TrafficPredict: Trajectory Prediction for Heterogeneous Traffic-Agents

교통 시스템 내에서 다양한 종류의 교통 주체들(자동차, 자전거, 보행자 등)의 궤적을 예측하는 시스템

Dept. of AI and Bigdata, SOON CHUN HYANG Univ.

민현식
minun001@sch.ac.kr

CONTENTS

1. Introduction

- To ensure the safety of self-driving, the system must analyze the movement patterns of other traffic agents and predict their future trajectories so that autonomous vehicles can make the appropriate search decisions.

- **Driving in urban settings is much more challenging than driving on highways**

- **Traffic-agents**

in an urban environment: cars, bicycles, buses, and pedestrians

- Traditional model

Kinematic and dynamic models (Toledo-Moreo and Zamora-Izquierdo 2009)
Bayesian filters (Kalman 1960)
Gaussian Processes (Ras-mussen and Williams 2006)

→ **It is difficult to analyze complex scenarios or make long-term predictions due to the lack of consideration of interaction between transport agents and the environment**

- Advanced situation

LSTM (Ma et al. 2017) -> predict trajectories of human crowds(Alahi et al. 2016) and vehicles trajectories (Lee et al. 2017)

↓

A common limitation of these tasks is to focus on the prediction of one type of group
(for example, pedestrians or vehicles only)
It may not work in heterogeneous traffic conditions

TrafficPredict: Trajectory Prediction for Heterogeneous Traffic-Agents

1. Introduction

the main result

- **TrafficPredict**

LSTM-based algorithm for trajectory prediction in heterogeneous traffic conditions.

- **4D Graph Structure**

Constructs a 4D graph from trajectory data, representing traffic agents and their interactions in both space and time.

- **LSTM Network Architecture**

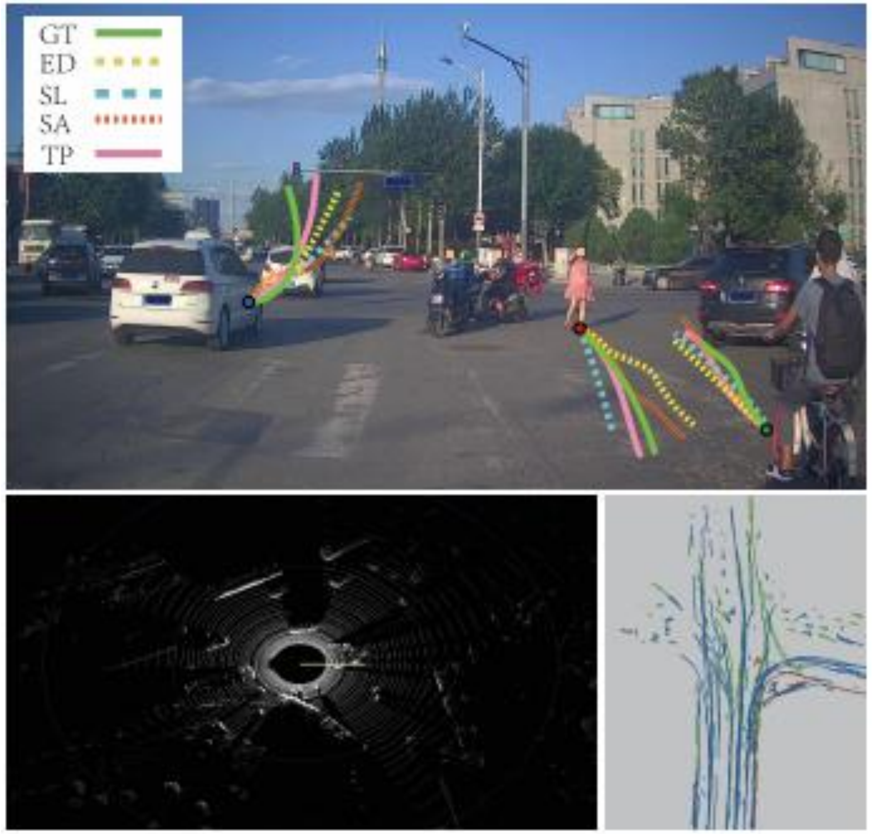
Divided into instance and category layers, focusing on agent interactions and behavior patterns within categories for prediction.

- **New Traffic Dataset**

Provides a dataset with varied traffic agent trajectories in urban environments, aiding in heterogeneous traffic prediction research.

- **Improved Performance**

TrafficPredict shows enhanced processing speed and 20% greater accuracy over previous methods.



