TTI Prediction in Urban Road Network Using Computer Vision Techniques

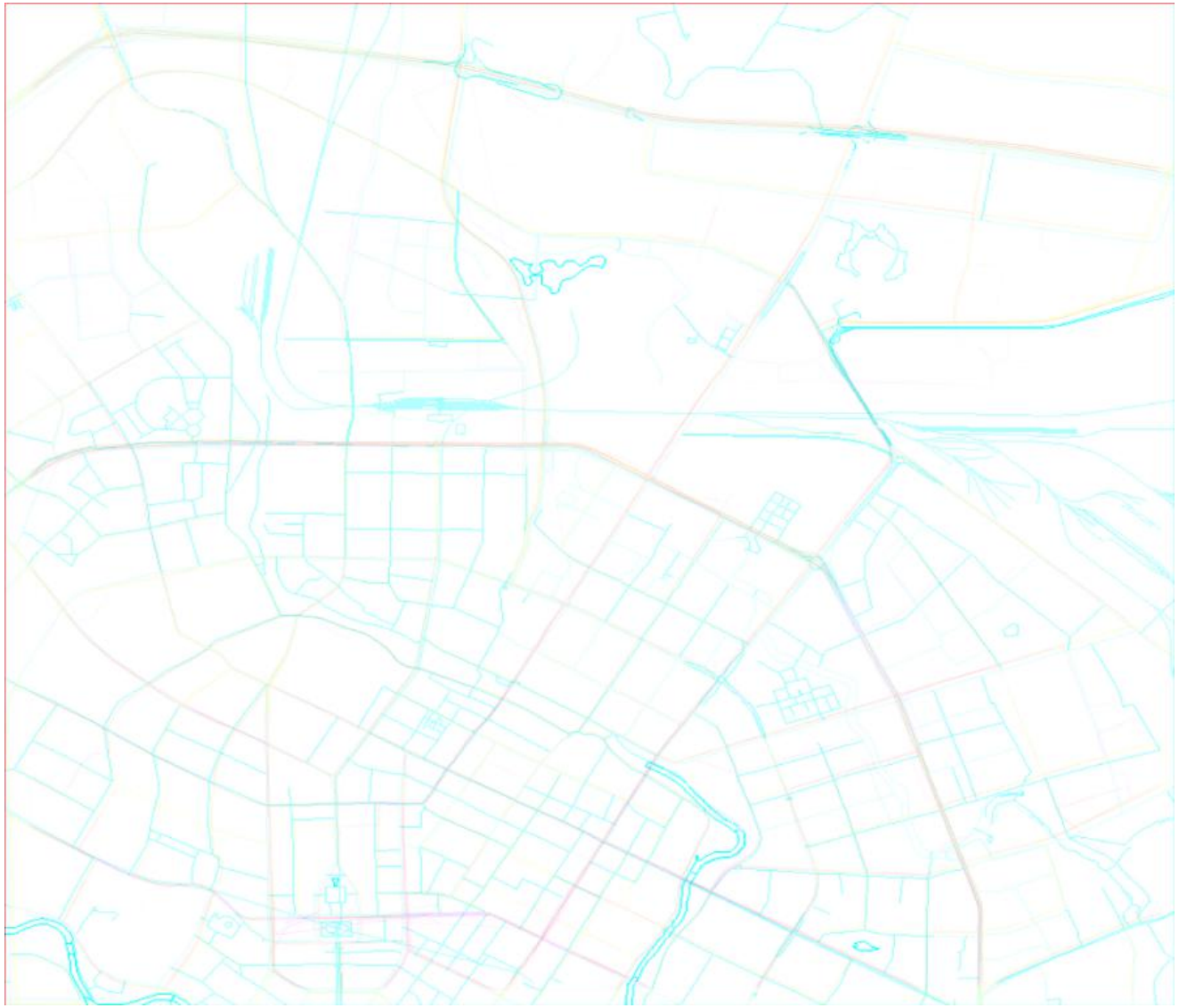
LI Xiahan

Modification

Renew the picture to white background

Renew the output from 1 to 21

Apply K-fold cross validation(now k=5)



Training Detail

- Dataset: Training and validation dataset: 20181013-20181017, 5 days. Len(Training):Len(Validation) = 4:1. Testing dataset: 20181018, 1 days
- Prediction: Predict 20 randomly selected individual road and 1 area TTI after 10 minutes
- Converge: because of the restriction of GPU usage of Kaggle (9 hours in maximum), none of neural network converge
- Neural Network: CNN, CNN+LSTM(series connection), CNN+LSTM(parallel connection), Historical Average(as baseline)

