalhos Relacionados	14
2	DESASTRES NATURAIS	16
2.1	Classificação	17
2.2	Inundações pelo Mundo	18
2.3	Inundações no Brasil	19
3	RADAR DE ABERTURA SINTÉTICA	20
3.1	Geração de Imagens	21
3.2	Retroespalhamento de Microondas	22
3.3	Polarimetria	24
4	REDES NEURAIIS CONVOLUCIONAIS	25
4.1	Camadas de processamento	25
5	METODOLOGIA	28
5.1	Ferramentas de Desenvolvimento	28
5.2	Conjunto de Dados	28
5.3	Pré-Processamento	31
5.4	Aumento de Dados	32
5.5	Segmentação Semântica e de Instâncias	33
5.6	Funções de Perda	36
5.7	Métricas de Desempenho	37
5.8	Treinamento do Modelo	38
6	RESULTADOS	39
7	CONCLUSÕES E TRABALHOS FUTUROS	43
	REFERÊNCIAS	44

1 INTRODUÇÃO

A Organização Meteorológica Mundial (WMO, 2023) relata o progressivo aumento global de desastres naturais ao longo dos anos. Foram registrados 11.778 eventos climáticos adversos em todo o mundo durante o período de 1970 a 2021, com as inundações correspondendo a 45% das notificações. É notório que 60% dos prejuízos financeiros totais e 90% das mortes relatadas ocorreram em países em desenvolvimento. Em particular, as inundações recorrentes na América do Sul foram a principal causa das mortes por desastres naturais, sendo uma preocupação para as autoridades de proteção e defesa civil. Apesar da implementação e o aprimoramento dos sistemas de alerta, que reduziram a taxa de mortalidade ao longo da última metade do século, detectar inundações é fundamental no gerenciamento de desastres e na adoção de medidas de respostas eficientes.

Embora a modelagem de fenômenos hidrológicos busque prever o comportamento desses eventos climáticos extremos, essa possui limitações em áreas extensas devido sua complexidade, custo dos recursos computacionais e a escassez de dados precisos em alta resolução (ALMEIDA *et al.*, 2018). Neste aspecto, o sensoriamento remoto é útil para monitorar e detectar inundações em grandes áreas. Particularmente, as imagens de satélites com Radar de Abertura Sintética (*Synthetic Aperture Radar* (SAR) em inglês) independem das condições atmosféricas e da iluminação solar ao obter informações da superfície terrestre (TANIM *et al.*, 2022). Tais imagens fornecem uma visão abrangente e detalhada das áreas afetadas por inundações, permitindo a identificação e a delimitação das regiões atingidas.

O aprendizado profundo permite que modelos computacionais aprendam representações complexas de dados em múltiplos níveis de abstração. Essa abordagem tem demonstrado eficácia na identificação de estruturas complexas em grandes volumes de dados, tendo relevância em diversas áreas, incluindo ciência, negócios e governo (LECUN *et al.*, 2015). Nesse contexto, esse trabalho apresenta a aplicação de técnicas de aprendizado profundo na detecção de inundações a partir de imagens SAR capturadas pelo satélite *Sentinel-1*. São estudadas as influências de diferentes funções de perda no desempenho das Redes Neurais Totalmente Convolucionais (*Fully Convolutional Neural Network* (FCNN) em inglês) U-Net (RONNEBERGER *et al.*, 2015) e U-Net++ (ZHOU *et al.*, 2018) para a tarefa de segmentação semântica (KONG *et al.*, 2021) de inundações. Uma vez que a estrutura dessas redes segue um modelo codificador-decodificador, será adotada a rede ResNet-18 (HE *et al.*, 2016) como codificador.

1.1 Objetivos

O objetivo geral desse estudo é segmentar inundações por meio de imagens de satélite SAR disponíveis publicamente, aplicando FCNNs com modelo codificador-decodificador, para o mapeamento preciso de áreas inundadas. Destacam-se os seguintes objetivos específicos:

- Desenvolver os modelos U-Net e U-Net++ para aplicação ao problema de segmentação semântica das áreas inundadas a partir de conjuntos de imagens SAR rotulados;
- Investigar como as funções de perda *Binary Cross-Entropy*, *Dice* e *Tversky* ($\alpha = \beta = 1$), comumente exploradas em tarefas de segmentação semântica, impactam o desempenho das FCNNs nesse problema;

1.2 Organização do Trabalho

- **Capítulo 1:** Apresenta a introdução ao tema do trabalho e seus objetivos;
- **Capítulo 2:** Descreve conceitos relacionados a desastres, classificando inundações e evidenciando seus danos.
- **Capítulo 3:** Trata do sensoriamento remoto por radar e suas características.
- **Capítulo 4:** Descreve o funcionamento dos modelos utilizados para a tarefa de segmentação semântica.
- **Capítulo 5:** Apresenta a metodologia proposta para a segmentação semântica de inundações a partir de imagens SAR por aprendizado profundo;
- **Capítulo 6:** Apresenta os experimentos realizados para avaliação da metodologia, bem como os resultados obtidos;
- **Capítulo 7:** Apresenta as conclusões e alternativas de trabalhos futuros.

1.3 Trabalhos Relacionados

Rambour *et al.* (2020) apresentam o conjunto de dados SEN12-FLOOD, composto por 336 séries temporais de imagens capturadas pelos satélites *Sentinel-1* e *Sentinel-2* da Agência Espacial Europeia (*European Space Agency* (ESA) em inglês). Essas séries abrangem eventos de inundação em várias regiões, como África Oriental, Sudoeste da África, Oriente Médio e Austrália, e incluem imagens ópticas e de radar com resoluções espaciais que variam de 10 a 60 metros. A rotulação dos dados foi obtida a partir do conjunto de dados *MediaEval 2019*, originado do Sistema de Gerenciamento de Emergências Copernicus (*Copernicus Emergency*

Management Service (CEMS) em inglês). Nesse estudo, os autores obtiveram uma acurácia de 0,957 para a classificação de inundações, utilizando imagens SAR e *Red Green Blue* (RGB) como entrada na rede ResNet-50 com Unidades Recorrentes com Portas Fechadas (*Gated Recurrent Unit* (GRU) em inglês).

Montello *et al.* (2022) apresentam o conjunto de dados MMFLOOD, composto por 1.748 imagens do *Sentinel-1* com resolução espacial de 20 metros. Os dados também foram obtidos a partir do CEMS e cobrem 95 eventos de inundação distribuídos por 42 países. Além das imagens de satélite, o conjunto de dados inclui Modelos Digitais de Elevação (MDE), mapas de hidrografia e de delimitação de inundações. Nesse estudo, os modelos ResNet-50 e TResNet apresentaram os melhores desempenhos na segmentação semântica de inundações, com *Intersection over Union* (IoU) médio de 0,63 e 0,59, respectivamente.

Tanim *et al.* (2022) combinaram dados de relatórios policiais de fechamento de estradas em San Diego, Califórnia, com imagens do *Sentinel-1* para criar modelos de aprendizado de máquina supervisionados e não supervisionados. Esses modelos incluíram Floresta Aleatória, Máquina de Vetor de Suporte, Classificador de Máxima Verossimilhança e um novo método de classificação e detecção de inundações não supervisionado baseado na abordagem de detecção de mudanças (*Change Detection* (CD) em inglês), que combina o algoritmo de Otsu, lógica fuzzy e métodos de *iso-clustering*, para detecção de pixels de inundações em áreas urbanas. O novo método de CD apresentou o melhor desempenho, requerendo menos dados e tempo de execução. Os resultados da avaliação de performance do modelo mostraram precisão de 0,81, sensibilidade de 0,9, pontuação F1 de 0,85 e acurácia de 0,87.

Portanto, esses estudos demonstram a relevância do tema proposto e a viabilidade de utilizar modelos e técnicas de aprendizado profundo na segmentação de inundações a partir de dados de sensoriamento remoto, como imagens do *Sentinel-1*.

2 DESASTRES NATURAIS

O escritório das Nações Unidas para Redução de Riscos e Desastres (UNDRR, 2020) define desastre como uma grave interrupção, em qualquer escala, no funcionamento de uma comunidade ou sociedade causada por uma combinação de eventos perigosos com fatores de exposição, vulnerabilidade e capacidade, que resultam em danos humanos, materiais, ambientais e/ou econômicos. Além disso, essa terminologia aborda características adicionais a essas situações, incluindo aspectos como a duração, regularidade e magnitude da ocorrência.

Desastres de início lento são caracterizados por uma evolução gradual ao longo do tempo, abrangendo fenômenos como secas, desertificação e elevação do nível do mar. Em contrapartida, os desastres de início súbito são desencadeados por eventos que emergem rapidamente, ou inesperadamente, como terremotos, erupções vulcânicas e inundações repentinas. A distinção entre desastres frequentes e infrequentes é baseada na probabilidade de ocorrência e no período de retorno das ameaças associadas. Desastres de pequena escala são aqueles que impactam comunidades locais, requerendo assistência de áreas vizinhas. Em contraste, desastres de grande escala englobam eventos que afetam toda uma sociedade, demandando auxílios nacionais ou internacionais.

Desastres naturais podem sobrecarregar, ou exceder, a capacidade de um grupo social de lidar com a situação pelos próprios recursos. Além disso, os impactos frequentes podem se acumular ao longo do tempo, tornando-se crônicos para uma comunidade ou sociedade. Esses danos atingem tanto nações desenvolvidas quanto em desenvolvimento, afetando, sem distinção, diversas classes sociais. A gravidade dos danos pode se intensificar na ausência de um planejamento urbano adequado e políticas públicas voltadas para a prevenção desses eventos, sobretudo quando considerados aspectos socioeconômicos. Isso se deve ao fato de que as comunidades mais vulneráveis economicamente são as que mais enfrentam desafios para se recuperar dos impactos causados, mesmo que historicamente contribuam menos para as mudanças climáticas (SAUSEN; LACRUZ, 2015; UNDRR, 2020).

Este capítulo apresenta a problemática dos desastres naturais, com foco particular em inundações. Será abordada a definição e as classificações desses eventos, além da avaliação do seu impacto nas populações em um contexto global e brasileiro.

2.1 Classificação

A Pesquisa Integrada sobre Risco de Desastres (IRDR, 2014) distingue os tipos de desastres em naturais, que envolvem processos ou fenômenos extremos da natureza, e tecnológicos, causados por ações humanas. A WMO (2021) e UNDRR (2015) ponderam que é possível estabelecer uma relação direta entre desastres naturais e as mudanças climáticas, bem como suas consequências sobre a dinâmica social e o meio ambiente. A Tabela 1 apresenta a classificação adotada no trabalho. Ela está organizada em grupos que representam a origem de cada fenômeno natural. Cada grupo é acompanhado por sua definição e pela segmentação dos desastres correspondentes.

Tabela 1 – Classificação de desastres naturais

Grupo	Definição	Subgrupos
Geológico	Um perigo causado por instabilidades na litosfera.	<ul style="list-style-type: none"> – Terremoto – Movimento de massa (seca) – Atividade vulcânica
Hidrológico	Um perigo causado por alterações no ciclo natural da água.	<ul style="list-style-type: none"> – Inundação – Deslizamento de terra – Ação de ondas
Meteorológico	Um perigo causado por fenômenos atmosféricos de curta duração.	<ul style="list-style-type: none"> – Temperatura extrema – Nevoeiro – Tempestade
Climatológico	Um perigo causado por variabilidades climáticas de longa duração.	<ul style="list-style-type: none"> – Seca – Transbordamento de lagos glaciais – Incêndio florestal
Biológico	Um perigo causado pela exposição a organismos vivos.	<ul style="list-style-type: none"> – Epidemia – Infestação de insetos – Acidente animal
Extraterrestre	Um perigo causado por asteroides, meteoros e cometas que colidem com a Terra.	<ul style="list-style-type: none"> – Impacto – Clima espacial

Fonte: Adaptado de IRDR (2014)

No grupo dos fenômenos hidrológicos, o termo inundação refere-se ao transbordamento de água, resultando na submersão de áreas que geralmente estão secas, podendo essas serem repentinas, fluviais ou costeiras. As repentinas são provocadas por chuvas intensas ou excessivas em um curto período de tempo, elevando o nível da água em minutos ou horas. Já as fluviais ocorrem quando chuvas persistentes, ou degelo de neve, fazem um rio ultrapassar sua capacidade e inundar áreas adjacentes. Por fim, as costeiras são resultado de tempestades, que podem durar dias ou semanas, associadas a mudanças nas marés (IRDR, 2014).