The green solid lines denote ground truth trajectories (GT), pink solid lines are for paper method (TP) and dashed lines are the predicted trajectories for other methods (ED, SL, SA)

(ground truth trajectories, GT), (Trajectory Prediction, TP), (Encoder-Decoder, ED), (Social LSTM, SL), (Social Attention, SA)

2. Related Work

2.1 Classical methods for trajectory prediction

Traditional trajectory prediction methods Trajectory prediction or path prediction problems have been extensively studied.

- Bayesian networks (Lefevre, Laugier, and J.Ibanez-Guzman 2011),
- Monte Carlo Simulation (Danielsson, Petersson, and Eidehall 2007), Hidden
- Markov Models (HMM) (Firl et al. 2012),
- Kalman Filters (Kalman 1960),
- linear and non-linear Gaussian Pro-cess regression models (Rasmussen and Williams 2006)

These methods focus on analyzing the inherent regularities of objects themselves based on their previous movements in simple traffic scenarios where there is little interaction between vehicles.

However, these methods may not work well when different types of vehicles and pedestrians appear simultaneously.

2. Related Work

2.2 Behavior modeling and interactions

- **Social Force Model**: Introduced by Helbing and Molnar in 1995, this model simulates pedestrian movement using concepts of attraction and repulsion and was later extended by Yamaguchi et al. in 2011.
- **Methodological Developments**: Other methods like continuum dynamics and Gaussian processes have been proposed, as seen in works by Treuille, Cooper, and Popovic in 2006, and Wang, Fleet, and Hertzmann in 2008.
- **Crowd Trajectory Prediction**: Bera and colleagues in 2016 and 2017 enhanced these concepts by combining an Ensemble Kalman Filter with a human motion model to predict crowd movements.
- **Application Scenarios**: These methods are beneficial for analyzing pedestrian behavior in environments like shopping malls, public squares, and pedestrian streets.
- **Emotion and Behavior Classification**: Approaches have also been developed to classify group emotions and driver behaviors, as noted by Cheung et al. in 2018.
- **General Traffic Application**: To apply these concepts to broader traffic scenarios, Ma, Manocha, and Wang in 2018 proposed a method that predicts the trajectories of multiple traffic agents while considering kinematic and dynamic constraints.
- **Model Assumptions**: The model by Ma, Manocha, and Wang assumes perfect sensing and complete information on the shape and dynamics of all traffic agents.

2. Related Work

2.3 RNN networks for sequence prediction

- **Deep Neural Network (DNN) Popularity:** DNNs, including Recurrent Neural Networks (RNNs), have gained immense popularity due to their performance across various domains.
- **RNN:** RNNs, a type of DNN architecture, are particularly effective for generating sequences in domains such as speech recognition, machine translation, and image captioning.
- **LSTM Innovations:** Long Short-Term Memory (LSTM) networks, a variant of RNNs, have led to advancements in maneuver classification and trajectory prediction.
- **Probabilistic Approaches:** Techniques have been developed to provide probabilistic information about future vehicle locations using encoder-decoder structures over grid maps or samples.
- **Sampling Method Limitations:** Sampling-based methods encounter accuracy issues due to the discretization limits.
- **Multi-modal Distributions:** Deo and Trivedi in 2018 introduced a model that produces multi-modal distributions for trajectory generation.
- **Road and Traffic Conditions:** Many of the existing methods perform well in conditions with clear road lanes and simple driving scenarios but struggle with complex traffic involving various types of agents.
- **LSTM-CNN Hybrid Networks:** Chandra et al. in 2018a used a hybrid network for predicting interactions and trajectories between different traffic agents in image-based models.
- **Human Interactions:** Alahi et al. in 2016, Gupta et al. in 2018, and Vemula et al. in 2017 have utilized LSTM to predict pedestrian movement in crowds, showing good results on public datasets.
- **Complex Traffic Scenarios:** However, these methods still face limitations when predicting trajectories in complex traffic scenarios with interactions between heterogeneous traffic agents, not just pedestrians.

2. Related Work

2.4 Traffic datasets

1. **Cityscapes Dataset**: Introduced by Cordts et al. in 2016, Cityscapes includes 2D semantic, instance-wise, dense pixel annotations across 30 classes.

2. **ApolloScape Dataset**: Created by Huang et al. in 2018, ApolloScape is a comprehensive dataset offering large-scale street views with high scene complexity, 2D/3D annotations, pose information, lane markings, and video frames.

3. **Lack of Trajectory Information**: Neither Cityscapes nor ApolloScape datasets provide trajectory information of the traffic agents.

