An Image Fusion Framework for Deep Learning in Traffic Forecasting

Xiahan LI

Technische Universität München

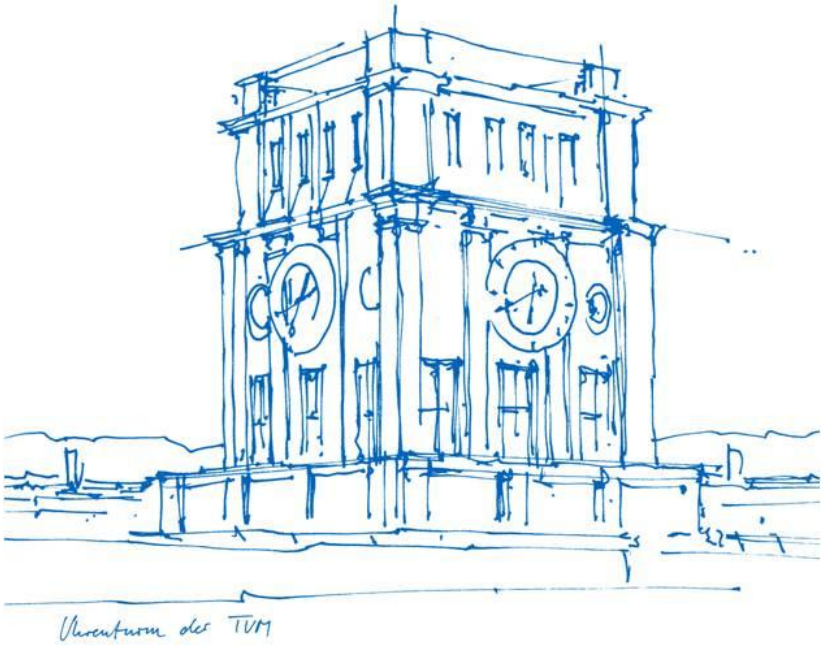
Master program in Rail, Transport and Logistics

Munich, 7th June 2021

Mentoring: Prof. Dr. Constantinos Antoniou

Dr. Tao Ma

Mr. Vishal Mahajan



State of Art in Traffic Forecasting

- Naïve Method

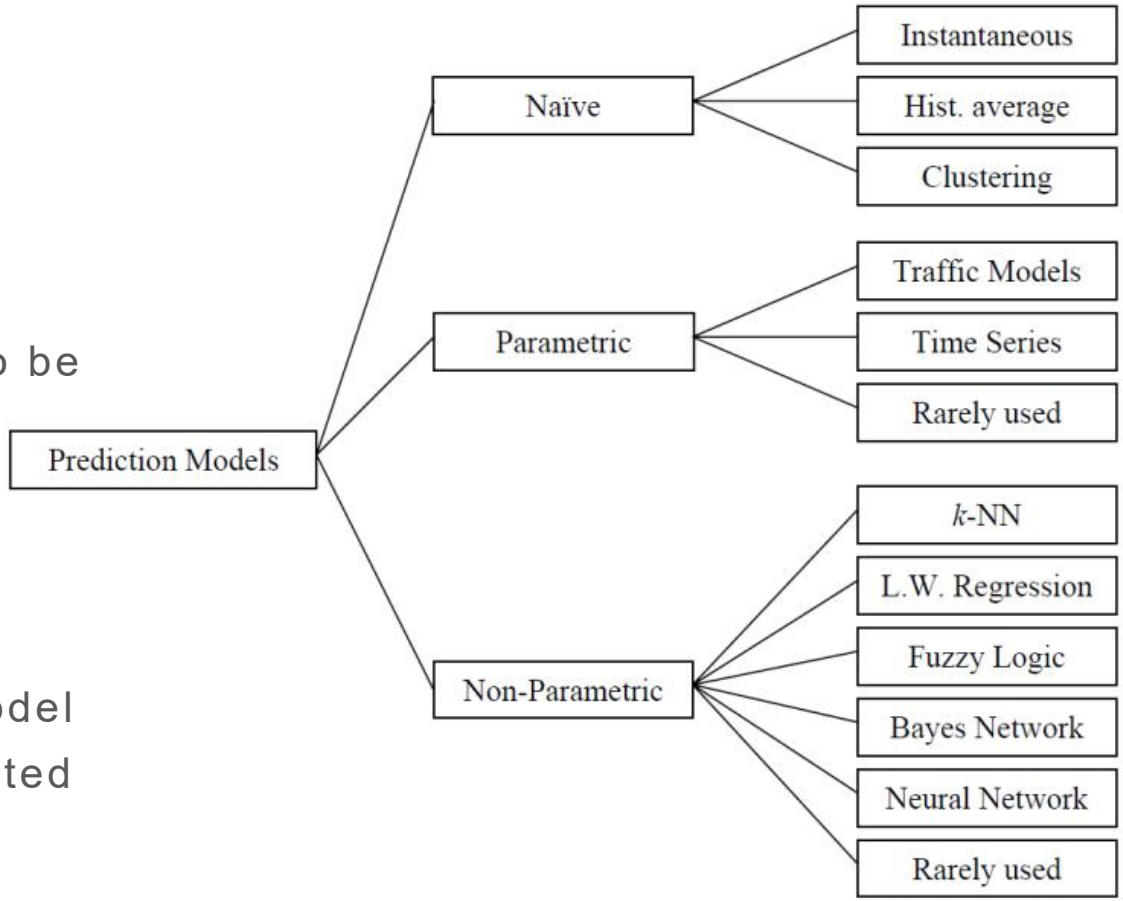
Without any model assumption

- Parametric Model

Parameters of model needed to be estimated.

- Non-Parametric Model

Both model parameters and model structures needed to be estimated



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.
- 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

Comparision between Input Data Area and TTI prediction Area

Data Structure

TTI & Average Speed Data Obtained from DIDI GAIA Dataset. Time interval: 10 minutes.

- 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.

