

Universidade Federal de Mato Grosso do Sul
Faculdade de Computação

**Atividade Orientada ao Ensino: Criação de Dataset de
Imagens Aéreas para Reconhecimento de Erosão em
Imagens de VANTs**

Aluno: Isabella Beatriz Montania Dominguez da Silva
Professor orientador: Wesley Nunes Gonçalves

Junho
2025

Universidade Federal de Mato Grosso do Sul
Faculdade de Computação

**Atividade Orientada ao Ensino: Criação de Dataset de
Imagens Aéreas para Reconhecimento de Erosão em
Imagens de VANTs**

Relatório de Atividade Orientada ao Ensino apre-
sentado ao Curso de Engenharia da Computação
da Universidade Federal de Mato Grosso do Sul
(UFMS)

Aluna: Isabella Beatriz Montania Dominguez da
Silva

Professor orientador: Wesley Nunes Gonçalves

Junho
2025

Conteúdo

1	Resumo	1
2	Apresentação	1
3	Descrição de atividades	1
3.1	Preparação das Imagens	2
3.2	Geração do Dataset Multiescala	3
3.3	Amostragem e Organização	3
3.4	Desafios Enfrentados	4
3.5	Relevância da Atividade	4
4	Análise dos Resultados	4
5	Trabalhos Futuros	6
	Bibliografia	7

1 Resumo

Esta Atividade Orientada ao Ensino apresenta o processo de criação de um dataset de imagens aéreas para o reconhecimento de áreas de erosão em plantações de cana-de-açúcar, com foco na segmentação semântica. As imagens foram capturadas por Veículos Aéreos Não Tripulados (VANTs), ortorretificadas e anotadas manualmente. O conjunto de dados foi estruturado em patches multiescala e organizado em conjuntos de treinamento, validação e teste, destacando desafios como o desbalanceamento de classes e o contraste visual, fundamentais para aplicações de inteligência artificial no monitoramento ambiental.

2 Apresentação

A erosão do solo é um fenômeno que compromete a sustentabilidade agrícola, causando perda de nutrientes, prejuízo ao desenvolvimento das plantas e impactos econômicos e ambientais, especialmente em culturas extensivas como a cana-de-açúcar. O monitoramento eficiente dessas áreas, por meio de imagens capturadas por Veículos Aéreos Não Tripulados (VANTs), aliado a técnicas de inteligência artificial, tem se mostrado uma solução promissora para a detecção automática de regiões afetadas.

Esta Atividade Orientada ao Ensino propõe a apresentação do processo de criação de datasets de imagens aéreas para o reconhecimento de áreas de erosão, com base no projeto desenvolvido no Trabalho de Conclusão de Curso (TCC) da autora, intitulado "Sistema de detecção de erosões em plantações de cana-de-açúcar utilizando métodos de Inteligência Artificial". A atividade descreve as etapas seguidas na preparação dos dados, treinamento de modelos e avaliação dos resultados, oferecendo um exemplo prático de aplicação de redes neurais profundas em problemas ambientais.

O foco está na geração e organização de dados capturados por VANTs (Veículos Aéreos Não Tripulados), incluindo a ortorretificação das imagens, a anotação manual das áreas de erosão e o recorte das ortofotos em patches de diferentes resoluções (256×256 , 512×512 e 1024×1024 pixels). O processo abrangeu também a divisão em conjuntos de treinamento, validação e teste, assegurando diversidade e representado por atividade das amostras.

3 Descrição de atividades

A atividade consiste na exposição do processo de criação de um dataset de imagens aéreas para reconhecimento de erosão em áreas agrícolas, destacando

o uso de VANTs na captura das ortofotos, a ortorretificação, a anotação manual das áreas erodidas e o recorte das imagens em patches. Também aborda a divisão dos dados em conjuntos de treinamento, validação e teste, e a aplicação de modelos de aprendizado profundo para segmentação semântica

3.1 Preparação das Imagens

O processo iniciou-se com a captura de imagens aéreas por drones equipados com câmeras de alta resolução sobre áreas de cultivo de cana-de-açúcar em diferentes regiões do Brasil. As imagens foram submetidas ao processo de ortorretificação, garantindo a correção de distorções geométricas e a uniformidade de escala. O resultado consistiu em 13 ortofotos com GSD (Ground Sample Distance) aproximado de 3 cm.

