

XIII Colóquio Brasileiro
de Ciências
Geodésicas • 2024

Universidade Federal do Paraná

25 Anos

*Conectando mentes e
provendo conhecimento*

ADAPTAÇÃO DE DOMÍNIO BASEADA EM APRENDIZADO PROFUNDO COM FOCO NO DELINEAMENTO GEOMÉTRICO DA REDE VIÁRIA

Gustavo R. Collegio¹, Antonio G. G. Filho², Aluir P. Dal Poz³

^{1,3}UNESP, Faculdade de Ciências e Tecnologia, campus de Presidente Prudente – SP

²DSG, Serviço Geográfico do Exército Brasileiro, Brasília – DF



Processo No. 2024/02040-4

Curitiba, 26 a 29 de novembro de 2024

Introdução

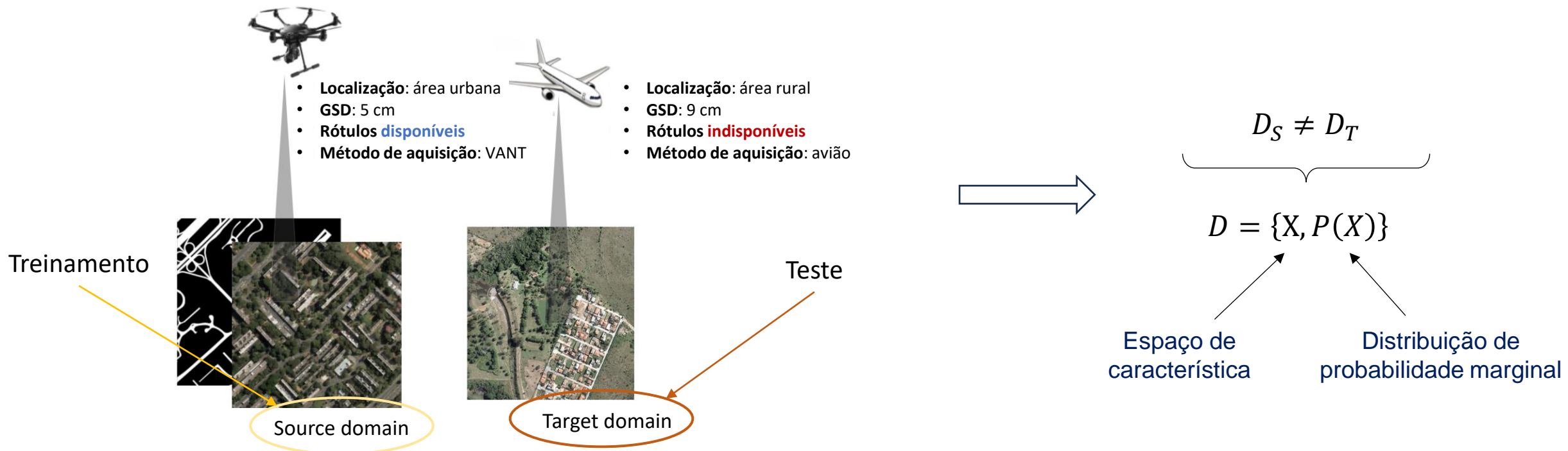
- **A extração automática de rodovias tem sido um tópico de pesquisa ativo e relevante nas últimas décadas;**
- **As principais motivações são:**
 - Mapeamento atualizado da malha viária;
 - Diversas aplicações: economia, infraestrutura, planejamento urbano, etc.;
- **No entanto, uma malha viária bem desenvolvida não é garantia de consistência e pleno uso.**



Fonte: [1]

Introdução

- Métodos baseados em *Deep Learning (DL)*, em especial as *Convolutional Neural Networks (CNNs)*, abriram uma nova era para a extração de rodovias. Entretanto, alguns desafios ainda existem, os quais se referem, principalmente, ao **dinamismo** e a **complexidade** das vias.