2.2 Inundações pelo Mundo

Entre os anos de 1970 e 2021, as catástrofes naturais provocaram danos econômicos em escala global avaliados em US\$ 4,3 trilhões e a perda de mais de 2 milhões de vidas humanas. Desses totais, atribuem-se que 32% dos prejuízos financeiros e 16% dos óbitos às inundações (WMO, 2023).

Inequivocadamente, as atividades humanas têm causado o aquecimento do planeta por meio da emissão de gases de efeito estufa e, conseqüentemente, as mudanças climáticas. Durante o período de 2011 a 2020 estas atividades produziram um aumento da temperatura global de 1,1°C, quando comparado ao período de 1850 e 1900. Esse efeito é proveniente de práticas insustentáveis de uso de energia, estilos de vida, padrões de consumo e produção (IPCC, 2023).

Uma consequência das mudanças climáticas é a intensificação do ciclo natural da água, que resulta em chuvas mais intensas em algumas regiões que, por conseguinte, aumentam os riscos de inundações graves, especialmente em áreas costeiras e outras regiões de baixa altitude (MONTELLO *et al.*, 2022). Além disso, pesquisas documentam a influência humana, combinada a outras influências climáticas, em eventos de precipitação extremos. Alguns exemplos incluem as chuvas extremas no leste da China, em junho e julho de 2016, e o furacão Harvey, que atingiu Huston no estado do Texas, em 2017 (WMO, 2021).

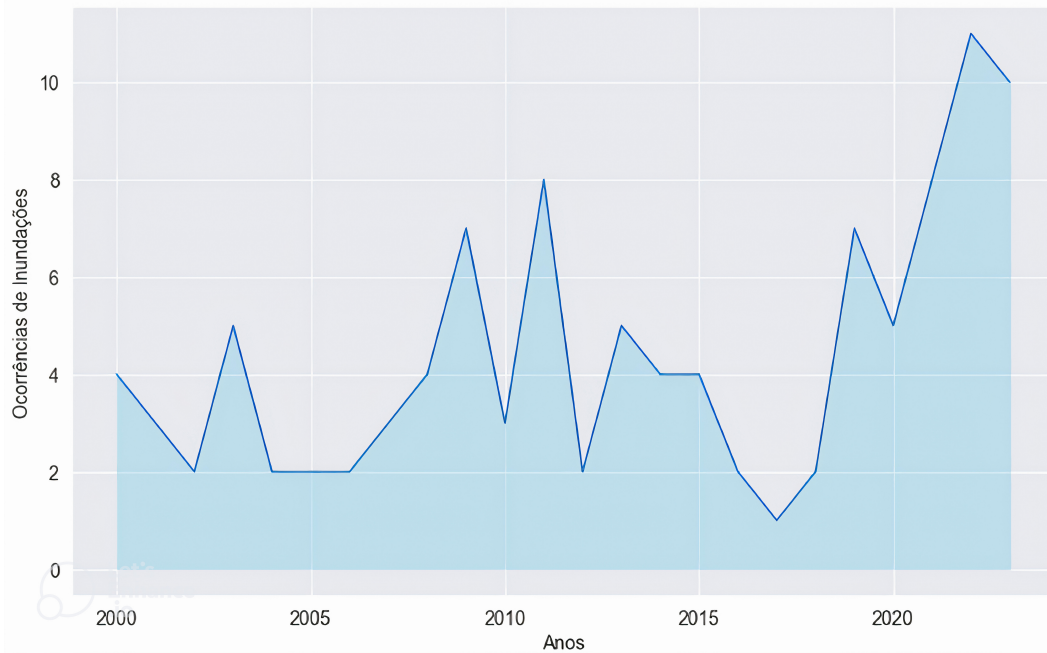
Segundo o relatório do Alto Comissariado das Nações Unidas para Refugiados (UNHCR, 2023), houveram deslocamentos populacionais em razão de desastres naturais em diversos países. Em 2022, registrou-se 32,6 milhões de deslocamentos internos decorrentes de catástrofes, dos quais 8,7 milhões de indivíduos permaneceram refugiados até o fim do ano. É notório que os deslocamentos internos resultantes de catástrofes naturais constituíram 54% do conjunto de novos deslocamentos ocorridos ao longo de 2022. O Paquistão registrou a maior magnitude de deslocamentos internos devido a estes eventos, com um total de 8,2 milhões de indivíduos impactados por extensas inundações em todo o país. Similarmente, nas Filipinas e na China, eventos climáticos como ciclones, inundações e tempestades tropicais ocasionaram o deslocamento de 5,4 milhões e 3,6 milhões de pessoas, respectivamente.

A Organização Mundial da Saúde (WHO, 2022) considera que a magnitude dos impactos econômicos e humanos causados por inundações pode ser reduzidos significativamente mediante a implementação de medidas adequadas de prevenção, preparação, resposta e recuperação a essas situações de emergência.

2.3 Inundações no Brasil

O Brasil registrou um total de 145 ocorrências de desastres naturais entre os anos de 2000 e 2023, sendo que 73% desses eventos estão relacionados a inundações (EM-DAT, 2020). Esse tipo de fenômeno natural é considerado um dos mais devastadores, pois são frequentes, extensos e afetam uma grande quantidade de pessoas (SAUSEN; LACRUZ, 2015). O gráfico da Figura 1 apresenta as ocorrências de inundações no Brasil durante o período mencionado. É perceptível um aumento das inundações ao longo do tempo em escala nacional, especialmente nos últimos anos. É importante ressaltar que, mesmo quando um desastre afeta múltiplos locais, as informações são registradas como uma única ocorrência nos dados internacional de desastres.

Figura 1 – Distribuição de ocorrências de inundações no Brasil de 2000 a 2023



Fonte: Adaptado de EM-DAT (2020).

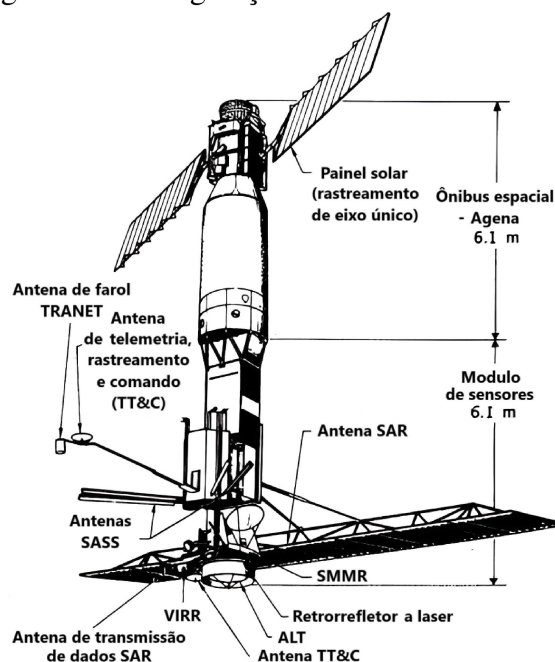
Segundo dados do Centro Nacional de Monitoramento e Alertas de Desastres Naturais (CEMADEN), órgão responsável por supervisionar e notificar riscos associados a eventos climáticos extremos, em 2023, o litoral norte do Estado de São Paulo registrou o maior índice pluviométrico do Brasil, totalizando 682 milímetros de chuva em um único dia. Esse valor superou o recorde estabelecido em 2022 por Petrópolis, no estado do Rio de Janeiro, quando um acumulado de 530 milímetros foi registrado em um único dia. Anteriormente, o recorde pertencia a Florianópolis, em Santa Catarina, que havia registrado, no ano de 1991, um acumulado de 400 milímetros em um único dia (SÃO PAULO, 2023).

3 RADAR DE ABERTURA SINTÉTICA

O termo Radar (*Radio Detection And Ranging*) é comumente utilizado para designar genericamente os sistemas que operam na faixa de frequência de microondas. Essa tecnologia foi inicialmente desenvolvida para propósitos militares durante a Segunda Guerra Mundial e, a partir da década de 1970, foi adaptada para aplicações civis. Esses sistemas podem ser classificados em imageadores e não imageadores. Entre os primeiros estão o Radar de Visada Lateral (*Side Looking Airborne Radar (SLAR)* em inglês) e o Radar de Abertura Sintética (SAR). Já os não imageadores incluem os escaterômetros, espectrômetros e altímetros (INPE, 2001). Segundo Montello *et al.* (2022), as inundações têm sido monitoradas com eficiência por meio do imageamento remoto com SAR.

O Seasat-1, lançado pela *National Aeronautics and Space Administration (NASA)*, em 1978, foi o primeiro satélite com o propósito de demonstrar a eficácia dos sensores de microondas para a observação de fenômenos oceanográficos da Terra a partir do espaço. Operando a uma altitude de 800 quilômetros em relação à superfície, esse satélite orbitava o planeta 14 vezes por dia, cobrindo 95% dos oceanos do mundo a cada 36 horas. Conforme ilustrado na Figura 2, a espaçonave tinha a bordo cinco sensores: *Radar Altimeter (ALT)*, *Scanning Multichannel Microwave Radiometer (SMMR)*, *Seasat-A Scatterometer System (SASS)*, *Visible and Infrared Radiometer (VIRR)* e SAR (BORN *et al.*, 1979).

Figura 2 – Configuração das antenas do Seasat-1



Fonte: Adaptado de BORN *et al.* (1979)

Embora esse satélite tenha se restringido a 105 dias de operação devido a um curto-circuito em seu sistema elétrico, seu instrumento SAR produziu imagens notáveis da superfície marítima, destacando o potencial dessa tecnologia para o sensoriamento remoto do planeta (ESA, 2013; BORN *et al.*, 1979). Neste capítulo, serão discutidos os aspectos do sensoriamento por radar, suas vantagens e desvantagens na detecção de inundações, bem como a geração de imagens SAR.

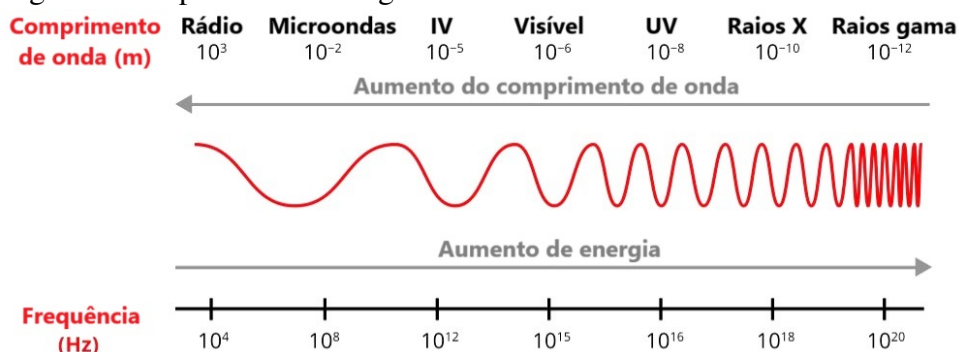
3.1 Geração de Imagens

Imaginar a superfície da Terra através de radar envolve medir a intensidade do sinal refletido para cada *pixel* que forma a imagem após a emissão da fonte de energia transportada pela plataforma de sensoriamento. À medida que o feixe de radiação da antena percorre o solo, os ecos do pulso de energia emitidos são dispersos até um sensor ativo para produzir imagens de alta resolução da área de interesse (RICHARDS, 2022).

Em sistemas tradicionais, a resolução espacial de uma imagem é definida pelo produto da resolução angular e da altitude do sensor em relação à superfície. A resolução angular é definida pela divisão do comprimento de onda irradiada pelo diâmetro da antena. Portanto, à medida que a altura do sensor aumenta, a resolução da imagem diminui, exceto se o tamanho físico da antena também aumentar (CURLANDER; MCDONOUGH, 1991).

A Figura 3 ilustra o espectro eletromagnético com os comprimentos de ondas e suas frequências correspondentes. Na faixa onde a luz é visível e nas proximidades do infravermelho (IV), é possível obter imagens de alta resolução com antenas de tamanho moderado mesmo a grandes altitudes. No entanto, nota-se que os comprimentos da região das microondas são significativamente maiores do que os da luz visível, impossibilitando a obtenção de imagens em alta resolução com antenas de tamanho razoáveis (CURLANDER; MCDONOUGH, 1991).

