An Image Fusion Framework for Deep Learning in Traffic Forecasting

Xiahan LI

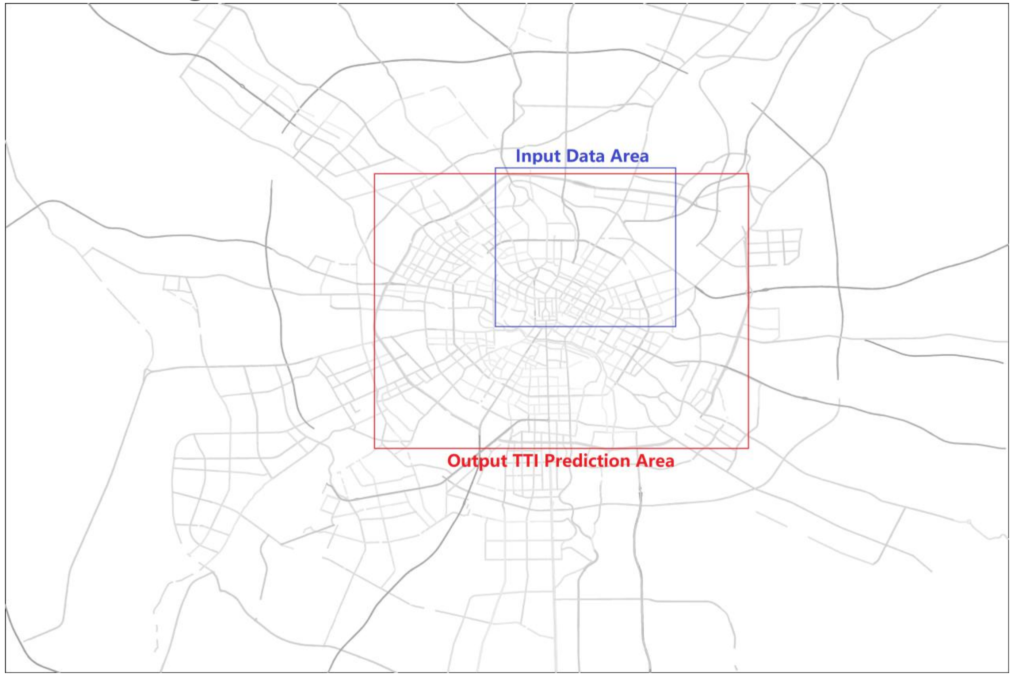
(Simplify Version)

Objectives and Background

- Predict Travel Time Index (TTI) after 10 minutes in 10 individual roads and an specific area in Chengdu City, China during one day (from 00:10:00 November 1st, 2018 to 00:10:00 November 2nd, 2018).

- TTI is an evaluation index of urban congestion degree. For one link, TTI = (Free Flow Speed) / (Actual Speed). For roads, use weight average to calculate.

Position and shape of roads in Chinese Gaode Map



Comparision between Input Data Area and TTI prediction Area

Data Structure

TTI & Average Speed Data Obtained from DIDI GAIA Dataset

- Object ID is the identification data of the road. Which is corresponding to the Object ID in Network Data.

Example: 281931.

- Datetime is the date and time when the travel time index and average speed are recorded.

Example: 2018/10/19 20:00:00.

- Travel Time Index(TTI)   Example: 2.45013.

- Average Speed is the average speed of all the vehicles in the specific road at the specific time.

Example: 14.7419

- 10 minutes time gap for data recording.