Introdução

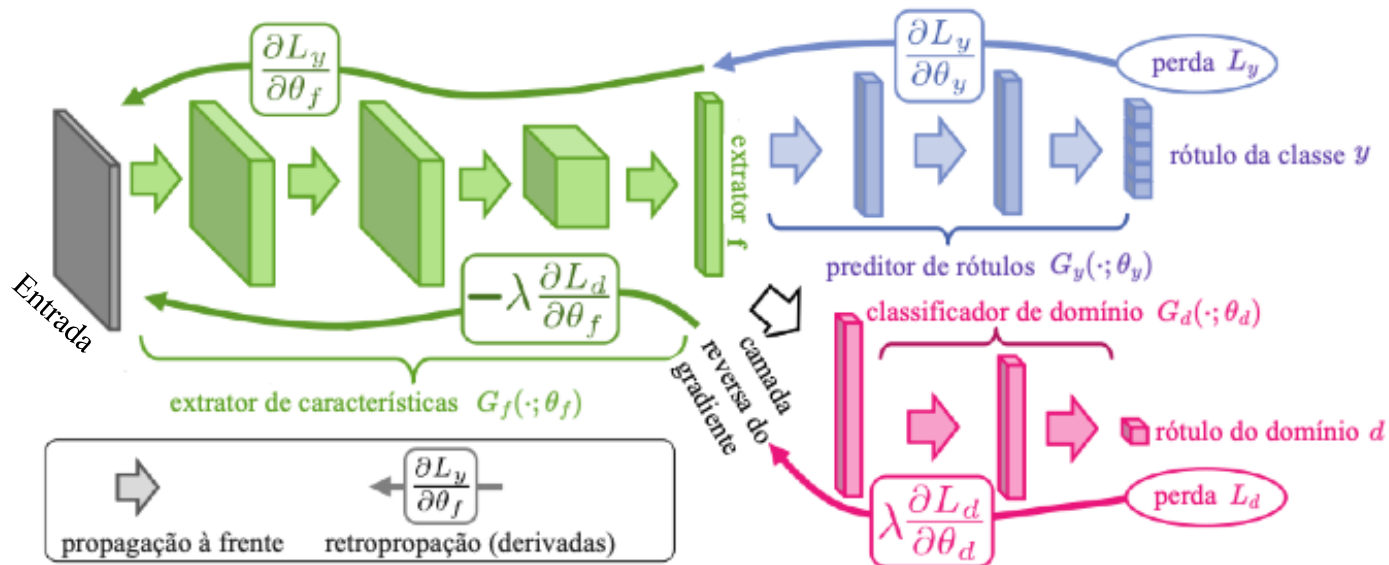
- Formalmente, modelos baseados em DL são treinados para **aprender a distribuição de probabilidade** de um conjunto de dados [2], mas como lidar quando $D_S \neq D_T$?
- No contexto de DL:
 - Transferência de aprendizado + ajuste fino; no entanto, **não é frequente a disponibilidade de dados rotulados** no domínio alvo.
- Uma vez que a anotação no conjunto teste deixa de ser um requisito, **Unsupervised Domain Adaptation (UDA)** torna-se uma solução atrativa.
 - Definição: transferir o aprendizado de um domínio fonte rotulado para um domínio alvo não rotulado na presença de variação entre domínios.

Introdução

- **Todavia, apesar das contribuições por meio de UDA, problemas relacionados ao delineamento geométrico da rede viária ainda permanecem (por ex.: continuidade);**
- **No estado da arte [3], o mecanismo denominado *attention* é alternativa para explorar diferentes informações contextuais das vias, mas não são usualmente integrados em UDA;**
- **Sob este aspecto, este trabalho propõe, portanto, incorporar *attention* em UDA;**
- **Especificamente, adota-se uma CNN que compõe o estado da arte: U-Net.**
 - Ao passo que para UDA tal abordagem é recente, isto torna-se a principal motivação de investigação deste trabalho.

Domain Adversarial Neural Network

- A **Domain Adversarial Neural Network (DANN)** [4] objetiva minimizar a divergência entre duas distribuições de probabilidade, parametrizada pelo **encoder** de uma rede neural profunda.