Figura 3 – Espectro eletromagnético



Fonte: Adaptado de Adobe Stock

A tecnologia SAR melhora a resolução da imagem sem aumentar o tamanho físico da antena. Isso é conseguido sintetizando-se uma abertura de antena extremamente longa através de processamento digital de sinais, onde se compensa as características de fase quadrática associadas a imagem formada. Desta forma, o sistema SAR alcança uma boa resolução independente da altitude do sensor (CURLANDER; MCDONOUGH, 1991).

O uso de radar de microondas também oferece a vantagem de permitir que a energia incidente penetre na superfície e subsuperfície da Terra, tornando possível observar propriedades que não eram detectáveis por meio de imagens ópticas, que dependem da radiação solar refletida para a geração de imagens (RICHARDS, 2022).

Embora as imagens SAR apresentem uma qualidade visual inferior em comparação com as ópticas ou multiespectrais, elas permitem monitorar uma região mesmo durante a noite, devido a sua capacidade de aquisição de dados com base no coeficiente de retroespalhamento. Além disso, como os satélites SAR possuem sua própria fonte de energia, as imagens são menos afetadas por condições atmosféricas adversas, como a intensa cobertura de nuvens frequentemente associada a inundações, o que possibilita o mapeamento preciso do solo em diferentes condições climáticas. Por fim, a sensibilidade inerente do SAR às propriedades físicas e dielétricas dos materiais presentes na superfície da Terra torna essas imagens valiosas para análise de água e umidade do solo (CURLANDER; MCDONOUGH, 1991).

Todavia, imagens SAR são afetadas por um ruído granular chamado *speckle*, proporcional a intensidade do sinal recebido pela plataforma de sensoriamento remoto. Esse fenômeno deriva do processo de radiação coerente, no qual ondas eletromagnéticas de mesma frequência e direção mantêm uma relação de fase constante entre si. Visualmente, o *speckle* produz uma textura granulada que dificulta a interpretação das imagens de radar, diminuindo a distinção entre os objetos na cena. Assim, a análise de imagens SAR requer um pré-processamento personalizado com base nas propriedades estatísticas dos dados *speckle* (FRERY, 2019; INPE, 2001).

3.2 Retroespalhamento de Microondas

O processo de retroespalhamento refere-se à reflexão do sinal de radar em direção à plataforma que o emitiu, resultado da interação entre o sinal e os objetos da superfície da Terra. A Figura 4 ilustra como as características dos materiais afetam o imageamento de áreas nos comprimentos de ondas mais utilizados para o imageamento por radar, apresentados na Tabela 2.

Tabela 2 – Bandas de microondas típicas no sensoriamento remoto

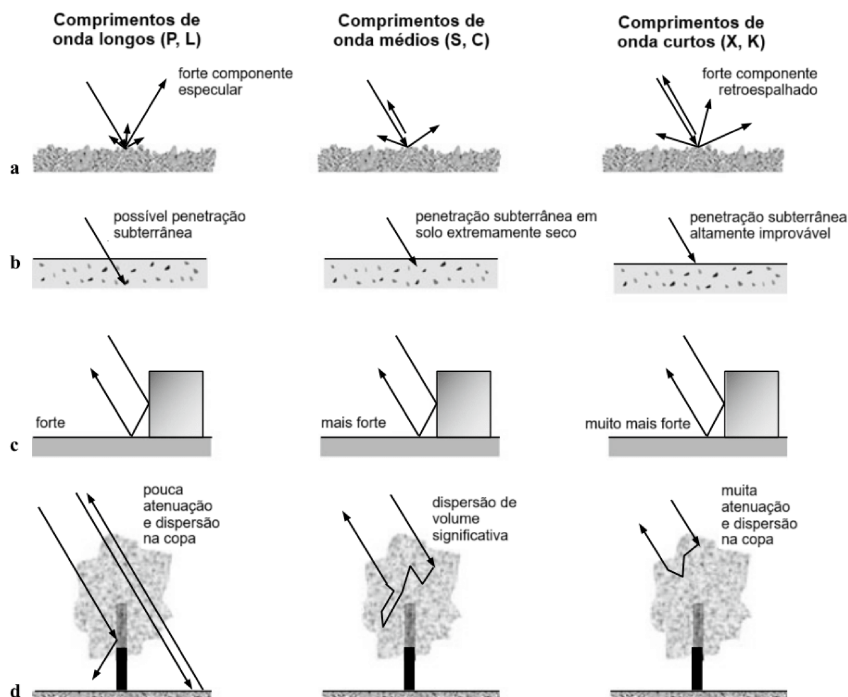
Banda	P	L	S	C	X	Ku
Comprimento de Onda (cm)	66,7	23,5	12,6	5,7	3,1	2,16
Frequência (GHz)	0,45	1,28	2,38	5,3	9,7	13,9

Fonte: Adaptado de Richards (2022)

Segundo Richards (2022), superfícies lisas atuam como refletores especulares, em que a direção do sinal é oposta à da incidência, e são representadas por *pixels* escuros na imagem. Superfícies rugosas agem como refletores difusos, espalhando a energia incidente em todas as direções, assim como de volta para a plataforma, e são representadas por *pixels* claros na imagem. Em comprimentos de onda maiores, muitas superfícies são consideradas lisas, enquanto as mesmas superfícies podem ser difusas em comprimentos de onda menores.

O mesmo autor aponta que, quando o material da superfície está seco (Figura 4b), a radiação de microondas incidente pode penetrar na superfície, especialmente em comprimentos de onda longos. Materiais como a copa de vegetações (Figura 4a) apresentam um comportamento de espalhamento de volume, ou seja, a energia retroespalhada emerge de múltiplos locais difíceis de definir dentro do volume, o que resulta em uma aparência clara nas imagens. Outro mecanismo de espalhamento da superfície que produz uma resposta brilhante é o efeito refletor de canto, frequentemente encontrado em estruturas verticais, como edifícios (Figura 4c) e troncos de árvores (Figura 4d), devido ao ângulo reto formado com a superfície.

Figura 4 – Retroespalhamento em função do comprimento de onda



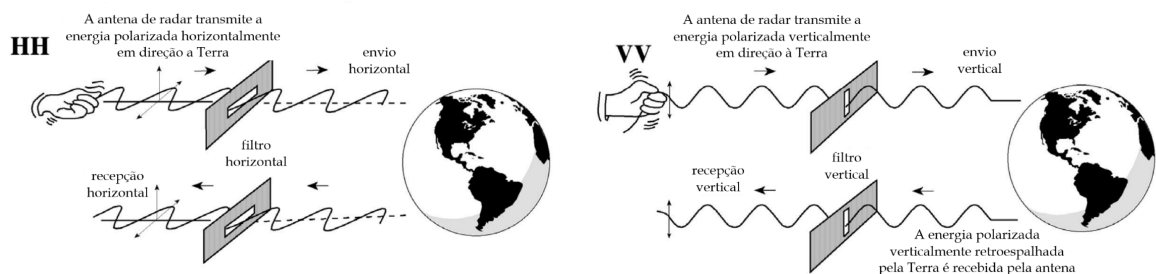
Fonte: Adaptado de Richards (2022)

3.3 Polarimetria

A polarização refere-se à orientação do vetor de deslocamento da onda eletromagnética, podendo ser classificada como vertical, horizontal ou circular. A direção da polarização é determinada pela orientação do campo elétrico da onda, que é sempre perpendicular ao campo magnético. Em sistemas SAR, são coletadas informações detalhadas sobre as propriedades polarimétricas da superfície observada. Estruturas com orientação linear, como edifícios, tendem a preservar a coerência do sinal polarimétrico, enquanto estruturas com orientação aleatória, como a copa de árvores, podem despolarizar o sinal (NISAR, 2020).

As imagens de radar podem ter diferentes configurações de polarização. Conforme ilustrado na Figura 5, um sistema de polarização única (*single-pol*) transmite e recebe em uma única polarização, geralmente na mesma direção. Isso resulta em imagens HH (horizontal-horizontal) ou VV (vertical-vertical), onde a primeira letra refere-se à direção da transmissão e a segunda à direção da recepção (NISAR, 2020). A direção de transmissão é representada na imagem pelas setas que apontam para a direita, enquanto a direção de recepção é indicada pelas setas que apontam para a esquerda.

Figura 5 – Sistema de polarização única



Fonte: Adaptado de Jensen (2006)

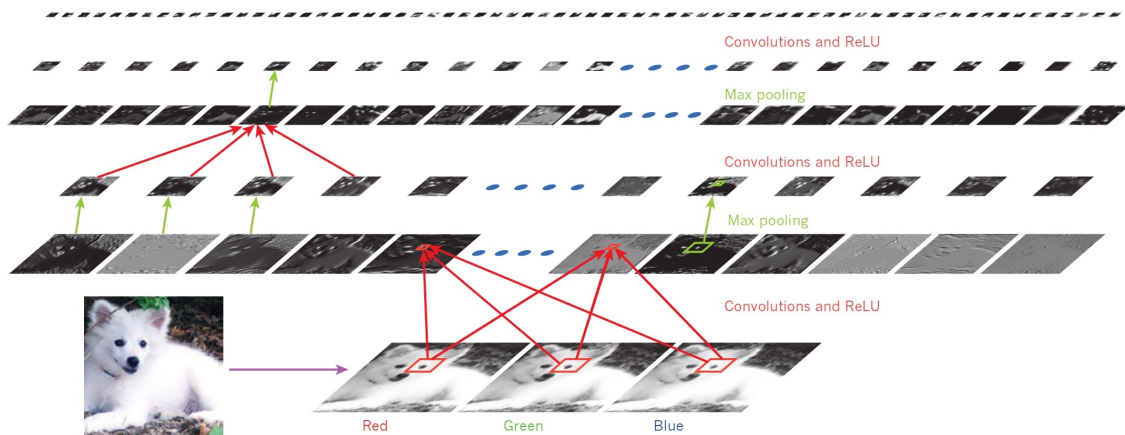
Um sistema de polarização dupla (*dual-pol*) se transmite em uma polarização, mas pode receber em duas. Isso resulta em imagens HH e HV, ou VH e VV. Essa configuração fornece informações adicionais sobre as características da superfície por meio de diferentes respostas polarimétricas. Um sistema de polarização quádrupla (*quad-pol*) alterna entre a transmissão e recepção de ondas horizontais e verticais. Isso resulta em imagens HH, HV, VH e VV. No entanto, para operar nesse modo, o radar deve pulsar duas vezes mais rápido do que um sistema único ou duplo, uma vez que a polarização de transmissão deve alternar entre H e V pulso a pulso. Quando se opera nesse modo, pode haver interferência entre os ecos recebidos devido à alternância de polarização. (NISAR, 2020).

4 REDES NEURAI CONVOLUCIONAIS

Técnicas de Inteligência Artificial (IA), especialmente as de aprendizado profundo, têm sido bem-sucedidas na obtenção de informações sobre a cobertura do solo. Dentre elas, destaca-se a Rede Neural Convolutiva (*Convolutional Neural Network* (CNN) em inglês) no reconhecimento de imagens, detecção de objetos e segmentação a partir de dados de sensoriamento remoto (ZHU *et al.*, 2017). Redes neurais artificiais são algoritmos bioinspirados. Compostas por camadas de neurônios interconectados, essas redes identificam padrões por meio de exemplos fornecidos, ajustando seus pesos por meio de funções de erro (ZHANG *et al.*, 2023).

Uma CNN é composta por dezenas, ou até centenas, de camadas de processamento projetadas para detectar diferentes características de uma imagem de entrada. Nelas, os valores dos pesos e vieses são compartilhados entre todos os neurônios ocultos da mesma camada. Isso permite que os neurônios processem características semelhantes em diferentes regiões da imagem. Conforme ilustrado na Figura 6, inicialmente cada camada (horizontal) detecta características genéricas de uma imagem, e à medida que avançam realizam operações mais complexas, identificando características mais específicas (MATHWORKS, 2017).

Figura 6 – Arquitetura típica de uma rede convolutiva



Fonte: LeCun *et al.* (2015)

4.1 Camadas de processamento

As camadas convolucionais possibilitam a extração de características, como bordas, linhas, texturas e formas, aplicando filtros de tamanhos fixos na imagem de entrada. Estes filtros são implementados discretamente por matrizes de pesos que realizam a operação de convolução percorrendo a imagem e produzindo um mapa de características. Cada operação de convolução

resulta em um novo mapa de características de saída, contendo detalhes, ou padrões, que a CNN aprendeu a identificar na imagem original (KUMAR, 2023).