4. **NGSIM Dataset**: The Simulation (NGSIM) dataset, released in 2005 by the Administration, includes trajectory data for cars on highways but is limited to simple road conditions.

5. **KITTI Dataset**: Developed by Geiger et al. in 2013, KITTI is a diverse dataset for computer vision tasks, which includes short tracklet data totaling about 22 minutes.

6. **Intersection Scarcity in KITTI**: KITTI has few intersections involving vehicles, pedestrians, and cyclists, making it less useful for studying motion patterns in complex traffic conditions.

7. **Pedestrian Trajectory Datasets**: Datasets like ETH and UCY focus exclusively on pedestrian trajectories and do not include vehicles, limiting their applicability for comprehensive traffic studies.

Senseable AI Lab

3. TrafficPredict

3.1 Problem Definition

- **Traffic Agent Features**

For any traffic agent A_i^t at time t, its feature $f_i^t = (x_i^t, y_i^t, c_i^t)$

x-axis coordinates

y-axis coordinate

Traffic-agent's category

- **Agent Categories**

$C_i \in \{1, 2, 3\}$

Pedestrians

Bicycles

Cars

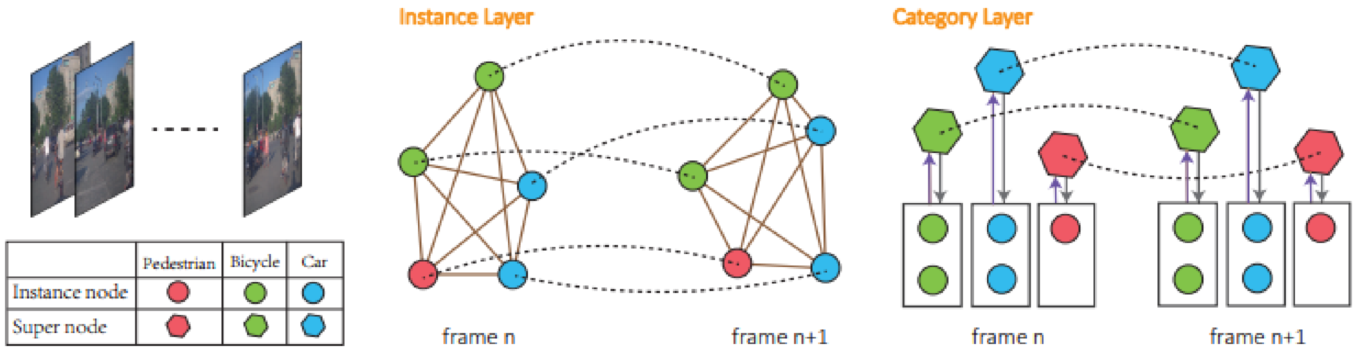
- **Task Objective**

Learning range $[1 : T_{obs}]$

Prediction range $[\,T_{obs} + 1 : T_{pred}]$

3. TrafficPredict

3.2 4D Graph Generation



(a) (b) (c)

- 4D Graph Generation models interactions among various traffic agents in urban traffic scenarios.
- Instance Layer: Traffic agents are represented as instance nodes, and their relationships are considered as edges to construct the graph in the instance layer
- Category Layer: Based on the observation that traffic agents of the same kind exhibit similar behavioral characteristics, super nodes are constructed for each category to learn and improve the prediction of their trajectory similarities

The edge between two instance nodes in one frame is called spatial edge (Jain et al. 2016; Vemula, Muelling, and Oh 2017)

$$(A_i^t, A_j^t)$$

The features of a spatial edge

$$f_{ij}^t = (x_{ij}^t, y_{ij}^t, c_{ij}^t)$$

method of calculation

$$x_{ij}^t = x_j^t - x_i^t, \; y_{ij}^t = y_j^t - y_i^t$$

3. TrafficPredict

3.2 4D Graph Generation

- The same kind of traffic- agents have similar behavior characteristics
- For example, the pedestrians have not only similar velocities but also similar reactions to other nearby traffic-agents.
- These similarities will be directly reflected in their trajectories.

$$C_u^t, u \in \{1, 2, 3\}$$

super node

for each kind of traffic-agent to learn the similarities of their trajectories and then utilize that super node to refine the prediction for instances.

4D Graph

two dimensions for traffic-agents and their interactions
+ one dimension for time series
+ one dimension for high-level categories.

