

Segmentação Semântica de Ambientes Internos

**Alexandre Soli Soares
Vinicius Cin**

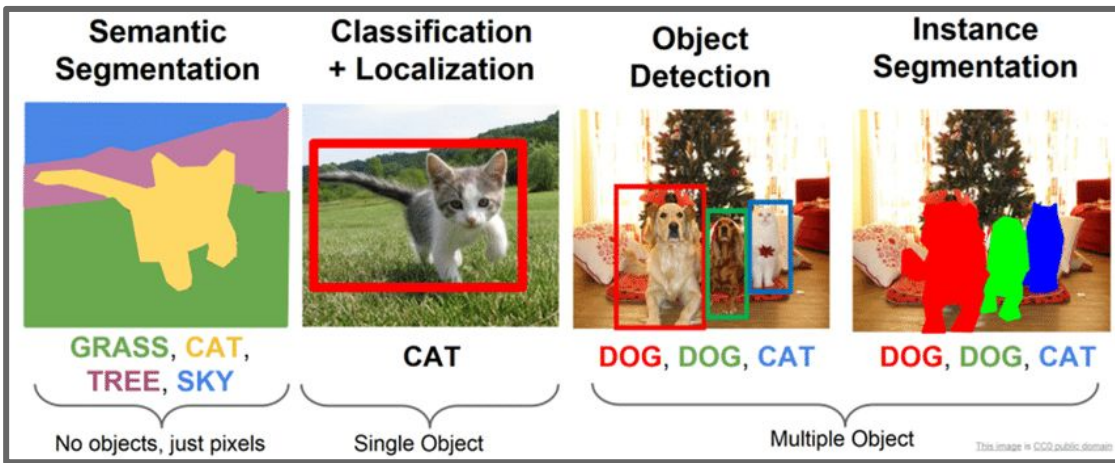


Motivações

- Desenvolvimento futuro de um projeto para guiar deficientes visuais
- Explorar novos conceitos não vistos em sala de aula
- Compreender a segmentação semântica



Fonte: Antonio Cruz, Wikipédia 2006



Fonte: Li, Johnson and Yeung, 2017

Motivações



Fonte: Instituto Magnus, Thiago Pereira

Instituto Magnus

“Esses animais foram escolhidos após uma análise criteriosa de origem, pedigree, genética e condições de saúde dos pais, em conformidade com o protocolo estabelecido pela Federação Internacional de Cães-Guias”, explica **Thiago Pereira**, principal responsável pelo instituto. Isso porque o processo de formação de um cão-guia é longo, caro e, **atualmente, muito abaixo da demanda**. Pereira estima que, até estar totalmente pronto, cada animal consuma entre **R\$ 50 mil e R\$ 80 mil**. Por isso, é necessário aumentar ao máximo o nível de aproveitamento dos filhotes.