A Figura 7 ilustra essa operação utilizando uma matriz de entrada de tamanho 3 x 3 e um filtro de tamanho 2 x 2 com passo unitário. Enquanto o filtro percorre a imagem, os elementos correspondentes da matriz de entrada, contidos na janela do filtro, são multiplicados elemento a elemento, e a soma dessas multiplicações produz um único valor na matriz de saída. Em alguns casos, o filtro de convolução pode não se ajustar ao tamanho da imagem de entrada. Nesse contexto, abordagens como o preenchimento da margem da imagem com zeros podem ser implementadas para garantir que a convolução possa ser aplicada de maneira consistente (ZHANG *et al.*, 2023).

Figura 7 – Operação de convolução

Entrada		Filtro		Saída				
0	1	2	*	0	1	=	19	25
3	4	5		2	3		37	43
6	7	8						

Fonte: Zhang *et al.* (2023)

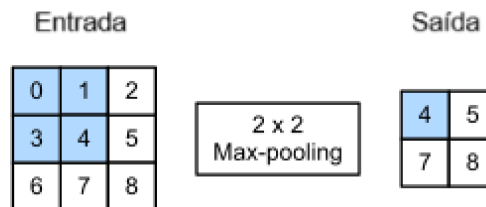
Os resultados produzidos na saída dos filtros de convolução seguem para a camada de ativação, que implementam as funções de ativação de sua saída. Estas funções são diferenciáveis e introduzem não linearidade ao processo de treinamento. Popularmente empregada, a função *Rectified Linear Unit* (ReLU) propaga entradas positivas para sua saída, enquanto que as entradas negativas resultam em saída nula. Isso é expresso matematicamente por $\max(x, 0)$, sendo $\max(\cdot)$ o operador de valor máximo entre dois valores (ZHANG *et al.*, 2023).

As saídas produzidas na camada de ativação são processadas na camada de *pooling*, diminuindo a dimensionalidade das matrizes enquanto preservam informações relevantes. Essa operação se dá a partir de uma janela de *pooling*, de tamanho fixo, que percorre a matriz de entrada com um passo definido, calculando um valor único para cada região percorrida. Diferente da convolução, as operações de *pooling* são determinísticas, computando o valor médio (*average pooling*), máximo (*max-pooling*) ou mínimo (*min-pooling*) dos elementos de entrada pertencentes à janela (ZHANG *et al.*, 2023).

Esse operador introduz invariância nos recursos extraídos da imagem, permitindo que a rede neural identifique um objeto independentemente de sua localização, orientação ou escala na imagem original. Enquanto as camadas de convolução buscam por características, as camadas

de *pooling* selecionam e preservam as características mais relevantes, tornando-as mais eficientes em termos de processamento computacional (KUMAR, 2023). A Figura 8 demonstra a operação de *max-pooling* para uma janela 2 x 2, onde apenas os valores máximos são considerados para representar as características na saída.

Figura 8 – Operação de *max-pooling*



Fonte: Zhang *et al.* (2023)

Por fim, as saídas das operações de *pooling* são processadas por uma ou mais camadas totalmente conectadas (*Fully Connected Layers* (FCL) em inglês), também conhecidas como camadas densas. Elas compõem as redes neurais convencionais de classificação de imagens. O papel principal das FCL é empregar as características extraídas anteriormente para fazer previsões sobre a classe ou rótulo da imagem de entrada, finalizando o processo de análise e interpretação. (KUMAR, 2023).

5 METODOLOGIA

Neste capítulo, são detalhados os métodos utilizados para abordar o problema proposto. A metodologia adotada abrange o treinamento de FCNNs para mapear áreas inundadas, utilizando imagens SAR do conjunto de dados *C2S-MS Floods* (STREET *et al.*, 2022). O experimento é estruturado em diferentes etapas. O pré-processamento das imagens inclui técnicas como calibração radiométrica, remoção de ruídos nas bordas, entre outras. Já o aumento de dados envolve a aplicação de técnicas de divisão de imagens (*image splitting*). A fase subsequente contempla o desenvolvimento dos modelos U-Net e U-Net++, variando as funções de perda *Binary Cross-Entropy*, *Dice* e *Tversky*. A avaliação da segmentação predita será conduzida utilizando métricas quantitativas, tais como Índice de *Jaccard* (IoU), pontuação F1 e sensibilidade.

5.1 Ferramentas de Desenvolvimento

Este trabalho foi conduzido no ambiente de nuvem do *Google Collaboratory*¹, em linguagem *Python*². Trata-se de um ambiente *Linux* que conta com o suporte da placa gráfica *Graphics Processing Units* (GPU) NVIDIA Tesla T4, com 15GB de memória dedicada e 12.7 GB de memória RAM, que acelerou o treinamento dos modelos e a manipulação de imagens. Foram empregadas as bibliotecas *PyTorch*³, *Lightning Fabric*⁴ e *Segmentation Models Pytorch*⁵ para a construção do modelo de aprendizado profundo, além das bibliotecas *Augmentation*⁶ e *Scikit-Image*⁷ para aumentar e diversificar o conjunto de dados. A visualização das imagens e gráficos foi produzida a partir da biblioteca *Matplotlib*⁸.

5.2 Conjunto de Dados

O *C2S-MS Floods* (STREET *et al.*, 2022) é um conjunto de dados de inundações rotulado pela *Floodbase*⁹, anteriormente conhecida como *Cloud to Street*, em colaboração com a *Microsoft Planetary Computer*¹⁰. Este conjunto foi parcialmente utilizado na competição *STAC*

¹ <https://colab.research.google.com/>

² <https://www.python.org/>

³ <https://pytorch.org/>

⁴ <https://lightning.ai/fabric>

⁵ https://github.com/qubvel/segmentation_models_pytorch

⁶ <https://github.com/albumentations-team/albumentations>

⁷ <https://scikit-image.org/>

⁸ <https://matplotlib.org/>

⁹ <https://www.floodbase.com>

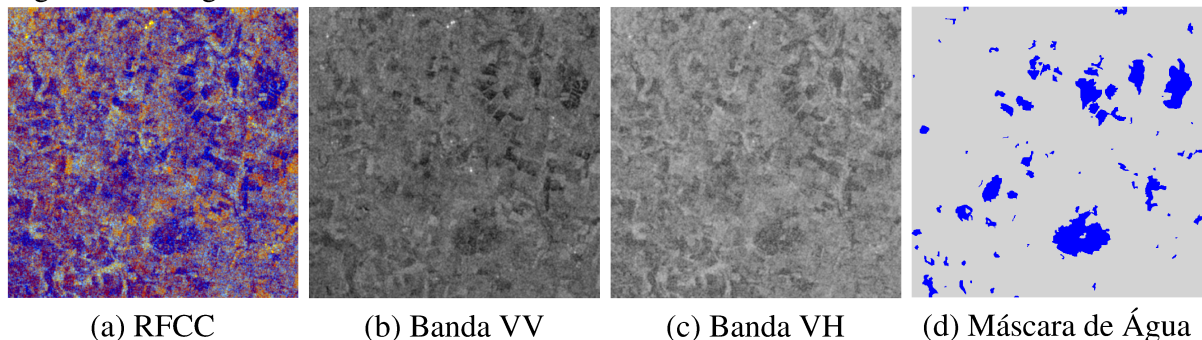
¹⁰ <https://planetarycomputer.microsoft.com/>

*Overflow: Map Floodwater from Radar Imagery*¹¹. Ele abrange 18 eventos de inundações em várias regiões do mundo, totalizando 900 pares de imagens capturadas pelos satélites *Sentinel-1* e *Sentinel-2* durante o período de 2016 a 2020. Cada imagem tem dimensão de 512 x 512 *pixels*.

O satélite *Sentinel-1* é equipado com um sensor SAR que fornece dados nas bandas VV e VH, que são armazenados no formato de imagem GeoTIFF, em ponto flutuante (*float32*). Cada *pixel* nas imagens representa a energia refletida de volta ao satélite, medida em decibéis (dB). Os valores dos *pixels* variam de negativos a positivos, com valor igual a zero indicando ausência de dados. Exemplos das imagens do *Sentinel-1* presentes na base de dados podem ser vistas na Figura 9.

São disponibilizadas imagens no formato PNG para permitir uma exploração rápida do conjunto de dados. Elas são criadas ao combinar as bandas VV (Figura 9b), VH (Figura 9c) e a relação VV/VH, resultando em uma imagem artificialmente colorida denominada RFCC (Figura 9a). Nessa representação, a água e outras superfícies lisas são identificadas em azul, enquanto a vegetação inundada é destacada em tons de laranja e vermelho. Os dados do *Sentinel-1* incluem rótulos que foram anotados manualmente, denominados máscaras de água (Figura 9d), os quais identificam a presença de inundações (STREET *et al.*, 2022).

Figura 9 – Imagens do satélite *Sentinel-1*



Fonte: Elaborado pelo autor (2023)

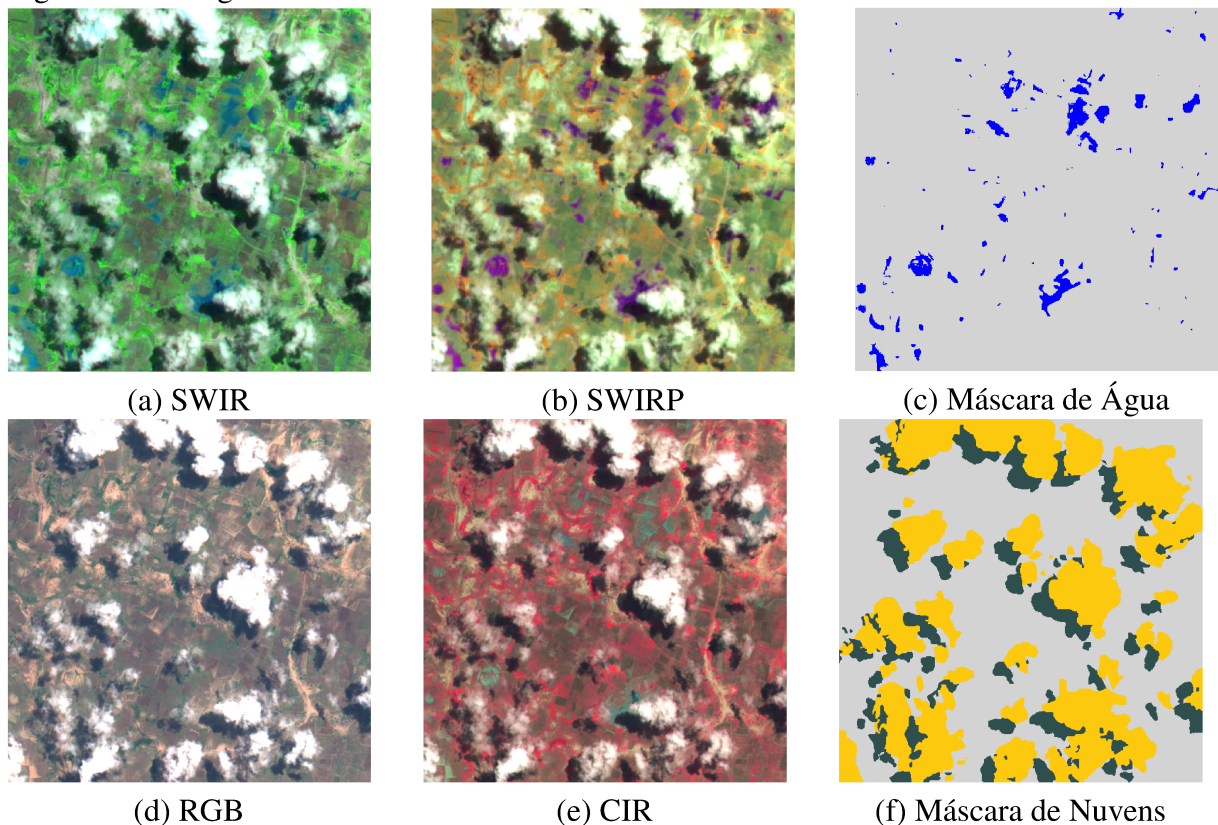
Os dados provenientes do *Sentinel-2* são compostos por 13 bandas armazenadas no formato GeoTIFF. Para facilitar a exploração do conjunto de dados, estão disponíveis quatro opções de visualização em PNG, denominadas *Short-Wave InfraRed* (SWIR) (Figura 10a), *Submm-Wave and long wave InfraRed Polarimeter* (SWIRP) (Figura 10b), RGB (Figura 10d) e CIR (Figura 10e). A imagem RGB exhibe a representação das cores reais, enquanto as demais proporcionam a visualização em cores falsas.

