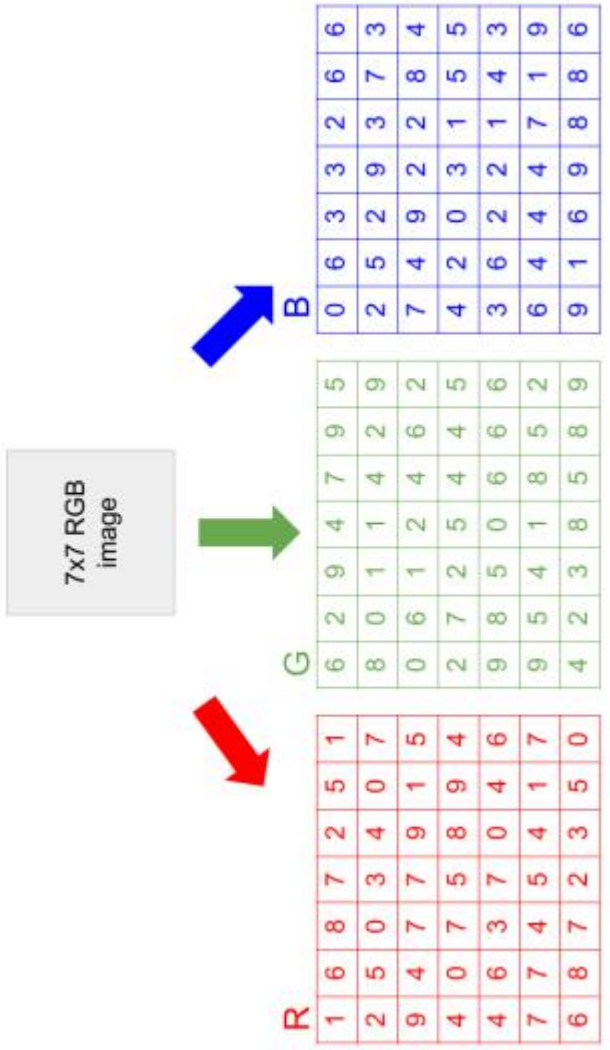
TTI Prediction in Urban Road Network Using Computer Vision Techniques

LI Xiahan

RGB 3 phase construction

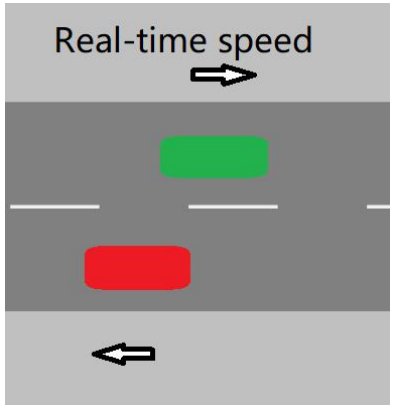


Accessibility from a given origin
Geographical identity

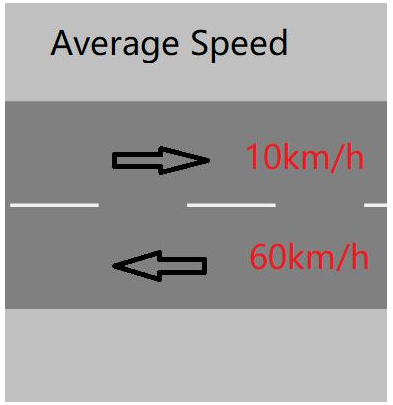
Trajectory of drivers
Real-time speed of drivers