Metodologia

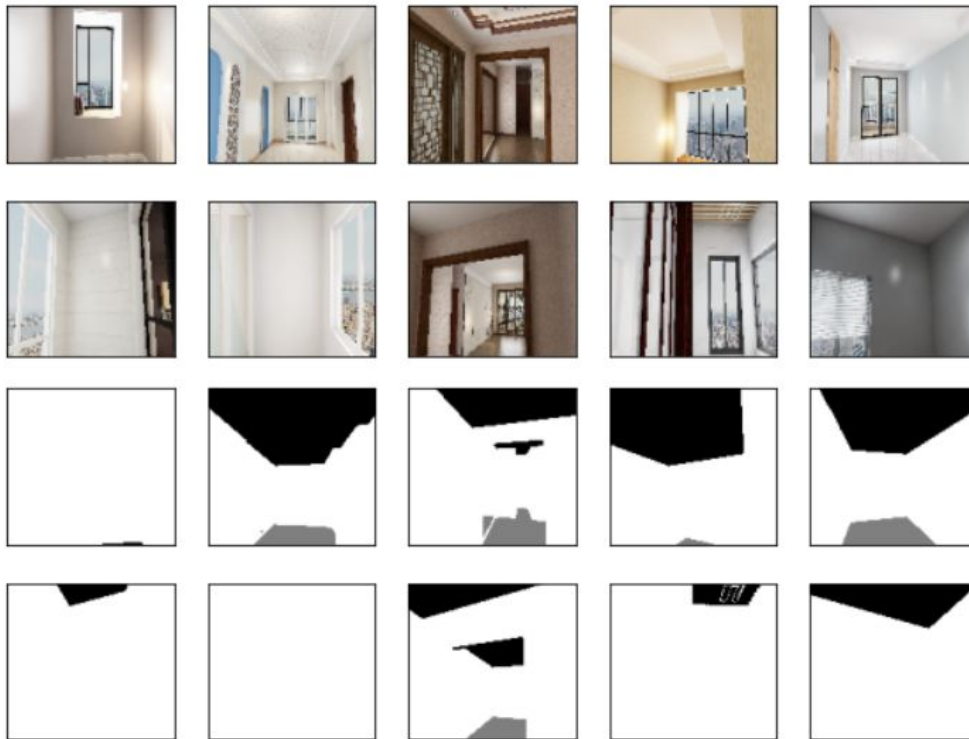


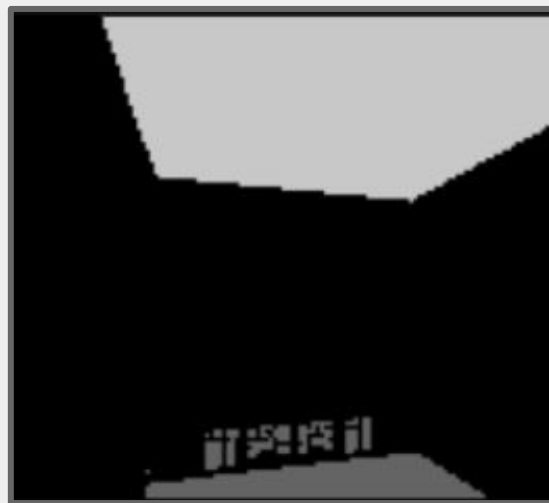
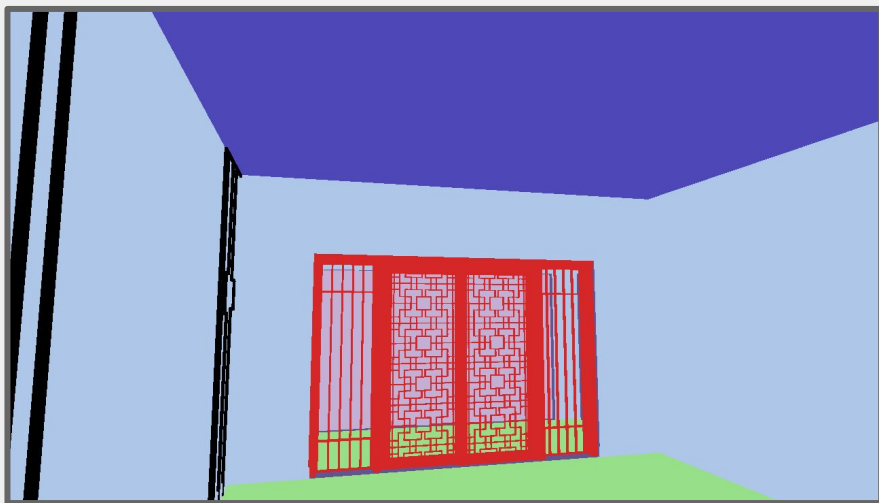
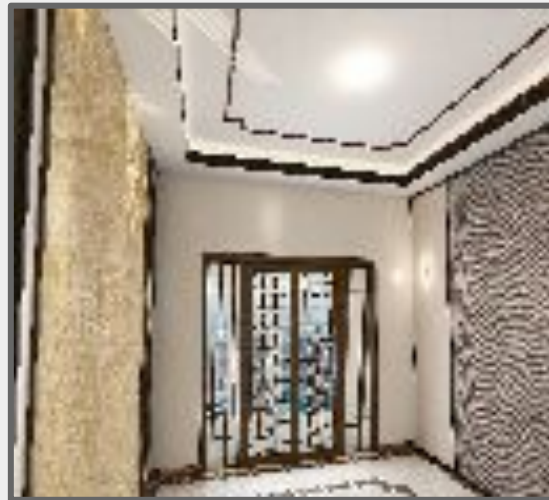
- Consulta bibliográfica
- Procura por um dataset
- Escolha de arquitetura de rede (U-Net)
- Experimentação
- Treinamento de redes
- Avaliações

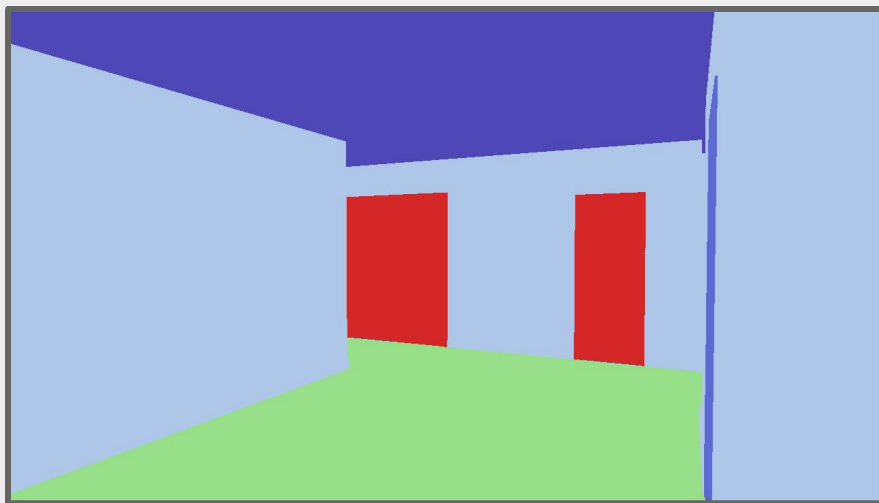
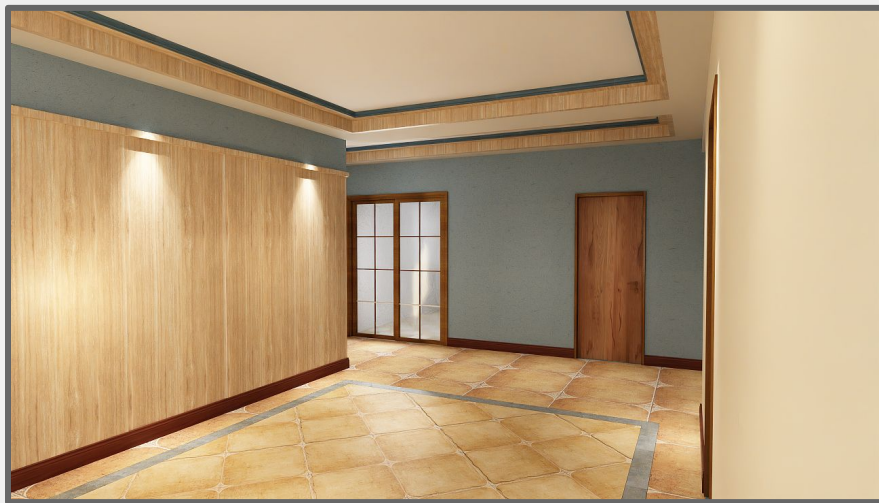
Dataset