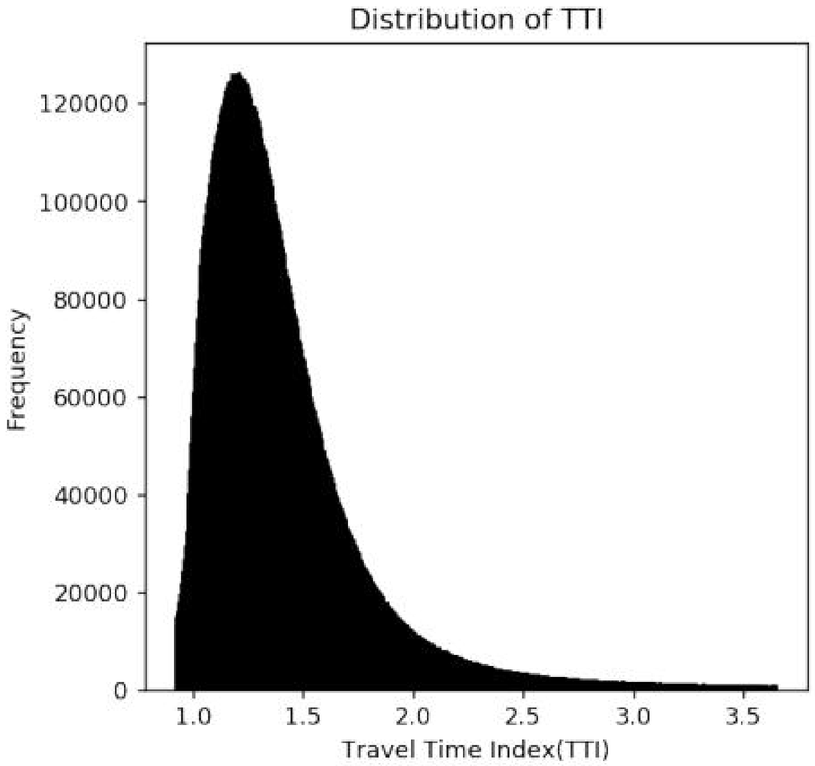
F6

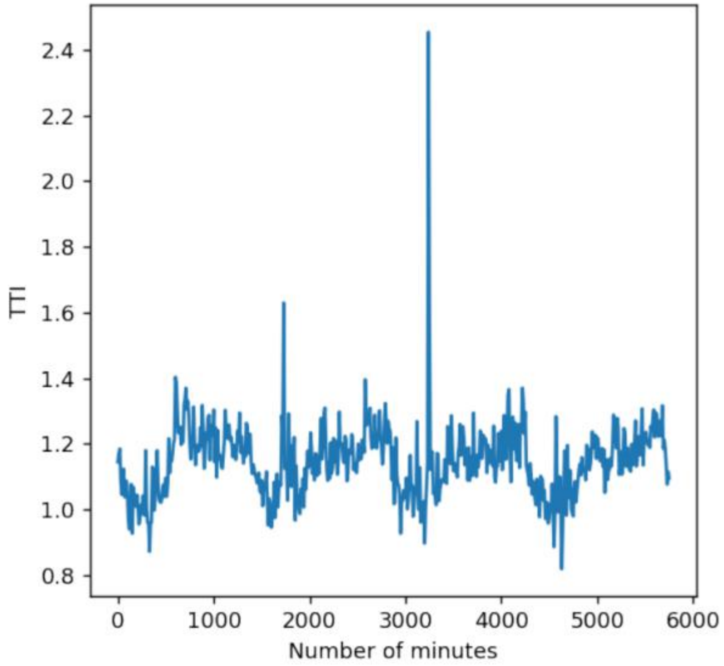
	A	B	C	D
1	Object ID	Datetime	TTI	Avg Speed
2	283404	2018/1/1 0:00	1.39778	39.389
3	283442	2018/1/1 0:00	1.07396	82.3449
4	283002	2018/1/1 0:00	1.09723	27.7525
5	282048	2018/1/1 0:00	1.74019	15.2246
6	283424	2018/1/1 0:00	1.32598	35.5202
7	282590	2018/1/1 0:00	1.23341	33.1325
8	283204	2018/1/1 0:00	1.10675	37.9812
9	283406	2018/1/1 0:00	1.73706	26.4112
10	283444	2018/1/1 0:00	1.11582	43.8828
11	283225	2018/1/1 0:00	1.30119	49.2513
12	283388	2018/1/1 0:00	1.10798	50.8591
13	283426	2018/1/1 0:00	1.1995	46.929
14	282070	2018/1/1 0:00	1.19949	47.0863
15	283408	2018/1/1 0:00	1.27594	36.8665
16	283446	2018/1/1 0:00	1.21089	32.9539
17	282230	2018/1/1 0:00	1.17645	32.1658
18	283390	2018/1/1 0:00	1.11703	55.4279
19	282440	2018/1/1 0:00	1.20862	36.5327
20	283169	2018/1/1 0:00	1.08221	56.8259
21	283428	2018/1/1 0:00	1.12467	40.2594
22	282791	2018/1/1 0:00	1.83667	13.8564
23	283228	2018/1/1 0:00	1.07945	46.6949
24	283410	2018/1/1 0:00	1.24032	42.3101
25	283448	2018/1/1 0:00	1.34308	33.7196
26	283209	2018/1/1 0:00	1.08677	43.3801
27	283392	2018/1/1 0:00	1.1171	40.6179
28	283430	2018/1/1 0:00	1.16427	82.5896
29	282291	2018/1/1 0:00	1.27808	26.4998
30	283412	2018/1/1 0:00	1.18142	42.8327

Statistical Data Description

- Analyze data from 00:00:00 October 1st, 2018 to 23:50:00 November 30nd 2018 (61 days).
- Considering the range is too large, only plot 1% to 99%.
- TTI vs time on road No.283509 from Oct. 1st 2018 to Oct. 3rd 2018.(Quasi Preiodic Pattern)

	Travel Time Index
Minimum	0.131779
Maximum	507.394
Range	507.262221
Mean	1.4474
Median	1.3152
25 percentile	1.15855
75 percentile	1.53686
Variance	0.7944
Skewness	2.10522
Kurtosis	6.49817



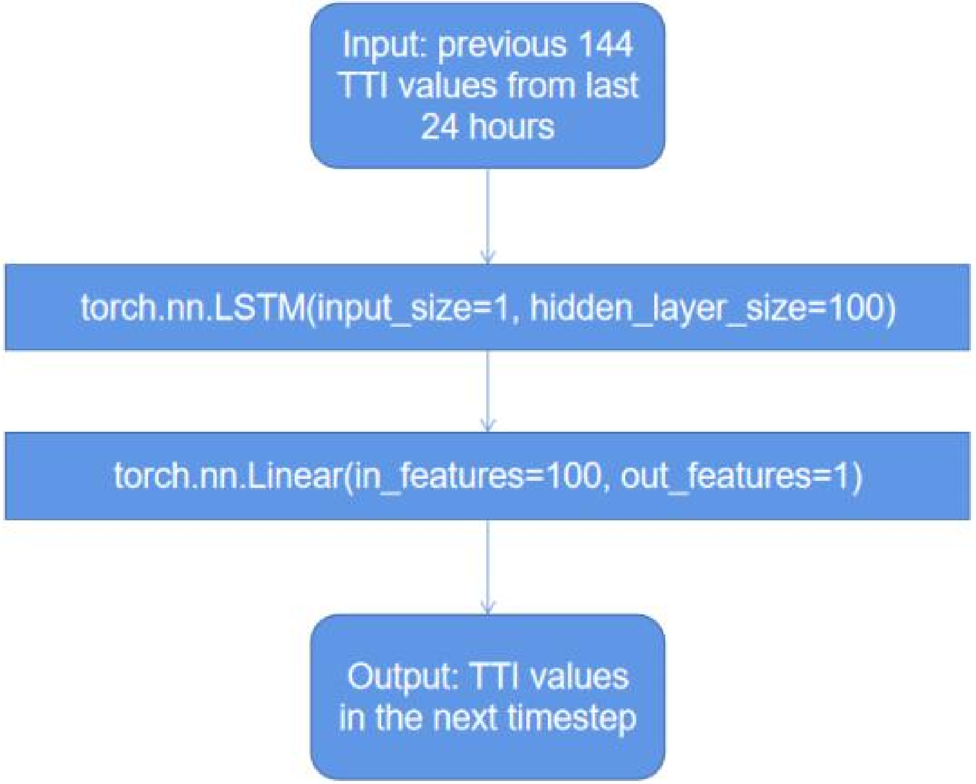


Pure LSTM Network

- Long Short Term Memory (LSTM) is an artificial recurrent neural network that is capable to learn more than 1000 discrete-time steps.

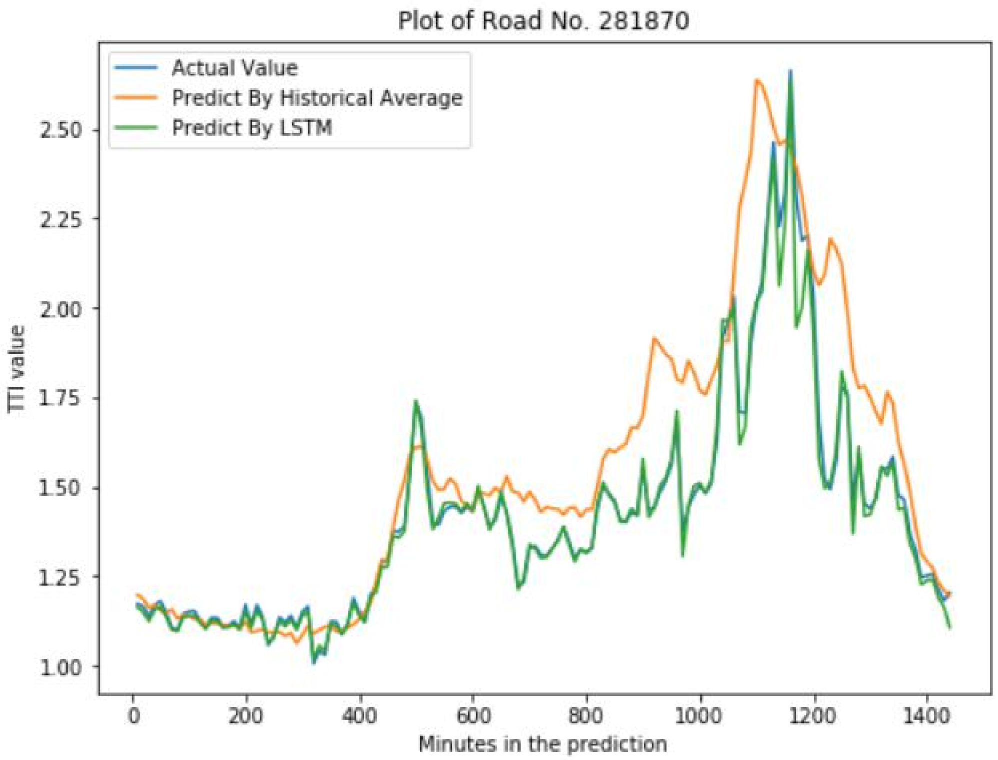
- In this task, the training and validation data is from October 8th to November 1st (24 days). The test data is from November 1st to 2nd (1 day).