Geometric shape of roads
Average speed on roads

78	77	77	78
77	76	76	77
76	75	75	76
75	74	74	75



0	0	0	0
0	32	33	35
54	60	64	0
0	0	0	0



0	0	0	0
10	10	10	10
60	60	60	60
0	0	0	0

Change of the Picture Size

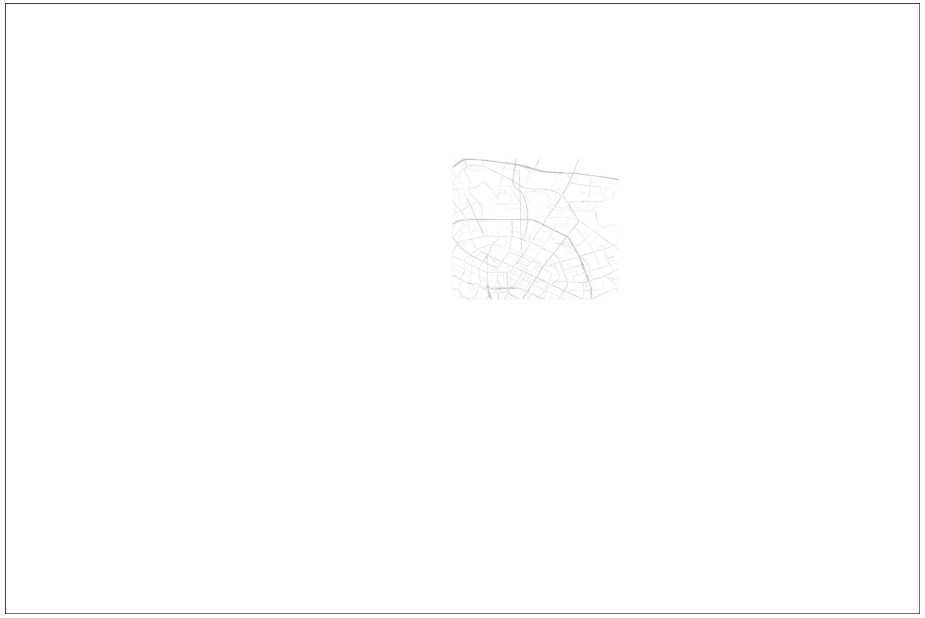
Resize the picture from 4800 * 3200 to 896 * 768 because the trajectory data is restricted in:

longtitude lowerbound = 104.0402
latitude lowerbound = 30.6516
longtitude upperbound = 104.1298
latitude upperbound = 30.7284

城市交通指数和轨迹数据

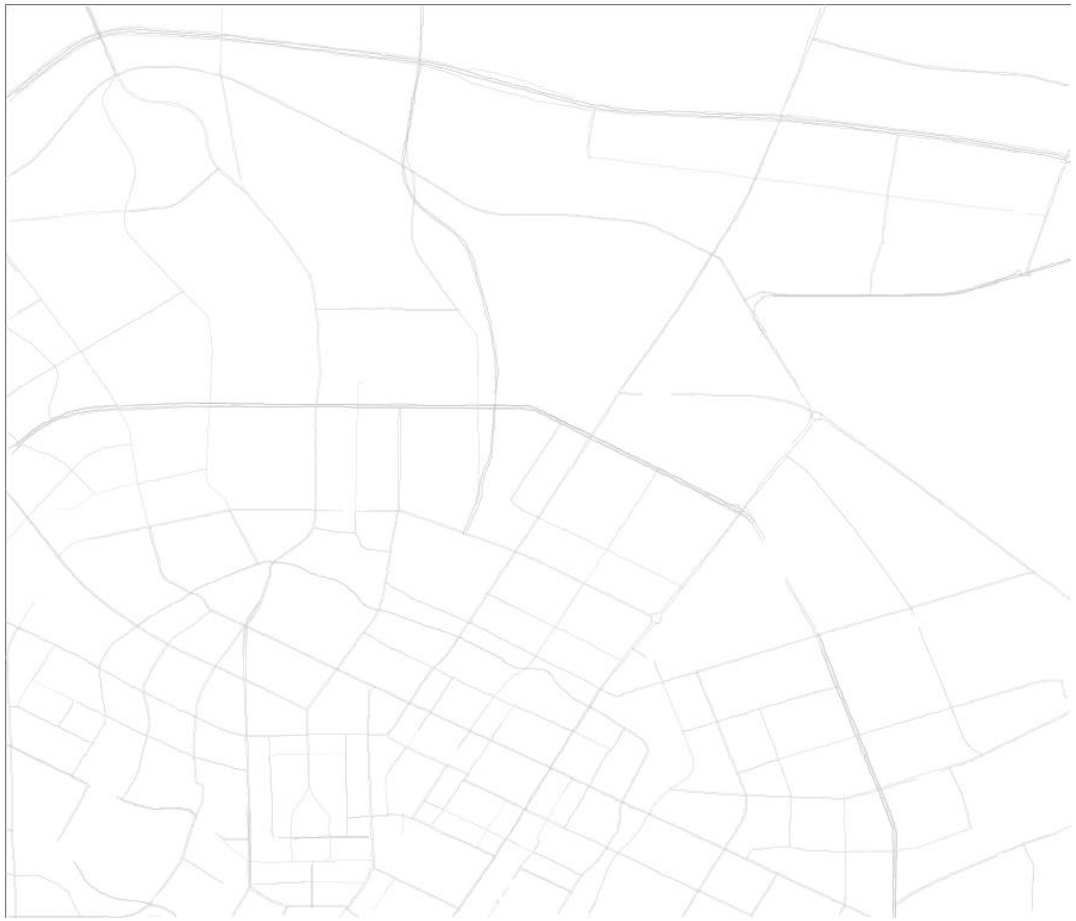
2018年10月11月成都和西安二环区域的滴滴快专车平台的轨迹数据及区域内道路级别的交通指数数据（Travel Time Index ,TTI）和平均行驶速度。
成都区域：
[30.727818,104.043333],[30.726490,104.129076]
[30.655191,104.129591],[30.652828,104.042102]