As ortofotos foram analisadas e anotadas manualmente por especialista, com a criação de máscaras poligonais para identificar e separar áreas de erosão (laminar, sulcos, ravinas, etc) das áreas intactas.

Um exemplo é apresentado na Figura 1, que exhibe a ortofoto completa e um destaque ampliado da região afetada pela erosão, anotada em vermelho pelo especialista.



Figura 1: Ortofoto completa da área analisada, destacando-se em detalhe (zoom) a região com erosão identificada e anotada em vermelho por especialista.

3.2 Geração do Dataset Multiescala

As ortofotos anotadas foram divididas em patches de três resoluções distintas — 256×256 , 512×512 e 1024×1024 pixels — para permitir a análise dos impactos da escala na segmentação. O número total de patches gerados por resolução foi distribuído conforme a Tabela 2.

Resolução	Total de Patches
256x256	39.051
512x512	13.058
1024x1024	2.848

Tabela 1: Distribuição total dos patches gerados para cada resolução.

Cada patch foi gerado juntamente com sua máscara de anotação correspondente.

3.3 Amostragem e Organização

Os patches de cada resolução foram embaralhados aleatoriamente para evitar viés na divisão dos conjuntos. Posteriormente, foram organizados da seguinte forma:

- **70%** para treinamento
- **10%** para validação
- **20%** para teste

A organização dos patches foi feita em diretórios específicos para cada resolução e conjunto, contendo as imagens RGB e suas respectivas máscaras. Os patches gerados para cada tamanho de imagem foram distribuídos conforme apresentado na Tabela 2.

Resolução	Treinamento	Validação	Teste	Total
256x256	27.759	3.764	7.528	39.051
512x512	8.144	1.700	3.214	13.058
1024x1024	1.774	376	698	2.848

Tabela 2: Distribuição dos patches nos conjuntos de treinamento, validação e teste para cada resolução de dataset.

3.4 Desafios Enfrentados

A criação do dataset apresentou desafios importantes, entre os quais se destaca o desbalanceamento de classes. A erosão apareceu em poucas regiões das ortofotos, resultando em um percentual muito baixo de pixels de erosão nos patches. Em média:

- **256×256 pixels:** 1,01% a 1,32% de erosão
- **512×512 pixels:** 1,09% a 1,25% de erosão
- **1024×1024 pixels:** 0,92% a 1,02% de erosão

Esse comportamento ocorre porque imagens maiores abrangem áreas mais amplas e diversas, incluindo tanto regiões erodidas quanto intactas, o que amplia a variabilidade e reduz a concentração relativa de erosão. Essa baixa proporção representa um desafio adicional para os modelos de aprendizado profundo, já que dificulta a correta identificação da classe minoritária.

Outros desafios observados incluem o baixo contraste visual entre solo intacto e áreas erodidas, principalmente nos casos de erosão laminar, e a cobertura vegetal esparsa associada à variação nas condições das imagens, como diferenças de ângulo e resolução, fatores que dificultam a consistência das anotações e comprometem a precisão da segmentação.

3.5 Relevância da Atividade

A criação do dataset representa uma aplicação prática dos desafios na preparação de dados georreferenciados para inteligência artificial, servindo como base para o treinamento de modelos de segmentação e para estudos voltados ao monitoramento ambiental com IA. A atividade reforça a importância de um processo criterioso na amostragem e organização dos dados.

4 Análise dos Resultados

A criação do dataset multiescala possibilitou análises quantitativas e qualitativas importantes sobre o impacto da resolução dos patches no desempenho de modelos de segmentação semântica. O processo de geração dos dados permitiu construir uma base robusta para a realização dos experimentos, oferecendo diferentes níveis de detalhamento e contexto espacial.