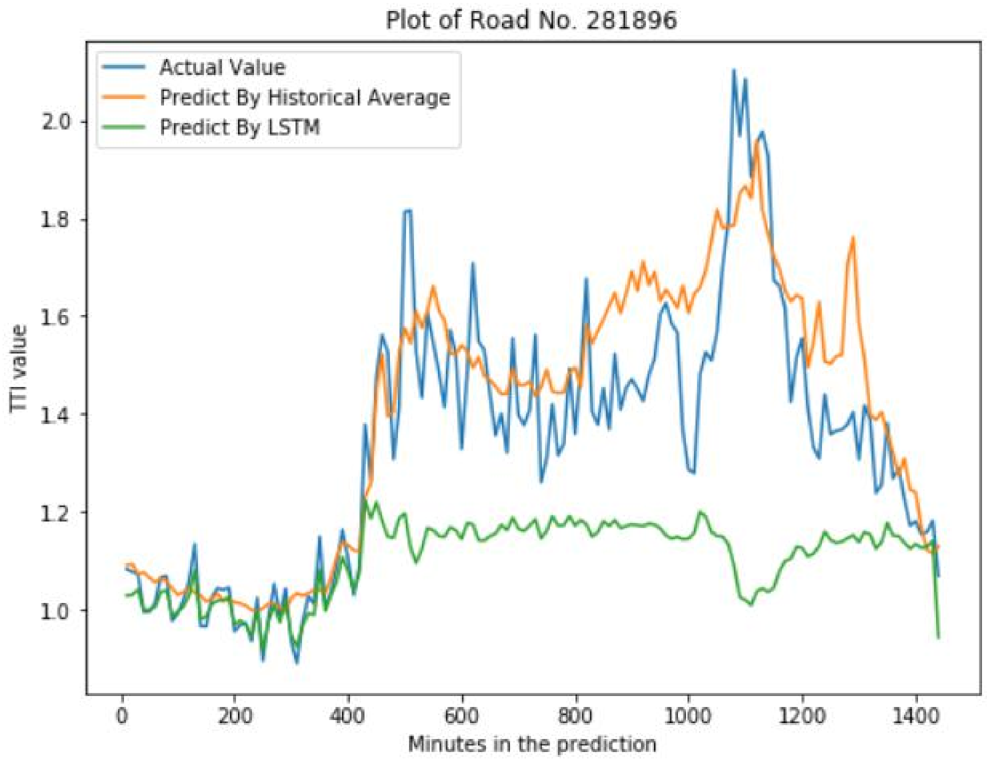
- In LSTM, the inputs are 144 successive historical TTI values (Example: from 00:00:00 to 23:50:00 in October 8th), and the output is the stepped 145th TTI value(Example, 00:00:00 in October 9th).



Results of Pure LSTM

- Take historical average as a baseline, the pure LSTM made excellent prediction on some roads but bad prediction on the other, which is unstable.





Problem of Pure LSTM

1. Performance of LSTM is unstable:

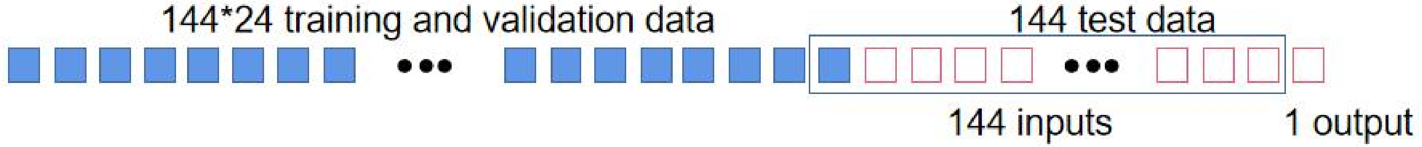
- Inputs of historical TTI values assume that the future is highly related to the history.
- However, alternative situation may affect the future.

Improvement: Expand the inputs, but Pure LSTM with multiple inputs get worse performance

2. Use test data to predict the test data:

- When predicting the last TTI value in the test dataset, the inputs should be the previous 144 TTI values from the last 24 hours, while 143 of them are in the test dataset.
- Cannot use any part of the test dataset to do the prediction of the test dataset.

Improvement: Use the first prediction result as the input of second prediction, but it get worse prediction because of error amplification.



Data Augmentation: Routing Data

- Routing Data Obtained from DIDI GAIA Dataset.
- Driver ID are the desensitization of drivers' personal information.

Example: 4ecea6b8473789e1fdbfea71adfb2451.

Trajectory is consist of 3 parts: longitude, latitude and timestamp in order.

Example: 104.12226 30.67012 1539952150, 104.12226 30.67012 1539952160.

Problem: No connection to any information in TTI & Average Speed Dataset.