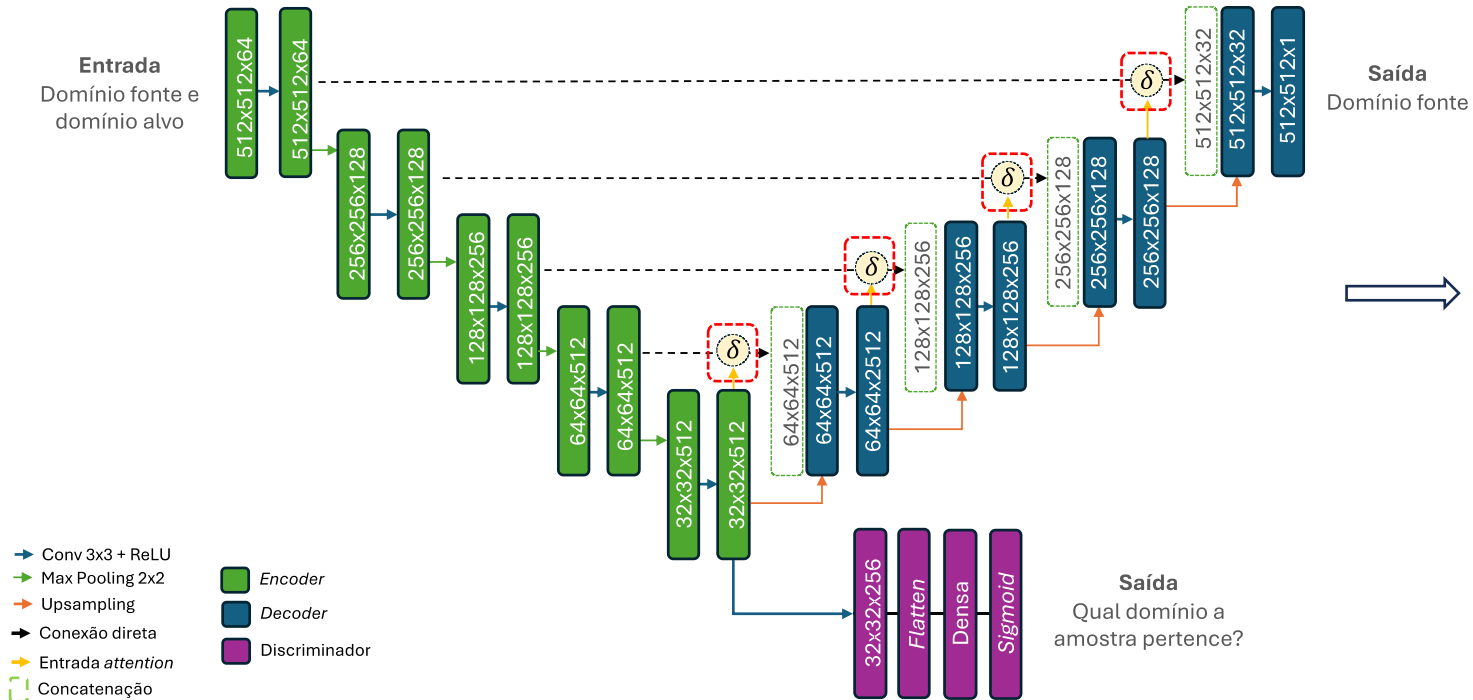
Conjunto de dados

- **O método proposta utiliza o conjunto de dados disponibilizado por [5];**
 - Imagens RGB com resolução espacial de 1,2 m.
- **Características:**
 - Diferentes pavimentos de rodovias;
 - Regiões com alta densidade residencial;
 - Alguns casos de oclusão;
 - Dispõe de comparação entre diversos trabalhos no estado da arte.
- **UDA: adaptação para diferentes tipos de dados;**
 - Imagens em tons de cinza para RGB.

