Semantic Segmentation of Buildings From Aerial Images

Sara Mohamed, Ali Mahmoud

CSCE 4604 – Advanced Machine Learning - Fall 2023

Prof. Moustafa Youssef

THE AMERICAN UNIVERSITY IN CAIRO

INTRODUCTION

Extracting or marking buildings from aerial images is a task required for many city building projects, ranging from road design and engineering by civil engineers, to urban planning and city design by architects. To facilitate the often-time-consuming process, this project aims to provide a hassle-free shortcut to the process, by inputting an aerial image to produce a mask of the buildings in the image.

Fig. 1 Example of marking buildings for urban planning. Image provided by friend from Department of Architectural Engineering- AUC

DATASET

For this task, we opted to change from our baseline's Massachusetts Dataset to the SpaceNET (Rio de Janeiro) dataset, which contains more images, is more balanced, and is easier to generalize.

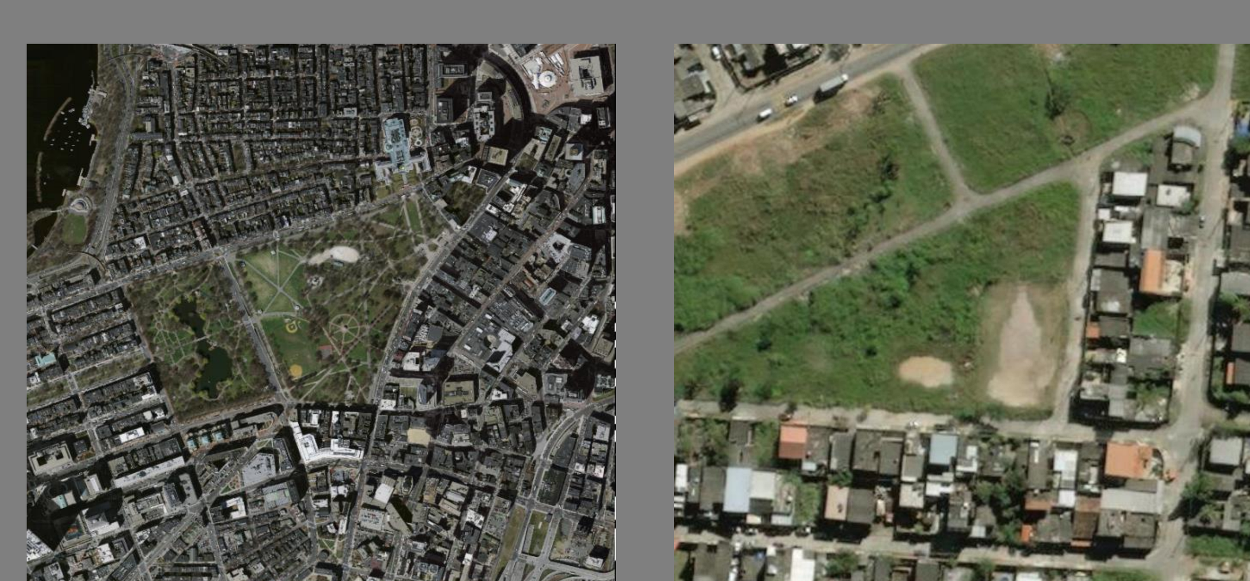

Fig. 2 On the left, an example image from the Massachusetts dataset. On the right, an example image from the SpaceNET dataset (chosen by us).

MODEL & ARCHITECTURE

To tackle this problem, we opted to employ CNNs using the U-Net architecture, initially introduced for image segmentation in biomedical engineering. It adopts an encoder-decoder architecture, with conv and pooling layers used in the encoder path for feature extraction, and conv and upconv layers used in the decoding path for feature localization. U-Net also uses skip connections between the decoder and the encoder paths to better localize information extracted by the decoder.