¹¹ <https://www.drivendata.org/competitions/81/detect-flood-water/>

Os dados do *Sentinel-2* incluem rótulos de água (Figura 10c) e de nuvens (Figura 10f), que indicam a presença de inundações e nuvens, respectivamente. Esses rótulos são armazenados como imagens GeoTIFF e podem ser visualizados em formato PNG. No contexto de detecção de inundações, as bandas SWIR e SWIRP são particularmente relevantes, ao passo que as imagens RGB e CIR auxiliam na identificação de nuvens e suas sombras (STREET *et al.*, 2022).

Na representação dessas imagens, a água é exibida em azul, as nuvens em amarelo, as sombras das nuvens em verde e o fundo em cinza. De forma semelhante ao processo de anotação das máscaras do *Sentinel-1*, os rótulos do *Sentinel-2* são gerados por marcação manual com base nas quatro imagens de referência (STREET *et al.*, 2022).

Figura 10 – Imagens do satélite *Sentinel-2*



Fonte: Elaborado pelo autor (2023)

Dados suplementares de elevação foram adquiridos da NASA para cada região através da API STAC¹². A plataforma *Microsoft AI for Earth*¹³ disponibiliza essa API para consultar o catálogo de dados hospedados no *Planetary Computer*. Este utiliza o padrão de Catálogo de Ativos Espaço-Temporais (*SpatioTemporal Asset Catalog* (STAC) em inglês) para simplificar a busca por dados que atendam a critérios espaciais e temporais.

¹² <https://stacspec.org/en>

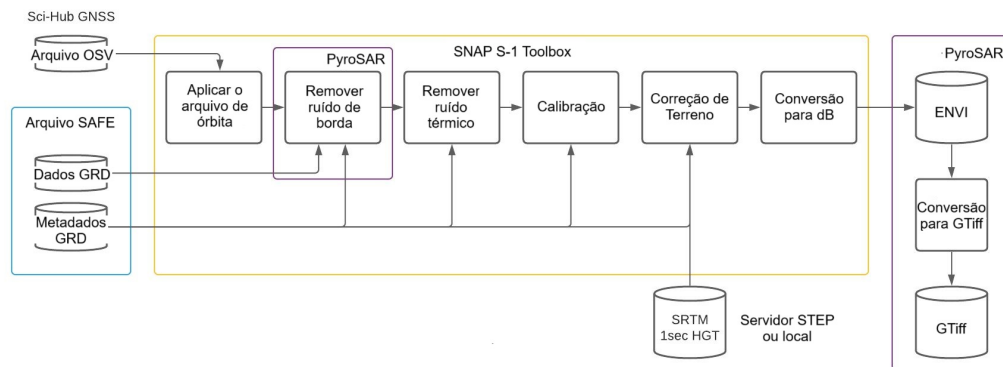
¹³ <https://www.microsoft.com/en-us/ai/ai-for-earth>

5.3 Pré-Processamento

O conjunto de dados *C2S-MS Floods* foi pré-processado pelos organizadores utilizando o *Sentinel-1 Toolbox* (S1TBX)¹⁴ em conjunto com o *PyroSAR*¹⁵, conforme o esquema ilustrado na Figura 11. Ambas são ferramentas de análise e processamento de imagens de radar. O S1TBX foi desenvolvido para dar suporte aos extensos conjuntos de dados provenientes das missões SAR da ESA, incluindo *Sentinel-1*, ERS-1 e 2 e ENVISAT, assim como os dados SAR de terceiros, como ALOS PALSAR, TerraSAR-X, COSMO-SkyMed e RADARSAT-2. Essa ferramenta oferece funcionalidades que abrangem calibração, filtragem de ruídos, ortorretificação, mosaico, conversão de dados, polarimetria e interferometria de imagens (ESA, 2015).

Primeiramente, o arquivo de órbitas é aplicado. Em seguida, o ruído nas bordas é removido para algumas cenas, dependendo da versão. As ferramentas de remoção de ruído térmico e calibração são aplicadas em seguida. Usando o SRTM, o S1TBX é utilizado para aplicar correção de terreno aos dados. Por fim, os dados são convertidos de unidades lineares para decibéis e salvos no formato de arquivo GeoTIFF. Todas as bandas são armazenadas com resolução de 10 metros por pixel (STREET *et al.*, 2022).

Figura 11 – Pré-processamento das imagens do *Sentinel-1*



Fonte: Adaptado de Street *et al.* (2022)

As bandas VH e VV, juntamente com os dados de elevação, compõem os canais de cada imagem de entrada do modelo. Inicialmente, esses dados são lidos e convertidos em *arrays NumPy* do tipo *float32*. Posteriormente, os *arrays* são empilhados de forma a compor três canais distintos e transformados em tensores. Da forma similar, cada máscara de inundação passa por esse processo de leitura e transformação.

¹⁴ <https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-1>

¹⁵ <https://github.com/johntruckenbrodt/pyroSAR>

Durante o treinamento de Redes Neurais Profundas, a mudança na distribuição das entradas em cada camada, devido à alteração dos parâmetros das camadas anteriores, desacelera o processo. Para amenizar esse problema, é comum aplicar métodos de normalização às entradas das camadas (IOFFE; SZEGEDY, 2015). Nesse contexto, os canais de cada imagem de entrada foram normalizados a partir do método *Z-Score*, descrito por:

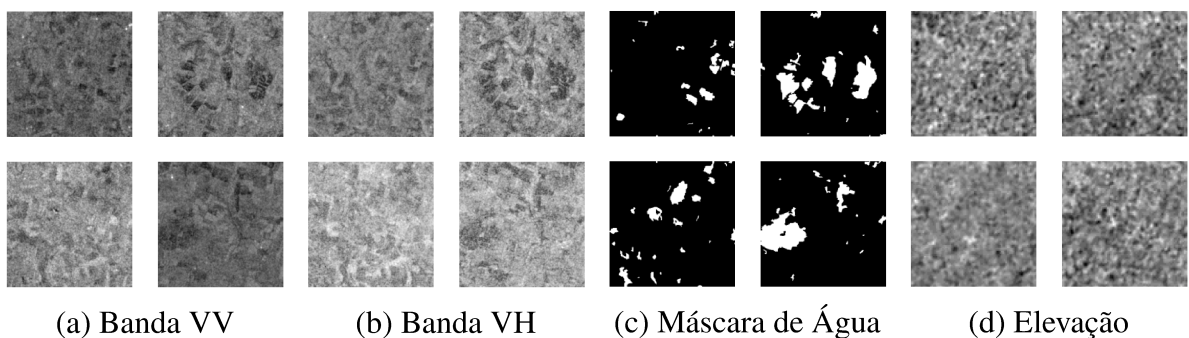
$$z = \frac{x - \mu}{\sigma} \quad (5.1)$$

onde z é a imagem normalizada, x a imagem original, μ a média e σ é o desvio padrão de todas as imagens de um mesmo canal. Esse método normaliza cada canal para que tenham média zero e desvio padrão unitário.

5.4 Aumento de Dados

O aumento de dados é uma técnica utilizada para elevar a quantidade de dados para o treinamento do modelo de aprendizagem de máquina, a fim de melhorar sua capacidade de generalização, reduzindo *overfitting* e aumentando a sua robustez. Nesse contexto, foi empregada a técnica *image splitting* para preservar a riqueza de informações. Cada canal e suas respectivas máscaras são divididas em 4 partes de tamanho 256 x 256, conforme ilustrado na Figura 12.

Figura 12 – Imagens de radar recortadas do conjunto *C2S-MS Floods*



Fonte: Elaborado pelo autor (2023)

Dessa técnica resultaram 3600 imagens de treinamento, das quais 70,42% contém inundações e 29,58% não apresentam inundações. Em seguida, as imagens foram separadas aleatoriamente em grupos de treinamento, validação e teste, preservando a distribuição do conjunto, na proporção de 70%, 15% e 15%, respectivamente. A fim de diversificar o conjunto de dados, foram aplicadas ainda as transformações a seguir:

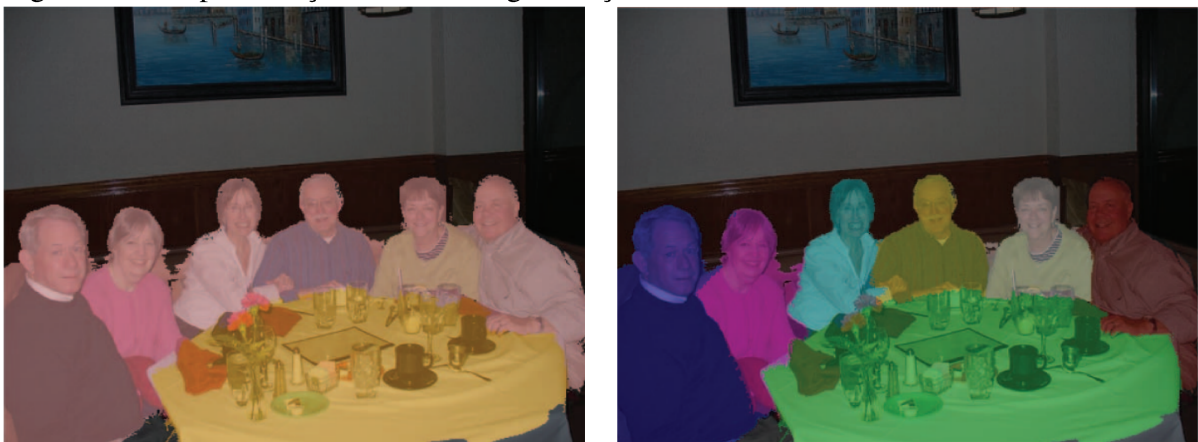
- 50% de chance de espelhar a imagem verticalmente;
- 50% de chance de espelhar a imagem horizontalmente.

5.5 Segmentação Semântica e de Instâncias

O uso convencional de CNNs concentra-se na classificação de imagens, em que um rótulo é atribuído à imagem inteira, todavia, tarefas mais avançadas em visão computacional exigem a classificação individual de cada pixel da imagem, segmentando-a em regiões associadas a categorias (RONNEBERGER *et al.*, 2015). Na segmentação semântica não há distinção entre diferentes ocorrências da mesma categoria, pois todas as instâncias classificadas na imagem recebem o mesmo rótulo.

Por outro lado, a segmentação baseada em instâncias tem como objetivo atribuir um identificador único a cada objeto segmentado na imagem. Portanto, cada instância de um objeto recebe um rótulo exclusivo e cada pixel é classificado de acordo com a instância na qual o objeto pertence. A Figura 13 exemplifica essa distinção. Na segmentação semântica, todas as pessoas a mesa estão agrupadas sob a mesma classe, enquanto na segmentação por instância cada pessoa é identificada separadamente (KUMAR, 2023; ARNAB *et al.*, 2018).