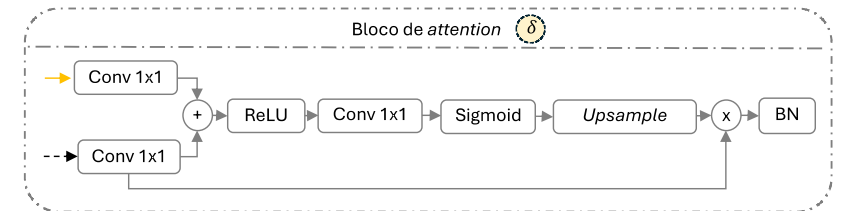


Modelo baseado em UDA

- Adaptação da arquitetura DANN para segmentação semântica: U-Net [6] com *attention* [7].



Attention: filtrar as características extraídas por meio das conexões diretas.



[6] Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597 [cs]*, 2015.

[7] Oktay, O. et al. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv:1804.03999*, 2018.

Otimização do modelo

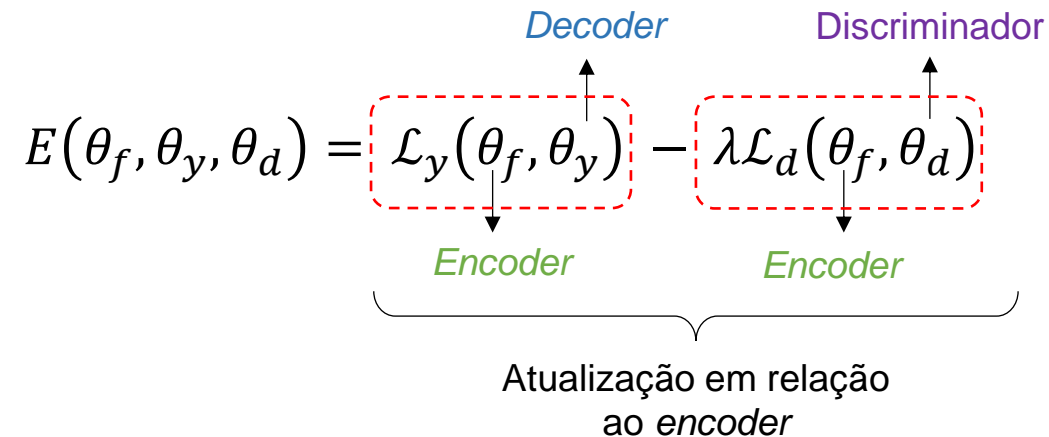
- **Função perda total para cálculo dos parâmetros ótimos:**

$$E(\theta_f, \theta_y, \theta_d) = \mathcal{L}_y(\theta_f, \theta_y) - \lambda \mathcal{L}_d(\theta_f, \theta_d)$$

Diagram illustrating the total loss function $E(\theta_f, \theta_y, \theta_d)$ for model optimization. The equation is shown as $E(\theta_f, \theta_y, \theta_d) = \mathcal{L}_y(\theta_f, \theta_y) - \lambda \mathcal{L}_d(\theta_f, \theta_d)$. The term $\mathcal{L}_y(\theta_f, \theta_y)$ is identified as the *Função perda decoder* (Decoder loss function), and the term $\lambda \mathcal{L}_d(\theta_f, \theta_d)$ is identified as the *Função perda discriminador* (Discriminator loss function).

Otimização do modelo

- **Função perda total para cálculo dos parâmetros ótimos:**

$$E(\theta_f, \theta_y, \theta_d) = \mathcal{L}_y(\theta_f, \theta_y) - \lambda \mathcal{L}_d(\theta_f, \theta_d)$$


Atualização em relação
ao *encoder*

- **Atualização em duas etapas:**
 - i) Gradientes do *encoder* e *decoder*;
 - ii) Gradientes do discriminador.

