PREDICTING TRAFFIC SPEED FOR NEW JERSEY FREEWAY WORK ZONES – A DEEP LEARNING APPROACH

ABDULLAH SHABAREK, PH.D., PE

NJDOT RESEARCH SHOWCASE
OCTOBER 2021
Background & Problem Statement

- Frequent Road Maintenance
- Accounts for 24% of non-recurrent congestion
- Accounts for 10% of the overall congestion
- Delay at Upstream segments including connected freeways
Objective

» Developing a sound model to Predict Traffic Speed under work zone conditions on both connected and mainline freeways
Work Scope

- New Jersey Freeways
- Work zone conditions
- 10-miles upstream work zone
- Data between 2014 and 2019
- CNN and Deep ANN
Data Collection

- Crash Records & Plan4Safety
- DEM
- OpenReach
- NJCMS
- NJS LD
- INRIX
- Google Earth
- Big Data Analysis
- Actual Work Zone Speed
Background & Problem Statement

- No well-designed model to predict traffic speed on the connected freeways. (Overfitting issues)
- CNN and Deep ANN models
## Pros & Cons of Modeling Approaches

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| **Parametric Models** | • Transferability  
• Scalability  
• Inexpensive | • Data distribution  
• Spatial-temporal data |
| **Simulation Models** | • High fidelity | • Specific work zone  
• High computation Power and Time  
• High calibration time |
| **Non-Parametric Models** | • Scalability  
• Extensibility  
• Less computational time  
• No data distribution | • Data dependency  
• Structure configuration |
## Tools for Work Zone Congestion Prediction

<table>
<thead>
<tr>
<th>Tool</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Modeling Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>FlagSim</td>
<td>Time and location of work zone</td>
<td>• Traffic volume&lt;br&gt;• Queue length&lt;br&gt;• Delay</td>
<td>Parametric</td>
</tr>
<tr>
<td>Web-based Work Zone Traffic</td>
<td></td>
<td>• Delay cost&lt;br&gt;• Queue length</td>
<td>Parametric</td>
</tr>
<tr>
<td>LCDSS</td>
<td></td>
<td>• Queue length</td>
<td>Parametric</td>
</tr>
<tr>
<td>WIMAP-P</td>
<td>Time, location of work zone, and values of time.</td>
<td>• Delay cost&lt;br&gt;• Queue length&lt;br&gt;• Predicted traffic speed</td>
<td>Non-parametric</td>
</tr>
<tr>
<td>RILCA</td>
<td>Time and location of work zone only for the GSP and NJTP.</td>
<td>• Queue length&lt;br&gt;• Delay</td>
<td>Parametric</td>
</tr>
</tbody>
</table>
Deep Learning

- Two or more hidden layers
- Number of neurons
- Overfitting
- Dropout is a regularization technique that is applied in hidden layers for the purpose of reducing the overfitting problem
Developed CNN

CNN Inputs

- Traffic Volume approaching Work Zone at time \( t \)
- Speed during Normal Condition for segment \( i \) at time \( t \)
- Traffic Volume Rate on the mainline downstream interchange/connector at time \( t \)
- Vertical Gradient of Segment \( i \)
- Work Zone Capacity
- Distance to Work Zone for Segment \( i \)
- Traffic Volume Rate at upstream segment \( i \) at time \( t \)

CNN Layers

- 1st Layer (128 Neurons)
  - Convolutional 1st Layer Dimension of Model Input * Neuron 1st Layer Dimension = \( 7 \times 128 \)
- 2nd Layer (256 Neurons)
  - Convolutional 2nd Layer Dimension of Previous Neuron Layer Input after 0.25 dropout ratio * Neuron 2nd Layer Dimension = \( 70 \times 256 \)
- 3rd Layer (512 Neurons)
  - Convolutional 3rd Layer Dimension of Previous Neuron Layer Input after 0.5 dropout ratio * Neuron 3rd Layer Dimension = \( 112 \times 512 \)
- 4th Layer (1024 Neurons)

CNN Output

- Filter Height = 3
  - \( f \otimes g_1 \) Matrix \( 7 \times 64 \)