Figura 13 – Representação visual da segmentação semântica e de instâncias



(a) Segmentação Semântica

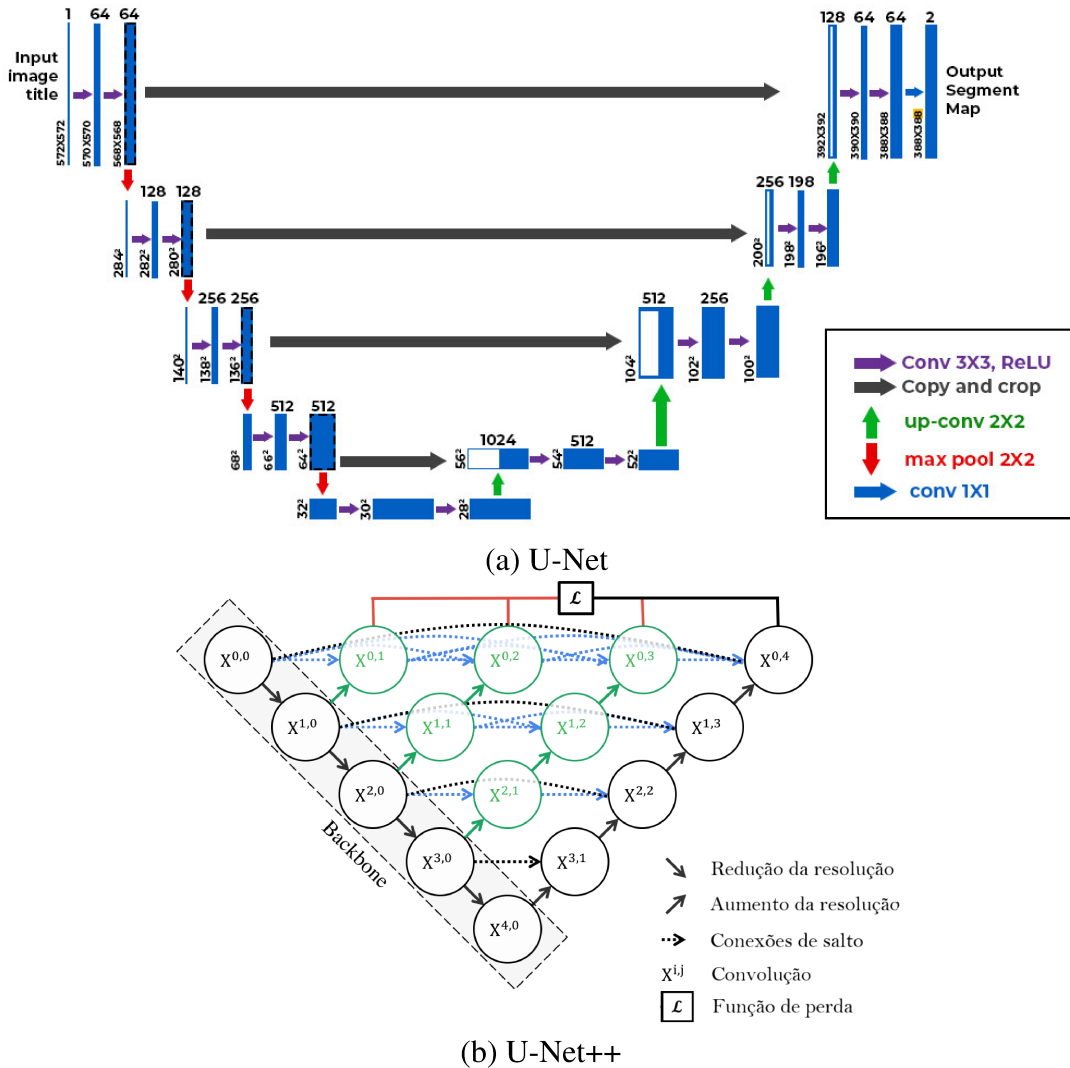
(b) Segmentação de Instâncias

Fonte: Arnab *et al.* (2018)

As FCNNs são empregadas na segmentação semântica de imagens. Elas substituem as camadas totalmente conectadas das arquiteturas convencionais por camadas convolucionais, que as tornam mais eficientes em uso de memória. Geralmente, as FCNNs adotam uma estrutura codificador-decodificador, com conexões de salto (*skip connections*, em inglês), como mostrado na Figura 14. O codificador (*down-sampling*) forma a primeira metade da rede, composto por camadas de convolução e *pooling*, reduzindo a resolução dos mapas de características e aumentando o número de canais.

O decodificador (*up-sampling*), na segunda metade da rede, incorpora camadas de amostragem ascendentes, aumentando a resolução dos mapas de características e reduzindo o número de canais para se aproximar da resolução da imagem original. As conexões de salto entre as sub-redes permitem que a rede incorpore informações de alto nível do codificador durante o processo de decodificação, melhorando a precisão da segmentação (KUMAR, 2023).

Figura 14 – Arquitetura das redes U-Net e U-Net++



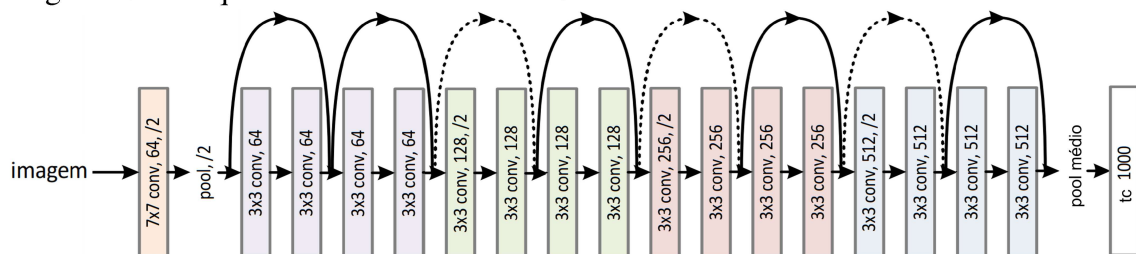
Para realizar a segmentação semântica de inundações nas imagens SAR do conjunto de dados *C2S-MS Flood*, foram utilizados os modelos U-Net (Figura 14a) e U-Net++ (Figura 14b) com a rede ResNet-18 (Figura 15) como codificador. O modelo U-Net foi especificamente concebido para realizar a segmentação de imagens médicas. Ela se tornou uma arquitetura popular devido à sua capacidade de preservar detalhes em imagens e à sua habilidade de produzir segmentações precisas mesmo com uma quantidade limitada de dados rotulados.

Posteriormente, o modelo U-Net++ foi proposto para aprimorar a precisão de segmentação nessa área. A principal inovação é preencher a lacuna semântica que existe entre os mapas de características do codificador e do decodificador da U-Net, por meio de caminhos de salto redesenhados, destacados em verde e azul na U-Net++, que conectam as duas sub-redes. Adicionalmente, a aplicação de supervisão profunda, ilustrada em vermelho, é incorporada ao modelo. Essa estratégia traz melhorias ao aprendizado, pois os mapas de características das sub-redes de codificação e decodificação são semanticamente similares (ZHOU *et al.*, 2018).

A Figura 15 ilustra a arquitetura de classificação ResNet-18, que incorpora o codificador das FCNNs utilizadas neste estudo. Esse modelo foi desenvolvido para lidar com o desafio de treinar redes profundas. Ele introduziu o conceito de blocos residuais, projetados para mitigar problemas associados ao aumento da profundidade da rede, como a redução gradual do gradiente durante o treinamento. Redes mais antigas muitas vezes sofrem degradação de desempenho devido à perda de informações conforme elas passam por diversas camadas. Essa degradação é resultado da complexidade crescente das representações aprendidas, levando a uma distorção ou perda de dados relevantes (HE *et al.*, 2016).

Na segmentação semântica, essa arquitetura é empregada como um extrator de características da imagem de entrada, posteriormente utilizadas por outras partes dos modelos U-Net e U-Net++. Ela é composta por várias camadas de convolução, subamostragem (*pooling* máximo e médio) e camadas totalmente conectadas. A biblioteca *Segmentation Model Pytorch* possibilita a utilização da técnica *Batch Normalization* após cada camada convolucional. Os autores Ioffe e Szegedy (2015) descrevem essa abordagem em detalhes.

Figura 15 – Arquitetura da rede ResNet-18



Fonte: Adaptado de He *et al.* (2016)

5.6 Funções de Perda

As funções de perda quantificam a diferença entre as máscaras reais e as previstas pelo modelo, permitindo que ele aprenda a segmentar corretamente áreas de interesse nas imagens. Essas funções oferecem uma métrica para avaliar o quão bem o modelo se ajusta às imagens. Valores menores indicam um ajuste mais preciso do modelo. Quando a perda é zero, significa que o modelo faz previsões perfeitas, estimando com precisão as máscaras verdadeiras (ZHANG *et al.*, 2023).

A entropia cruzada (SHRUTI, 2020) é uma medida da teoria da informação frequentemente empregada como função de perda em tarefas de segmentação de imagens. Ela indica diferença entre classificações produzidas pelo modelo e suas contrapartes verdadeiras, sendo sua função perda expressa para a tarefa de segmentação binária como:

$$BCE(y, \hat{y}) = -(y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})) \quad (5.2)$$

Na Equação 5.2, y é o rótulo original, \hat{y} é o previsto pelo modelo e $\log(\cdot)$ é a função logarítmica na base 2. Em cenários com desbalanceamento de dados, essa abordagem tende a gerar viés de representação (VINOD, 2020).

O coeficiente Sørensen-Dice (SHRUTI, 2020) é uma métrica de desempenho que avalia a sobreposição entre duas imagens, assumindo valores de 0 a 1. Ele é definido como duas vezes a interseção entre a máscara prevista pelo modelo e a verdadeira, dividido pela soma entre eles. A função de perda *Dice* é definida a partir do coeficiente *Dice* como:

$$DiceLoss(y, \hat{y}) = 1 - Dice = 1 - \frac{2 \times |y \cap \hat{y}|}{|y| + |\hat{y}|} \quad (5.3)$$

Outra função de perda importante é a *Tversky* (SHRUTI, 2020), que baseia-se numa medida de similaridade assimétrica denominada índice de *Tversky*. Esse último generaliza o coeficiente *Dice* e o Índice de *Jaccard*, sendo a perda de *Tversky*:

$$TverskyLoss(y, \hat{y}) = 1 - Tversky = 1 - \frac{|y \cap \hat{y}|}{|y \cap \hat{y}| + \alpha \cdot |y - \hat{y}| + \beta \cdot |\hat{y} - y|} \quad (5.4)$$

Nessa equação, α e β regulam o peso dos falsos positivos e dos falsos negativos na métrica, possibilitando ajustar sua sensibilidade. Quando ambos são iguais a 0,5, a fórmula se reduz ao coeficiente *Dice*, e quando iguais a 1 à perda de *Jaccard*. Ao definir um valor maior para β em comparação a α é possível penalizar os falsos negativos na comparação entre os rótulos y e \hat{y} (VINOD, 2020).

5.7 Métricas de Desempenho

As métricas de desempenho são comumente utilizadas em tarefas de IA para avaliar o quão bem um modelo está performando em uma determinada tarefa. No contexto de segmentação semântica, avalia-se o quão similares são as máscaras verdadeiras e previstas. As métricas de avaliação utilizadas durante a validação do modelo foram obtidas através do módulo *Metrics* da biblioteca *Segmentation Models Pytorch*.

O índice de *Jaccard*, conhecido como IoU, é utilizado para comparar a similaridade entre as máscaras verdadeiras e previstas. Seu resultado varia de 0 a 1, onde 0 indica nenhuma sobreposição entre os conjuntos e 1 indica uma sobreposição total (ZHANG *et al.*, 2023). Conforme expresso pela Equação 5.5, ela é definida como a divisão da área da interseção pelo tamanho da união dos conjuntos.

$$Jaccard(y, \hat{y}) = \frac{|y \cap \hat{y}|}{|y \cup \hat{y}|} \quad (5.5)$$

A sensibilidade, conhecida como revocação (*Recall*, em inglês) ou taxa dos verdadeiros positivos, mede a capacidade do modelo em identificar corretamente todos os exemplos positivos. Ela é descrita como:

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (5.6)$$

Na Equação 5.6, os Verdadeiros Positivos (VP) são os casos em que o modelo previu corretamente os *pixels* de inundações, enquanto os Falsos Negativos (FN) são os casos em que o modelo não previu onde havia *pixels* de inundações.

O F1 Score, descrita na Equação 5.8, é uma métrica que combina precisão e sensibilidade em uma única métrica. Quando há um desequilíbrio entre as classes, essa métrica pode ser mais confiável para avaliar o desempenho de um modelo. A precisão é descrita como:

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (5.7)$$

Na Equação 5.7, os Falsos Positivos (FP) são os casos em que o modelo categorizou *pixels* como inundações onde na verdade não havia.

$$\text{F1 Score} = \frac{2 \times \text{Precisão} \times \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}} \quad (5.8)$$

5.8 Treinamento do Modelo

Foram realizadas quatro execuções utilizando diferentes divisões aleatórias do conjunto de dados em treinamento, validação e teste, para reduzir o risco de obter um resultado enviesado. Ao repetir o treinamento do modelo, é possível obter uma visão mais abrangente do seu comportamento com diferentes amostras. Isso auxilia na avaliação da consistência do desempenho, indicando se o modelo é estável e capaz de generalizar bem para novos conjuntos de dados. Quando o desempenho se mantém bom em diferentes conjuntos, isso sugere uma maior robustez do modelo.

