

**XII** Colóquio Brasileiro  
de Ciências  
Geodésicas

**V** Simpósio  
Brasileiro de  
Geomática



# FUSÃO MULTIESPECTRAL EM DEEP LEARNING: MELHOR DELINEAMENTO E IDENTIFICAÇÃO DE REDE VIÁRIA EM AMBIENTE URBANO

*Antonio Filho<sup>1</sup>, Milton Shimabukuro<sup>1</sup>, Aluir Dal Poz<sup>1</sup>*

<sup>1</sup> Universidade Estadual Paulista (Unesp), Faculdade de Ciência e Tecnologia, Câmpus Presidente Prudente - SP

Curitiba, 8 a 11 de novembro de 2022

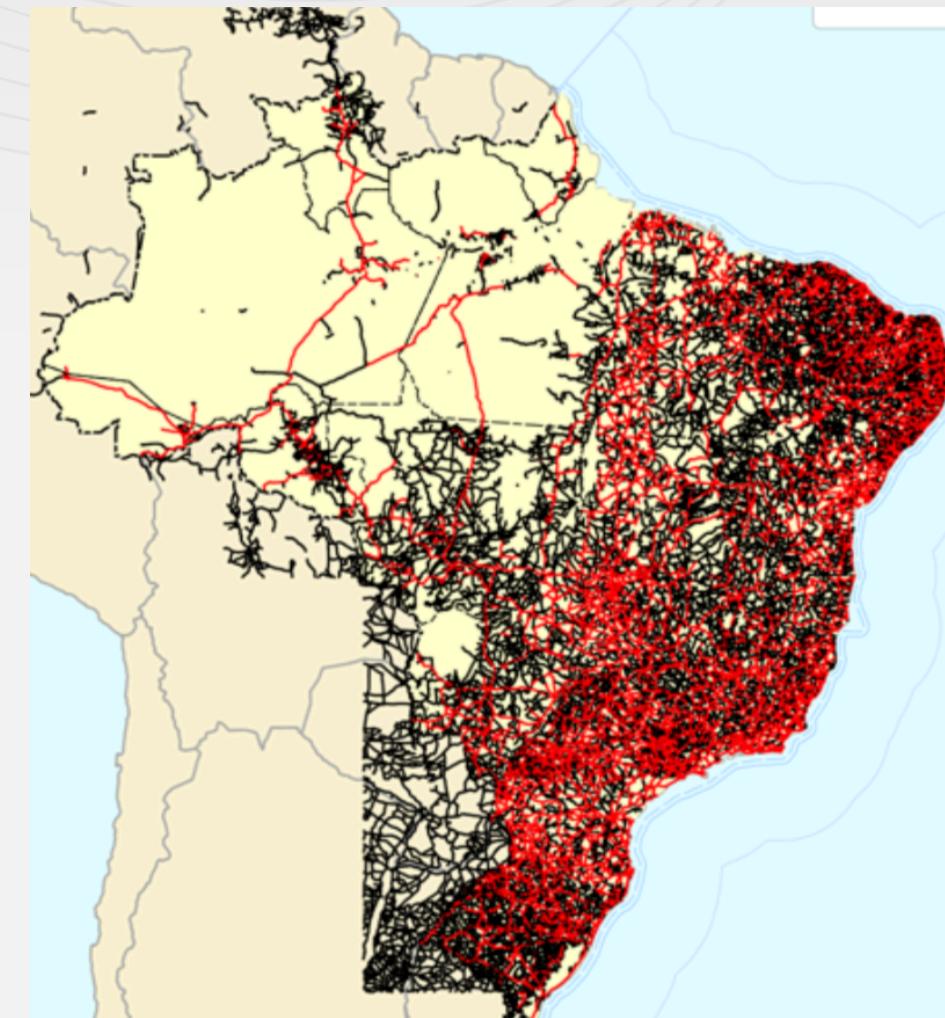




# Introdução

- O Brasil é um país que possui 1,7 milhões de quilômetros de **rodovias distribuídas de forma irregular** pelo território nacional [1]
- **Diferentes ambientes/contextos** influenciam as características das vias.
- Alto dinamismo
- **Informações sobre rodovias são relevantes** para planejamento e gestão urbana, bem como para a economia.
- Tal dinamismo e complexidade tornam a obtenção de informações da rede viária, de forma ágil, um grande desafio.