	A	B	C	D	E
1	Driverid_1	Driverid_2	Trajectory		
2	4ecea6b8473789e1fdbfea71adfb2451	8ac962c8b85724f4602dbc9e50c3ff46	[104.12226 30.67012 1539952150,		
3	3018c0c10f32ec1aab592168bb6f6787	7535711890e34f9476ad616fba2648e4	[104.05973 30.68469 1539949303,		
4	876853121ec96be317b065c286081976	cf9d01b5bf5de4415c3a56f0c6815b53	[104.09237 30.67939 1539949359,		
5	414b5e790bdc74bc23ce18694ce04909	065bc24b5194fa06b7d87c592c4d9582	[104.09797 30.65295 1539949558,		
6	bb53ed3292cbccc4faca7a20effd221b	f8f1a4a9f0ccc3af0c7a5db7b91e4fc5	[104.04503 30.70138 1539949951,		
7	d284a1b9d8b2a718fe64d78a6e8cb955	9423200264957ffa8cbad6f45936e46a	[104.05008 30.65319 1539949988,		
8	1c27d8f18f793eae9594768b0ed919f9	ce3ba07009895b3ffa567b0c866e9177	[104.06992 30.72481 1539950979,		
9	e178060e8293067c4cb55740afbb3778	8f9e45d0a939d1ee06f9c81cd5b250ad	[104.11174 30.67238 1539951002,		
10	7994b20e1abb70a0a133f0ad8116433a	b2c51f11b3616600be85dd99395c003a	[104.11498 30.65305 1539952444,		
11	bb5964ef941fd68905de4092a8c8e705	cf72f4425102bef43c529f61e55902d3	[104.06933 30.65311 1539953300,		
12	807d819a622a4a73d235d58ee498e3b3	4013d331f7eed7b5a0f3ef7883657523	[104.08998 30.65284 1539952103,		
13	3cdc46fbb2708031cfebcc50d5fab7fc	474753f40157a5755c25c5beb0a49599	[104.11453 30.68045 1539951876,		
14	e218b45a8707e24720821b3900c2fb2c	4691f4ac8245bca69388f65b03374612	[104.10951 30.71557 1539952311,		
15	30dc174e050ab460f37364a1b755f9c8	10295bcca2c8351dfa533cdf700eef5e	[104.10964 30.66423 1539952577,		
16	d7a25408642df8dafd44b9aeeee8e9e0	0edfc263c065eb5760eafaa577b172a2	[104.04893 30.68223 1539952517,		
17	9079d767649fe104d44cb1159fefc311	d13b2cfca7ddef86d2095b0777141b0e	[104.05231 30.66831 1539952899,		
18	7b11248aa8952e726e7564d54eff5188	a9246255ad24b1b4386755cc4917dc15	[104.0424 30.66884 1539952961, 1		
19	89ef8309454ef68e319e573089298dfe	a2302a696bc6050515fcdd3894625765	[104.12446 30.67114 1539952766,		
20	a8f12cf7cca16ac70a93413683a62dad	8964abb4eaca957fad1a0c257e152288	[104.10349 30.66936 1539953744,		
21	22ee13c2cf38b5b2ff29699328b89d87	8a831d24361f8b1fb1a244b90cb1362b	[104.05783 30.65886 1539953930,		
22	37daa34263e40b51c224ac4f707e3802	28b477cb385de9bfdc60d462ffc2b9fa	[104.07864 30.65286 1539949911,		
23	a420751e6b5fc2d3903cb4c313282d16	6add8ca276335722bac4e59193113fa2	[104.06312 30.65285 1539952369,		
24	3eb218753dbb0878aa3f51ab1a7fabe3	49de7258e2d8959a4285cda1bee9890c	[104.04819 30.67848 1539950840,		
25	870ad27dad4b4411f4639da06564b9e3	688c1811d07f79b2784eece1183629d7	[104.09933 30.68228 1539950641,		
26	46bdf54e2d209e26f2a9398437470975	39cf7dfa539b98799dc3d1a145f78887	[104.05041 30.67871 1539951229,		
27	055c3538f23f7deaf6d354750d2583ba	04a60b77f0c23f018063467e7d54aca9	[104.08714 30.65401 1539951898,		
28	79c51310de9b407b144122d4dd7994c4	d7c39d8f269e1ff1b1acf794fa858d51	[104.04226 30.66525 1539953684,		
29	fbd9bc32eaf08aacc10ca0811be171a7	43cb3d50e22d2c7637f00426101f6b4e	[104.09839 30.70254 1539951950,		
30	80b2871fc09f717e32aff5c3d4565da7	89baecf30a96389ca54f31eeeab085e1	[104.04944 30.66565 1539952066,		

Data Augmentation: Network Data

Network data is consist of Object ID, Object Name, and WKT Geometry.

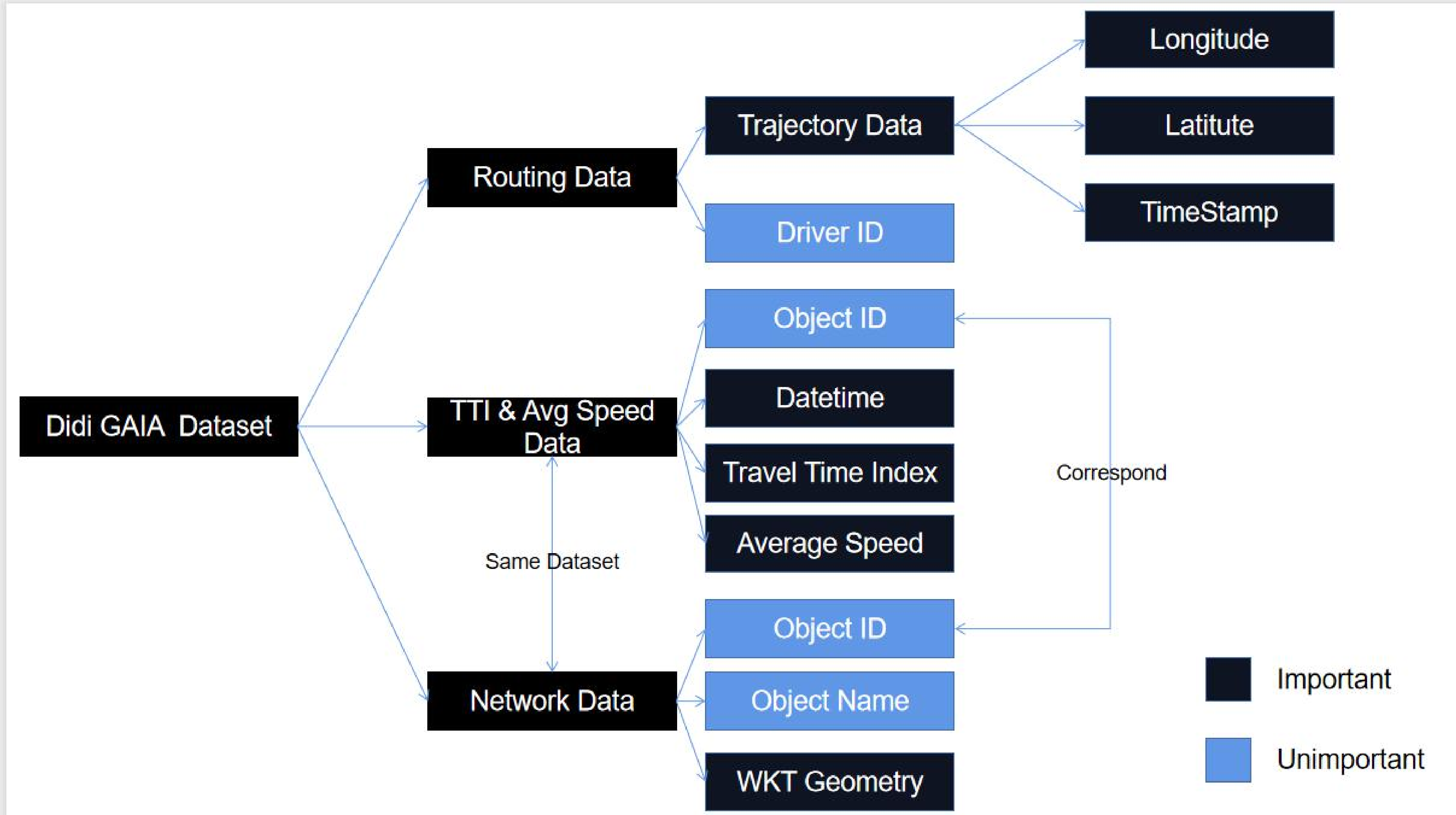
- Object ID is the identification data of the road. Example: 281931. Corresponding to the Object ID in TTI & Average Speed Data.
- Object Name is the Chinese name of the road, whose detailed information can be further investigated on Maps.
- WKT Geometry can be used to coordinates to represent various geometric objects like points, linestrings, polygons, networks and geometry collections, etc.