- Structured3D Empty
- Original: 1280x720
- Pré-processamento
- 6515 imagens (80%, 10%, 10%)









Treinamento de Redes

- Rede “from scratch” e Transfer learning
- Obter um modelo com capacidade
- Perda Soft Dice
- Refinar os hiperparâmetros
- Aplicar regularização

Resultados

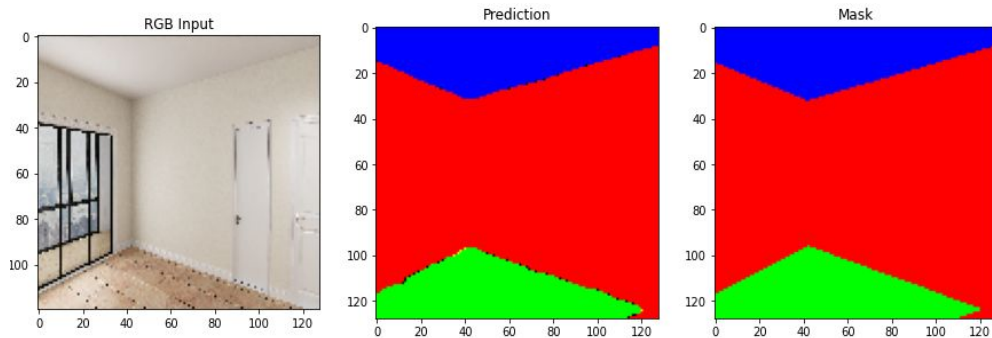


- Comparação entre Redes
- Análise de predições ruins
- Predições em imagens reais

Comparação entre Redes

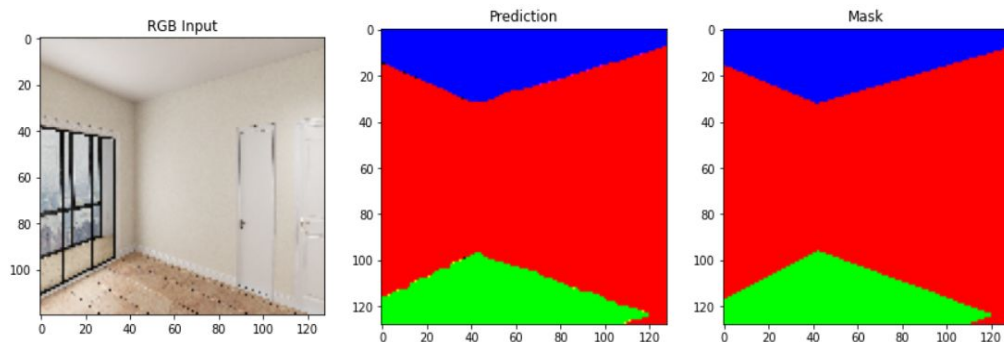


VGG16



	Soft Dice	Mean iou	MCC
VGG16	0.9931	0.9862	0.9921

U-Net:

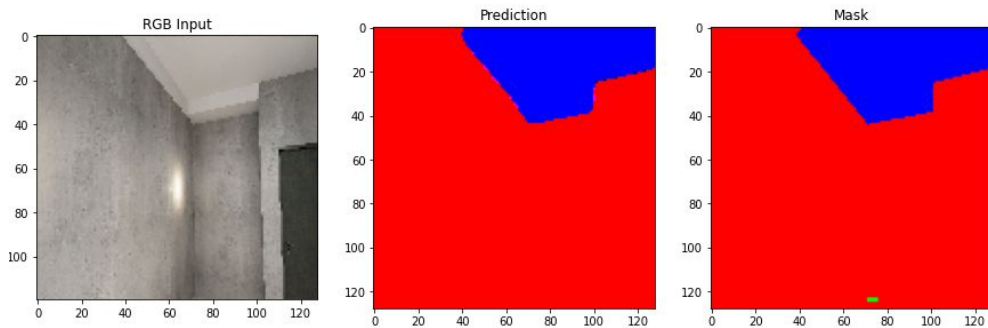


	Soft Dice	Mean iou	MCC
U-Net	0.9909	0.9821	0.9899

Comparação entre Redes

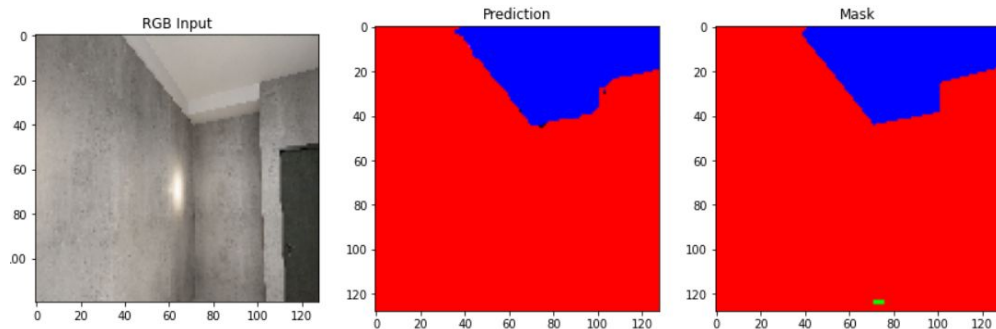


VGG16



	Soft Dice	Mean iou	MCC
VGG16	0.6640	0.6614	0.9956

U-Net:

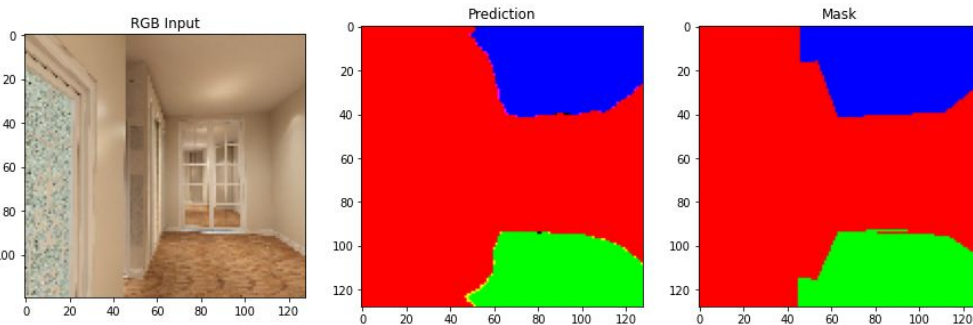


	Soft Dice	Mean iou	MCC
U-Net	0.6602	0.6539	0.9912

Comparação entre Redes

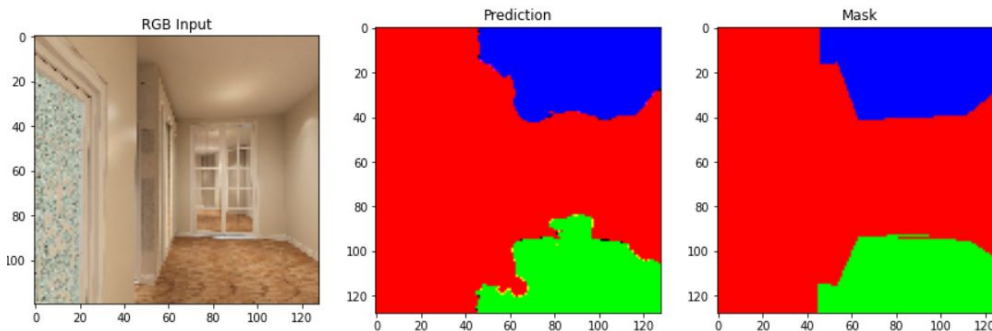


VGG16



	Soft Dice	Mean iou	MCC
VGG16	0.9651	0.9327	0.9584

U-Net:



	Soft Dice	Mean iou	MCC
U-Net	0.9540	0.9131	0.9477

Predições Ruins VGG16



Imagem 1

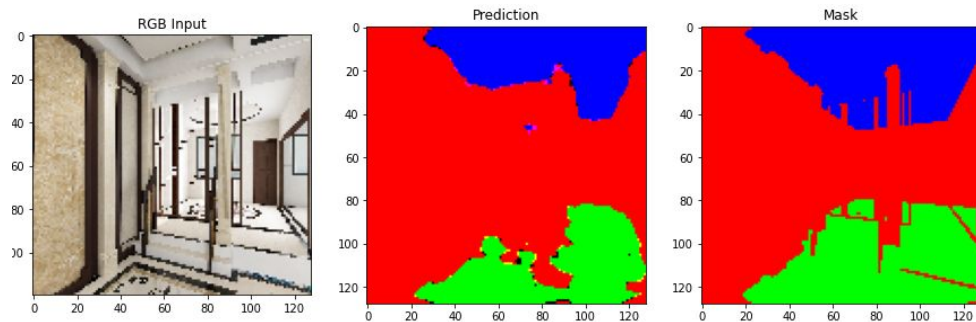
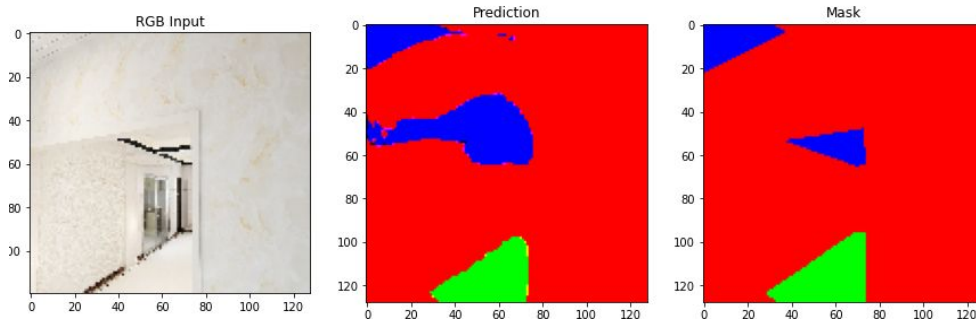


Imagem 2



	Soft Dice	Mean iou	MCC
Imagem 1	0.8673	0.7679	0.8300
Imagem 2	0.8400	0.7621	0.8965

