# CLASSIFICAÇÃO DE INUNDAÇÕES NO RIO GRANDE DO SUL UTILIZANDO TEMPCNN EM SÉRIES TEMPORAIS DE IMAGENS SENTINEL-2

*Cesar Augusto de Moraes Costa*[1], *Brenda Oliveira Rocha*[1], *Alisson Cleiton de Oliveira*[1], *Karine Reis Ferreira*[1], *Lubia Vinhas*[1], *Gilberto Ribeiro Queiroz*[1] *e Thales Sehn Körting*[1]

[1]Instituto Nacional de Pesquisas Espaciais - INPE, São Jose dos Campos - SP, cesar.moraes@inpe.br, brenda.rocha@inpe.br, alisson.oliveira@inpe.br, karine.ferreira@inpe.br, lubia.vinhas@inpe.br, gilberto.queiroz@inpe.br, thales.korting@inpe.br

**RESUMO**

Este artigo apresenta uma análise da classificação do evento de inundação ocorrido em 2024 no estado do Rio Grande do Sul, Brasil, utilizando a Rede Neural Convolucional Temporal (TempCNN). O objetivo principal foi avaliar a capacidade da rede de processar séries temporais de quatro imagens, considerando que estudos anteriores focam predominantemente em pares de imagens, capturadas antes e logo após o desastre. Com três imagens de antes e uma de depois do evento de inundação, a rede foi encarregada de identificar as mudanças ocorridas. Os resultados indicam que a TempCNN, em sua configuração sem camadas de *Pooling*, apresentou desempenho semelhante tanto para as séries temporais compostas por quatro imagens, quanto para imagens bitemporais, utilizando as bandas BGR-NIR e as bandas GR-NIR, sem a banda *Blue*. No conjunto de inferência, a métrica da média Interseção sobre União (mIoU) alcançou um valor de 72,8% para os dados balanceados.

***Palavras-chave –*** *Inundação, TempCNN, Sensoriamento Remoto, Séries Temporais de Imagens de Satélite, Sentinel-2.*

***ABSTRACT***

*This paper presents an analysis of the classification of the 2024 flood event in the state of Rio Grande do Sul, Brazil, using the Temporal Convolutional Neural Network (TempCNN). The main objective was to evaluate the network's ability to process time series of four images, considering that previous studies predominantly focus on pairs of images, captured before and immediately after the disaster. With three images before and one after the flood event, the network was tasked with identifying the changes that occurred. The results indicate that TempCNN, in its configuration without Pooling layers, presented similar performance for both the time series composed of four images and for bitemporal images, using the BGR-NIR bands and the GR-NIR bands, without the Blue band. In the inference set, the mean Intersection over Union (mIoU) metric reached a value of 72.8% for the balanced data.*

***Key words –*** *Flood, TempCNN, Remote Sensing, Satellite Image Time Series, Sentinel-2.*

## 1. INTRODUÇÃO

As imagens de Sensoriamento Remoto (SR) tornaram-se essenciais para o monitoramento de desastres ambientais, permitindo respostas rápidas para resgates emergenciais, auxiliando órgãos competentes e agências humanitárias. Em particular, a análise do uso e cobertura da Terra (*Land Use and Land Cover* - LULC) identifica diferentes tipos de cobertura física da superfície terrestre. Atividades humanas, como urbanização, mineração, agricultura e extração de madeira, feitas de maneira inadequada, ocasionam a degradação do meio ambiente, resultando na perda de biodiversidade e impactos severos nos ecossistemas, que contribuem para a ocorrência de desastres ambientais [1].

Neste trabalho, foi avaliado o desempenho da rede TempCNN (*Temporal Convolutional Neural Network*) na classificação de inundações no estado do Rio Grande do Sul, utilizando Séries Temporais de Imagens de Satélite (SITS) do Sentinel-2. As classificações foram avaliadas tanto em uma série temporal de quatro imagens quanto em imagens bitemporais, utilizando as bandas BGR-NIR e GR-NIR, excluindo a banda do *Blue*.

A TempCNN é usada com o propósito de classificar uma série temporal de imagens de LULC [2]. Embora sua aplicação seja mais comum em classificações multitemporais, ela também foi eficaz na classificação de imagens bitemporais, especialmente em situações de desastres, conforme foi aplicada para este estudo.

Ao realizar classificações com as CNNs (*Convolutional Neural Networks*), a rede captura padrões espaciais (por exemplo, um formato oval) e padrões espectrais (por exemplo, feições derivadas da combinação de bandas dos sensores dos satélites) [3]. Neste estudo, será realizada uma análise com a TempCNN que utiliza CNNs para capturar a dimensão temporal em SITS.