Example: LINESTRING(104.10583 30.67989, 104.10649 30.67871)

Problem: Difficult to mapping positional relationship

	A	B	C	D	E	F	G	H
1	obj_id	obj_name	geom					
2	281863	八里桥路:北站东二路,三环路	MULTILINESTRING((104.07437 30.71442,104.0742 30.71483,1					
3	281864	八里桥路:三环路,北站东二路	MULTILINESTRING((104.07452 30.71324,104.0745 30.71338,1					
4	281865	二环路西段:广福路,金牛大道	MULTILINESTRING((104.02746 30.64001,104.02707 30.64039),					
5	281866	二环路西段:金牛大道,广福路	MULTILINESTRING((104.02536 30.64168,104.02488 30.64216,1					
6	281867	二环路南段:郭家桥西街,广福路	MULTILINESTRING((104.079 30.62056,104.07975 30.62059),(1					
7	281868	二环路南段:广福路,郭家桥西街	MULTILINESTRING((104.07771 30.62034,104.0773 30.62033,1					
8	281869	二环路东段:桃蹊路,郭家桥西街	MULTILINESTRING((104.10583 30.67989,104.10649 30.67871),					
9	281870	二环路东段:郭家桥西街,桃蹊路	MULTILINESTRING((104.10843 30.67562,104.10846 30.67555,1					
10	281871	二环路北段:金牛大道,桃蹊路	MULTILINESTRING((104.03997 30.69011,104.04119 30.69117),					
11	281872	二环路北段:桃蹊路,金牛大道	MULTILINESTRING((104.04793 30.69457,104.04809 30.69458,1					
12	281873	一环路北段:花牌坊街,建设北路	MULTILINESTRING((104.04799 30.68351,104.0478399 30.6834					
13	281874	一环路北段:建设北路,花牌坊街	MULTILINESTRING((104.05008 30.68407,104.05036 30.68413),					
14	281875	一环路东段:建设北路,红瓦寺街	MULTILINESTRING((104.09827 30.67224,104.09829 30.67221),					
15	281876	一环路东段:红瓦寺街,建设北路	MULTILINESTRING((104.09979 30.6688,104.10007 30.66811),					
16	281877	一环路西段:燃灯寺东街,花牌坊街	MULTILINESTRING((104.04134 30.64506,104.04135 30.6447),					
17	281878	一环路西段:花牌坊街,燃灯寺东街	MULTILINESTRING((104.041 30.65171,104.04102 30.65147),(1					
18	281879	一环路南段:红瓦寺街,燃灯寺东街	MULTILINESTRING((104.08273 30.63689,104.08254 30.6368),					
19	281880	一环路南段:燃灯寺东街,红瓦寺街	MULTILINESTRING((104.07628 30.63372,104.07686 30.63398),					
20	281881	高攀路:中环路火车南站东路段,二环路	MULTILINESTRING((104.08087 30.60913,104.08069 30.60853),					
21	281882	高攀路:二环路,中环路火车南站东路段	MULTILINESTRING((104.08155 30.6121,104.08128 30.61099),					
22	281883	三官堂街:二环路,龙舟路	MULTILINESTRING((104.09415 30.62572,104.09415 30.62552),					
23	281884	三官堂街:龙舟路,二环路	MULTILINESTRING((104.09414 30.62543,104.09426 30.62537),					
24	281885	建设北路:府青路,中环路二仙桥东路段	MULTILINESTRING((104.09077 30.67054,104.09038 30.67077),					
25	281886	建设北路:中环路二仙桥东路段,府青路	MULTILINESTRING((104.09547 30.67143,104.0953 30.67143,1					
26	281887	东城根南街:西御街,蜀都大道	MULTILINESTRING((104.06081 30.6564,104.0608 30.65656),(1					
27	281888	东城根南街:蜀都大道,西御街	MULTILINESTRING((104.06061 30.65744,104.0606 30.65767,1					
28	281889	东城根街:蜀都大道,八宝街	MULTILINESTRING((104.06078 30.65831,104.06077 30.65895),					
29	281890	东城根街:八宝街,蜀都大道	MULTILINESTRING((104.06056 30.66187,104.06057 30.6612),					
30	281891	万和路:八宝街,新华大道	MULTILINESTRING((104.06035 30.67213,104.06035 30.67197,1					