Resultados finais

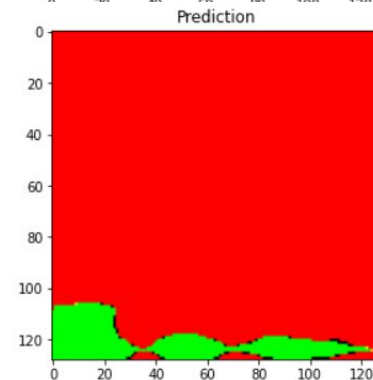
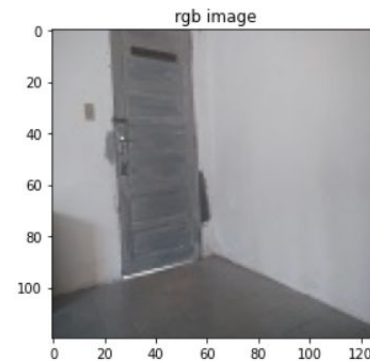
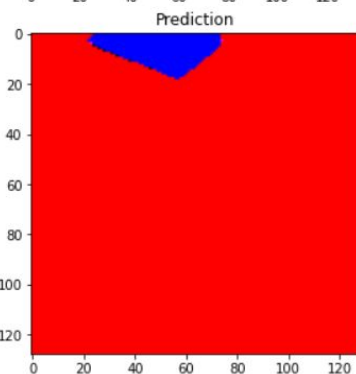
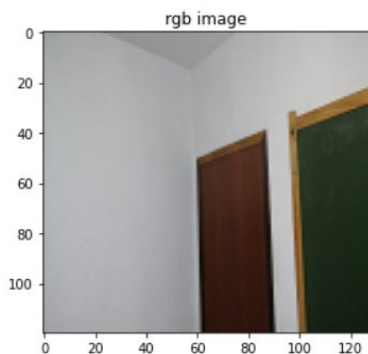
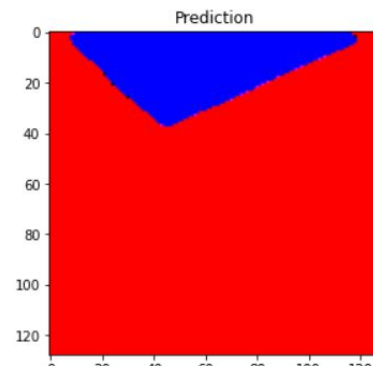
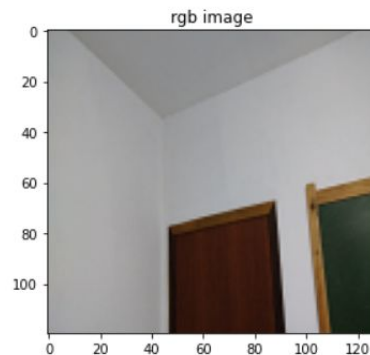
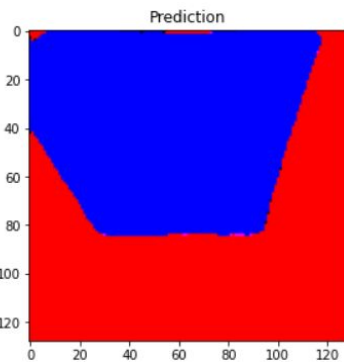
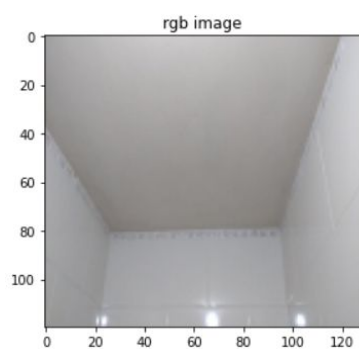


Métrica	U-Net Treino	VGG16 Treino	U-Net Validação	VGG16 Validação	U-Net Teste	VGG16 Teste
Soft Dice	0.9194	0.9875	0.9330	0.9539	0.9320	0.9562
Mean-IoU	0.8206	0.9535	0.8405	0.8897	0.8392	0.8922
Accuracy	0.9737	0.9955	0.9735	0.9848	0.9749	0.9849



Predições em imagens reais

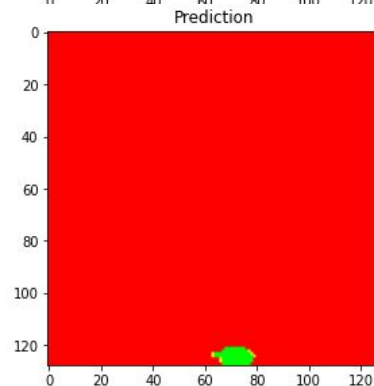
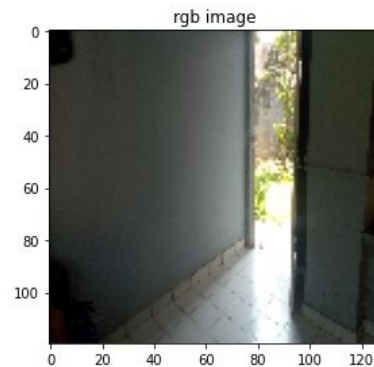
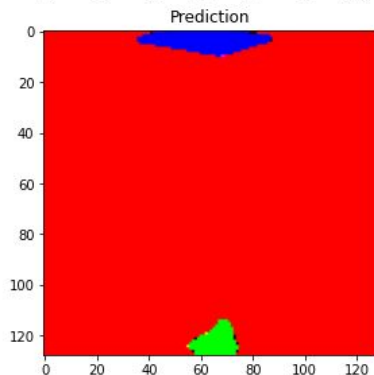
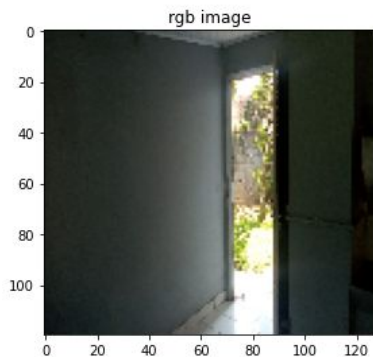
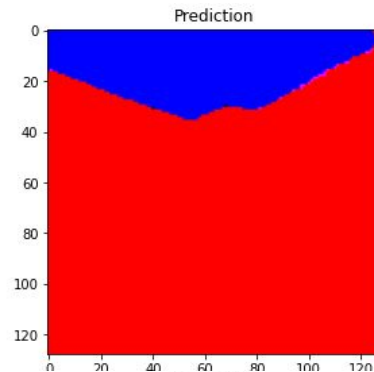
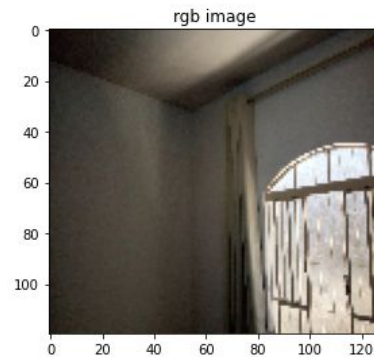
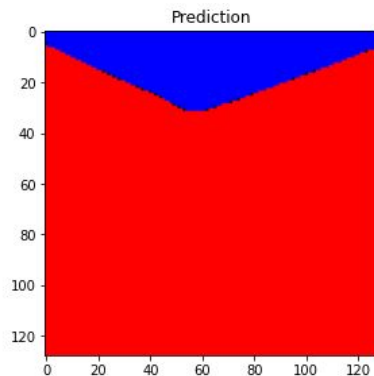
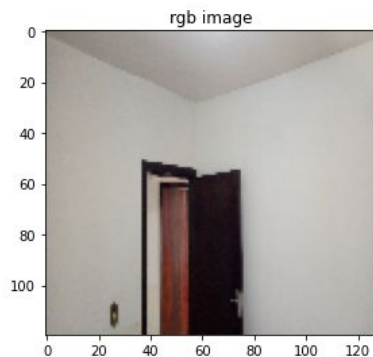
Imagens reais obtidas pelos autores





