

Aprendizado de Representações Via Auto-Supervisão em Imagens de Sensoriamento Remoto de Áreas Rurais

Lucas V. dos Santos e Hugo N. Oliveira

Departamento de Informática, Universidade Federal de Viçosa (UFV), Viçosa-MG, Brasil

ODS 2: Fome Zero e Agricultura Sustentável

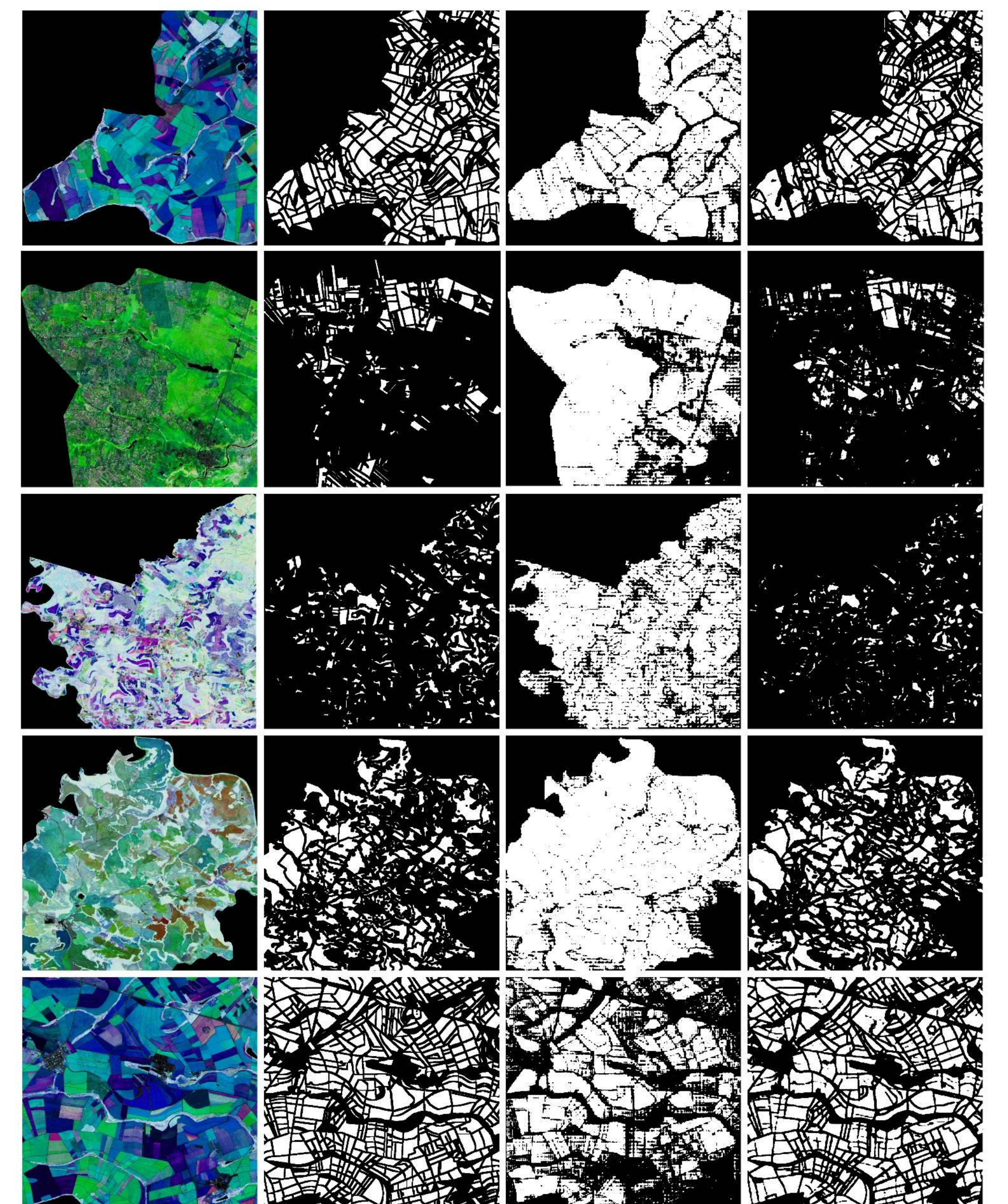
Pesquisa de Iniciação Científica

Introdução

A segmentação semântica é uma técnica central em visão computacional, pois permite identificar e delimitar objetos em imagens, sendo aplicada em áreas como saúde, análise ambiental e, em especial, na agricultura de precisão. No entanto, embora eficazes, métodos tradicionais dependem de grandes volumes de dados anotados manualmente, um processo caro e que exige especialistas. Nesse cenário, os *foundation models* surgem como alternativa promissora por aliar treinamento massivo a uma maior capacidade de generalização. O *Segment Anything Model* (SAM) [1] é um marco nesse sentido, ao incorporar a noção de *prompts*, informações simples fornecidas pelo usuário, que orientam a segmentação e viabiliza cenários *zero-shot* e *few-shot*, reduzindo a necessidade de bases extensas. Apesar do êxito em imagens naturais, aplicar o SAM ao sensoriamento remoto envolve desafios relacionados às diferenças espectrais, às escalas e à complexidade das fronteiras agrícolas. Diante disso, o objetivo deste trabalho é verificar a aplicabilidade do SAM ao domínio agrícola, explorando sua capacidade para a segmentação de talhões em imagens de cidades brasileiras e propondo melhorias que permitam apoiar a anotação automática de dados nesse contexto. Ao avançar nessa direção, buscamos fortalecer aplicações de agricultura de precisão e corroborar iniciativas estratégicas como o Censo Agro do IBGE.