Roadmap data in 4800 * 3200

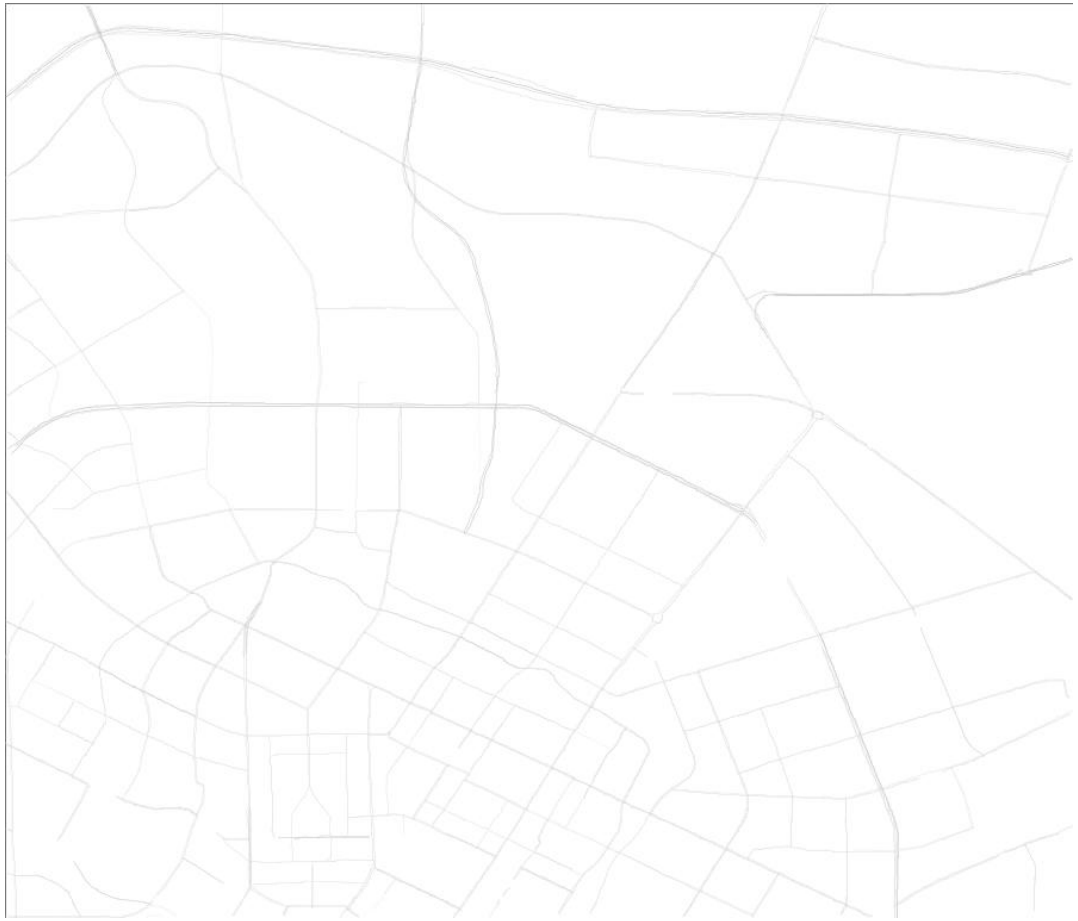


Trajectory data in 4800 * 3200

Roadmap Data Visualization

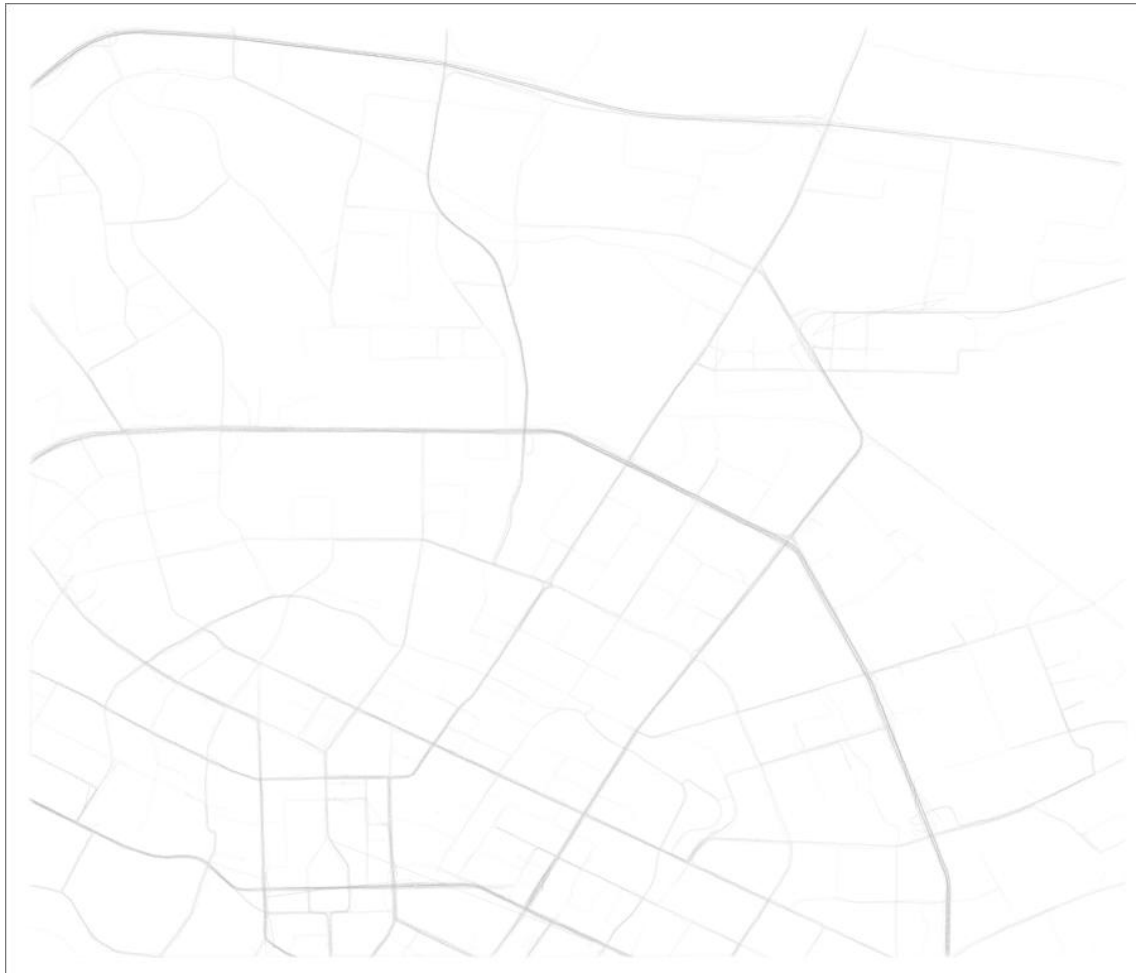


2018-10-10 03:00:00 (138KB)



2018-10-10 08:00:00 (127KB)