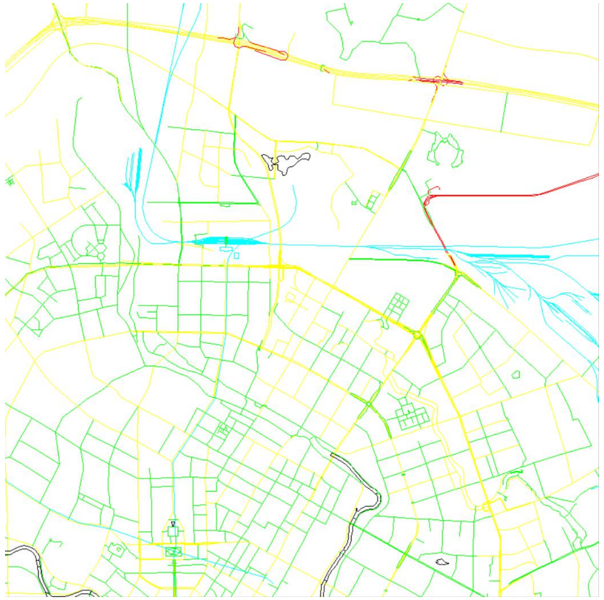
Data Structure after Augmentation



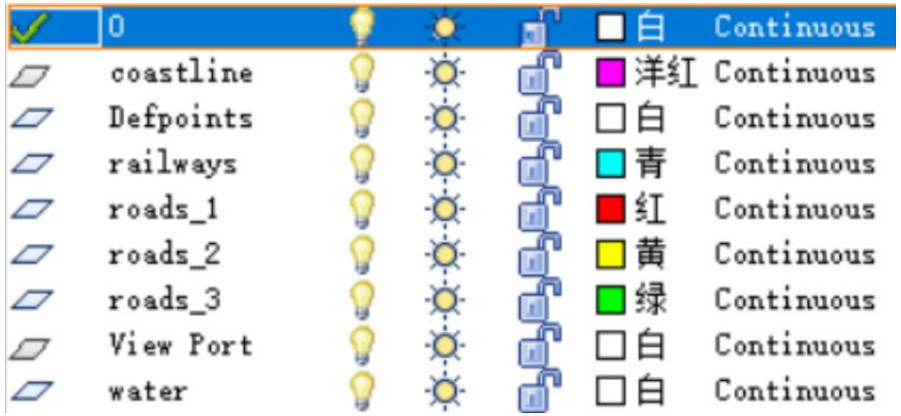
Innovative Image Fusion Framework

Geographical Image

- The geographical 2D CAD picture of Chengdu City is downloaded from a free opensource website called Cadmapper.
- The geographical image mainly categorizes transport network by different colors.
- During data fusion, the colored image will be transformed to gray image.



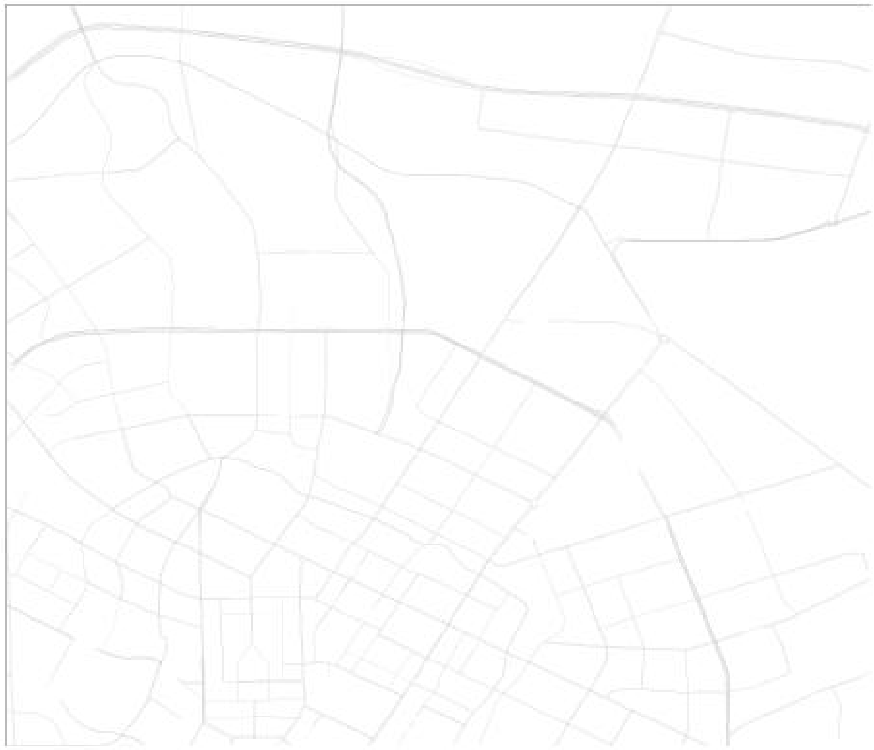
CAD picture of the input data area



Labels of the CAD graph

Roadmap Image

- Use the longitude and latitude values of roads in Network Data to plot the network.
- Use the Average Speed of roads in TTI & Average Speed data to decide the pixel. Value of pixel = 255 - integer of speed.

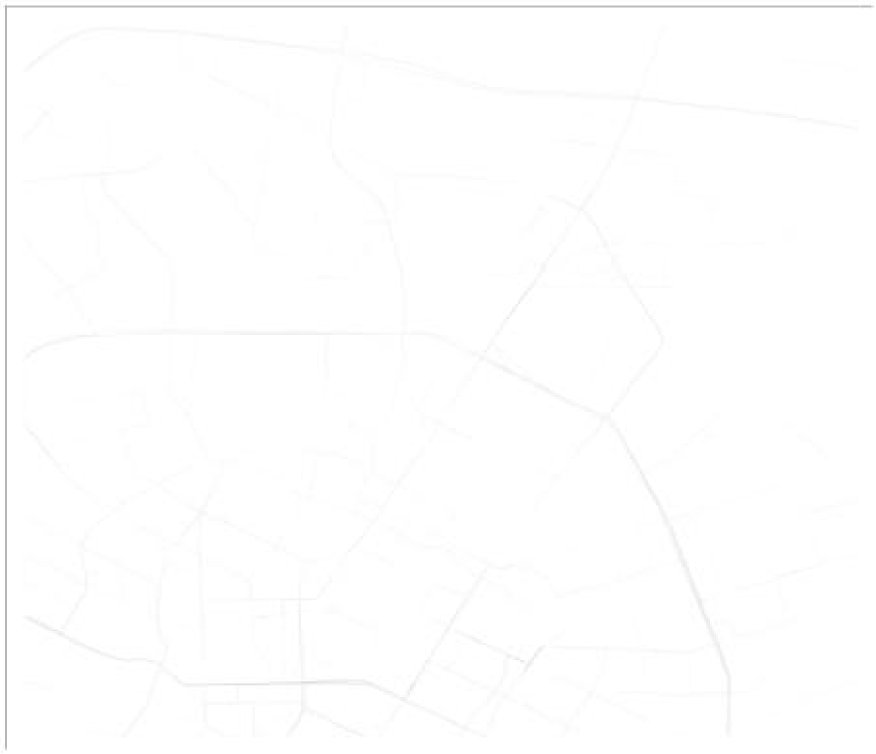


Roadmap image of November 8th, 2018 at 03:00:00

Roadmap image of November 8th, 2018 at 09:00:00

Trajectory Image

- Use the longitude and latitude values of roads in Routing Data to plot the network.
- Compute the speed of roads by timestamp, longitude and latitude values. Value of pixel = 255 - integer of speed.