Os patches foram utilizados para treinar e avaliar os modelos SegFormer, FCN e DeepLabV3+. A Tabela 3 apresenta as médias das métricas F1-Score

Modelo	SegFormer		FCN		DeepLabV3+	
Resolução Patches	F1	IoU	F1	IoU	F1	IoU
256	0.9713	0.9646	0.9786	0.9713	0.9773	0.9698
512	0.9780	0.9681	0.9728	0.9596	0.9693	0.9562
1024	0.9489	0.9254	0.9334	0.9089	0.8630	0.8411

Tabela 3: Média das métricas F1 e IoU para os modelos SegFormer, FCN e DeepLabV3+, considerando todas as classes nos patches, nos tamanhos de dataset 256, 512 e 1024.

e IoU, considerando todas as classes nos patches, para os três modelos em cada resolução avaliada:

Os resultados quantitativos reforçam a relação entre a resolução dos patches e o comportamento dos modelos. Em resoluções menores (256×256 pixels), o FCN apresentou o melhor F1-Score geral (0.9786), beneficiado pelo maior nível de detalhamento local, o que facilitou a identificação de pequenos padrões de erosão em áreas pontuais. Na resolução intermediária (512×512 pixels), o SegFormer alcançou o maior F1-Score (0.9780), resultado de sua capacidade de equilibrar detalhamento e contexto global, integrando informações espaciais mais amplas sem perder precisão em bordas e pequenos detalhes. Em resoluções maiores (1024×1024 pixels), o SegFormer manteve o melhor desempenho (F1 de 0.9489), enquanto o DeepLabV3+ apresentou queda significativa (F1 de 0.8630), reflexo das limitações das arquiteturas convolucionais em lidar com o contexto global em imagens mais extensas.

A análise qualitativa reforçou esses padrões e demonstrou como o dataset preparado permitiu observar erros típicos, como falsos positivos (ampliação indevida de áreas detectadas como erosão) e falsos negativos (omissão de pequenas regiões erodidas). Além disso, a menor concentração relativa de pixels de erosão em imagens maiores (cerca de 0,92% a 1,02%) tornou o desbalanceamento de classes ainda mais desafiador e dificultou a segmentação precisa em cenários reais, exigindo maior capacidade dos modelos para distinguir corretamente as áreas de interesse em meio ao predomínio do fundo.

Esses resultados ilustram a relevância do processo de criação do dataset na construção de bases consistentes para o treinamento e a avaliação de modelos. Além de permitir estudos comparativos entre diferentes arquiteturas, o dataset evidenciou como escolhas relacionadas ao tamanho dos patches e ao processo de amostragem impactam diretamente na dificuldade da tarefa e nos tipos de erros observados na segmentação.

5 Trabalhos Futuros

Como continuidade do trabalho apresentado, recomenda-se a ampliação do dataset com a inclusão de novas regiões geográficas e diferentes condições ambientais e climáticas. Essa diversificação pode ampliar a solidez do conjunto de dados e permitir o treinamento de modelos mais generalizáveis, adequados para aplicações em diferentes contextos agrícolas e ambientais.

Outra direção para trabalhos futuros envolve a adoção de estratégias de balanceamento de classes, como a superamostragem da classe minoritária, a geração de dados sintéticos ou o uso de funções de perda ponderadas no treinamento, a fim de reduzir o impacto do desbalanceamento e aprimorar a detecção das áreas de erosão, sobretudo daquelas com baixa representatividade. Além disso, recomenda-se o uso de técnicas de pós-processamento para o refinamento das bordas das regiões segmentadas e a redução de erros, como falsos positivos e negativos, especialmente em patches de maior resolução onde há maior variabilidade nas anotações.

Por fim, os dados criados neste trabalho podem servir como base para o desenvolvimento de ferramentas integradas para monitoramento ambiental, combinando diferentes fontes de dados geoespaciais e ampliando a aplicabilidade de soluções baseadas em inteligência artificial no manejo sustentável do solo.

Bibliografia

S. A. Bezerra, J. R. B. Cantalice, Interrill erosion under different conditions of soil cover of sugarcane, *Revista Brasileira de Ciencia do Solo* 30 (3) (2006) 565–573. doi:10.1590/S0100-06832006000300016. URL <https://www.rbcjournal.org/article/interrill-erosion-under-different-cond>

V. Carabassa, P. Montero, J. M. Alcañiz, J.-C. Padró, Soil erosion monitoring in quarry restoration using drones, *Minerals* 11 (9) (2021). doi:10.3390/min11090949.