Todos os codificadores da biblioteca *Segmentation Models Pytorch* possuem pesos pré-treinados para melhorar o desempenho durante o treinamento. Foram utilizados os pesos *imagenet* para segmentação binária a partir de uma imagem de entrada com 3 canais. Além disso, foi utilizado o otimizador *Adam* com taxa de aprendizado de 0,001 durante, no máximo, 100 épocas. A escolha do otimizador e da taxa de aprendizado foi resultado de experimentação e análise dos resultados ao longo do treinamento. Para ajustar dinamicamente a taxa de aprendizado durante o processo de treinamento foi utilizado o método *ReduceLROnPlateau()*, que monitora a métrica IoU do conjunto de validação por 4 épocas e reduz a taxa em 50% se uma estagnação for detectada.

A técnica de parada antecipada (*Early Stopping*, em inglês) foi empregada para evitar o sobreajuste (*overfitting*, em inglês) durante o treinamento dos modelos. Seu objetivo é interromper o treinamento quando não há melhorias no desempenho do conjunto de validação após um determinado número de épocas. À medida que o modelo aprende com os dados de treinamento, há o risco de se ajustar demais a esses dados, perdendo a capacidade de generalização. Portanto, o *Early Stopping* busca encontrar um ponto ideal, onde o modelo tenha aprendido o suficiente sem ficar excessivamente ajustado aos dados de treinamento. Nesse contexto, o IoU do conjunto de validação foi observado por 13 épocas.

6 RESULTADOS

Os experimentos realizados com as funções de perda *Binary Cross-Entropy*, *Dice* e *Tversky* são apresentados nas Tabelas 3, 4 e 5, respectivamente. São empregadas as métricas IoU, pontuação F1 e sensibilidade para avaliar os conjuntos de validação e teste modelos U-Net e U-Net++. Cada linha final exibe as médias aritméticas e desvios-padrão entre quatro execuções.

Tabela 3 – Desempenhos de validação e teste utilizando a função de perda *Binary Cross-Entropy*

		U-Net		U-Net++	
		Validação	Teste	Validação	Teste
Exec. 1	IoU	83,16%	84,91%	84,10%	85,46%
	Pontuação F1	90,73%	91,80%	91,26%	92,13%
	Sensibilidade	87,94%	89,22%	89,80%	91,07%
Exec. 2	IoU	82,85%	80,00%	81,25%	81,50%
	Pontuação F1	90,55%	88,64%	89,54%	89,69%
	Sensibilidade	88,63%	86,28%	87,95%	88,19%
Exec. 3	IoU	81,53%	82,76%	82,26%	82,63%
	Pontuação F1	89,79%	90,44%	90,21%	90,40%
	Sensibilidade	87,98%	90,01%	89,29%	89,33%
Exec. 4	IoU	82,48%	81,94%	81,79%	81,86%
	Pontuação F1	90,29%	89,97%	89,88%	89,91%
	Sensibilidade	88,10%	89,28%	89,78%	91,46%
Média	IoU	82,50% ± 0,71	82,40% ± 2,00	82,35% ± 1,20	82,86% ± 1,80
	Pontuação F1	90,34% ± 0,41	90,21% ± 1,30	90,22% ± 0,74	90,53% ± 1,10
	Sensibilidade	88,16% ± 0,32	88,70% ± 1,70	89,21% ± 0,87	90,01% ± 1,50

Fonte: Elaborado pelo autor (2023)

Tabela 4 – Desempenhos de validação e teste utilizando a função de perda *Dice*

		U-Net		U-Net++	
		Validação	Teste	Validação	Teste
Exec. 1	IoU	83,61%	84,96%	85,46%	86,13%
	Pontuação F1	91,02%	91,84%	92,09%	92,52%
	Sensibilidade	90,09%	90,92%	91,88%	92,61%
Exec. 2	IoU	79,57%	80,30%	79,85%	80,09%
	Pontuação F1	88,47%	88,90%	88,71%	88,80%
	Sensibilidade	87,29%	87,17%	87,69%	87,50%
Exec. 3	IoU	81,11%	81,52%	80,30%	80,90%
	Pontuação F1	89,52%	89,69%	88,99%	89,29%
	Sensibilidade	89,63%	89,40%	87,25%	87,54%
Exec. 4	IoU	81,48%	81,06%	80,51%	80,10%
	Pontuação F1	89,72%	89,45%	89,08%	88,84%
	Sensibilidade	89,38%	90,09%	88,90%	89,34%
Média	IoU	81,44% ± 1,70	81,96% ± 2,10	81,53% ± 2,60	81,81% ± 2,90
	Pontuação F1	89,68% ± 1,00	89,97% ± 1,30	89,72% ± 1,60	89,86% ± 1,80
	Sensibilidade	89,10% ± 1,20	89,39% ± 1,60	88,93% ± 2,10	89,25% ± 2,40

Fonte: Elaborado pelo autor (2023)

Tabela 5 – Desempenhos de validação e teste utilizando a função de perda *Tversky*

		U-Net		U-Net++	
		Validação	Teste	Validação	Teste
Exec. 1	IoU	83,76%	85,74%	84,27%	84,97%
	Pontuação F1	91,08%	92,30%	91,34%	91,84%
	Sensibilidade	90,28%	91,89%	91,14%	91,75%
Exec. 2	IoU	79,53%	80,07%	81,80%	80,47%
	Pontuação F1	88,43%	88,77%	89,93%	88,99%
	Sensibilidade	87,08%	87,33%	88,32%	87,36%
Exec. 3	IoU	81,17%	81,41%	80,59%	81,36%
	Pontuação F1	89,48%	89,55%	89,19%	89,54%
	Sensibilidade	87,43%	87,65%	86,86%	86,88%
Exec. 4	IoU	80,20%	79,62%	80,48%	80,38%
	Pontuação F1	88,91%	88,56%	89,09%	89,03%
	Sensibilidade	87,93%	88,04%	87,91%	89,05%
Média	IoU	81,17% ± 1,90	81,71% ± 2,80	81,79% ± 1,80	81,80% ± 2,20
	Pontuação F1	89,48% ± 1,20	89,80% ± 1,70	89,89% ± 1,00	89,85% ± 1,40
	Sensibilidade	88,18% ± 1,40	88,73% ± 2,10	88,56% ± 1,80	88,76% ± 2,20









Fonte: Elaborado pelo autor (2023)

Esses resultados revelam que diferentes funções de perda têm um impacto sutil no desempenho das redes utilizadas. A função de perda *Binary Cross-Entropy* mostrou consistência, exibindo variações mínimas entre as execuções de ambas as redes. No entanto, a função de perda *Dice* apresentou certa variabilidade entre as execuções, especialmente em termos de sensibilidade. Isso se deve à sua fragilidade à sobreposição entre as áreas segmentadas e as áreas de referência. No caso da função de perda *Tversky*, os resultados médios foram semelhantes aos da perda *Dice*, mas com sensibilidade ligeiramente inferior. Essa variabilidade também é influenciada pela fragilidade aos limites da sobreposição entre as regiões.

A Tabela 6 possibilita comparar qualitativamente as máscaras de inundações reais e previstas pelas redes, para cada função de erro. A imagem predita é proveniente do conjunto de teste das primeiras execuções. A classe 1, representada em branco nas imagens, corresponde às áreas inundadas, enquanto a classe 0, em preto nas imagens, ilustra as áreas não inundadas. Os IoUs da rede U-Net para as funções de perda *Binary Cross-Entropy*, *Dice* e *Tversky* são, respectivamente, 60,90%, 94,88% e 93,73%. Já a rede U-Net++ produziu IoUs de 95,83%, 95,23% e 94,90% para as mesmas funções, respectivamente.

A rede U-Net++, usando a função de perda *Binary Cross-Entropy* para classificação de cada pixel, alcançou os melhores desempenhos no conjunto de testes, apresentando as maiores médias e os menores desvios-padrão para cada métrica avaliada.

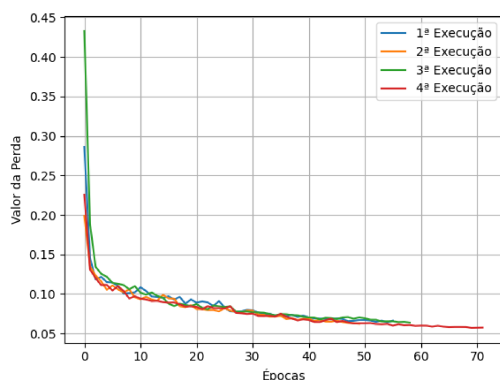
Tabela 6 – Predições dos modelos U-Net e U-Net++ para cada função de perda

	Máscara Real	Binary Cross-Entropy	Dice	Tversky
U-Net				
U-Net++				

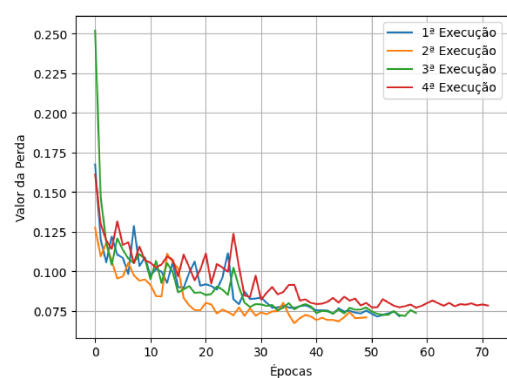
Fonte: Elaborado pelo autor (2023)

A Figura 16 exibe os gráficos com a evolução das perdas dos conjuntos de treinamento e validação em cada execução desse modelo. Ao longo das iterações, é observada uma tendência de redução nas perdas dos conjuntos de treinamento e validação, sugerindo uma melhoria progressiva do modelo. Essa tendência foi consistente em todos os experimentos realizados com as redes avaliadas para cada função de perda. Em alguns casos, contudo, observou-se uma disparidade entre as perdas de treinamento e validação, indicando possível sobreajuste do modelo. Nesses contextos, a aplicação da técnica de parada antecipada se mostrou útil.

Figura 16 – Evolução das perdas de treinamento e validação do melhor modelo



(a) Perdas de treinamento.



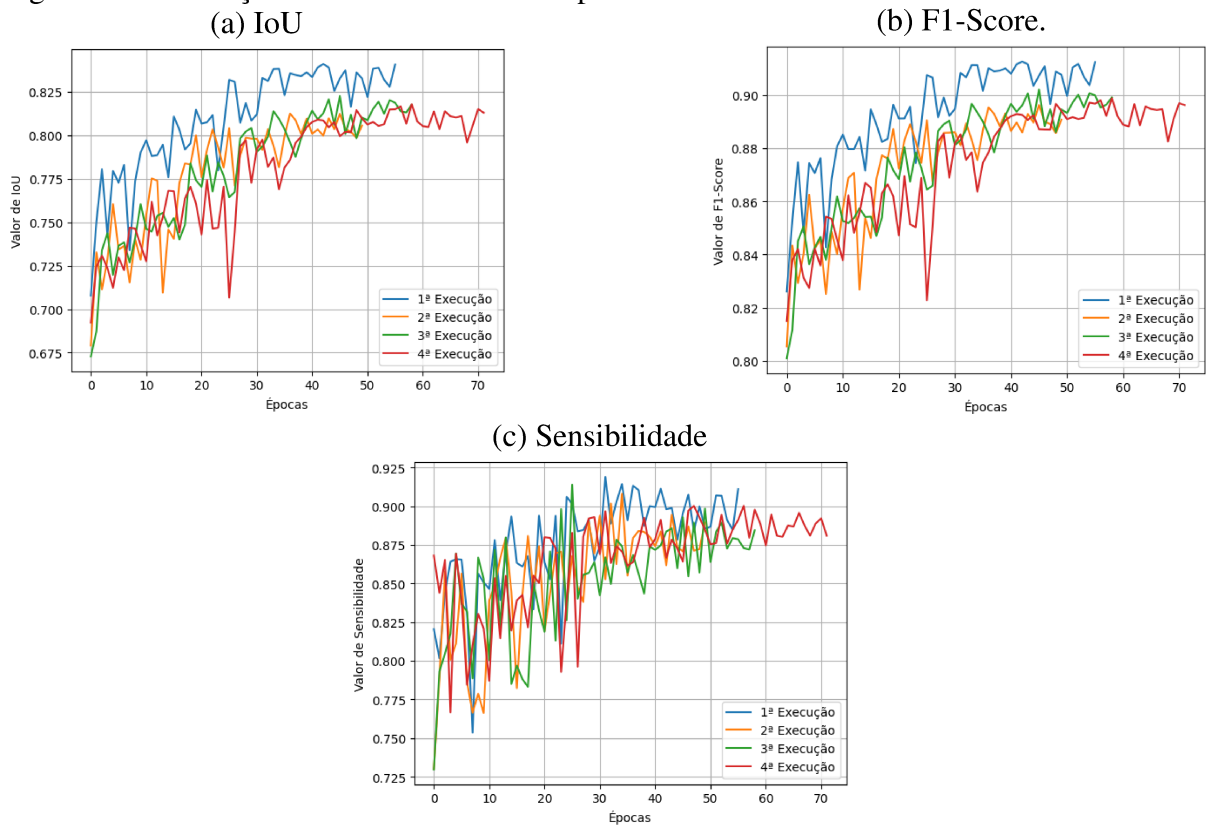
(b) Perdas de validação.