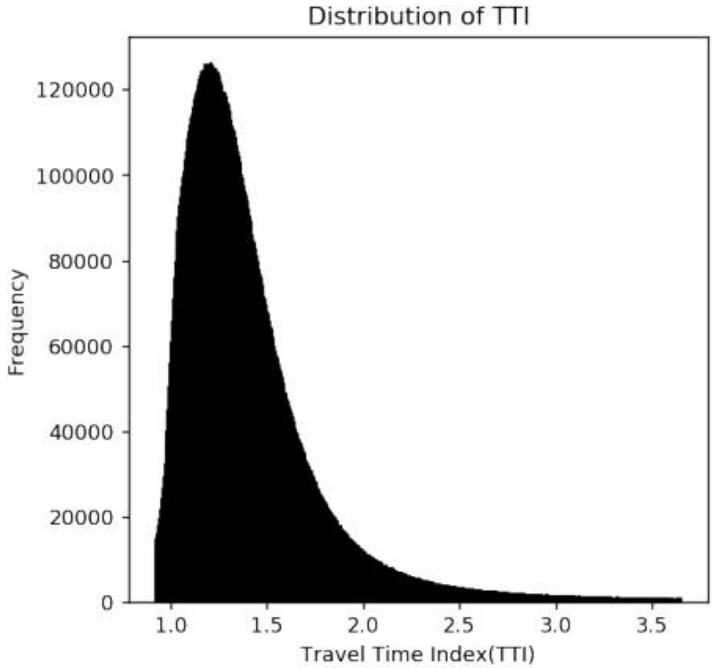
- Travel Time Index(TTI) Example: 2.45013.
- Average Speed is the average speed of all the vehicles in the specific road at the specific time. Example: 14.7419

Object ID	Datetime	TTI	Avg Speed
283404	2018/1/1 0:00	1.39778	39.389
283442	2018/1/1 0:00	1.07396	82.3449
283002	2018/1/1 0:00	1.09723	27.7525
282048	2018/1/1 0:00	1.74019	15.2246
283424	2018/1/1 0:00	1.32598	35.5202

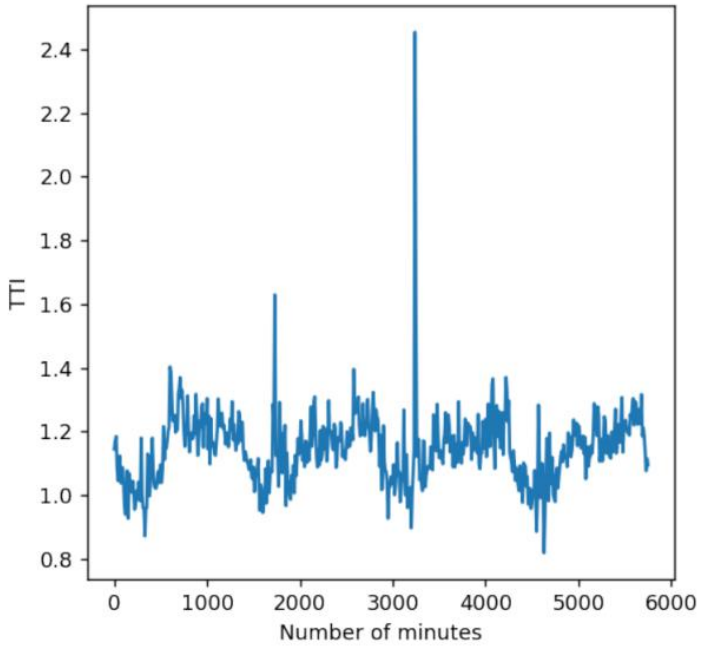
An Example of TTI & Average Speed Data

Statistical Data Description

- Analyze data from 00:00:00 October 1st, 2018 to 23:50:00 November 30nd 2018 (61 days).
- To eliminate abnormalties, plot the Distribution of TTI in range of 1% to 99%.
- TTI vs time on road No.283509 from Oct. 1st 2018 to Oct. 3rd 2018. (Quasi Preiodic)



Distribution of TTI in range of 1% to 99%



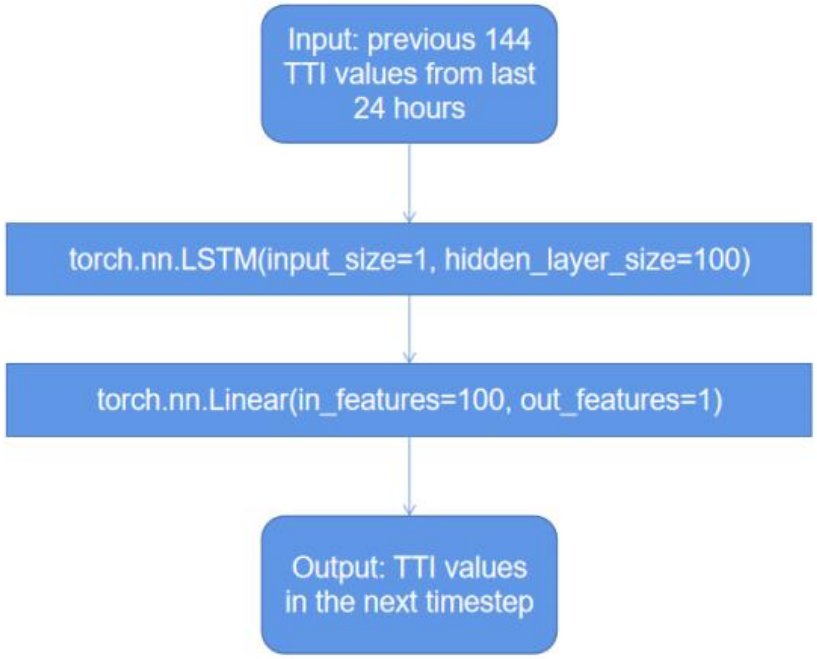
Plot of TTI vs time on road No.283509

Structure of Pure LSTM Network

- Test dataset: From 00:10:00 Nov. 1st, 2018 to 00:10:00 Nov. 2nd, 2018.(1 day)
- Traning and validation dataset: From 00:10:00 Oct. 8th, 2018 to Nov. 1st, 2018.(24 days)