Predições em imagens reais

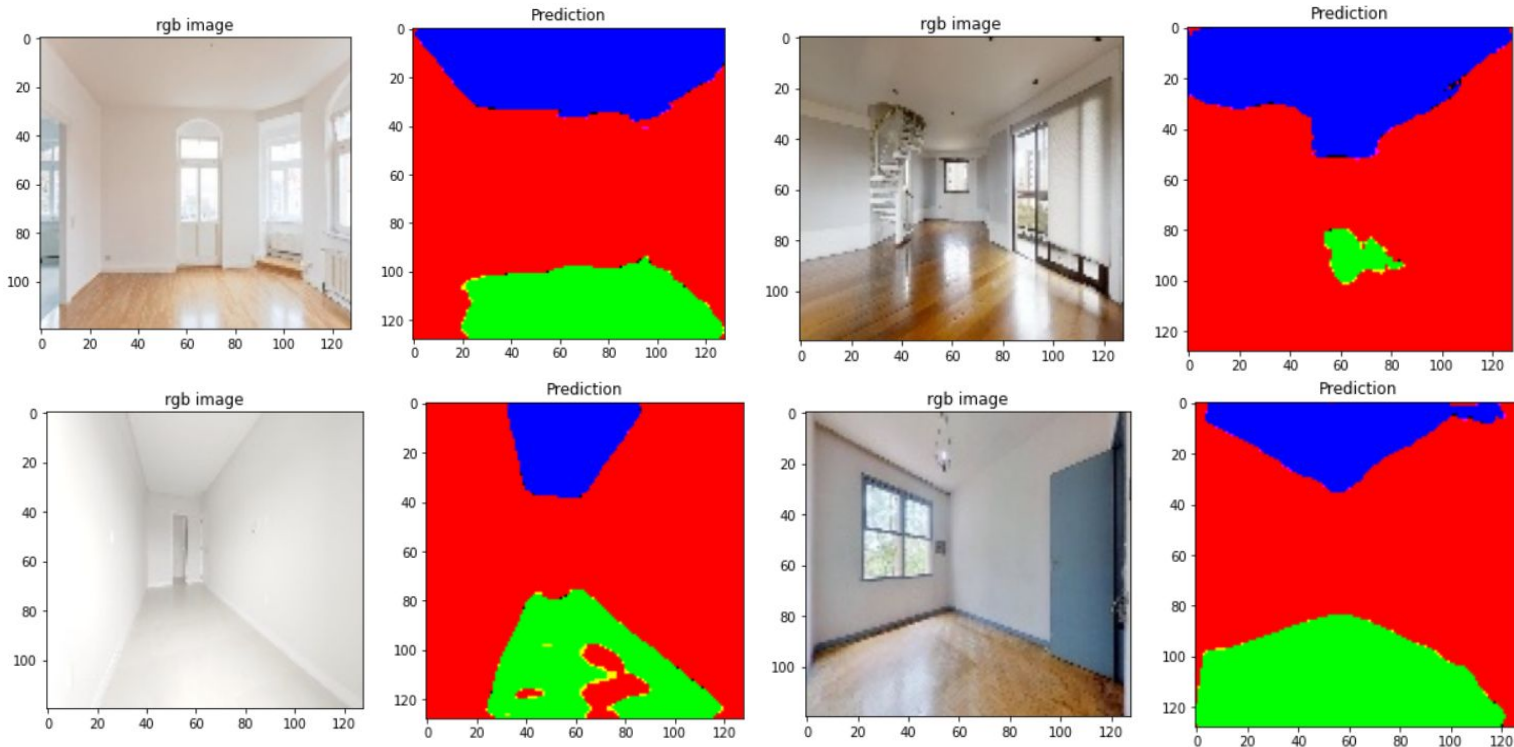
Imagens reais obtidas pelos autores





Predições em imagens reais

Imagens reais obtidas da internet

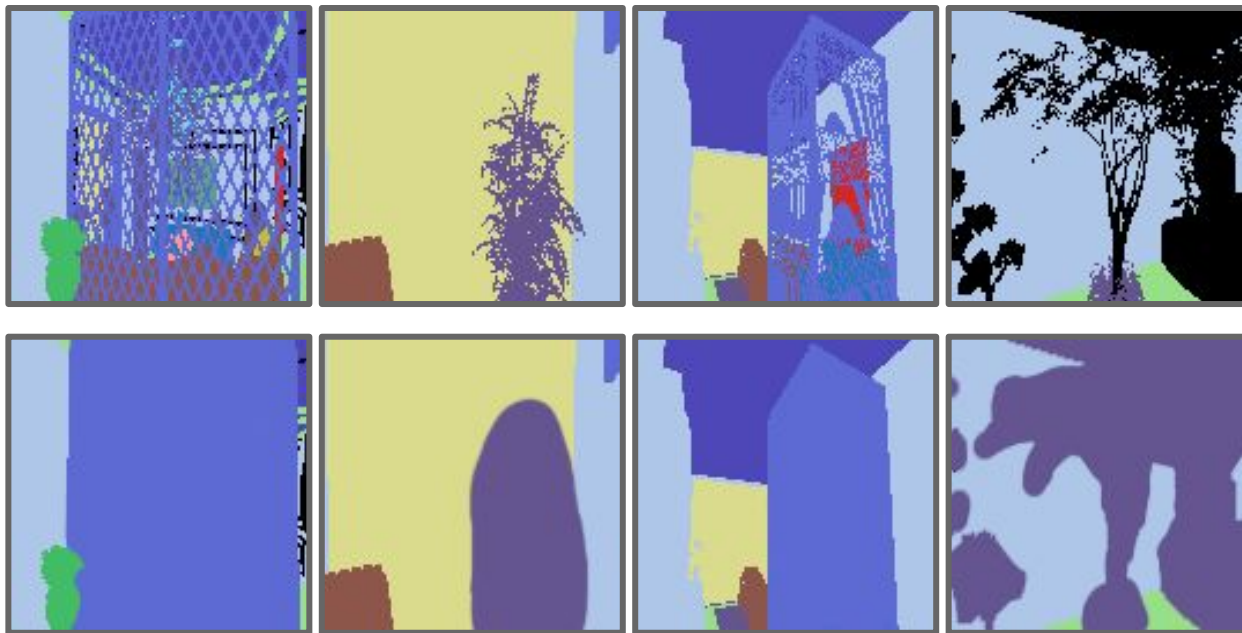


Dificuldades



- Dataset perspective full
- Iluminação
- Desbalanceamento do dataset

Dificuldades



Fonte: Structured3D Dataset, Modificado de Structured3D Dataset