Na seção de Materiais e Métodos, detalhamos o processo de seleção, tratamento e processamento das imagens, assim como a implementação do modelo de classificação para a detecção de áreas inundadas e não inundadas.

## 2. MATERIAL E MÉTODOS

A base de dados foi construída a partir da visualização dos produtos da iniciativa Charter (Carta Internacional Espaço e Grandes Desastres, do termo em inglês: *International Charter Space and Major Disasters*), com foco nas inundações ocorridas próximas à cidade de Porto Alegre–RS [4].

Com base nos mapas fornecidos pelo Charter, as imagens foram coletadas durante quatro períodos entre dezembro de 2023 e maio de 2024, abrangendo tanto o período anterior quanto posterior ao desastre. Esses dados foram processados

para servir de entrada na rede TempCNN, que realizou a classificação das áreas inundadas e não inundadas. A escolha das imagens foi auxiliada por produtos de mapas do Charter, disponibilizados com os dados dos rótulos das classes, executada por profissionais especialistas para a ativação número 875 referente ao desastre do Rio Grande do Sul, Brasil. O processamento das imagens, em composições BGR-NIR (*Blue, Green, Red, Near-infrared*) e GR-NIR, foi implementado em Python, em ambiente local.

As Figuras 1 e 2 apresentam a área de estudo com uma imagem de antes e outra de depois das inundações, respectivamente, capturadas pelo satélite Sentinel-2, enquanto a Figura 3 mostra o mapa das classes extraídas, onde as áreas inundadas foram vetorizadas a partir da imagem pós-desastre.

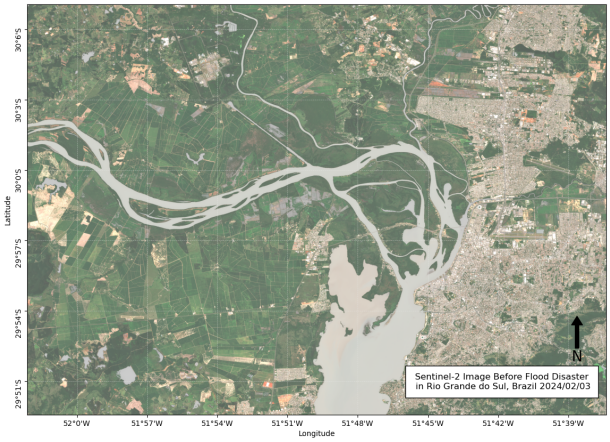



**Figura 1: Composição colorida da imagem do satélite Sentinel-2, capturada antes das inundações nos municípios de Porto Alegre e arredores.**

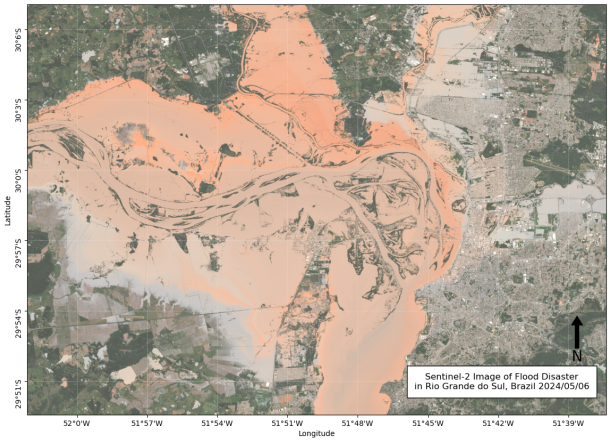



**Figura 2: Composição colorida da imagem do satélite Sentinel-2, capturada após as inundações que afetaram os municípios de Porto Alegre e seus arredores.**

No recorte das imagens consideradas no estudo, de tamanho 3200 px × 4000 px, o número de *pixels* com o valor 1 (inundação) foi de 3.797.135 e com o valor 0 (não inundação) foi de 9.002.865. Para melhorar o balanceamento dos dados, foi considerada uma área menor, com isso os