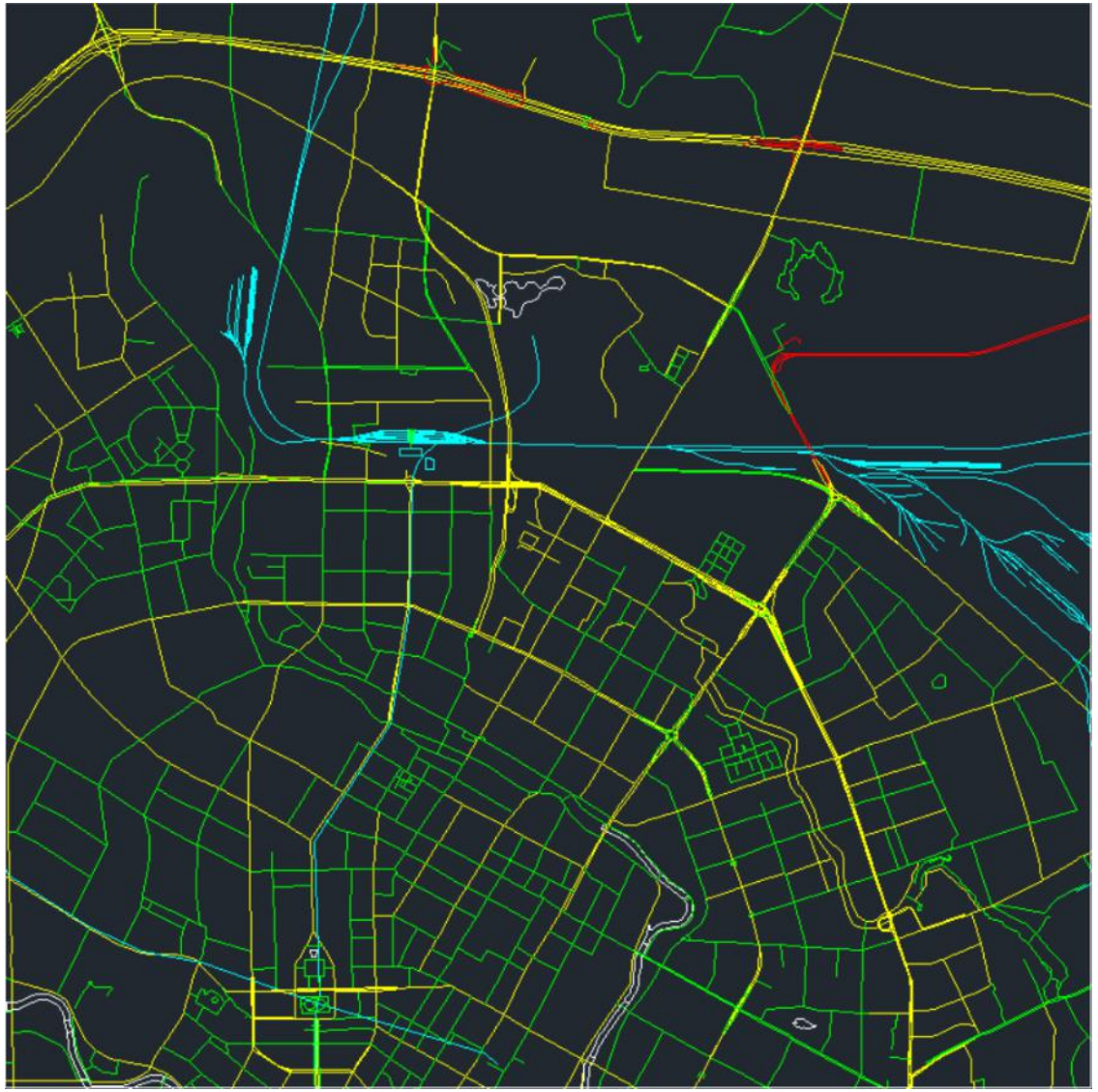
Trajectory Data Visualization



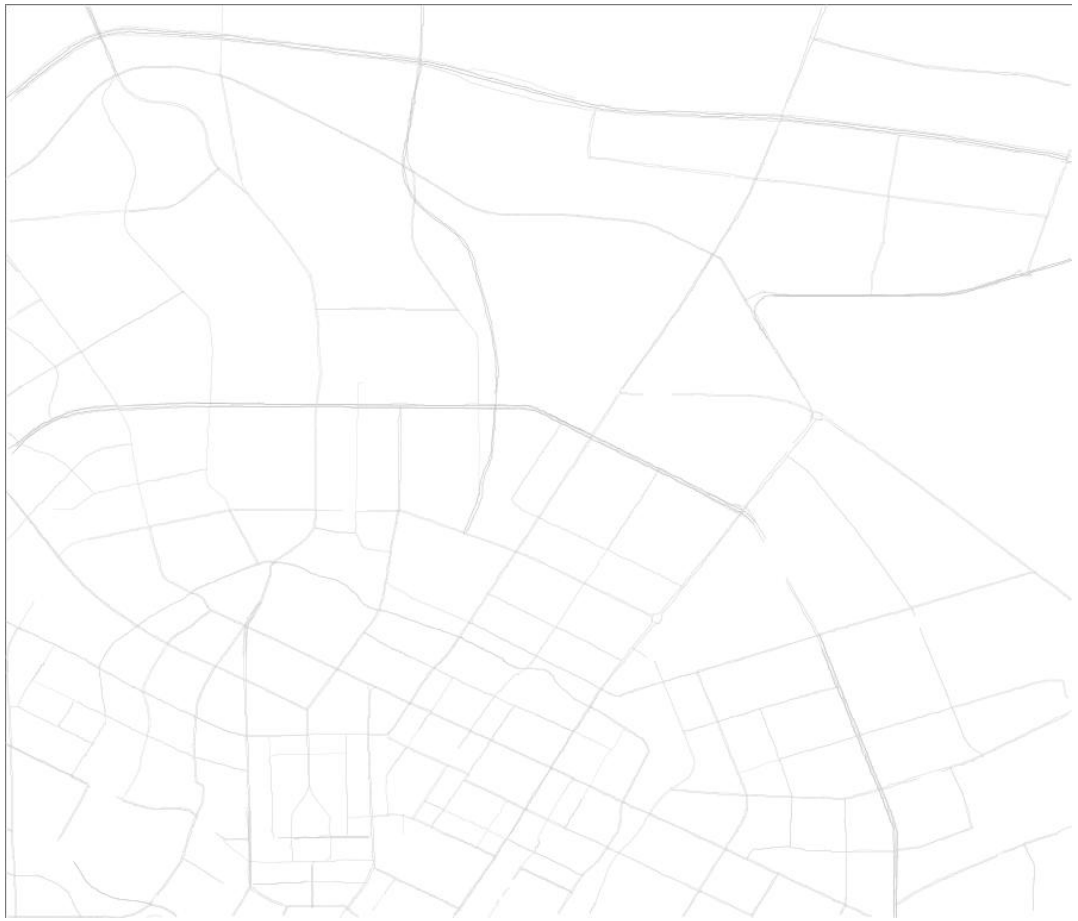
2018-10-10 02:25:00-02:35:00

2018-10-10 12:05:00-12:15:00

Geographical Data Visualization



Roadmap data



One roadmap picture represent for:

- The shape of the road
- The average speed of vehicles on the road
- Above information in every 10 minutes, e.g:

20:00, 20:10, 20:20, etc.