Fonte: Elaborado pelo autor (2023)

A Figura 17 exibe os gráficos com a evolução dos desempenhos do conjunto de validação em cada execução do melhor modelo. As diferentes execuções mostram resultados variados, indicando que a performance do modelo pode ser sensível a fatores como inicialização aleatória, hiperparâmetros ou à natureza estocástica do treinamento de redes neurais.

A tendência de aumento do IoU (Figura 17a) em algumas execuções sugere melhorias na precisão da segmentação ao longo do treinamento. A pontuação F1 (Figura 17b) e a sensibilidade (Figura 17c) demonstram melhorias em algumas execuções, indicando uma melhor capacidade do modelo em capturar corretamente a classe de inundação.

Figura 17 – Evolução das métricas de desempenho do melhor modelo



Fonte: Elaborado pelo autor (2023)

7 CONCLUSÕES E TRABALHOS FUTUROS

A análise comparativa de funções de perda teve como foco aprimorar a segmentação semântica de áreas inundadas em imagens SAR, a partir das FCNNs. O aumento de dados distribuiu a quantidade excessiva de informações das imagens de satélite em subimagens, expandindo os dados de um total de 900 para 3600 imagens. Tal estratégia preservou a riqueza de informações sem comprometer o desempenho computacional do modelo. Além disso, dados de elevação dos terrenos foram integrados para conferir contexto espacial às análises.

O desempenho das redes U-Net e U-Net++ na tarefa de mapeamento de inundações foi avaliado considerando as funções de perda *Binary Cross-Entropy*, *Dice* e *Tversky* ($\alpha = \beta = 1$). Isso permitiu comparar o impacto dessas funções no desempenho dos modelos. A arquitetura U-Net++, combinada com a função de perda *Binary Cross-Entropy*, exibiu as maiores médias e os menores desvios-padrão, entre quatro execuções distintas, para cada métrica avaliada.

Especificamente, a U-Net++ com *Binary Cross-Entropy* alcançou um IoU de 82,80%, pontuação F1 de 90,53% e sensibilidade de 90,01%. Essa arquitetura demonstrou um padrão de segmentação com menos Falsos Positivos em comparação com as áreas previstas pela U-Net. Esses resultados consolidam a importância da escolha adequada da função de perda para alcançar bons resultados na segmentação semântica de áreas inundadas em imagens de satélite por meio de FCNNs.

Planeja-se, para futuros trabalhos, a exploração da fusão de dados provenientes de SAR com outros conjuntos de inundações rotulados, a fim de realizar análises mais abrangentes. Isso incluirá a incorporação de séries temporais para uma detecção dinâmica e contínua de mudanças em áreas inundadas ao longo do tempo. Além disso, há a intenção de investigar outras funções de perda ou adaptações das funções existentes, levando em conta o ruído *Speckle*, visando avaliar seu impacto na precisão das segmentações realizadas. Por fim, pretende-se explorar técnicas de aprendizado profundo para lidar com conjuntos de dados em que a anotação manual seja limitada ou onerosa.

REFERÊNCIAS

- ALMEIDA, G. A. M. de; BATES, P.; OZDEMIR, H. Modelling urban floods at submetre resolution: challenges or opportunities for flood risk management? **Journal of Flood Risk Management**, CIWEM, v. 11, p. S855–S865, 2018.
- ARNAB, A.; ZHENG, S.; JAYASUMANA, S.; ROMERA-PAREDES, B.; LARSSON, M.; KIRILLOV, A.; SAVCHYNSKYI, B.; ROTHER, C.; KAHL, F.; TORR, P. H. Conditional random fields meet deep neural networks for semantic segmentation: Combining probabilistic graphical models with deep learning for structured prediction. **IEEE Signal Processing Magazine**, v. 35, n. 1, p. 37–52, 2018.
- BORN, G. H.; DUNNE, J. A.; LAME, D. B. Seasat mission overview. **Science**, v. 204, n. 4400, p. 1405–1406, 1979.
- CURLANDER, J. C.; MCDONOUGH, R. N. **Synthetic Aperture Radar: Systems and signal processing**. Cham: Wiley-Interscience, 1991. (ISBN 978-0-471-85770-9).
- EM-DAT. **EM-DAT Public**. Emergency Events Database, 2020. The International Disaster Database of Centre for Research on the Epidemiology of Disasters (CRED). Disponível em: <https://public.emdat.be/>. Acesso em: 23 nov. 2023.
- ESA. **Seasat**. European Space Agency, 2013. Disponível em: <https://earth.esa.int/eogateway/missions/seasat>. Acesso em: 06 set. 2023.
- ESA. **The Sentinel-1 Toolbox**. European Space Agency, 2015. Disponível em: <https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-1>. Acesso em: 02 nov. 2023.
- FRERY, A. C. Statistical information theory and geometry for sar. image analysis: Invited paper. **IEEE Recent Advances in Geoscience and Remote Sensing : Technologies, Standards and Applications (TENGARSS)**, p. 1–4, 2019.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S. l.: s. n.], 2016. p. 770–778.
- INPE. **Sistemas de Radar**. Instituto Nacional de Pesquisas Espaciais, 2001. Disponível em: <https://www.dpi.inpe.br/spring/portugues/tutorial/radar.html>. Acesso em: 06 nov. 2023.
- IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: **Proceedings of the 32nd International Conference on International Conference on Machine Learning**. Lille, France: JMLR, 2015. (ICML'15, v. 37), p. 448–456.
- IPCC. **Climate Change 2023: Synthesis Report: Summary for policymakers**. Geneva, Switzerland: Intergovernmental Panel on Climate Change, 2023. 1-34 p.
- IRDR. **IRDR Peril Classification and Hazard Glossary**. Integrated Research on Disaster Risk, 2014. Disponível em: https://www.irdrinternational.org/knowledge_pool/publications/173. Acesso em: 30 ago. 2023.
- JENSEN, J. R. **Remote Sensing of the Environment: An earth resource perspective**. [S. l.]: Pearson, 2006. (ISBN-13: 9780131889507).

KONG, Y.; LIU, Y.; YAN, B.; LEUNG, H.; PENG, X. A novel deeplabv3+ network for sar imagery semantic segmentation based on the potential energy loss function of gibbs distribution. **Remote Sensing**, v. 13, n. 3, 2021. ISSN 2072-4292.

KUMAR, A. **CNN Basic Architecture for Classification Segmentation**. Analytics Yogi, 2023. Disponível em: <https://vitalflux.com/cnn-basic-architecture-for-classification-segmentation/>. Acesso em: 30 set. 2023.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, p. 436–444, 2015.

MATHWORKS. **What Is a Convolutional Neural Network?** 2017. Disponível em: <https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html>. Acesso em: 23 set. 2023.

MONTELLO, F.; ARNAUDO, E.; ROSSI, C. Mmflood: A multimodal dataset for flood delineation from satellite imagery. **IEEE Access**, v. 10, p. 96774–96787, 2022.

NISAR. **Polarimetry**. NASA-ISPRO SAR Mission, 2020. Disponível em: <https://nisar.jpl.nasa.gov/mission/get-to-know-sar/polarimetry/>. Acesso em: 30 set. 2023.

RAMBOUR, C.; AUDEBERT, N.; KOENIGUER, E.; SAUX, B. L.; CRUCIANU, M.; DATCU, M. Flood detection in time series of optical and sar images. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, XLIII-B2-2020, p. 1343–1346, 2020.

RICHARDS, J. A. **Remote Sensing Digital Image Analysis**. Cham: Springer, 2022. (ISBN 978-3-030-82326-9).

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015**, Springer, Cham, v. 9351, p. 234–241, 2015.

SAUSEN, T. M.; LACRUZ, M. S. P. **Sensoriamento Remoto para Desastres**. São Paulo: Oficina de Textos, 2015. (ISBN 978-85-7975-175-2).

SHRUTI, J. A survey of loss functions for semantic segmentation. In: **2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)**. [S. l.: s. n.], 2020. p. 1–7.

STREET, C. to; MICROSOFT; FOUNDATION, R. E. **A global flood events and cloud cover dataset**. Radiant MLHub, 2022. Disponível em: <https://beta.source.coop/repositories/c2sms/c2smsfloods/description/>. Acesso em: 25 set. 2023.

SÃO PAULO. **Litoral Norte de SP registrou maior acumulado de chuva da história**. Governo do Estado de São Paulo, 2023. Disponível em: <https://www.saopaulo.sp.gov.br/ultimas-noticias/litoral-norte-de-sp-registrou-maior-acumulado-de-chuva-da-historia-2/>. Acesso em: 26 mai. 2023.

TANIM, A. H.; MCRAE, C. B.; TAVAKOL-DAVANI, H.; GOHARIAN, E. Flood detection in urban areas using satellite imagery and machine learning. **Water**, MDPI, v. 14, n. 7, p. 1140, 2022.

UNDRR. **Sendai Framework for Disaster Risk Reduction 2015 - 2030**. Sendai, Japan: United Nations Office for Disaster Risk Reduction, 2015. Disponível em: <https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030>. Acesso em: 18 mai. 2023.

UNDRR. **Disaster**. United Nations Office for Disaster Risk Reduction, 2020. Disponível em: <https://www.undrr.org/terminology/disaster>. Acesso em: 02 mai. 2023.

UNHCR. **Global Trends: Forced Displacement In 2022**. United Nations High Commissioner for Refugees, 2023. Disponível em: <https://www.unhcr.org/global-trends-report-2022>. Acesso em: 28 ago. 2023.

VINOD, R. **Dealing with class imbalanced image datasets using the Focal Tversky Loss**. Towards Data Science, 2020. Disponível em: <https://towardsdatascience.com/dealing-with-class-imbalanced-image-datasets-1cbd17de76b5>. Acesso em: 26 out. 2023.

WHO. **Floods**. World Health Organization, 2022. Disponível em: <https://www.who.int/health-topics/floods>. Acesso em: 03 set. 2023.

WMO. **WMO Atlas of Mortality and Economic Losses from Weather, Climate and Water-related Hazards (1970 - 2019)**. Geneva: World Meteorological Organizations, 2021. (ISBN 978-92-63-11267-5).

WMO. **Status of Mortality and Economic Losses due to Weather, Climate and Water Extremes (1970 - 2021)**. World Meteorological Organization, 2023. Disponível em: <https://public.wmo.int/en/resources/atlas-of-mortality>. Acesso em: 22 ago. 2023.

ZHANG, A.; LIPTON, Z. C.; LI, M.; SMOLA, A. J. **Dive into Deep Learning**. [S. l.]: Cambridge University Press, 2023. <https://D2L.ai>.

ZHOU, Z.; SIDDIQUEE, M. M. R.; TAJBAKSH, N.; LIANG, J. Unet++: A nested u-net architecture for medical image segmentation. **Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support – DLMIA ML-CDS 2018**, Springer, Cham, v. 11045, p. 3–11, 2018.

ZHU, X. X.; TUIA, D.; MOU, L.; XIA, G.-S.; ZHANG, L.; XU, F.; FRAUNDORFER, F. Deep learning in remote sensing: A comprehensive review and list of resources. **IEEE Geoscience and Remote Sensing Magazine**, v. 5, n. 4, p. 8–36, 2017.