Figura 1 - Rodovias no Brasil



Fonte: INDE, 2022



# Introdução

- Nesse contexto, a **extração de vias** a partir de imagens de Sensoriamento Remoto é **complexa** - tipo de sensor, resolução espectral e espacial, clima, variação de luz e característica do solo.
- **A partir de 2014**, o interesse sobre **aprendizado profundo** (*Deep Learning* - DL) aumentou na comunidade de Sensoriamento Remoto pela capacidade de **extrair características a partir dos dados** sem intervenção humana para construir regras [2].
- Desde então, **algoritmos de DL** têm obtido **notáveis resultados na extração de rede viária** a partir de imagens multiespectrais (espectro do visível) [3].



# Introdução

- Paralelamente, **avanços tecnológicos nas décadas de 80/90** impulsionaram pesquisas sobre técnicas de **fusão de dados multisensores** [4-6].
- Os Dados de observação da Terra não estão limitados aos sensores ópticos que atuam na faixa do visível [7].
- Portanto, este estudo pretende **analisar a influência de dados do Infravermelho (IR)** e da **escolha da função perda (Loss)** na **extração de rede viária** na área urbana por meio de DL.

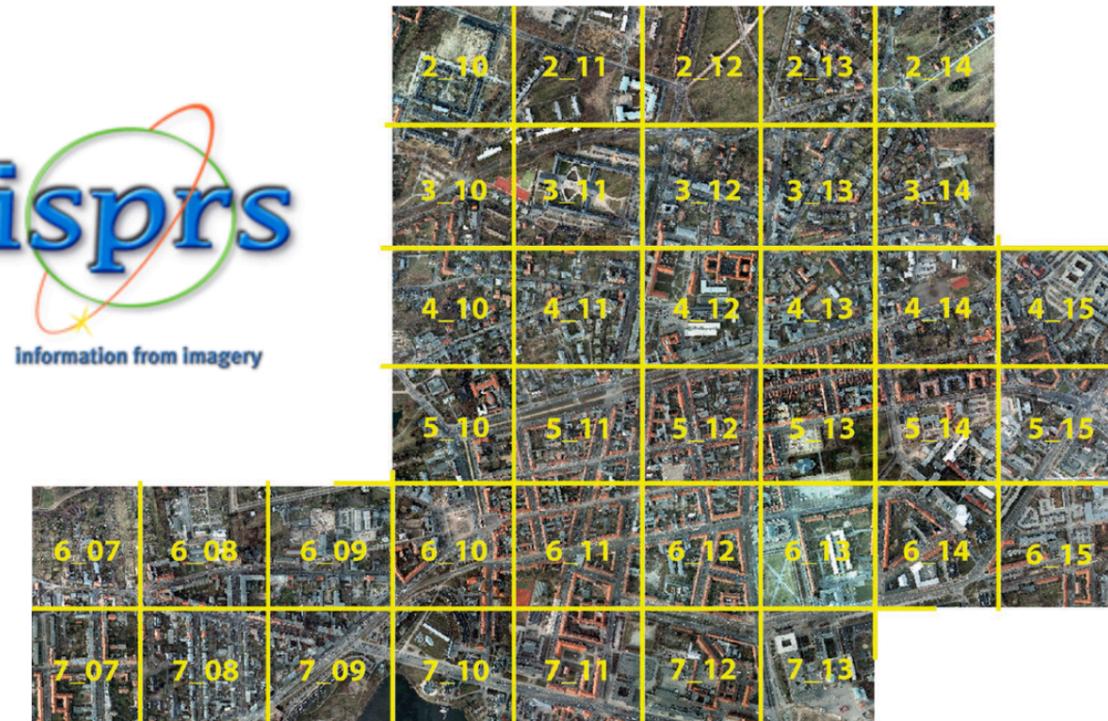


# Metodologia

- Para isso, o conjunto de dados referência **ISPRS Potsdam** [8] foi utilizado junto com a **arquitetura UNet**;
- uso de *encoders* VGG-16 e ResNet-50,
- inserção de *Batch Normalization* (BN);
- imagens de entrada com tamanho e canais diferentes da original;
- recortes aleatórios (*random crop*), espelhamento horizontal, vertical e rotação de 90°;