Protocolo experimental

- **Recursos computacionais:** GPU NVIDIA RTX 3080 10GB DDR6X;
- **Biblioteca:** TensorFlow (versão 2.15);
- **Iterações:** cerca de 25.000 (200 épocas e *batch* igual a 2);
- **Dados de entrada:** recortes de 512x512 px (sem sobreposição) com aumento artificial;
- **Funções perda:**
 - **Decoder:** Focal Tversky Loss [8];
 - **Discriminador:** Entropia Cruzada Binária;
 - Rótulos: 0 e 1.
- **Para verificar a contribuição de UDA e do uso de attention, o método é comparado com:**
 - **Sem UDA:** treinamento no domínio fonte e teste direto no domínio alvo;
 - **Treinamento supervisionado:** treinamento e teste no domínio alvo.
 - **Métricas:** F1-Score e *Intersection-over-Union* (IoU).

Resultados

Modelo	<u>Sem attention</u>		Com attention	
	F1-Score (%)	IoU (%)	F1-Score (%)	IoU (%)
Supervisionado (domínio alvo)*	<u>89,56</u>	<u>81,86</u>	89,97	81,97
Sem UDA	80,02	67,36	86,25	76,39
UDA	<u>85,04</u>	<u>74,31</u>	88,66	79,87

*Treinamento supervisionado sem *attention*: U-Net modificada por [5].

- **UDA é superior em relação a referência (sem UDA);**
- **UDA contribui em aproximadamente 5% de F1-Score e 9% de IoU.**
- **No entanto, é inferior quanto ao treinamento supervisionado, conforme esperado.**