The roadmap picture is similar to heatmap picture:

- Darker area means higher speed
- White area means 0 speed (pixel = 255)

Strava Heatmap

https://www.strava.com/heatmap#12.00/104.01781/30.66105/hot/all

How we plot Roadmap

282659	2018/10/19 20:00	1.89428	21.0006
282964	2018/10/19 20:00	1.47032	24.84
281952	2018/10/19 20:00	1.35722	43.2719
281933	2018/10/19 20:00	1.50704	29.3833
282289	2018/10/19 20:00	2.61735	14.4739
282860	2018/10/19 20:00	1.38475	29.7054
282966	2018/10/19 20:00	1.08128	96.1934

Extract TTI & Speed data in the specific boundary

Extract Network data in the specific boundary

Merge the two data by pandas.merge Criteria: same ID

Fill in missing value by pandas.fillna Criteria: median

(Missing value: few or no vehicles recorded in the specific road)

Longitude -> X coordinate of point

Latitude -> Y coordinate of point

Speed -> number of pixel(value of color) of line

Use matplotlib to plot the picture

181581	双流区	POLYGON((104.11863 30.22707,104.11661 30.22718,104.11644 30.22719,104.11491 3
181582	简阳市	POLYGON((104.42813 30.08827,104.42665 30.08851,104.42629 30.08857,104.42363 3
281863	八里桥路:北站东二路,三环路	MULTILINESTRING((104.07437 30.71442,104.0742 30.71483,104
281864	八里桥路:三环路,北站东二路	MULTILINESTRING((104.07452 30.71324,104.0745 30.71338,104
281865	二环路西段:广福路,金牛大道	MULTILINESTRING((104.02746 30.64001,104.02707 30.64039),(
281866	二环路西段:金牛大道,广福路	MULTILINESTRING((104.02536 30.64168,104.02488 30.64216,10

Trajectory Data

One trajectory picture represent for:

- The trajectory of the drivers
- The speed of the drivers
- Above information collected within +- 2min for every 10 minutes, e.g. 00:08:00 - 00:12:00, 00:18:00 - 00:22:00
- Darker area means higher speed
- White area means 0 speed (pixel = 255)
- Denser lines means more drivers during the time

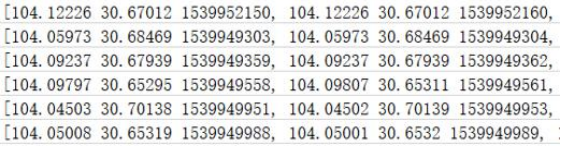
Time Interval Decision:

- When deciding the time interval of plotting, we have two extreme intervals:

1. +- 5 min, e.g. 00:05:00-00:15:00. Advantages: keep the time-related vehicle transport continuous, no lost of raw data.
2. +- 1 s, e.g. 00:09:59 - 00:10:01. Advantages: High relevance to the Roadmap data(discrete data every 10 min)

Final decision: compromise interval +- 2 min. In feature engineering, there's no absolutely right or wrong.

How we plot Trajectory



Extract Trajectory data in the specific boundary

Longitude -> X coordinate of point

Latitude -> Y coordinate of point

Timestamps -> Timestamps in the +- 2 min interval

Compute speed by Euclidean distance and time interval

Discard missing values, modify error values like v=364

Use matplotlib to plot the picture

(Unnecessary to fill in these missing values. Set maximum velocity to v = 150 to avoid error speed)

Plotting Detail

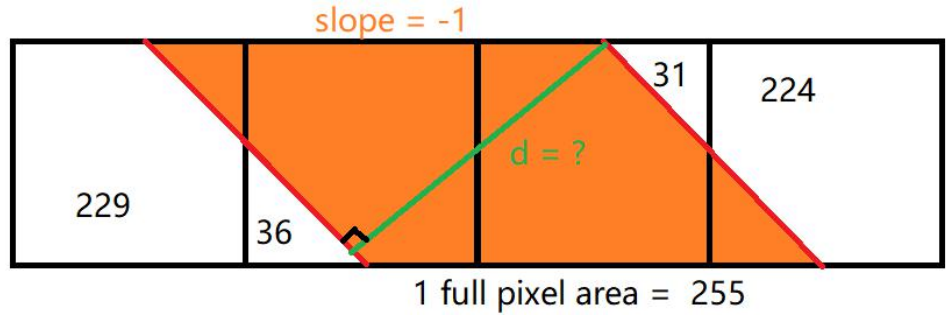
- In python matplotlib plot command, parameters that mainly affect the pixel values are: 'color' value and 'linewidth' value.
- 'color' consist of three hexadecimal value ranging from 00 to ff (0 to 255). we use: 255 - speed = pixel value.
- To figure out the representation of 'linewidth' value , we plot a linewidth = 1, slope = -1 line and check the pixel of the crossing line. By geometric calculation, we find that linewidth = 1 represent for 1.5 pixel occupation (or 165m)
- Assuming the actual road width = 50m, we use linewidth=0.3 to plot roadmap
- Assuming the vehicle occupancy= 5m, we use linewidth=0.03 to plot trajectory

```
x = [0,1]
y = [1,0]
plt.rcParams['figure.dpi']=100
plt.rcParams['figure.figsize']=(1,1)
plt.plot(x,y,color = 'black', linewidth = 1.0)
plt.axis('off')
plt.savefig('test0.png')
```

```
from PIL import Image
import numpy as np
import os
img = Image.open('test0.png')
ig = np.array(img)
ig[:,50,0]
```

```
array([255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255,
       255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255,
       255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255,
       255, 255, 255, 255, 255, 255, 255, 229,  36,  31, 224, 255,
       255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255,
       255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255,
       255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255, 255,
       255, 255, 255, 255, 255, 255, 255, 255], dtype=uint8)
```