Figura 2 - Conjunto de dados referência ISPRS Potsdam



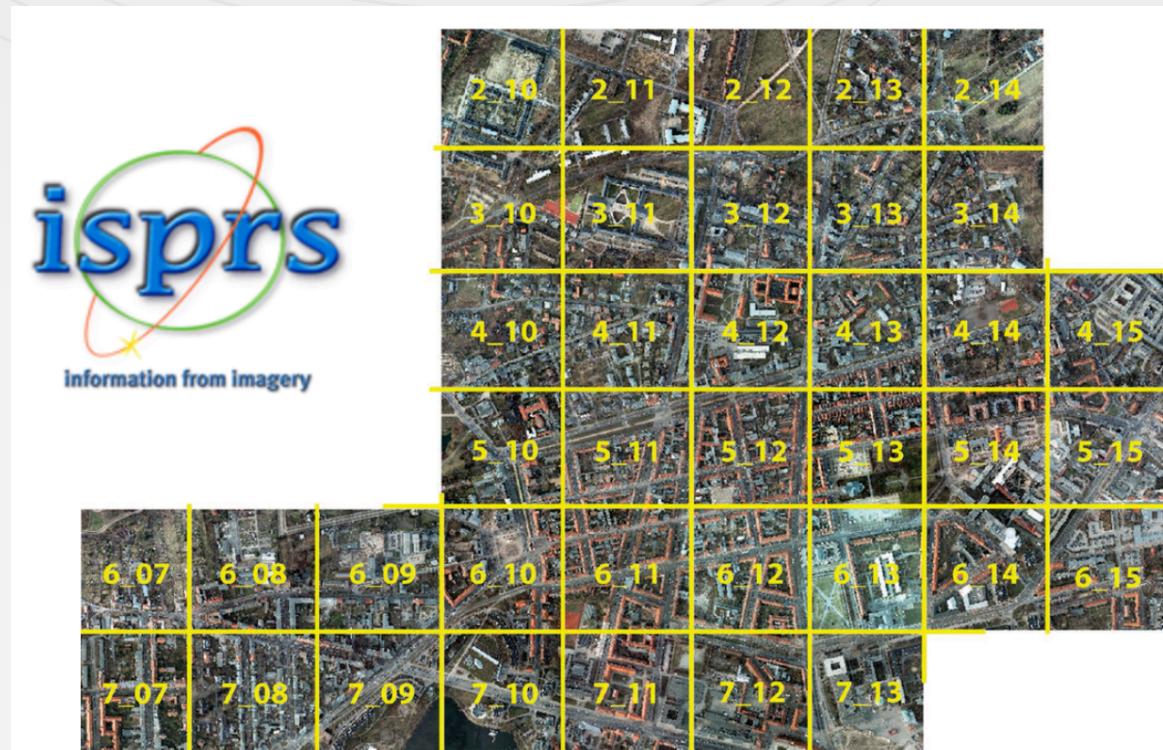
Fonte: [8]



# Metodologia

- utilização de **Binary CrossEntropy**, **Dice Loss**, **Tversky** ou **Focal Tversky Loss** para o desbalanceamento com a classe não rodovia e
- utilização de transferência de aprendizado (**Transfer Learning**) ImageNet .

Figura 2 - Conjunto de dados referência ISPRS Potsdam



Fonte: [8]



# Metodologia

- O conjunto de dados - **38 amostras de ortofotos**,
- divididos em **24 amostras** para conjunto treinamento e **14** para teste;
- *Ground Sample Distance (GSD) de 5 cm*;
- **RGB, RG-IR ou RGB-IR** com resolução radiométrica de 8 bits por banda;
- máscaras de segmentação para **6 classes** - Superfícies Impermeáveis (*Impervious surfaces*).