Deep Learning Result

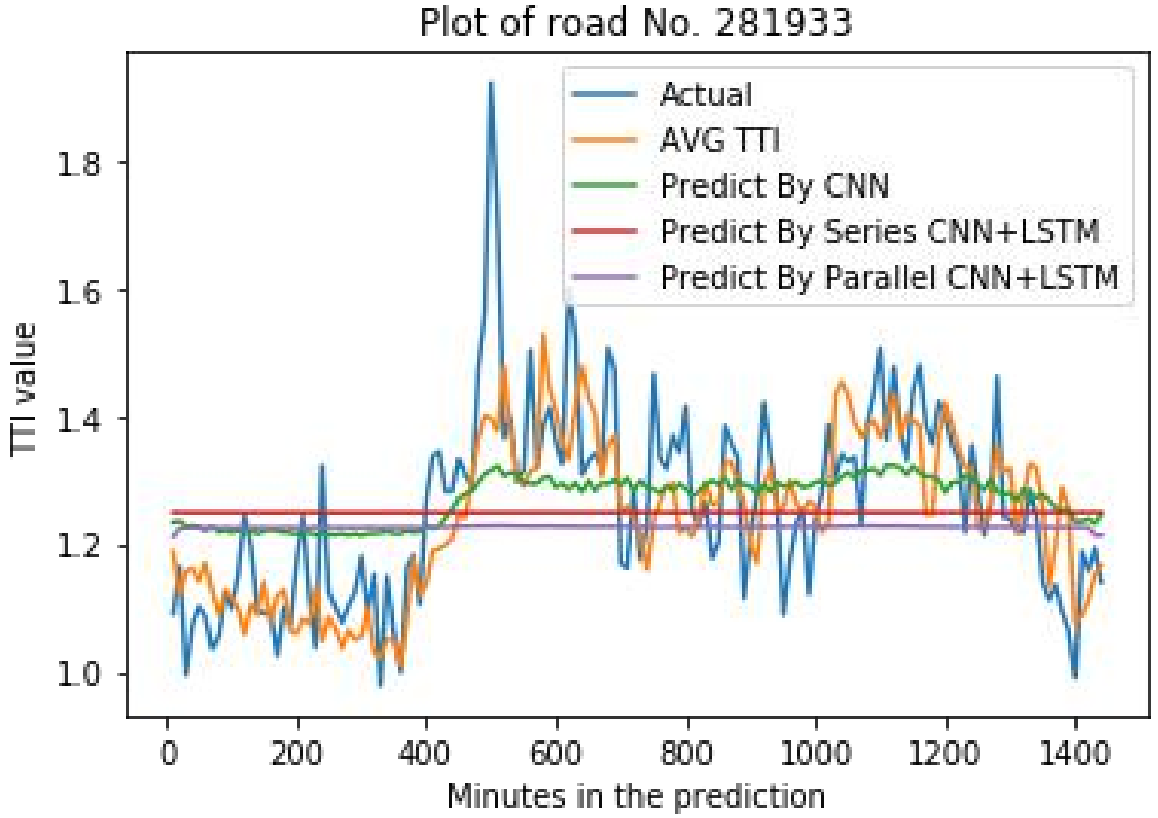
Actual: Actual TTI Value of the road

AVG TTI: Historical Average

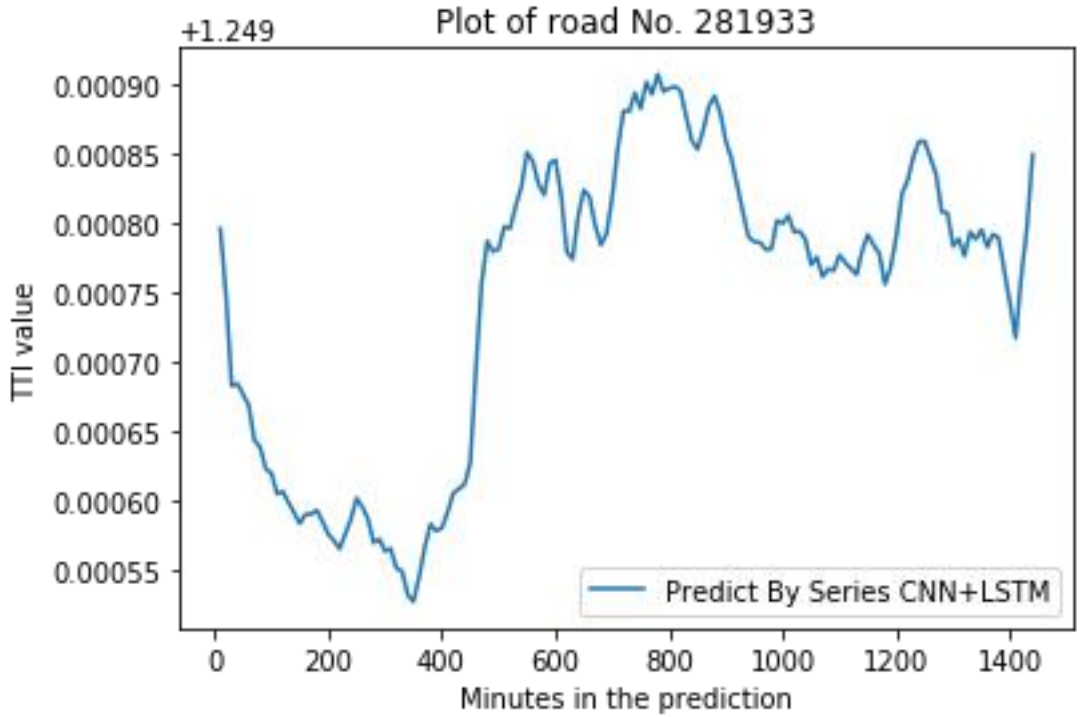
CNN: Resnet34

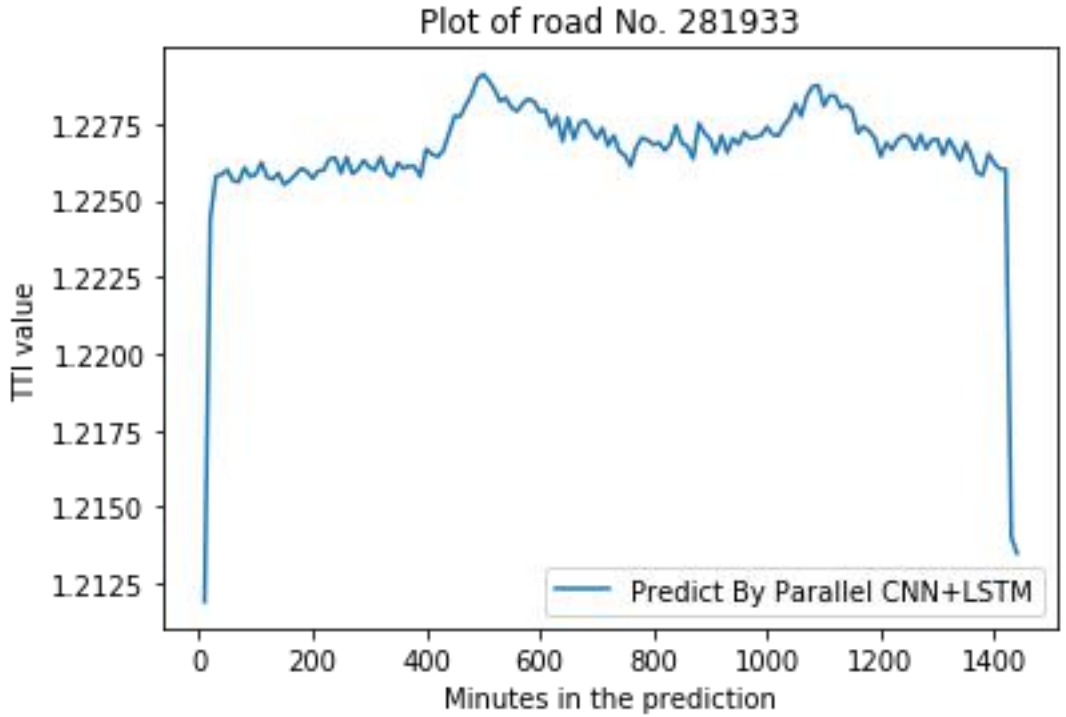
Series CNN+LSTM: Input of LSTM is the output of Resnet34, similar to the series connection in circult

Parallel CNN+LSTM: Combine the output of LSTM and output of Resnet34, similar to the parallel connection in circult



Deep Learning Result





Possible reason: LSTM is good for continuous data, but our data have 10 minutes interval, which is discrete.

Statistical Result

- Goodness of prediction:

CNN >CNN+LSTM(Series) = CNN+LSTM(Parallel) > Historical Average

According to MSELoss

```
In [3]:
def Loss(df1,df2):
    total_loss = 0.0
    for item in df1.columns.tolist()[1:]:
        for i in range(144):
            total_loss += abs(np.power(df1.at[i,item]-df2.at[i,item],2))
    print('Total Loss in Average: ',total_loss)

In [4]:
cnndf = pd.read_csv(os.path.join(path,'CNNresult.csv'))

In [5]:
actualdf = pd.read_csv(os.path.join(path,'Actual.csv'))

In [6]:
cnnlstmdf = pd.read_csv(os.path.join(path,'CNN+LSTM_result.csv'))

In [7]:
cnnlstmattdf = pd.read_csv(os.path.join(path,'CNN+LSTM+Parallel_result.csv'))

In [8]:
avgresultdf = pd.read_csv(os.path.join(path,'Avg_TTI.csv'))

In [9]:
Loss(cnndf,actualdf)#MSE Loss of results predicted by CNN

Total Loss in Average:  291.94913895356495

In [10]:
Loss(cnnlstmdf,actualdf)#MSE Loss of results predicted by CNN+LSTM series connection

Total Loss in Average:  316.96470999038377

In [11]:
Loss(avgresultdf,actualdf)#MSE Loss of results predicted by Historical Average

Total Loss in Average:  453.12626849645585

In [12]:
Loss(cnnlstmattdf,actualdf)#MSE Loss of results predicted by CNN+LSTM parallel connection

Total Loss in Average:  316.9975012693749
```

Further Plan

- Begin to write the thesis
- Add exploration data analysis part, we need to statistically explain the data to reader.