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

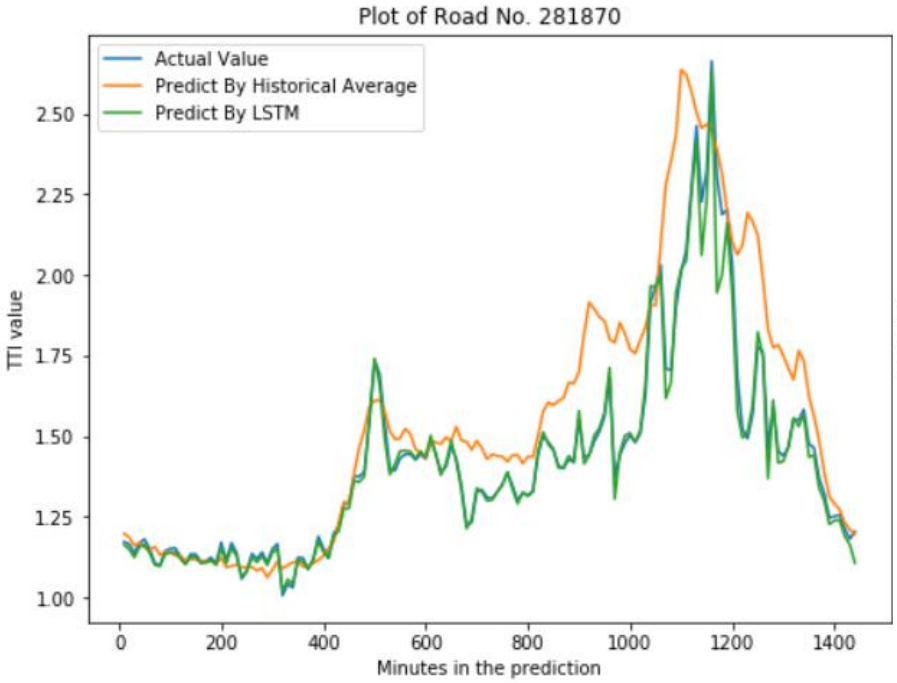
- In LSTM, the inputs are 144 successive historical TTI values (Example: from 00:00:00 to 23:50:00 in October 8th), and the output is the stepped 145th TTI value(Example, 00:00:00 in October 9th).

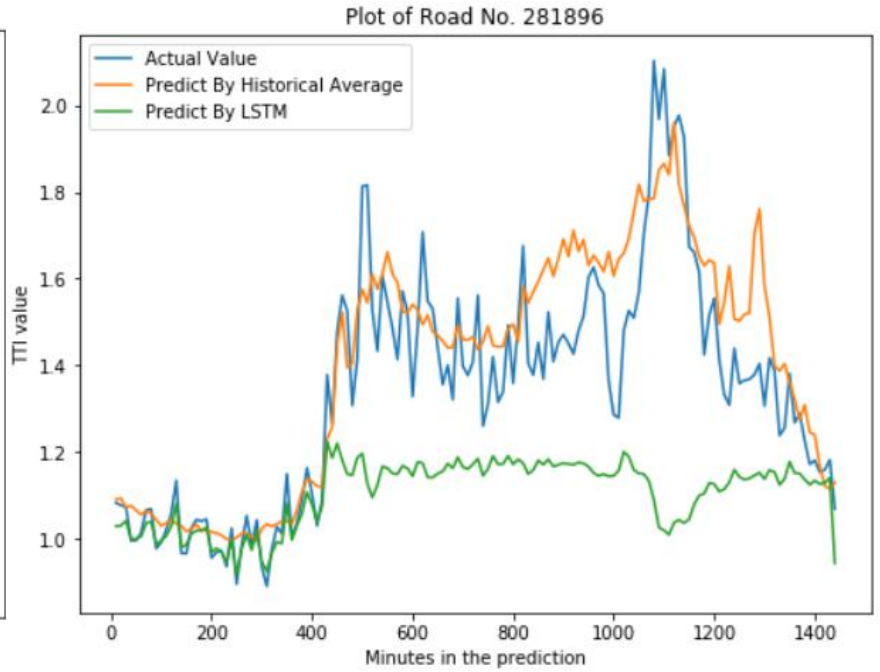


Structure of Pure LSTM Network

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.

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.

We can not use any part of the test dataset to do the prediction of the test dataset

Data Augmentation: Routing Data

Routing Data Obtained from DIDI GAIA Dataset. Time interval: several seconds.

- 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.

Driverid_1	Driverid_2	Trajectory
4ecea6b8473789e1fdbfea71adfb2451	8ac962c8b85724f4602dbc9e50c3ff46	[104.12226 30.67012 1539952150,
3018c0c10f32ec1aab592168bb6f6787	7535711890e34f9476ad616fba2648e4	[104.05973 30.68469 1539949303,
876853121ec96be317b065c286081976	cf9d01b5bf5de4415c3a56f0c6815b53	[104.09237 30.67939 1539949359,
414b5e790bdc74bc23ce18694ce04909	065bc24b5194fa06b7d87c592c4d9582	[104.09797 30.65295 1539949558,
bb53ed3292cbccc4faca7a20effd221b	f8f1a4a9f0ccc3af0c7a5db7b91e4fc5	[104.04503 30.70138 1539949951,

An Example of Routing Data

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.
- WKT Geometry is a representation of geometrics.