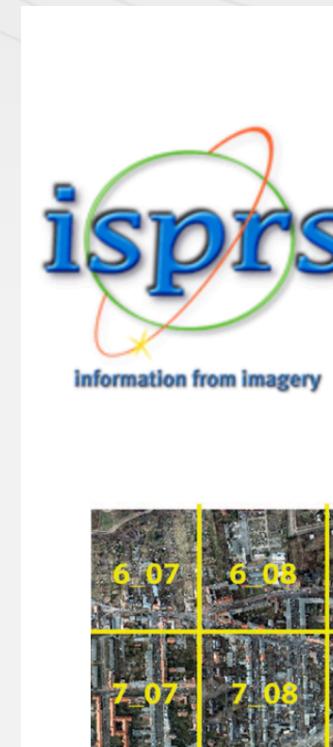


Figura 2 - Conjunto de dados referência ISPRS Potsdam



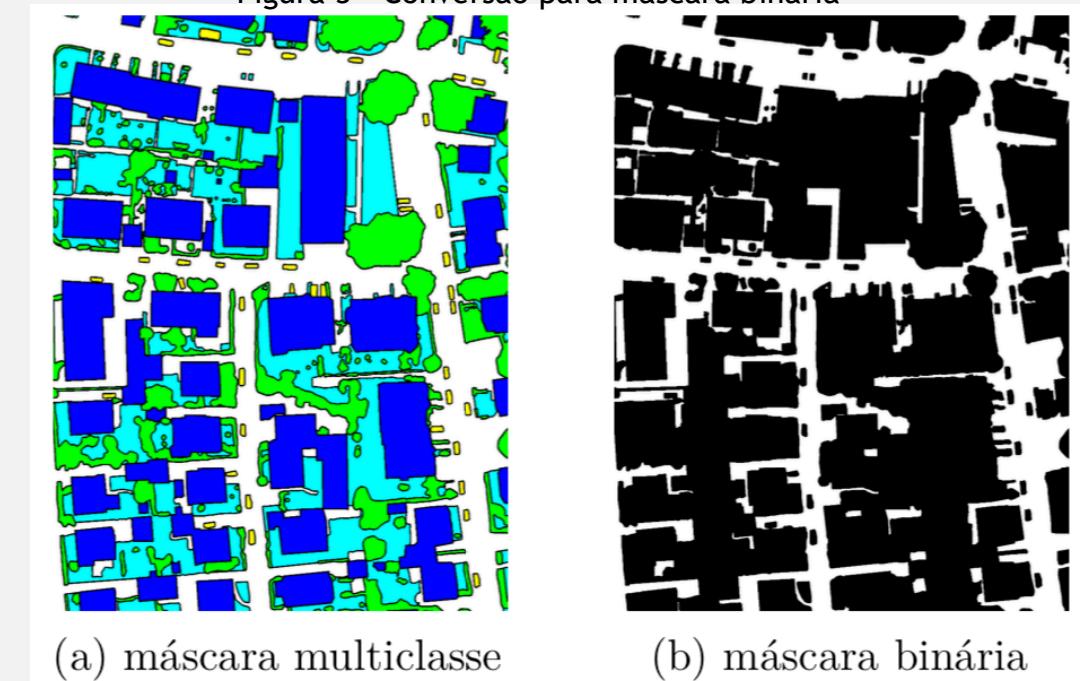
Fonte: [8]



# Metodologia

- matemática de bandas foi realizado para separação da classe - **máscara binária**.
- reamostragem para **10cm** de modo a inserir contexto
- **recortes de tamanho 600x600 pixels com sobreposição vertical e horizontal de 50% e normalização** para valores de pixel entre **[0,1]**.
- ISPRS Potsdam foi utilizado devido:
  1. **disponibilidade de IR e**
  2. **altíssima resolução espacial.**