Dificuldades



Imagem 1 (VGG16)

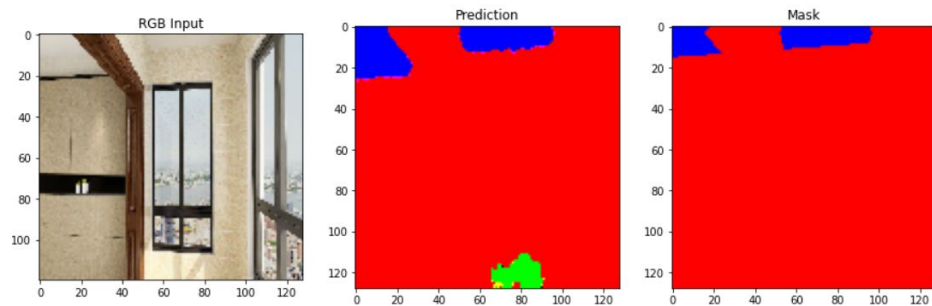
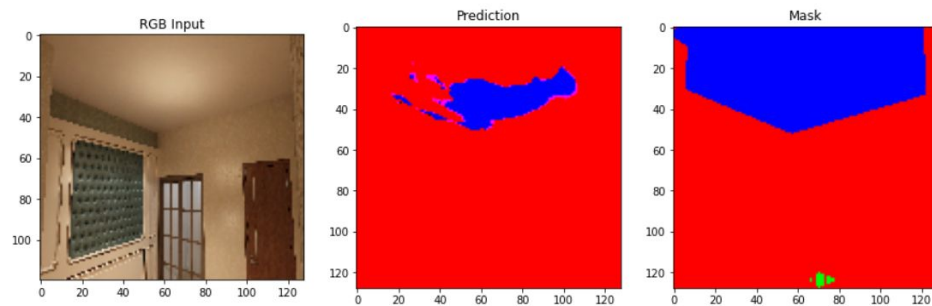


Imagem 2 (VGG16)



	Soft Dice	Mean iou	MCC
Imagem 1	0.5871	0.5336	0.9375
Imagem 2	0.4064	0.3219	0.6393

Dificuldades



Proporções

Parede (0, 263954217)

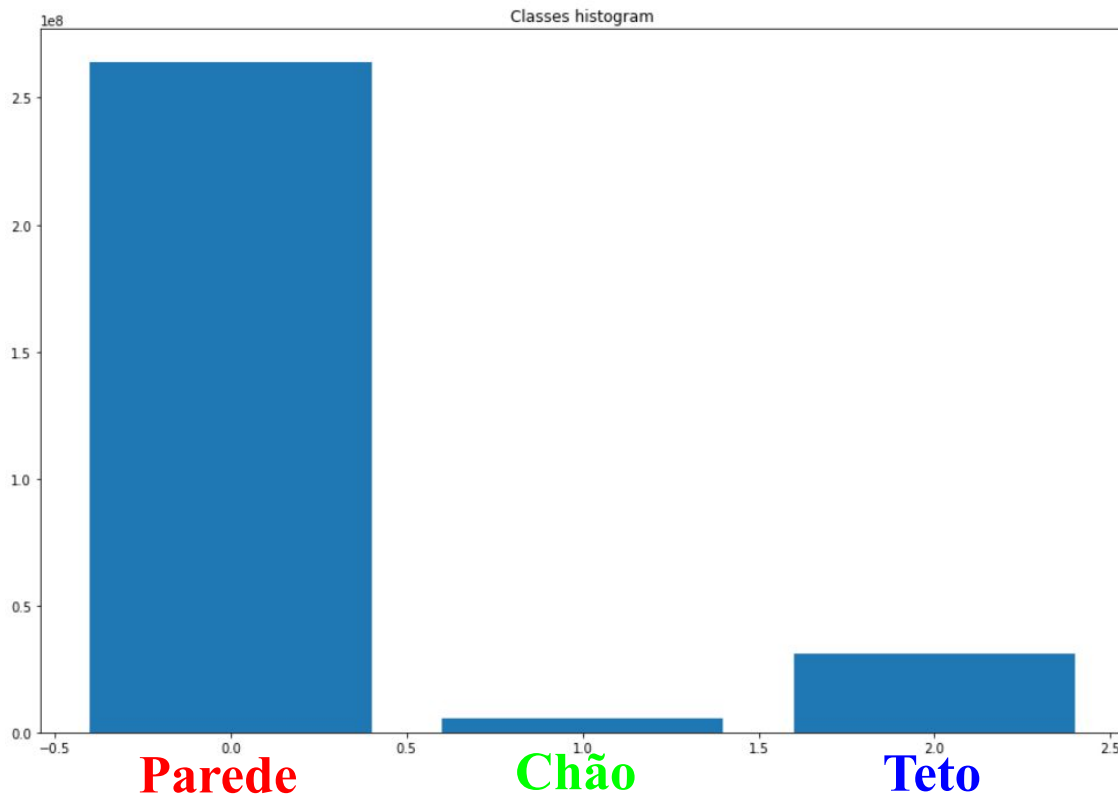
Chão (1, 5387763)