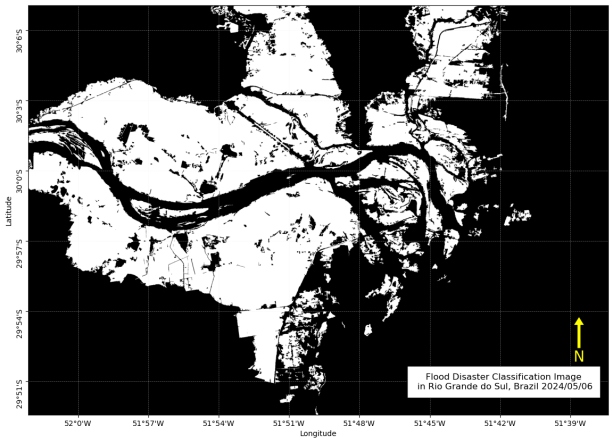



**Figura 3: Mapa com as classes de inundação (em branco) e não inundação (em preto) extraídas da imagem do Sentinel-2 após a ocorrência do desastre.**

valores ficaram da seguinte forma: área recortada de tamanho 2200 px × 3036 px, número de *pixels* com o valor 1 (inundação) igual a 3.339.817, e número de *pixels* com o valor 0 (não inundação) igual a 3.339.383. Os dados foram divididos em 60% para treinamento e 40% para teste.

As amostras de inundação foram obtidas manualmente ao comparar as imagens capturadas antes e após o evento de inundação, utilizando mapas elaborados por especialistas durante a ativação do Charter. A segmentação da imagem pós-evento (Figura 2) foi realizada com a ajuda de módulos da biblioteca RSGISLib para posteriormente realizar a amostragem dos dados por meio do *software* QGIS.[1]

### 2.1. Arquitetura TempCNN

A TempCNN é uma arquitetura proposta para classificação de séries temporais de imagens de satélite (SITS). O uso das convoluções no modelo TempCNN opera diretamente em SITS, considerando-as como imagens 3D, onde a terceira dimensão representa o tempo. A arquitetura TempCNN emprega convoluções com capacidade de extrair características espaço-temporais valiosas das SITS [2]. Para repetir o experimento deste trabalho, disponibilizamos o código em `https://github.com/CesarAugusto88/geoinformatic_floods`.

O modelo usado neste manuscrito seguiu a configuração da TempCNN adicionando três camadas convolucionais 3D, sem a utilização de camadas de *Pooling*, conforme sugerido no artigo de Pelletier *et al.* (2019). Foi visto que em estudos semelhantes para classificação de águas de inundação, usando imagens do Sentinel 1 e Sentinel-2, a métrica utilizada para avaliar o desempenho dos algoritmos de redes neurais foi a média *Intersection over Union* (mIoU) [5]. O presente estudo observou que a TempCNN obteve mIoU entre 84,4% a 86,9% para dados desbalanceados e mIoU de 72,7% com dados balanceados, na classificação de inundações, validando seu desempenho para imagens de média resolução espacial

[1] Disponível no Notebook `segmentation2vectorization_RSGISLib.ipy` no repositório do GitHub: `https://github.com/CesarAugusto88/geoinformatic_floods`.

(10 metros).

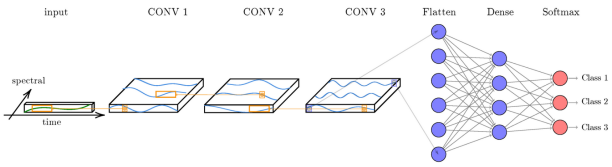



**Figura 4: Arquitetura da rede TempCNN. A entrada da rede é uma série temporal. Três filtros convolucionais são aplicados consecutivamente, uma camada densa e, por fim, a camada Softmax, que fornece a classe de previsão [2].**

O *Max-Pooling*, é empregado em tarefas de visão computacional para acelerar os cálculos e tornar as características aprendidas mais robustas. O *Average Pooling Global* é usado para reduzir a resolução da série temporal produzida pelos neurônios. Como relatado no artigo de implementação da TempCNN para séries temporais de imagens de SR, os efeitos de camadas das *Pooling* levaram à redução da acurácia, levando o modelo a uma grande redução do número de parâmetros treináveis, que resulta em uma redução extrema na dimensionalidade, consequentemente reduzindo o desempenho das métricas para o modelo [2].

Conforme mencionam os autores Pelletier *et al.* (2019) [2], ao integrar convoluções temporais e espectrais na mesma arquitetura, têm-se vantagens significativas em termos de precisão e capacidade de generalização na classificação da TempCNN.

## 3. RESULTADOS

Avaliando visualmente os dados da máscara de referência (*Ground truth*), é possível perceber que as classes não foram bem representadas. Isso é explicado porque imagens de média resolução espacial exigem um trabalho mais detalhado para rotular adequadamente todas as áreas. Porém, o modelo TempCNN demonstrou boas previsões, tanto para a série temporal composta por quatro imagens quanto para as imagens bitemporais, como apresenta a Figura 5.