Figura 3 - Conversão para máscara binária



(a) máscara multiclasse

(b) máscara binária



# Experimentos

- **Analisar a influência da banda do infravermelho (IR) na extração de via.**
- imagens de treino - proporção de **90% para treinamento e 10% para validação**;
- **SGD com momentum de 0,9, taxa de aprendizado inicial de 0,01 e final de 0,0001** (decaimento polinomial taxa de 0.9) , **regularização por Gradiente Centralization (GC)**, **nº subamostras = 4** e aumento artificial dos dados;
- **1701** recortes para treinamento, **108** para validação e **504** para teste.
- **180 épocas (76.545 iterações)**, **tempo médio de 4 horas** de treinamento por modelo em computador equipado com placa GPU NVIDIA RTX 3080 10GB DDR6x, 128 GB RAM, 1TB HD SSD.



# Experimentos

Tabela 1 – Resultado no conjunto teste de ISPRS Potsdam.

Modelos	IoU		Precision		Recall		F1	
	RGB	RGIR	RGB	RGIR	RGB	RGIR	RGB	RGIR
ResUnet_ftv	79,65	<b>79,95</b>	86,79	<b>87,19</b>	<b>90,60</b>	90,54	88,61	<b>88,78</b>
ResUnet_dice	<b>79,85</b>	79,83	88,68	<b>88,94</b>	<b>88,93</b>	88,60	<b>88,73</b>	88,71
ResUnet_tv	<b>79,04</b>	77,95	<b>86,10</b>	83,98	90,58	<b>91,55</b>	<b>88,23</b>	87,53
ResUnet_bce+dice	79,78	<b>81,12</b>	89,40	<b>90,77</b>	88,08	<b>88,39</b>	88,67	<b>89,52</b>
Unet_ftv	77,10	<b>78,47</b>	82,62	<b>84,35</b>	<b>92,10</b>	91,81	87,00	<b>87,88</b>
Unet_dice	79,75	<b>80,03</b>	<b>89,84</b>	89,34	87,67	<b>88,45</b>	88,65	<b>88,83</b>
Unet_tv	78,96	<b>79,24</b>	<b>86,23</b>	85,26	90,34	<b>91,75</b>	88,18	<b>88,33</b>
Unet_bce+dice	78,85	<b>80,01</b>	<b>89,88</b>	88,97	86,61	<b>88,75</b>	88,11	<b>88,81</b>

Fonte: Autor (2022). **Em negrito:** melhores resultados do modelo na métrica por tipo de imagem de entrada. **Em negrito e sublinhado:** melhor resultado na métrica. ftv: *focal tversky* ; tv: *tversky* ; bce+dice: *binary cross entropy* e *dice loss*



# Experimentos

- modelos otimizados com funções de perda que penalizam falsos negativos (tversky e focal tversky) proporcionam os valores mais altos na métrica *recall*;
- **Unet\_ftv** apresenta *recall* 92,10% no conjunto de imagens RGB. Entretanto, houve perda significativa de **precision** (82,62%);
- As redes otimizadas pelas funções de perda **dice** e **bce+dice** apresentaram os **melhores resultados nas métricas**, principalmente F1 e IoU;
- O modelo **ResUnet\_bce+dice** obteve o melhor resultado no conjunto teste de Potsdam com **IoU = 81,12** e **F1 = 89,52**;