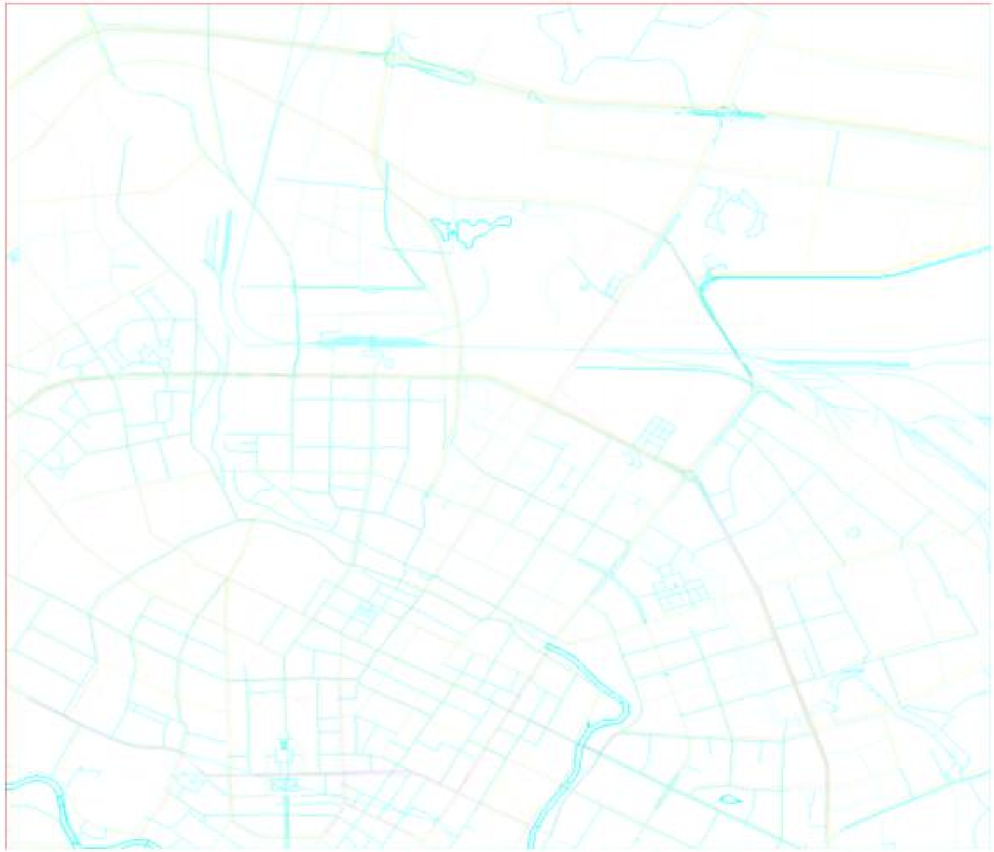


Trajectory image of November 8th, 2018 at 03:00:00,

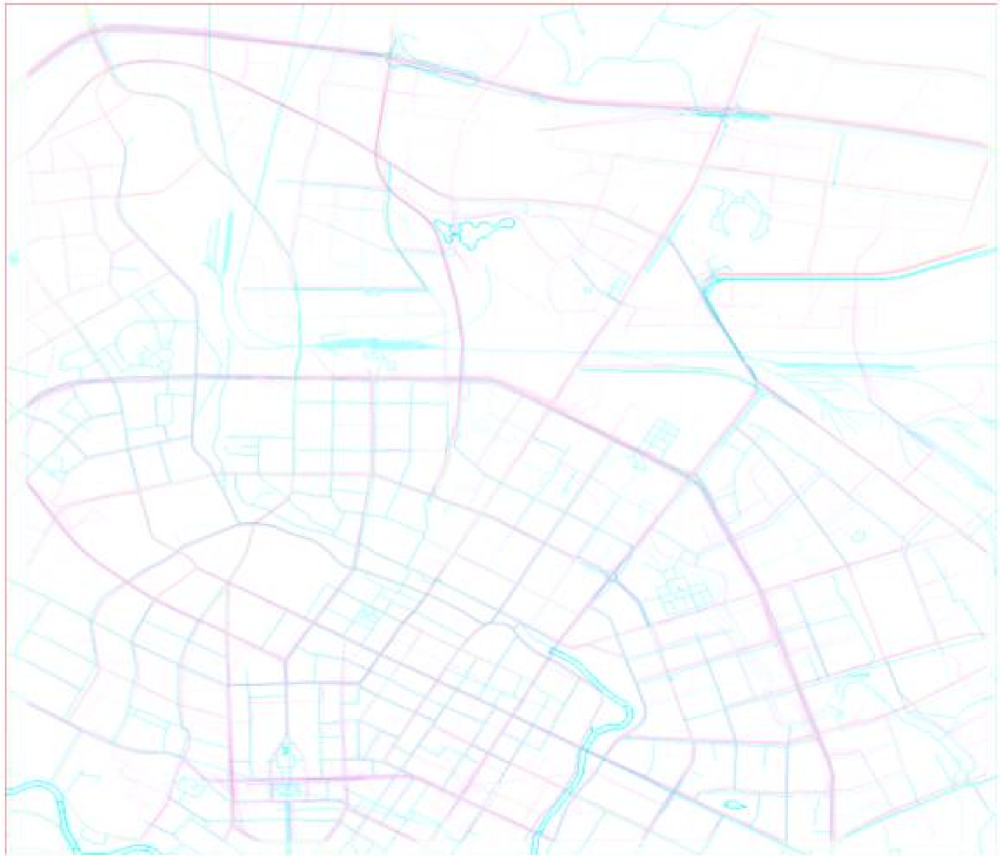
Trajectory image of November 8th, 2018 at 09:00:00,

Image Fusion

- Classify the three kinds of images according to the same time.
- Fuse the 3 phases using the image fusion framework.



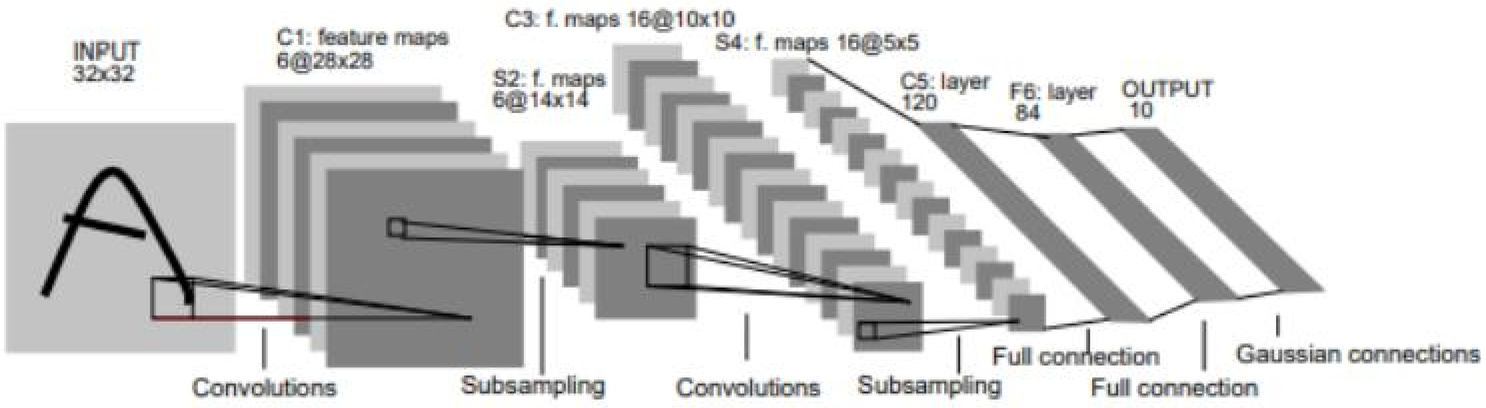
Full 3 RGB image of November 8th, 2018 at 03:00:00



Full 3 RGB image of November 8th, 2018 at 09:00:00

CNN and Resnet

- Convolutional neural networks (CNN) are used in pattern recognition.
- Deep Residual Neural Networks (Resnet) is a kind of CNN that are created to against degradation of training accuracy.
- Considering both the computation cost and model accuracy, we use Resnet34.