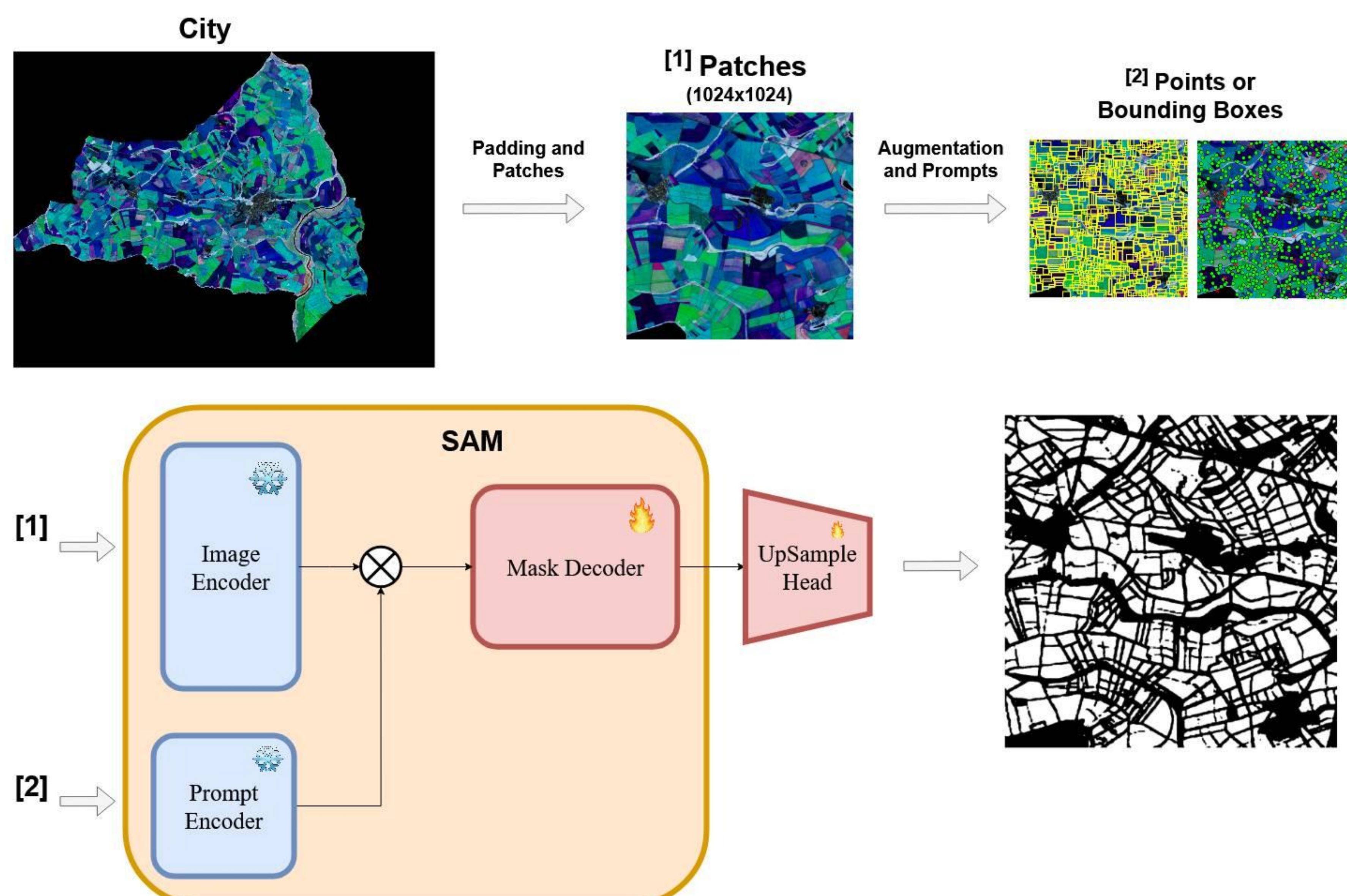
Experimentos e Resultados

Os experimentos feitos mostraram que o SAM em regime *zero-shot* com prompts apresentou segmentações instáveis e pouco adaptadas ao domínio agrícola. Após o fine-tuning, percebe-se ganhos expressivos, com melhora consistente em métricas a nível de pixel e em nível de instâncias, além de segmentações visivelmente bem mais precisas, sobretudo na delimitação de talhões menores onde há áreas complexas. A figura ao lado ilustra esses avanços qualitativos entre os dois regimes.



Metodologia

Fluxo metodológico adotado: a imagem de satélite é dividida em patches e, junto aos *prompts*, é processada pelos encoders do SAM. As representações são combinadas no *mask decoder* e refinadas por uma *upsample head*, resultando na máscara segmentada.



Conclusões

Os resultados mostraram que, embora o modelo apresente limitações em regime *zero-shot*, sua adaptação com fine-tuning seletivo, incluindo uma *upsample head* aprendida e o uso de *prompts*, trouxe ganhos significativos. Assim, o SAM mostra potencial para apoiar a anotação de dados no domínio agrícola, contribuindo para aplicações em agricultura de precisão e iniciativas como o Censo Agro.

Referências

[1] Kirillov, Alexander, et al. "Segment anything." Proceedings of the IEEE/CVF international conference on computer vision. 2023.

Apoio Financeiro