# Experimentos

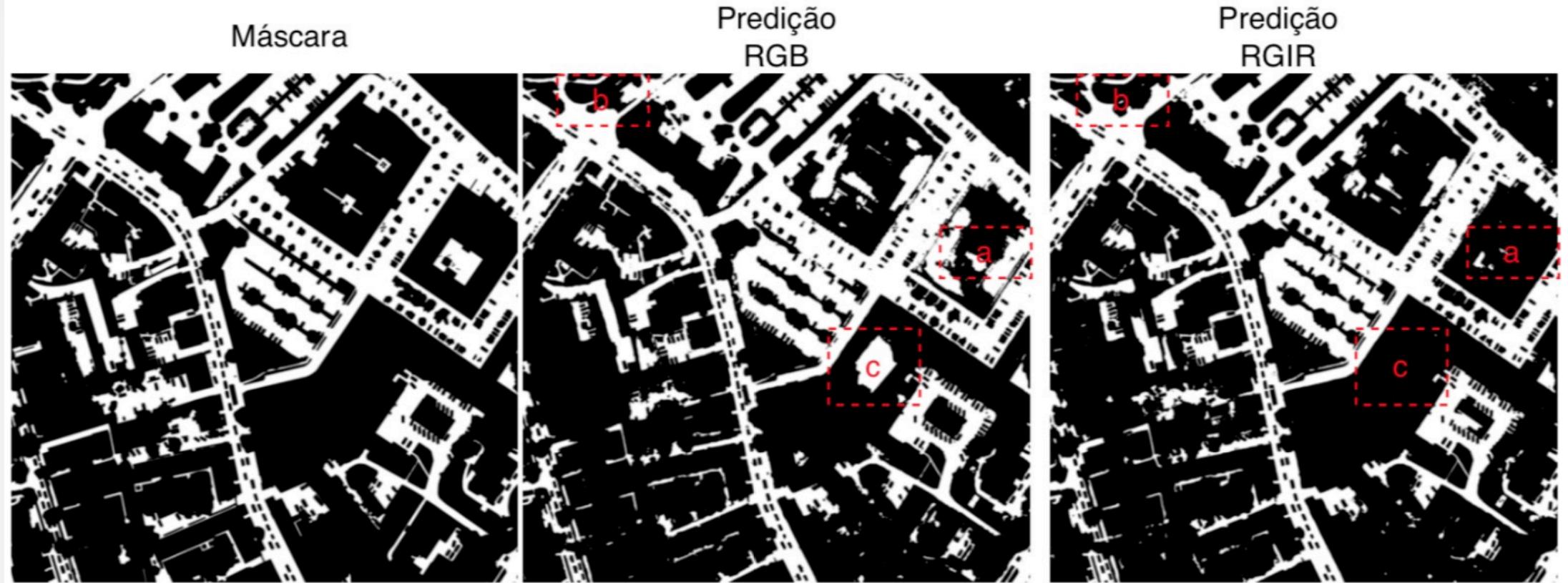
- Em uma mesma arquitetura, a escolha da função de perda para otimização pode acarretar melhores resultados nas predições finais.
- **ResUnet\_bce+dice** possui acréscimo de 4% em IoU e 2,3% em F1, assim como o **Unet\_dice** tem 2% em IoU e 1,1% em F1 em relação aos otimizados pela função **tversky**. Ambos os casos no conjunto com dados IR;
- Outro aspecto está relacionado ao uso da imagem IR no conjunto de dados;
- Praticamente todos os modelos apresentaram desempenho superior utilizando os dados com IR nas métricas F1 e IoU (exceto ResUnet\_dice e ResUnet\_tv);
- **ResUnet\_bce+dice**, o acréscimo foi de 1,8% em IoU e 1% em F1 em relação ao treinado com RGB.
- Com isso, o **canal IR agrega informação na separação de classes** (especificamente na métrica *precision*), principalmente com a vegetação e na diferença de vias e telhados. Essa separação acarreta um **melhor delineamento da via**.