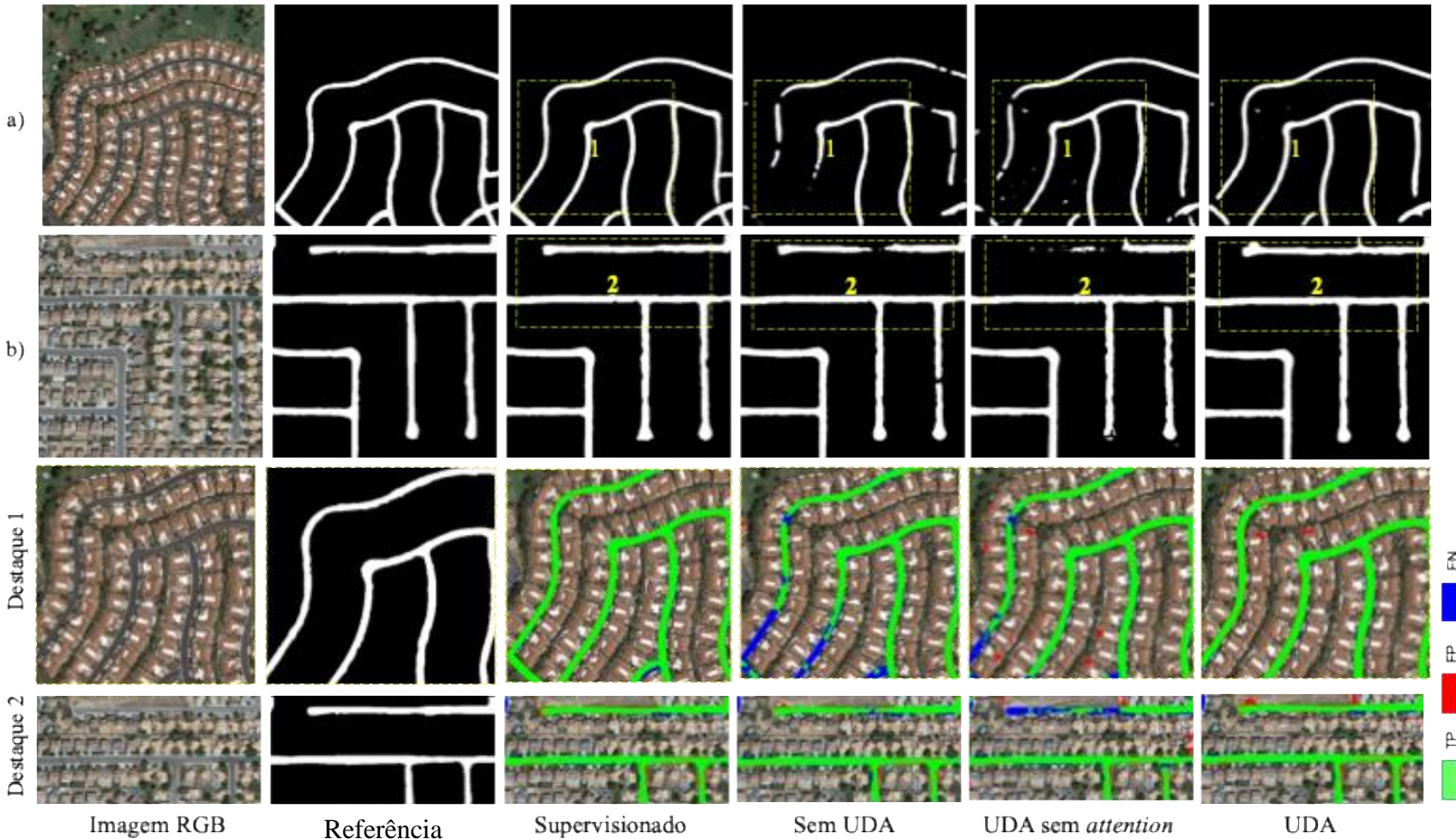
Resultados

Modelo	Sem <i>attention</i>		Com <i>attention</i>	
	F1-Score (%)	IoU (%)	F1-Score (%)	IoU (%)
Supervisionado (domínio alvo)*	89,56	81,86	<u>89,97</u>	<u>81,97</u>
Sem UDA	80,02	67,36	86,25	76,39
UDA	85,04	74,31	<u>88,66</u>	<u>79,87</u>

*Treinamento supervisionado sem *attention*: U-Net modificada por [5].

- **Melhor performance é obtida via *attention*;**
- **Em relação ao treinamento sem UDA, a contribuição é de, aproximadamente, 3% e 2% para IoU e F1-Score;**
- **Especificamente:**
 - *Attention* contribui em aproximadamente 5% de IoU e 3% de F1-Score em UDA;
 - O método proposto atinge maior performance que [5] no treinamento supervisionado.

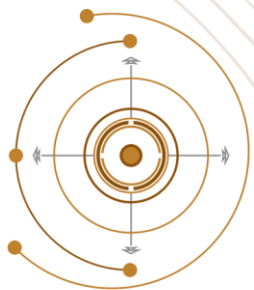
Resultados



- **UDA** incorpora informação complementar do domínio alvo.
- **Attention**: obtém informação contextual em casos de possível oclusão → melhor discriminação entre classes na cena.
- A abordagem proposta permite mitigar o problema de **descontinuidade da rede viária**.

Conclusão

- **Este trabalho objetivou investigar a implementação de *attention* em UDA para extração da rede viária no caso de imagens em tons de cinza para RGB;**
- **UDA contribui para a segmentação da rede viária, assim como *attention*;**
 - Filtragem das conexões direta da U-Net permite melhor delineamento geométrico da rede viária;
- **Outros estudos são sugeridos a partir da abordagem proposta, como o uso de transformadores (*transformers*) e complementaridade de dados além do espectro do visível.**



XIII Colóquio Brasileiro de Ciências Geodésicas • 2024

Universidade Federal do Paraná

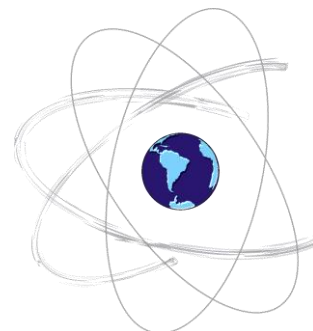
25 Anos

*Conectando mentes e
provendo conhecimento*

AGRADECIMENTOS



Processos No. 2024/02040-4 e
2021/03586-2



PPG em Ciências Cartográficas
(FCT/UNESP)

REALIZAÇÃO



Curitiba, 26 a 29 de novembro de 2024