As métricas de desempenho obtidas pelo modelo TempCNN com os dados desbalanceados estão apresentadas nas Tabelas 1 e 2.

| (**BGR-NIR**) | Inferência 4T | Inferência 2T |
|---|---|---|
| F1-*score* (*weighted*) | 0,96747 | 0,96673 |
| Acurácia | 0,96719 | 0,96631 |
| mIoU | 0,84623 | 0,84616 |

**Tabela 1: Métricas de F1-*score*, Acurácia e mIoU obtidas no conjunto de inferência para o modelo TempCNN para série temporal de quatro imagens (4T) e bitemporal (2T) em composição BGR-NIR.**

Com as métricas apresentadas nas Tabelas 1 e 2, observa-se uma leve melhoria utilizando os dados de entrada com as bandas GR-NIR com quatro tempos. Por ser uma série de imagens com revisitas próximas do Sentinel-2, conseguiu-se outras duas imagens semelhantes à de logo antes do evento de inundação, ajudando a identificar a mudança abrupta ocorrida na última imagem que contém inundações. Considerando

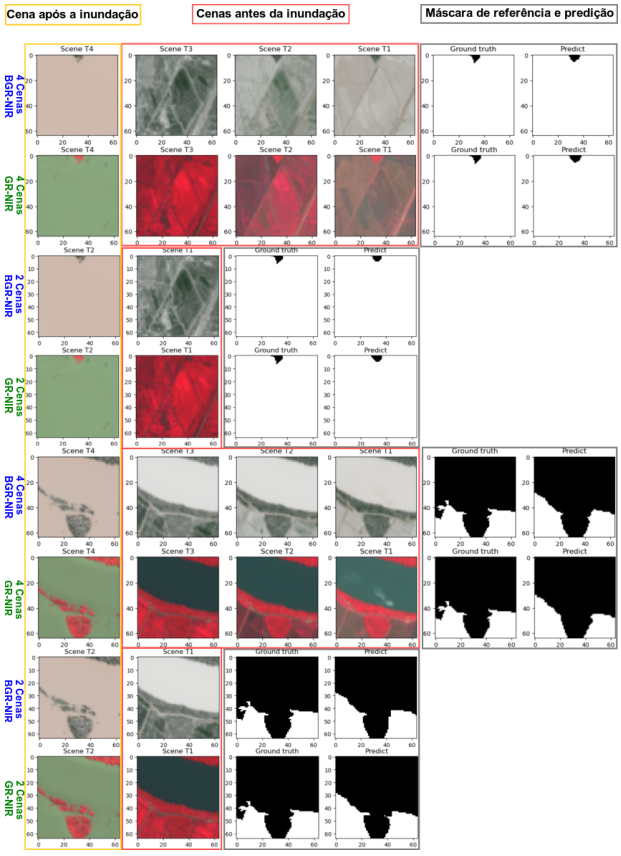



**Figura 5: Resultados das classificações com a TempCNN na inferência do modelo com a série temporal de quatro imagens (T4,T3,T2,T1) e com imagens bitemporais (T2,T1). De cima para baixo e da esquerda para direita, são apresentadas quatro cenas (BGR-NIR) de dimensões 64 px × 64 px, uma cena após a inundação - em amarelo, e três cenas antes da inundação - em vermelho, a máscara de referência (*Ground Truth*) e a predição (*Predict*) do modelo - em cinza. Na próxima linha são apresentadas cenas com as bandas GR-NIR, seguindo a mesma lógica. Depois, nas duas linhas abaixo, são apresentadas duas cenas (bitemporal), com as bandas BGR-NIR e GR-NIR, seguidas da máscara de referência e a predição. As demais linhas seguem o mesmo formato.**

| (**GR-NIR**) | Inferência 4T | Inferência 2T |
|---|---|---|
| F1-*score* (*weighted*) | 0,9697 | 0,9656 |
| Acurácia | 0,9698 | 0,9653 |
| mIoU | 0,8694 | 0,8440 |

**Tabela 2: Métricas de F1-*score*, Acurácia e mIoU obtidas no conjunto de inferência para o modelo TempCNN para série temporal de 4 imagens (4T) e bitemporal (2T) em composição GR-NIR.**

isso, a Tabela 3 apresenta as mesmas métricas para os dados balanceados descritos no final da seção 2.