# Experimentos

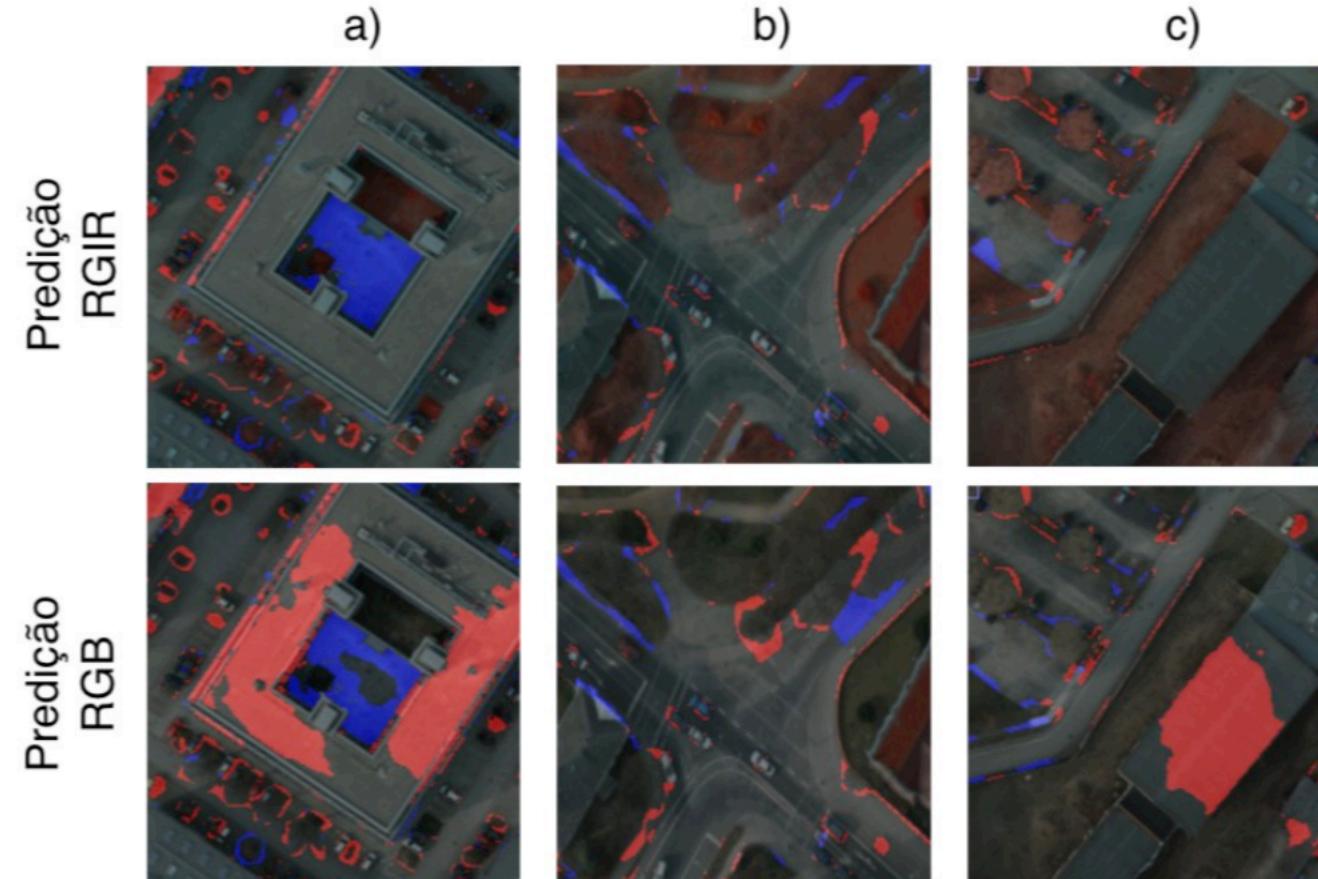
- A predição da classe superfícies impermeáveis com IR possui contornos mais definidos (exemplos 1a), 1b) e 1c)) e também diminui a confusão com texturas diferentes (exemplos 1a) e 1c))

Figura 4 – Influência canal IR na predição.





# Experimentos



Fonte: Autor (2022). Predições realizadas com entrada RGB e RGIR do modelo de maior desempenho na métrica IoU e F1. Recortes (a), (b) e (c) em azul: falsos negativos; em vermelho: falsos positivos sobrepostos na imagem.