Teto (2, 30869220)

Parede : 87,92%

Chão : 1,79%

Teto : 10,28%



Aprendizado



- Arquitetura da U-Net
- Processo de segmentação
- Importância de um bom dataset
- Novas métricas e funções de perda
- Utilizar uma rede pré treinada como encoder
- Lidar com Tensores



Possíveis melhorias

- Obter um dataset mais adequado para a nossa proposta
- Experimentar outras redes
 - MobileNet, ResNet34
 - SegNet, Mask R-CNN
- Conjunto de dados a partir de filmagens
- Realizar testes com a perda MCC [5]
- Remover amostras negativas do dataset e utilizar ResNet34 como Encoder [7]
- Experimentar uma ponderação da função perda sugerida por [8]
- Experimentar outras resoluções de imagem

Referências Bibliográficas

- [1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, **U-Net: Convolutional Networks for Biomedical Image Segmentation**, University of Freiburg, Germany, Disponível em: <https://arxiv.org/pdf/1505.04597.pdf>
- [2] Jia Zheng, Junfei Zhang, Jing Li, Rui Tang, Shenghua Gao and Zihan Zhou, **Structured3D: A Large Photo-realistic Dataset for Structured 3D Modeling**, Artigo disponível em: https://www.ecva.net/papers/eccv_2020/papers_ECCV/papers/123540494.pdf
Dataset disponível em: <https://structured3d-dataset.org/>
- [3] Fabien H. Wagner, Alber Sanchez, Yuliya Tarabalka, Rodolfo G. Lotte, Matheus P. Ferreira, Marcos P. M. Aidar, Emanuel Gloor, Oliver L. Phillips & Luiz E. O. C. Aragão, **Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images**, Disponível em: https://www.researchgate.net/publication/331549658_Using_the_U-net_convolutional_network
- [4] Jeroen Bertels, Tom Eelbode, Maxim Berman, Dirk Vandermeulen, Frederik Maes, Raf Bisschops, and Matthew B. Blaschko, **Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory & Practice**, Disponível em: <https://arxiv.org/pdf/1911.01685v1.pdf>
- [5] Kumar Abhishek and Ghassan Hamarneh, **Matthews correlation coefficient loss for deep convolutional networks: Application to skin lesion segmentation**, School of Computing Science, Simon Fraser University, Canada, Disponível em: <https://arxiv.org/pdf/2010.13454.pdf>
- [6] Yinda Zhang, Shuran Song, Ersin Yumer, Manolis Savva, Joon-Young Lee, Hailin Jin, Thomas Funkhouser, **Physically-Based Rendering for Indoor Scene Understanding Using Convolutional Neural Networks**, Disponível em: https://joonyoung-cv.github.io/assets/paper/17_cvpr_physically_based.pdf
- [7] Xiaoling Xia, Qinyang Lu, Xin Gu, **Exploring An Easy Way for Imbalanced Data Sets in Semantic Image Segmentation**, College of Computer Science, Donghua University Shanghai, China, Disponível em: <https://iopscience.iop.org/article/10.1088/1742-6596/1213/2/022003/pdf>
- [8] Patrik Olá Bressan, José Marcato Junior, José Augusto Correa Martins, Diogo Nunes Gonçalves, Daniel Matte Freitas, Lucas Prado Osco, Jonathan de Andrade Silva, Zhipeng Luo, Jonathan Li, Raymundo Cordero Garcia, Wesley Nunes Gonçalves, **Semantic segmentation with labeling uncertainty and class imbalance**, Disponível em: <https://arxiv.org/pdf/2102.04566.pdf>
- [9] Wei Song, Nan Zheng, Xiangchun Liu, Lirong Qiu, and Rui Zheng, **An Improved U-Net Convolutional Networks for Seabed Mineral Image Segmentation**. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8740923>
- [10] Bhakti Baheti, Shubham Innani, Suhas Gajre Sanjay Talbar, **Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment**, Disponível em: https://openaccess.thecvf.com/content_CVPRW_2020/papers/w22/Baheti_Eff-UNet_A_Novel_Architecture_for_Semantic_Segmentation_in_Unstructured_Environment_CVPRW_2020_paper.pdf
- [11] Henry Yau, **Image segmentation using U-Net**, Disponível em: <https://github.com/henyau/Image-Segmentation-with-U-Net>

Obrigado pela atenção!

Alexandre Soli
Vinicius Cin

solisoares3@gmail.com
viniciuscimbr@gmail.com