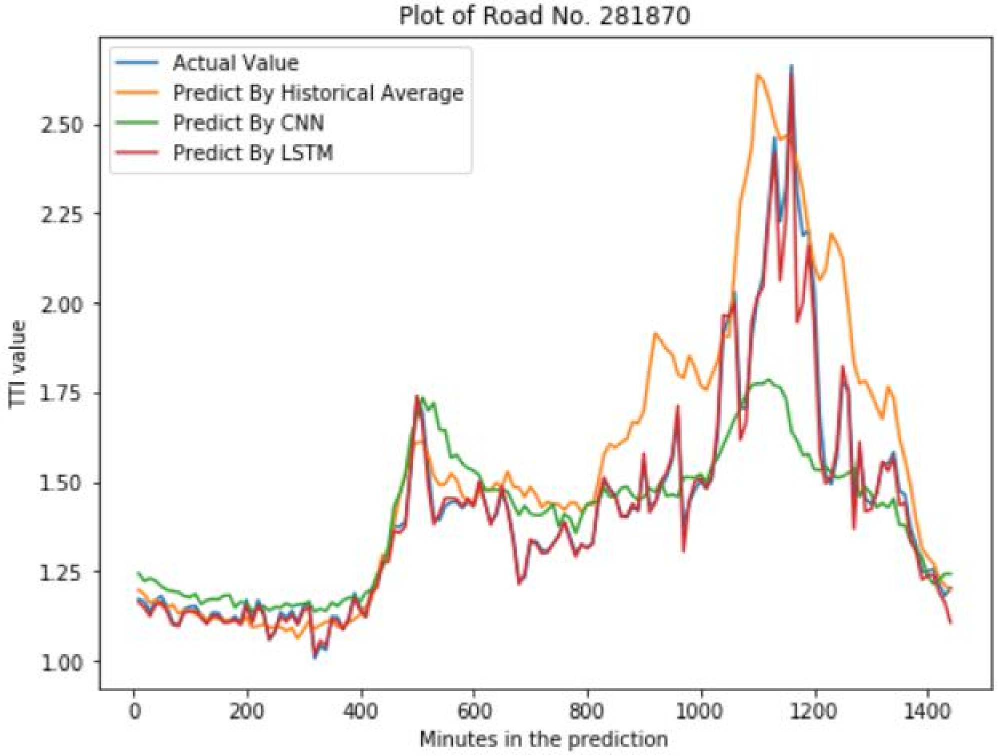
An example of digits recognition by using convolutional neural network

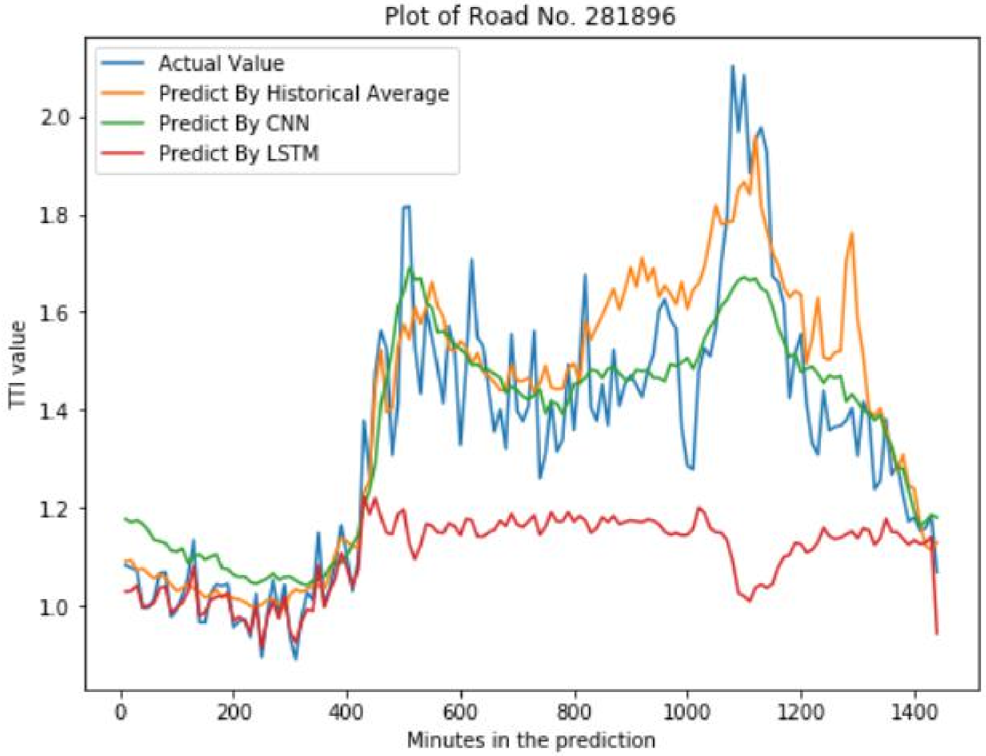
layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Resnets with different layers

Result of Resnet with Image Fusion

- Take historical average as a baseline, and compare to the LSTM model, Resnet with image fusion have a stable results. Besides, Resnet with image fusion can solve all the problems of LSTM model that mentioned before.





Results in Detail

- Set the loss of historical average model as 100%
- Resnet34: 91.1%.
- LSTM model 102.8%.

Resnet with Image fusion can improve the accuracy by 9%.

Single LSTM reaches nearly similar accuracy with baseline and encounter a series of problems.

Road ID	Historical Average	Resnet34	LSTM (Single Input)	LSTM (Multiple Input)	LSTM
281870	0.13751	0.11284	0.23952	0.69540	0.02241
281874	0.07473	0.06299	0.22158	0.58185	0.04958
281876	0.08684	0.04913	0.18057	0.48999	0.03003
281885	0.06096	0.06370	0.37082	0.60964	0.03108
281896	0.10174	0.09482	0.28598	0.61231	0.24025
282014	0.06234	0.05960	0.40116	0.53485	0.11554
282228	0.17517	0.17066	0.51937	0.80210	0.42007
282242	0.11649	0.11439	0.48036	0.72747	0.14938
282252	0.10599	0.10186	0.15552	0.75103	0.06292
282272	0.19019	0.19143	0.38607	0.73633	0.05428
Area	0.04659	0.03400	0.16136	0.25482	0.15864
Total	0.10532	0.09595	0.30930	0.61780	0.1083

Table 5.1.: Average Loss Comparison of Different Models

Thanks!

Q & A