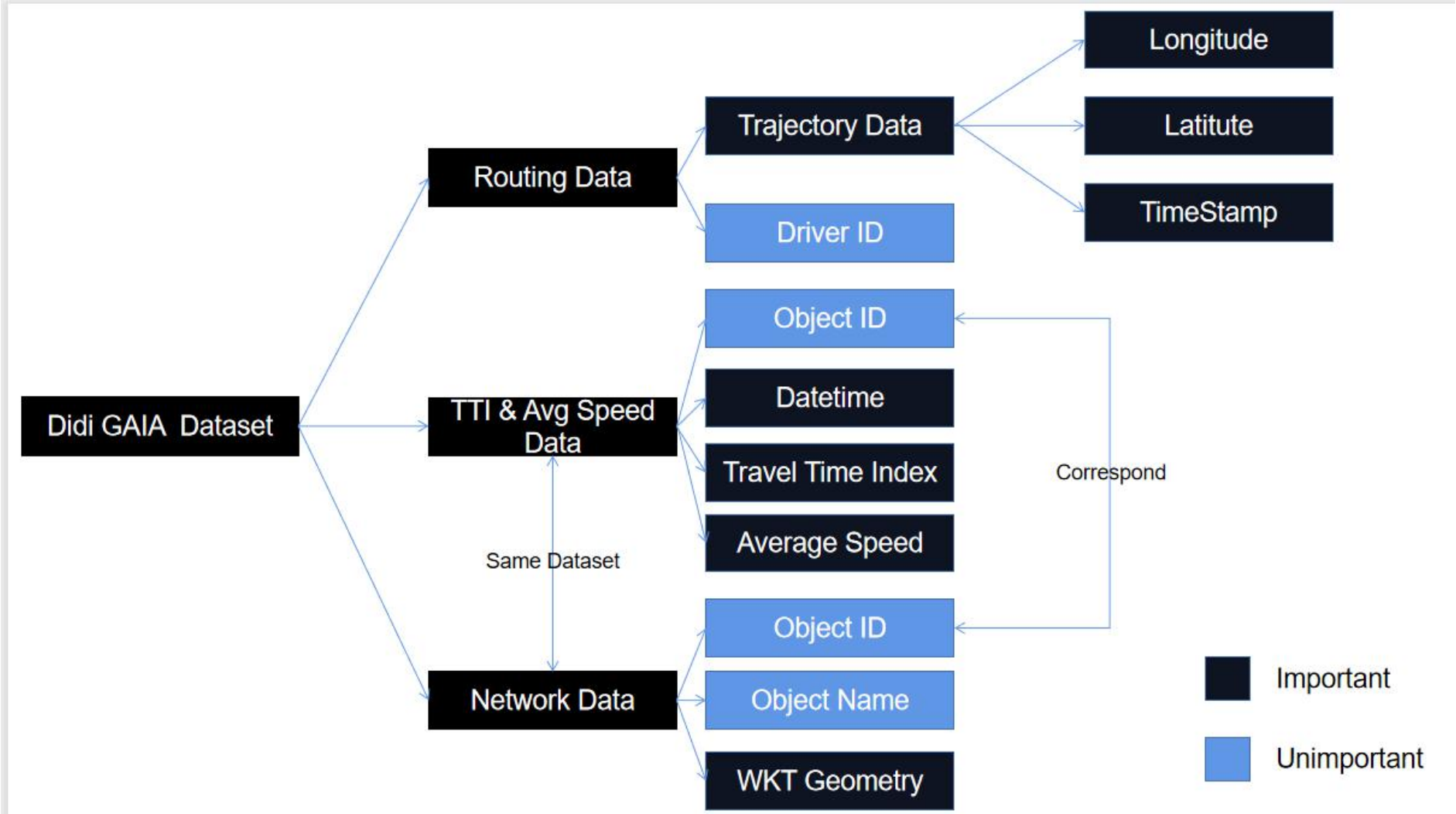
Example: LINESTRING(104.10583 30.67989, 104.10649 30.67871)

- Problem: Difficult to mapping positional relationship to TTI & Average Speed Dataset

obj_id	obj_name	geom
281863	八里桥路:北站东二路,三环路	MULTILINESTRING((104.07437 30.71442,104.0742 30.71483,104.07
281864	八里桥路:三环路,北站东二路	MULTILINESTRING((104.07452 30.71324,104.0745 30.71338,104.07
281865	二环路西段:广福路,金牛大道	MULTILINESTRING((104.02746 30.64001,104.02707 30.64039),(104
281866	二环路西段:金牛大道,广福路	MULTILINESTRING((104.02536 30.64168,104.02488 30.64216,104.0
281867	二环路南段:郭家桥西街,广福路	MULTILINESTRING((104.079 30.62056,104.07975 30.62059),(104.0