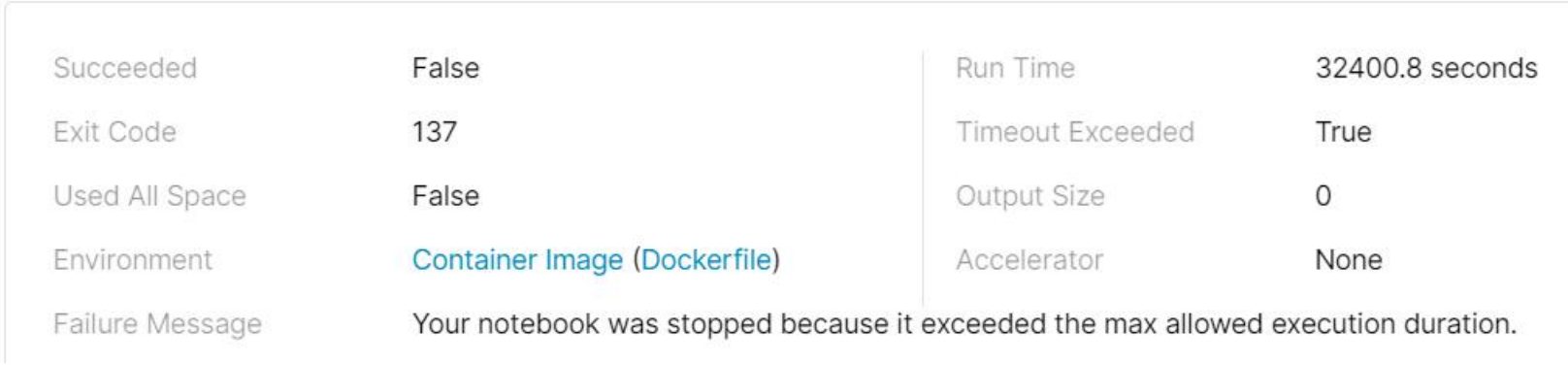
(229,36,31,224)



calculate 'd'