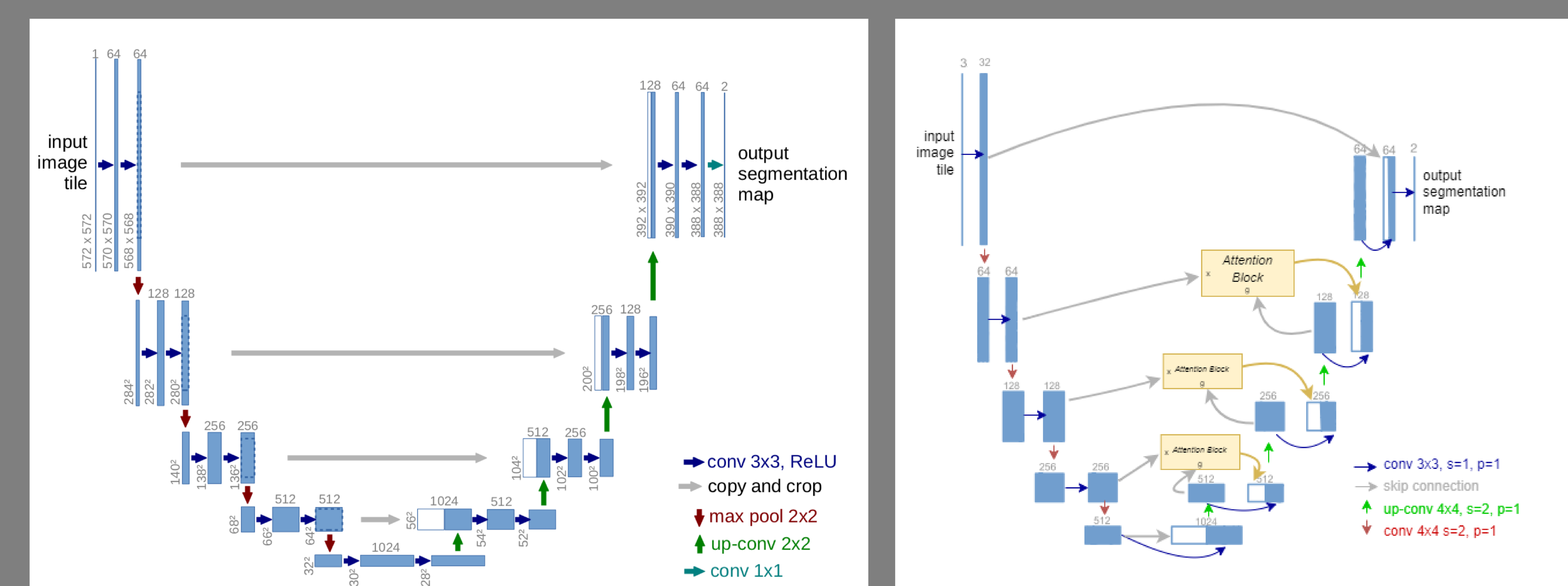

Fig. 3 On the left, the U-Net architecture proposed in "U-Net: Convolutional Networks for Biomedical Image Segmentation" by Ronneberger et. Al. On the right, our altered U-Net architecture

The alterations to the architecture outlined above yielded better results on all evaluation metrics. They include adding batch normalization after all upconv layers, replacing pooling layers with4x4 convs with stride 2, adding padding to all layers, altering skip connections to occur before upconvs and adding attention blocks at the skip connections.

EVALUATION & RESULTS

Evaluation metrics for our problem include IoU score, pixel accuracy, and F1 score. Outlined below are the changes overwent throughout the project and the corresponding results. Baseline loss function was Dice Loss.

	IoU	Pixel Accuracy	F1 Score
Baseline Model with Massachusetts dataset	0.4648	0.4832	0.4898
Baseline Model with SpaceNET dataset	0.7095	0.847	0.8
Altered U-Net Architecture with SpaceNET dataset and BCE with Logits loss	0.9138	0.945	0.951
Altered U-Net Architecture with SpaceNet dataset, BCE with Logits loss, and Attention Blocks	0.9144	0.9472	0.9539

REFERENCES

- Chen Q, Wang L, Wu Y, Wu G, Guo Z, Waslander SL. Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings. ISPRS Journal of Photogrammetry and Remote Sensing. 2019;147:42-55.
- Balraj98. (n.d.). UNet for Building Segmentation [PyTorch]. Kaggle. Retrieved October 21, 2023, from https://www.kaggle.com/code/balraj98/unet-for-building-segmentation-pytorch/notebook#Training-UNet
- Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al.. Generative Adversarial Networks; 2014. Available from: https://arxiv.org/abs/1406.2661
- Mo Y, Wu Y, Yang X, Liu F, Liao Y. Review the state-of-the-art technologies of semantic segmentation based on deep learning. Neurocomputing. 2022;493:626-46. Available from: https://www.sciencedirect.com/science/article/pii/S0925231222000054.
- Oktay et al. Attention U-Net: Learning Where to Look for the Pancreas; 2018. MIDL'18. Available from: https://arxiv.org/abs/1804.03999
- Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lecture Notes in Computer Science. 2015 May:234–241
- Ulmas P, Liiv I. Segmentation of Satellite Imagery using U-Net Models for Land Cover Classification. IEEE Access. 2020 Mar
- ;. Available from: https://spacenet.ai/datasets/.
- Volodymyr Mnih. Machine Learning for Aerial Image Labeling & The Massachusetts Dataset. PhD Thesis for the University of Toronto. Available from: https://www.cs.toronto.edu/~vmnih/docs/Mnih_Volodymyr_PhD_Thesis.pdf

CONTACT

Sara Mohamed
sara_mohamed@aucegypt.edu
www.github.com/saraa-mohamedd

Ali Mahmoud
ali_90082@aucegypt.edu
www.github.com/alibeep1