An Example of Network Data

Relationship of Augmented Multiple Datasets



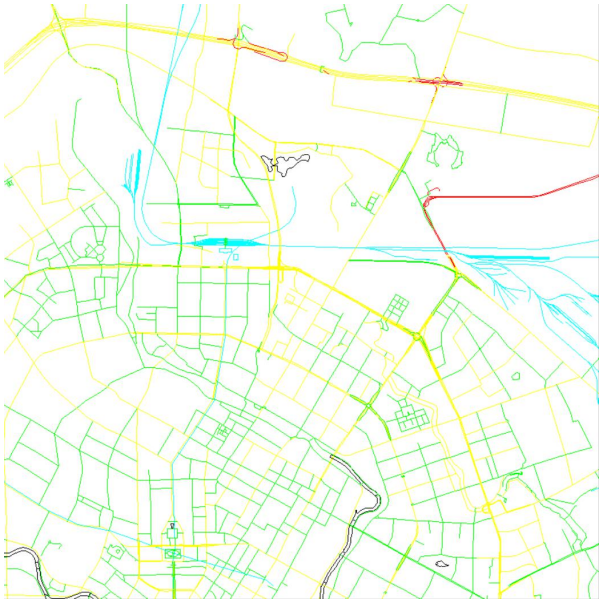
Innovative Image Fusion Framework

Geographical Image

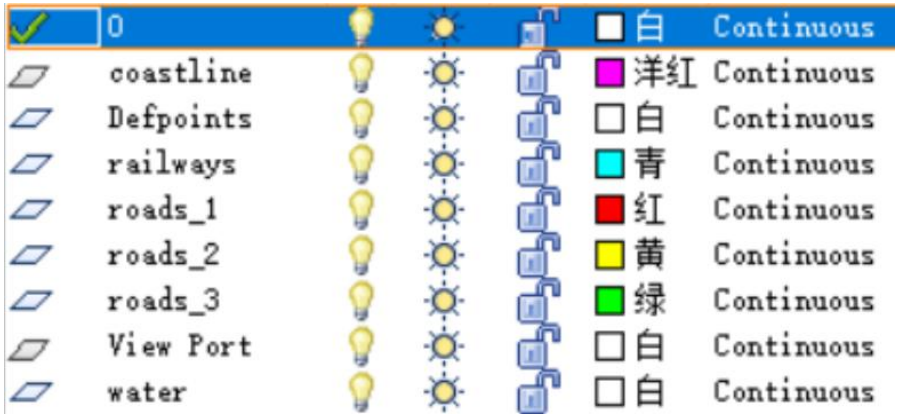
The geographical 2D CAD picture of Chengdu City is downloaded from a free opensource website called Cadmapper.

The geographical image mainly categorizes transport network by different colors.

The geographical image represent for the geographical information of the road network.



CAD picture of the input data area



Labels of the CAD graph

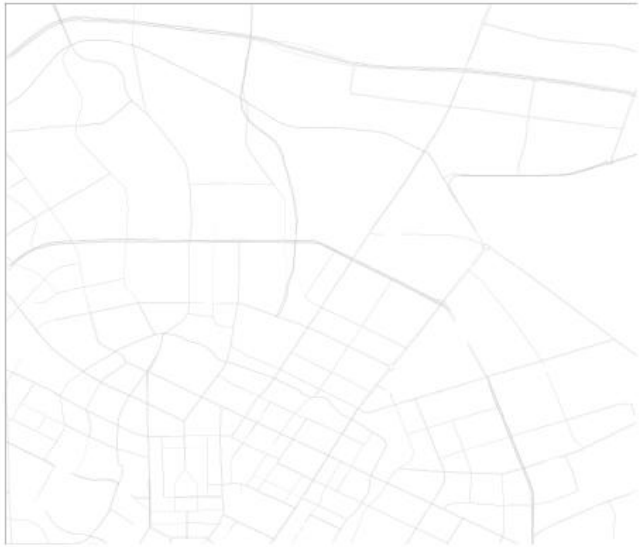
Roadmap Image

Use the longitude and latitude values in WKT Geometry in Network Data to plot the road network.

Use the Average Speed of roads in TTI & Average Speed data to decide the pixel.

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, etc.



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, and use the speed to decide the pixel.

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

One trajectory picture represent for:

- The trajectory of the drivers within +- 2min time interval.
- The speed of the drivers within +- 2min time interval.

Above information collected within +- 2min for every 10 minutes, e.g.:

19:58:00 - 20:02:00 represent for 20:00

20:08:00 - 20:12:00 represent for 20:10

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

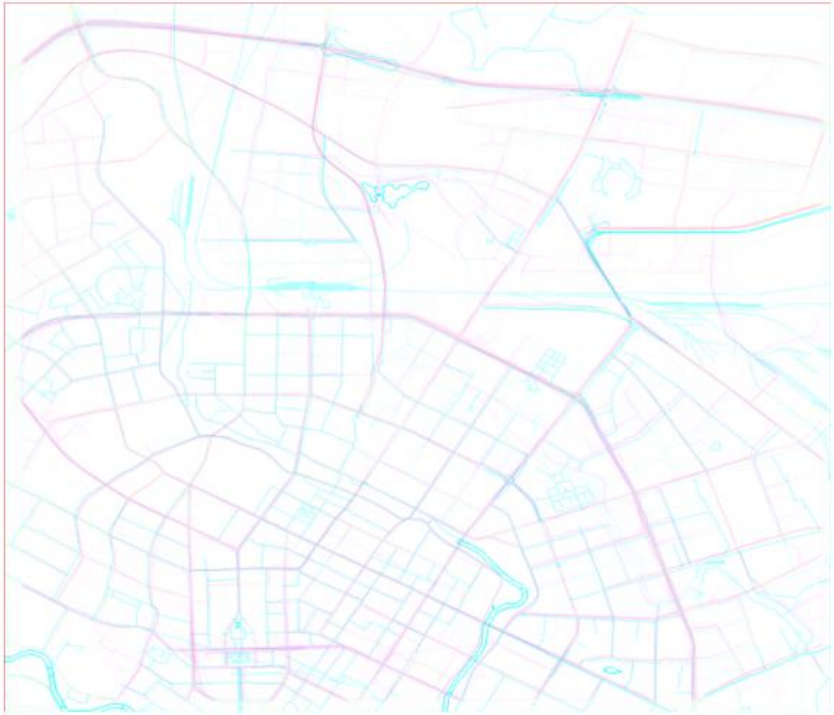
Image Fusion

Fuse the 3 kinds of images according to the same time using the image fusion framework.

Time interval of adjacent images is 10 minutes. 144 images in one day.

One image represent for the traffic operation procedure at every 10 minutes.

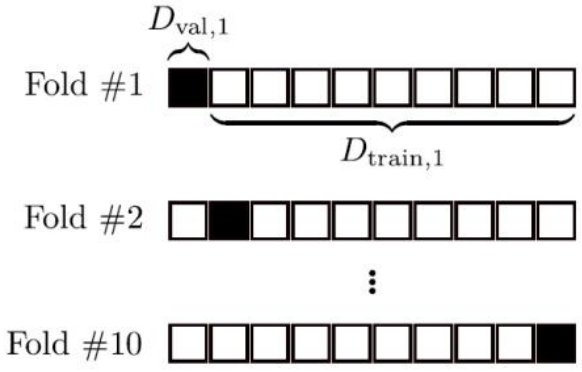
Full 3 RGB image of November 8th, 2018 at 03:00:00



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

K-Fold Cross-Validation

- Test dataset: From 00:10:00 Nov. 1st, 2018 to 00:10:00 Nov. 2nd, 2018.(1 day)
- Traning and validation dataset: From 00:10:00 Oct. 8th, 2018 to 00:10:00 Nov. 1st, 2018.(24 days)
- In 6-fold cross-validation, train 6 models seperately and fuse the results by averaging the performances of 6 models.