- When \( \tau = 1 \)
  - \( f \otimes g_1 \) Matrix
- When \( \tau = 2 \)
- When \( \tau = 64 \)

Matrix of Hidden Layer

\( 7 \times 128 \)
Developed CNN Results
Case Study

- 1-mile work zone on I-287 SB
- One Lane closure over 4 lanes
- Milepost 39 and Milepost 38
- From 3:00 PM till 09:00 PM on 07/08/2015
- I-80 as a connected route
Case Study (Location)
Case Study

Heat map of (a) passenger cars and (b) trucks distribution for I-287 SB
Source: New Jersey Congestion Management Systems
Methodology (CNN)
Heat map of (a) passenger cars and (b) truck volumes of I-80 Westbound

Heat map of (a) passenger cars and (b) truck volumes of I-80 Eastbound
Case Study

Heat map of I-287 SB of (a) Actual speed reported from INRIX (b) predicted speed from the CNN Model (c) predicted speed from the model of WIMAP-P

Heat map of traffic speed without work zone conditions for (a) I-287 SB (b) I-80 WB and (c) I-80 EB
Case Study

Heat map of traffic speed on I-80 WB from (a) the CNN prediction model (b) the actual traffic speed reported from INRIX.

Heat map of traffic speed on I-80 EB from (a) the CNN prediction model (b) the actual traffic speed reported from INRIX.
Case Study

Heat map of absolute error of the CNN results again the actual speed for (a) I-287 SB (b) I-80 WB and (c) I-80 EB
Case Study

Comparison of total delay cost for both the mainline (i.e., I-287 SB) and the connectors segments (i.e., I-80 EB and I-80 WB) to the actual work zone delay.
Case Study

The RMSE values in variation of distance to work zone
Case Study

Two main categories of TMC segments are distinguished: Type 1, which is the TMC segments on the mainline immediate upstream to the on-ramp and Type 2, which is all the other TMC segments.

<table>
<thead>
<tr>
<th>Type of TMC segment</th>
<th>Model</th>
<th>Number of Lanes</th>
<th>RMSE (mph) (% of testing data)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Shoulder Closure</td>
<td>One Lane Closure</td>
<td>Two Lane Closure</td>
<td></td>
</tr>
<tr>
<td>Type 1</td>
<td>Deep ANN</td>
<td>2</td>
<td>11.2 (5%)</td>
<td>9.5 (13%)</td>
<td>NA (0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>12.3 (8%)</td>
<td>9.1 (12%)</td>
<td>10.5 (6%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>14.9 (4%)</td>
<td>11.0 (10%)</td>
<td>11.3 (3%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CNN</td>
<td>2</td>
<td>10.0 (5%)</td>
<td>9.2 (13%)</td>
<td>NA (0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>11.6 (8%)</td>
<td>8.2 (12%)</td>
<td>9.9 (6%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>14.1 (4%)</td>
<td>10.3 (10%)</td>
<td>10.6 (3%)</td>
<td></td>
</tr>
<tr>
<td>Type 2</td>
<td>Deep ANN</td>
<td>2</td>
<td>6.4 (95%)</td>
<td>5.5 (87%)</td>
<td>NA (0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>5.9 (92%)</td>
<td>5.4 (88%)</td>
<td>7.3 (94%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>7.0 (96%)</td>
<td>5.7 (90%)</td>
<td>7.7 (97%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CNN</td>
<td>2</td>
<td>6.0 (95%)</td>
<td>5.3 (87%)</td>
<td>NA (0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>5.4 (92%)</td>
<td>4.8 (88%)</td>
<td>7.2 (94%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>6.4 (96%)</td>
<td>5.8 (90%)</td>
<td>7.5 (97%)</td>
<td></td>
</tr>
</tbody>
</table>
Applications

- Quantify the congestion costs (i.e., spatio-temporal)
- User delay costs
- User delay Vs. agencies costs
- Queue warning systems
Conclusions

- Connected roadways
- CNN outperforms Deep ANN and WIMAP-P
- Congestion mitigation plans
- Proximity to the mainline links immediate upstream segments
- Database
Future Research

- Optimal work zone scheduling with rerouting plans
- Work zone staging optimization
- Combination of work zone and accidents prediction modulus.
thank you