Construct an information network for the entire traffic

3. TrafficPredict

3.3 Model Architecture

1. Instance Layer
2. Category Layer

1. Instance Layer → Goal: Capture Exercise Patterns

Three LSTM → car, bicycle, and pedestrian

$$(A_i^t, A_j^t) \xrightarrow{\text{Allocation}} \text{LSTM } L_{ij}$$

$$e_{ij}^t = \phi(f_{ij}^t; W_{spa}^e),$$
$$h_{ij}^t = LSTM(h_{ij}^{t-1}; e_{ij}^t; W_{spa}^r)$$

Each of the other instances has different impacts on the node's behavior

Soft attention mechanism (Vemula, Muelling, and Oh 2017)

$$w(h_{ij}^t) = softmax(\frac{m}{\sqrt{d_e}} Dot(W_i h_{ii}^t, W_{ij} h_{ij}^t))$$

$$e_i^t = \phi(f_i^t; W_{ins}^e),$$
$$a_i^t = \phi(concat(h_{ii}^t, H_i^t); W_{ins}^a),$$
$$h1_i^t = LSTM(h2_i^{t-1}; concat(e_i^t, a_i^t); W_{ins}^r)$$

3. TrafficPredict

3.3 Model Architecture

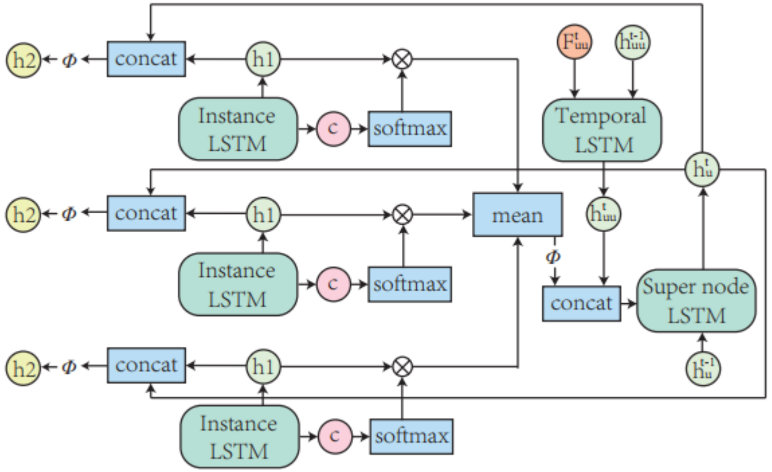
1. Instance Layer
2. Category Layer

better predict the trajectories for the entire instances.

2. Category Layer

Usually traffic-agents of the same category have similar dynamic properties, including the speed, acceleration, steering

learn the movement patterns from instances of the same category



Position estimation

Assumption: The future position of a traffic agent is presumed to follow a bivariate Gaussian distribution, following the method established by Alahi et al. (2016).

Back-Propagation Technique: The training employs joint back-propagation through various components of the model – instance nodes, super nodes, and both spatial and temporal edges.

4. Experiments

4.1 Dataset

Count		NGSIM	KITTI	Our Dataset
duration (min)		45	22	**155**
frames ($\times 10^3$)		11.2	13.1	**93.0**
total ($\times 10^3$)	pedestrian	0	0.09	**16.2**
	bicycle	0	0.04	**5.5**
	vehicle	2.91	0.93	**60.1**
average (1/f)	pedestrian	0	1.3	**1.6**
	bicycle	0	0.24	**1.9**
	vehicle	**845**	3.4	12.9
device	camera	yes	yes	yes
	lidar	no	yes	yes
	GPS	no	yes	yes

4. Experiments

4.2 Evaluation Metrics and Baselines

1. Evaluation Metrics

•**Average displacement error**: The mean Euclidean distance between all predicted and actual positions during prediction.

•**Final displacement error**: The mean Euclidean distance between the final predicted positions and the true locations.

2. Compared Methods (Baselines)

1.**RNN ED (ED)**: An RNN encoder-decoder model extensively used for motion and trajectory prediction of vehicles.

2.**Social LSTM (SL)**: An LSTM network with social pooling (Alahi et al. 2016), outperforming traditional methods like linear models, Social Force models, and Interacting Gaussian Processes.

3.**Social Attention (SA)**: An attention-based S-RNN architecture (Vemula, Muelling, and Oh 2017) focused on understanding relative influences in crowds for pedestrian trajectory prediction.

4.**TrafficPredict-NoCL (TP-NoCL)**: The proposed method, excluding the category layer.

5.**TrafficPredict-NoSA (TP-NoSA)**: The proposed method, omitting the self-attention mechanism in the category layer.