Graphical illustration in K-fold cross-validation

Fold	Training Dataset	Validation Dataset
1	Oct. 8th - Oct. 27th	Oct. 27th - Nov. 1st
2	Oct. 8th - Oct. 23rd Oct.27th - Nov. 1st	Oct. 23rd - Nov. 27nd
3	Oct. 8th - Oct. 19th Oct.23rd - Nov. 1st	Oct. 19th - Oct. 23rd
4	Oct. 8th - Oct. 15th Oct.19th - Nov. 1st	Oct. 15th - Oct. 19th
5	Oct. 8th - Oct. 11th Oct.15th - Nov. 1st	Oct. 11th - Oct. 15th
6	Oct. 11th - Nov. 1st	Oct. 8th - Oct. 11th

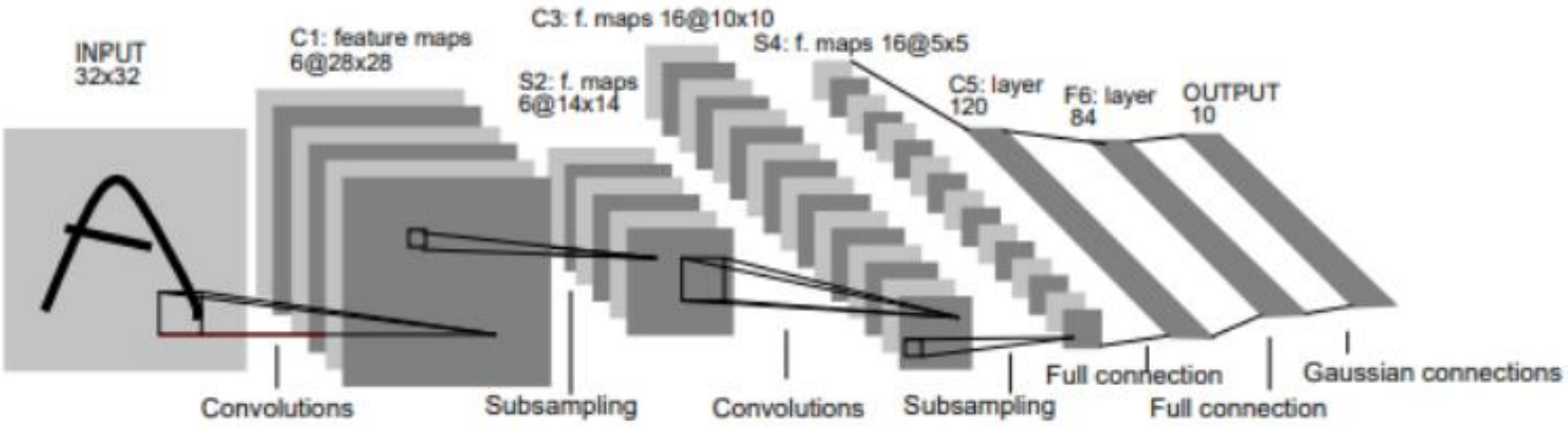
Dataset Split of 6-fold Cross-Validation

CNN and Resnet Training

Repeat(Convolution -> Subsampling->) Full Connection -> Output.

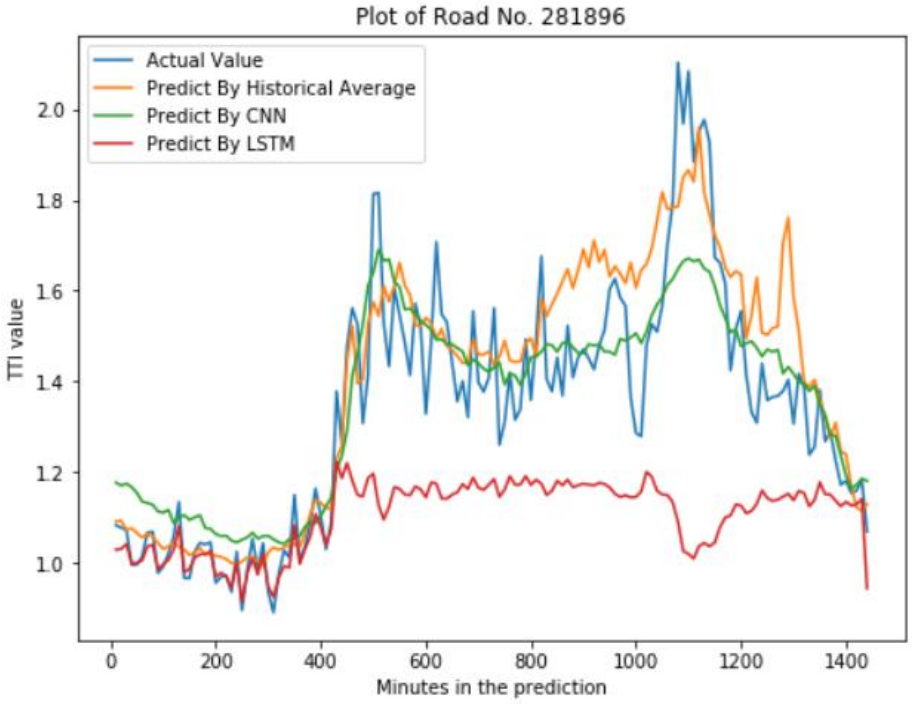
Deep Residual Neural Networks (Resnet) is a kind of Convolutional neural networks (CNN) that are created to against degradation of training accuracy.