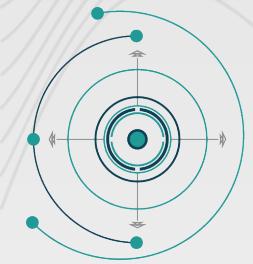
# Conclusão e Trabalhos futuros

- testes realizados mostram que a **escolha da função de perda e o uso do IR influenciam na predição das vias e das outras feições** dentro da classe de superfícies na área urbana.
- Embora os **modelos treinados com funções de perda diferentes** tenham influência no resultado, **não foi encontrada uma relação entre a escolha e a predição final**.
- **Outros trabalhos** na linha de **fusão de dados de sensores/fontes** e **escolha de função Loss** são **relevantes na extração de objetos** bem como **melhor compreensão da superfície imageada**.



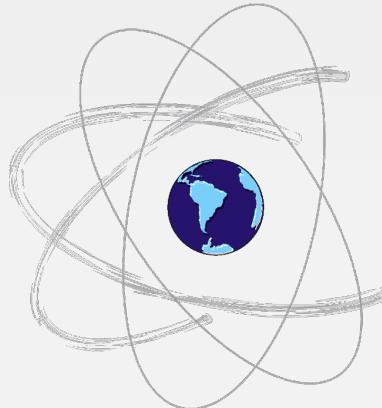
# Referências Bibliográficas

- [1] CNT. Pesquisa CNT de rodovias 2019. Brasília: Sest Senat, 2019. 236 p. Disponível em: [https://pesquisarodovias.cnt.org.br/Downloads/Edicoes//2019/Relatório%20Gerencial/ Relatorio\\_Gerencial\\_Pesquisa\\_CNT\\_Rodovias\\_2019.pdf](https://pesquisarodovias.cnt.org.br/Downloads/Edicoes//2019/Relatório%20Gerencial/ Relatorio_Gerencial_Pesquisa_CNT_Rodovias_2019.pdf). Acesso em: 08 set. 2020.
- [2] MA, L. et al. ISPRS Journal of Photogrammetry and Remote Sensing Deep learning in remote sensing applications : A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 152, n. March, p. 166–177, 2019. ISSN 0924-2716. Disponível em: <<https://doi.org/10.1016/j.isprsjprs.2019.04.015>>.
- [3] ABDOLLAHI, A. et al. Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction: A State-Of-The-Art Review. *Remote Sensing*, 2 maio. 2020. v. 12, n. 9, p. 1444. Disponível em: <<https://www.mdpi.com/2072-4292/12/9/1444>>.
- [4] SCHENK, T.; CSATHÓ, B. Fusion of Lidar Data and Aerial Imagery for a More Complete Surface Description. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, v. 34, n. 3/A, p. 310–317, 2002. Disponível em: <[http://rsl.geology.buffalo.edu/documents/schenk\\\_\\\_csatho\\\_\\\_gr](http://rsl.geology.buffalo.edu/documents/schenk\_\_csatho\_\_gr)>.
- [5] MENA, J. B. State of the art on automatic road extraction for GIS update: A novel classification. *Pattern Recognition Letters*, v. 24, n. 16, p. 3037–3058, 2003. ISSN 01678655.
- [6] LIU, Y. et al. Deep learning for pixel-level image fusion: Recent advances and future prospects. *Information Fusion*, Elsevier, v. 42, n. September 2017, p. 158–173, 2018. ISSN 15662535. Disponível em: <<https://doi.org/10.1016/j.inffus.2017.10.007>>.
- [7] AUDEBERT, N.; SAUX, B. L.; LEFÈVRE, S. Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 140, p. 20–32, 2018. ISSN 09242716.
- [8] ROTTENSTEINER, F. et al. ISPRS Test Project on Urban Classification and 3D Building Reconstruction. *ISPRS - Commission III - Photogrammetric Computer Vision and Image Analysis Working Group III / 4 - 3D Scene Analysis*, p. 1–16, 2013. ISSN 18770428.



# XII Colóquio Brasileiro de Ciências Geodésicas

## AGRADECIMENTOS



# V Simpósio Brasileiro de Geomática



processo nº 2021/03586-2

Fundação de Amparo à Pesquisa do Estado de São Paulo

- Implementação e resultados disponíveis em [https://github.com/tunofilho/CBCG\\_2022\\_article](https://github.com/tunofilho/CBCG_2022_article)