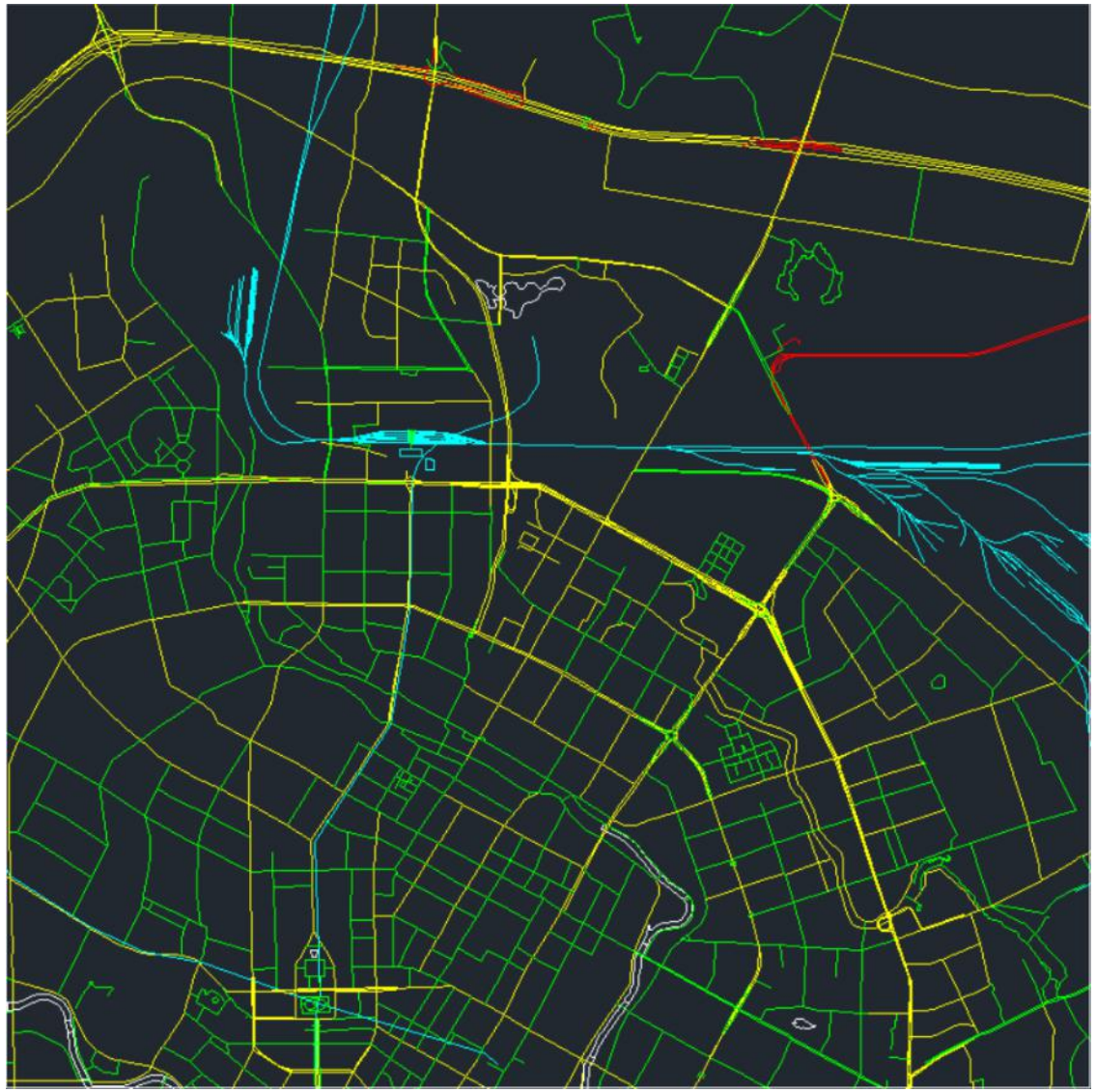
RAM Shortage Problem

- The bottleneck of progress now is the generation of trajectory data.
- I can only plot 1 trajectory picture every time because of the high RAM usage when plotting. Time usage: ranging from 10 min to 2 hours for 1 pic.
- In Kaggle, I can open 10 sessions meanwhile to plot, since plotting only use CPU



Succeeded	False	Run Time	32400.8 seconds
Exit Code	137	Timeout Exceeded	True
Used All Space	False	Output Size	0
Environment	Container Image (Dockerfile)	Accelerator	None
Failure Message	Your notebook was stopped because it exceeded the max allowed execution duration.		

Geographical Data



In mechanical engineering, we use AutoCAD to plot and modify 2-D file (e.g laser cutting graph)

In https://cadmapper.com/#metro website, we can download chengdu.dxf(AutoCAD readable) file.

The file include urban road, expressway, rail, water, etc. By different colors.

How we generate geo map

Open Gaode map (lbs.amap.com)

Record GCj-02 coordinates of 30+ recognizable places

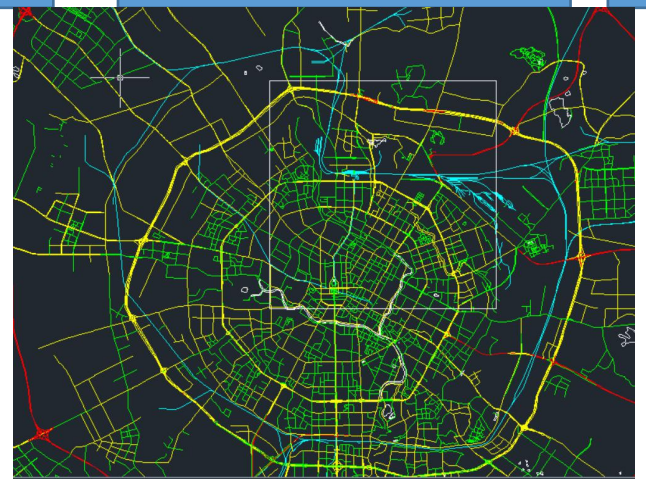
Download chengdu.dxf

Record Autocad coordinates of the same 30+ places

Coordination of two maps using python linear regression

Predict the coordinates of plot boundary in Autocad

Extract the geo map from the boundary in Autocad



Further Plan

- Keep generating the pictures and try to merge pictures in different phase.
- Consider about how to generate TTI output, as a measureable value.
- Enter Phase II, begin to search suitable deep learning neutral network.