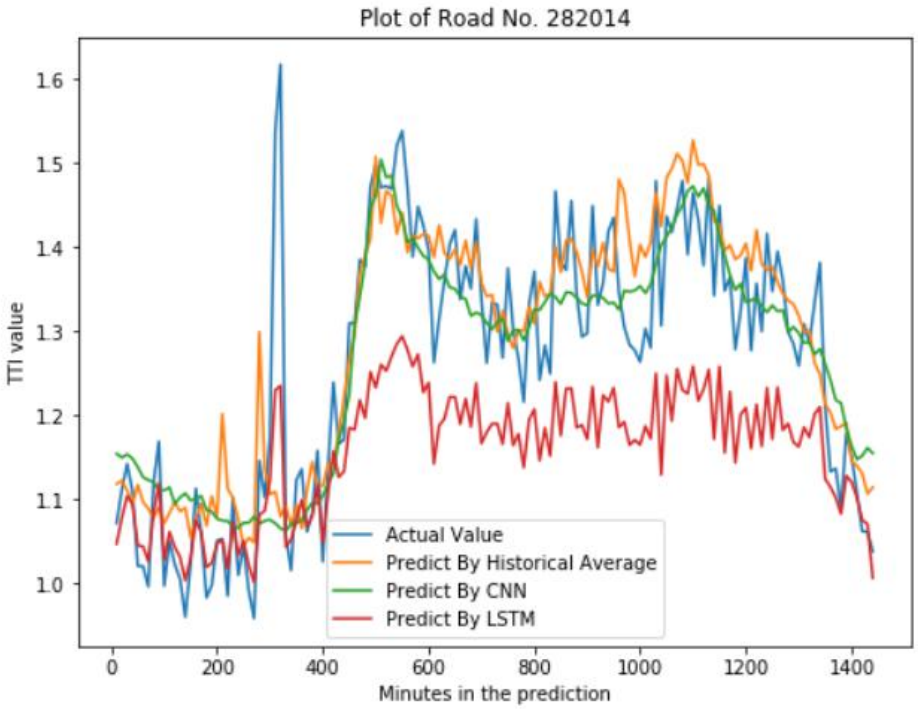
One input is a 3 phases RGB image, and the output is the TTI value 10 minutes after the time that the image represent for.



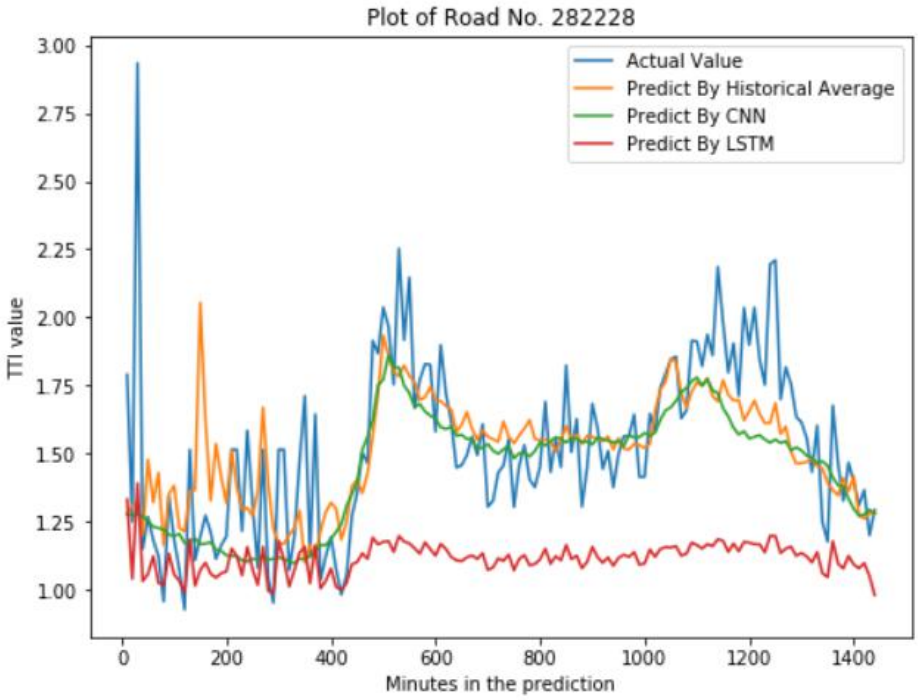
An example of digits recognition by using convolutional neural network