Senseable AI Lab

4. Experiments

4.3 Implementation Details

- **Hidden State Dimensions**: 128 for edge cells, 64 for node cells.
- **Input & Attention Layer**: Fixed at 64.
- **Optimization**: Adam with β1 = 0.9, β2 = 0.999, learning rate 0.001.
- **Training**: On Tesla K40 GPU, batch size 8, gradient clipping -10 to 10.
- **Prediction**: Observes 2 seconds, predicts next 3 seconds of trajectories.

TrafficPredict: Trajectory Prediction for Heterogeneous Traffic-Agents

4. Experiments

4.4 Analysis

Metric	Methods	ED	SL	SA	TP-NoCL	TP-NoSA	TrafficPredict
Avg. disp. error	pedestrian	0.121	0.135	0.112	0.125	0.118	**0.091**
	bicycle	0.112	0.142	0.111	0.115	0.110	**0.083**
	vehicle	0.122	0.147	0.108	0.101	0.096	**0.080**
	total	0.120	0.145	0.110	0.113	0.108	**0.085**
Final disp. error	pedestrian	0.255	0.173	0.160	0.188	0.178	**0.150**
	bicycle	0.190	0.184	0.170	0.193	0.169	**0.139**
	vehicle	0.195	0.202	0.189	0.172	0.150	**0.131**
	total	0.214	0.198	0.178	0.187	0.165	**0.141**

•**Error Measurement**: Average and final displacement errors calculated for all traffic agents.

•**Model Comparison**: TrafficPredict method shows higher accuracy compared to SA, RNN ED, Social LSTM.

•**Best Performance**: TrafficPredict achieves approximately 20% improvement in accuracy.

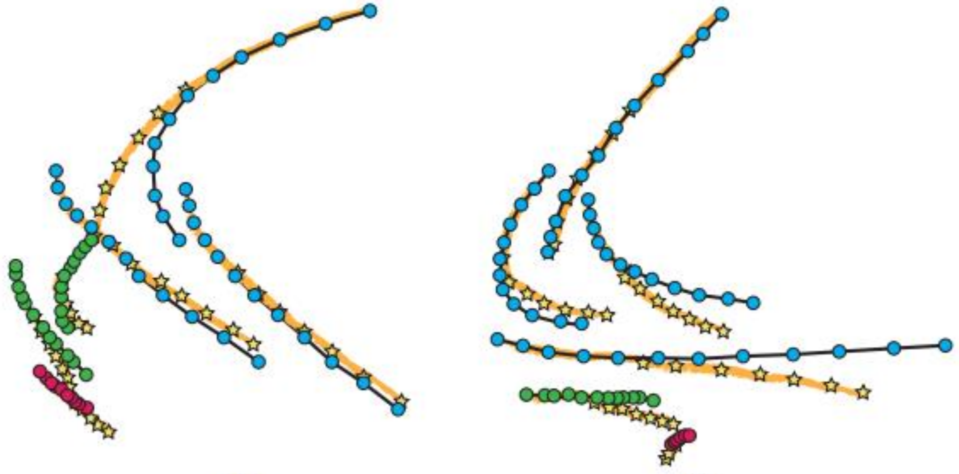
•**Applicability**: Effective prediction in diverse traffic conditions.

•**Visualization:** Accurate trajectory predictions in complex scenarios like crossroads.

•**Trajectory Colors: Vehicles (blue), bicycles (green), pedestrians (red).**

•**Predictions: Marked with yellow stars.**

•**Observation**
First five points of each trajectory shown, with some overlaps in pedestrian trajectories.



(a) (b)

4. Experiments

4.4 Analysis

•Six scenarios with different road conditions and traffic situations.

•Only select instances' trajectories shown in each scenario.

•Ground truth trajectories (green) and predictions from other methods (ED, SL, SA) displayed with dashed lines.

•TP algorithm's predicted trajectories (pink lines) are closest to the ground truth in most cases.

5. Conclusion

1. Developed TrafficPredict, an LSTM-based algorithm for predicting trajectories of diverse traffic agents in urban settings.

2. Utilizes an instance layer for trajectory and interaction capture, and a category layer for summarizing movement pattern similarities.

3. The 4D Graph efficiently processes spatial and temporal information.

4. TrafficPredict surpasses previous methods in trajectory prediction accuracy on a new heterogeneous traffic dataset.

5. Demonstrates good accuracy in dense urban traffic scenarios and operates in real-time without assumptions about traffic conditions or agent count.

6. Future improvements will focus on incorporating lane directions, traffic signals, and rules, and testing in denser traffic scenarios.

6. How to apply

Thank You