## 4. DISCUSSÃO

Conforme o artigo de Pelletier *et al.* (2019), as camadas de *Pooling* descartam informações valiosas quando se trata de dados de séries temporais de imagens de satélites. A perda

| (**GR-NIR**) | Inferência 4T |
|---|---|
| F1-*score* (*weighted*) | 0,92083 |
| Acurácia | 0,92132 |
| mIoU | 0,72810 |

**Tabela 3: Métricas de F1-*score*, Acurácia e mIoU obtidas no conjunto de inferência para o modelo TempCNN gerado nos dados balanceados para série temporal de quatro imagens (4T).**

de resolução temporal pode reduzir a capacidade do modelo em capturar com precisão a dinâmica dos dados temporais. Por esse motivo, a arquitetura TempCNN não inclui tais camadas, permitindo que seja mantida a resolução temporal completa em todas as camadas convolucionais, essencial para o contexto de classificação de séries temporais [2]. O artigo também apresenta que para séries temporais curtas o método *Average Pooling* pode ter um ganho.

No problema de detecção de água em desastre de inundação, pode ser viável o uso de camadas de *Pooling*, conforme é usada em outras arquiteturas, como, por exemplo, ResNet-FCN, Deeplab v3, $U^2$Net e BASNet [5]. Porém, para este trabalho foi visto que a rede TempCNN usada sem camadas de *Pooling*, tanto para uma série temporal de imagens (quatro tempos), quanto para uma série temporal pequena, com imagens bitemporais (dois tempos), a rede TempCNN teve desempenho semelhante para o contexto de classificação de desastres de inundações.

Para a implementação da TempCNN deste artigo, foram apresentadas como entrada da rede séries temporais de quatro imagens e imagens bitemporais, com as bandas BGR-NIR e GR-NIR. A banda *Blue* foi descartada, devido à sua sensibilidade a artefatos atmosféricos [2]. Apesar dessa banda apresentar essa discrepância, os resultados apresentaram variações pequenas para o problema de inundações.

## 5. CONCLUSÕES

Os resultados obtidos pela TempCNN podem contribuir significativamente para a gestão de desastres de inundações utilizando como entradas uma série temporal de 4 imagens (três anteriores e uma após o desastre), com melhor capacidade de detecção, conforme apresentado nos resultados, usando as bandas GR-NIR, e assim classificando melhor áreas inundadas. As SITS nos permitem detectar processos complexos difíceis de identificar usando abordagens tradicionais de detecção de mudanças em tempos distintos [6]. Nesse viés, podemos gerar previamente um modelo mais acurado para detecção das áreas inundadas e não inundadas para auxiliar o salvamento de vidas em futuros desastres, cuja frequência tem aumentando.

Em trabalhos futuros será avaliada uma aplicação de classificação de desastre de inundação usando as redes CNN+RNN, que implementa a união da rede convolucional com a rede recorrente, enquanto a primeira captura características espectrais-espaciais, a última analisa a sequência temporal em imagens bitemporais [7]. Essa rede é treinável de ponta a ponta para detecção de mudanças em imagens de SR para classificação de LULC.

## 6. AGRADECIMENTOS

Os autores agradecem a CAPES, a FAPESP (Processo 2023/09118-6) e ao CNPq (Processos 302205/2023-3 e 302517/2023-5).

## 7. REFERÊNCIAS

[1] Zhujun Gu and Maimai Zeng. The use of artificial intelligence and satellite remote sensing in land cover change detection: review and perspectives. *Sustainability*, 16(1):274, 2023.

[2] Charlotte Pelletier, Geoffrey I Webb, and François Petitjean. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sensing*, 11(5):523, 2019.

[3] Aurélien Géron. *Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow*. Alta Books, 2019.

[4] Disasters Charter. Flooding in brazil , available in: `https://disasterscharter.org/web/guest/activations/-/article/flood-in-brazil-activation-875-`. access in: 05 aug., 2024.

[5] Yanbing Bai, Wenqi Wu, Zhengxin Yang, Jinze Yu, Bo Zhao, Xing Liu, Hanfang Yang, Erick Mas, and Shunichi Koshimura. Enhancement of detecting permanent water and temporary water in flood disasters by fusing sentinel-1 and sentinel-2 imagery using deep learning algorithms: Demonstration of sen1floods11 benchmark datasets. *Remote Sensing*, 13(11):2220, 2021.

[6] Valerie J Pasquarella, Christopher E Holden, Les Kaufman, and Curtis E Woodcock. From imagery to ecology: leveraging time series of all available landsat observations to map and monitor ecosystem state and dynamics. *Remote Sensing in Ecology and Conservation*, 2(3):152–170, 2016.

[7] Lichao Mou, Lorenzo Bruzzone, and Xiao Xiang Zhu. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2):924–935, 2018.