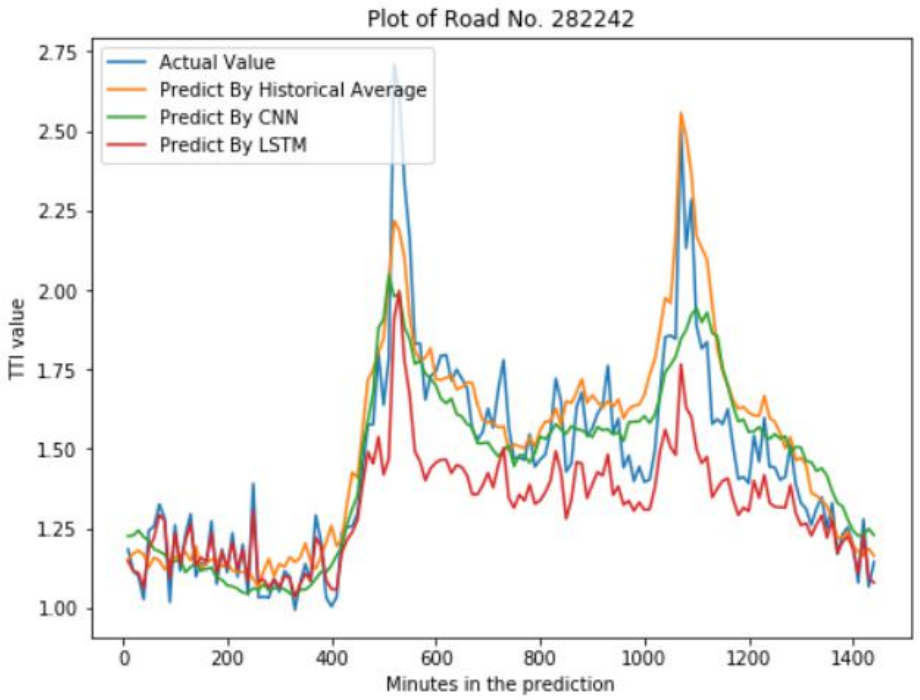
Result of Resnet with Image Fusion



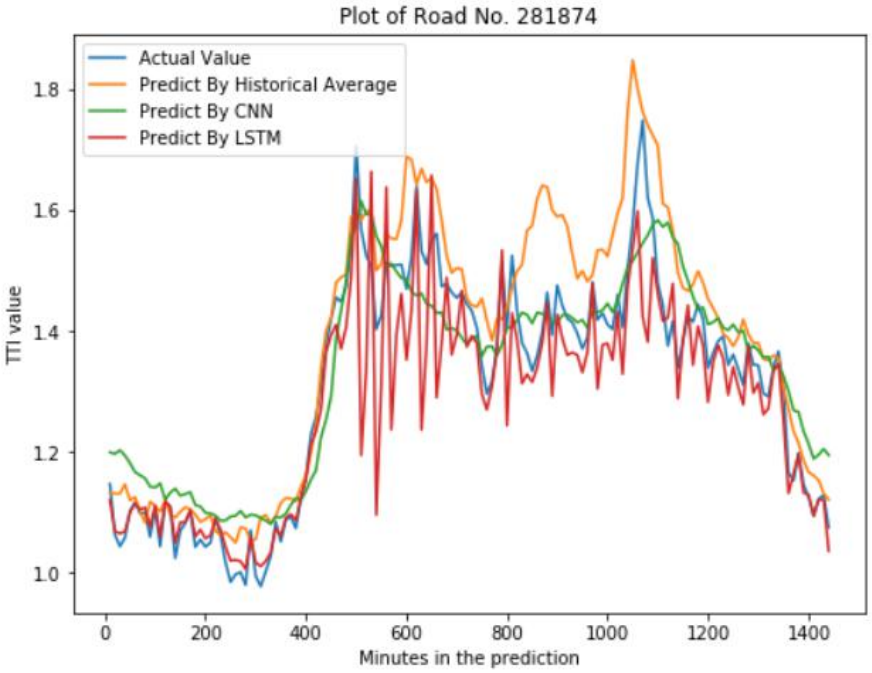


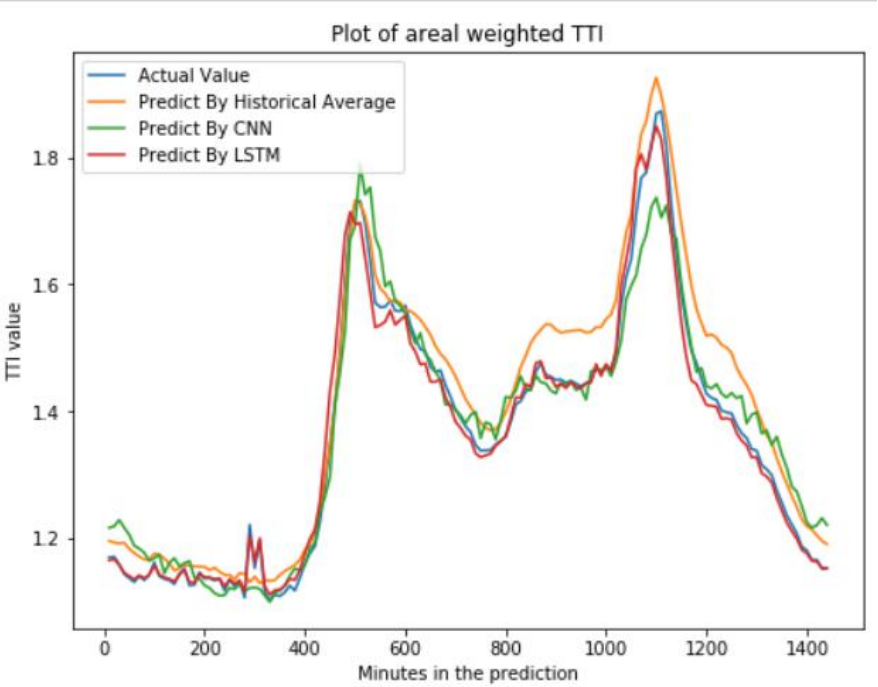
Result of Resnet with Image Fusion





Result of Resnet with Image Fusion





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%.

Resnet with image fusion have a stable results.

Resnet with image fusion can solve all the problems of LSTM model that mentioned before.

Road ID	Historical Average	Resnet34	LSTM
281870	0.13751	0.11284	0.02241
281874	0.07473	0.06299	0.04958
281876	0.08684	0.04913	0.03003
281885	0.06096	0.0637	0.03108
281896	0.10174	0.09482	0.24025
282014	0.06234	0.0596	0.11554
282228	0.17517	0.17066	0.42007
282242	0.11649	0.11439	0.14938
282252	0.10599	0.10186	0.06292
282272	0.19019	0.19143	0.05428
Area	0.04659	0.034	0.01586
Total	0.10532	0.09595	0.1083

Average Loss Comparison of Different Models

Discussion

Why Image Fusion Framework with Resnet wins in the task?

Comparing to single dataset with limited information, if multiple relevant datasets can be utilize in a model, the probability of improving it's performance increase.

- Single LSTM can only utilize simple historical dataset.
- Image Fusion Framework can fuse multiple relevant datasets from different sources.
- Resnet can find the relationship of dataset within different phases in computer vision.

Thanks!

Q & A