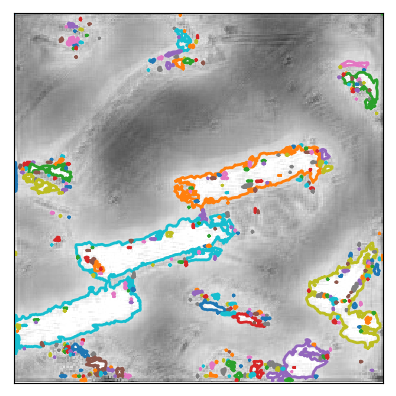
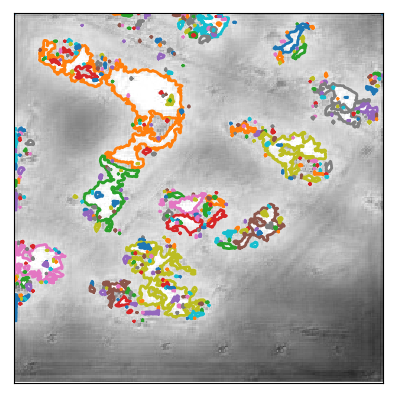
1. **Formalization:** To extract foodprint of buildings from satellite image, I first as most common way form it as a semantic segmentation problem. So I try to use densly deep neural network to address this problem.

1. **Data:** Here I use satellite data from Spacenet challenge, due to the limitation of time and computational source. I only use the sub-dataset for Paris region which contains 1147 images (918 training, 229 test) , I futher devided training into trainning and validation(20%)
2. **Method:** Here I tried two different s pipelines for this problem, (1) first I tried to solve this problem with an end to end netwrok (actually state of art winner for this challenge use similar idea and network). (2) Then, instead of end to end method I also try to preprocess groundtruth into distance map(pixel-wise distance between it’s nearest footprint boundry) and train the network to learn distance map and use post-processing(contour finding, clustering) to find footprint polygon. Here is the networking I am using (Res-Unet)

**A close up of text on a white background

Description automatically generated**

1. **Experiement details:** Input is 3 channels RGB image, ground truth is binary mask or distance map. Here I use Adam optimizer and Iou-coefficient loss, and trained 200 epochs. Here I only use IOU as metric (precision, recall, F1 score can be easily added). For detail codes and reproduction, please check: <https://github.com/LzVv123456/SpaceNet>
2. **Results:** Here I randomly showed results for 3 images, and do not contain quantative result. (From left to right: (1)original image, (2)end to end method, (3)distance-map method, (4)ground truth)

**A picture containing animal, reptile

Description automatically generated**

A view of a city

Description automatically generated**A close up of a logo

Description automatically generated**

**A view of a large building

Description automatically generatedA picture containing animal, reptile

Description automatically generatedA picture containing indoor

Description automatically generated**

1. **Analysis and further work:** (1)For the end to end method: Use whole dataset, more data and larger network can definitely improve performance. Besides, here I only use 3 channels RGB image, with 8 channel tiff image (also provided by challenge) as input data or concatenation of 3 both channels and 8 channels can largely improve performance.(2) my second pipeline works extremely bad, that’s probably due to my bad distance map prediction. Many pre-processing or post-processing can be used to address this problem. So my work here can not fully express the potential of this method.

This work done in a hurry, a lot of improvements can be made and many other methods can be try.

1. **Reference:**

<https://spacenetchallenge.github.io>

<https://medium.com/the-downlinq/getting-started-with-spacenet-data-827fd2ec9f53>

<http://cs231n.stanford.edu/reports/2017/pdfs/550.pdf>

<https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-1-W1/481/2017/isprs-archives-XLII-1-W1-481-2017.pdf>

<https://www.tandfonline.com/doi/full/10.1080/22797254.2017.1416676>