DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics: Robotics, Cognition, Intelligence

Anomaly Detection for the behavior of drivers based on Structural Temporal Graph Neural Networks

Hanxi Jiang

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics: Robotics, Cognition, Intelligence

Anomaly Detection for the behavior of drivers based on Structural Temporal Graph Neural Networks

Erkennung von Anomalien im Verhalten von Autofahrern auf der Grundlage struktureller temporaler neuronaler Netze

Author:	Hanxi Jiang
Supervisor:	Supervisor
Advisor:	Advisor
Submission Date:	Submission date

I confirm that this master's thesis in informatics: robotics, cognition, intelligence is my own work and I have documented all sources and material used.

Munich, Submission date Hanxi Jiang

Acknowledgments

Abstract

Kurzfassung

Contents

1. Introduction

With the development of autonomous driving technology, the concern of driving safety is gradually drawing an issue.

1.1. contribution

1.2. Structure

1.2.1. Subsection

See Table 1.1, Figure 1.1, Figure 1.2, Figure 1.3, Figure 1.4, Figure 1.5.

Table 1.1.: An example for a simple table.

A	B	C	D
1	2	1	2
2	3	2	3

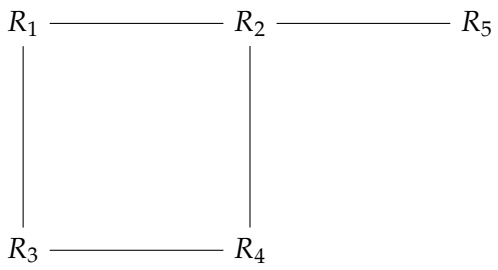

Figure 1.1.: An example for a simple drawing.

This is how the glossary will be used.

Donor dye, ex. Alexa 488 (D_{dye}), Förster distance, Förster distance (R_0), and k_{DEAC}. Also, the TUM has many computers, not only one Computer. Subsequent acronym usage will only print the short version of Technical University of Munich (TUM) (take care of plural, if needed!), like here with TUM, too. It can also be –> hidden[1] <–.

[(TODO: Now it is your turn to write your thesis.

[1] Example for a hidden TUM glossary entry.

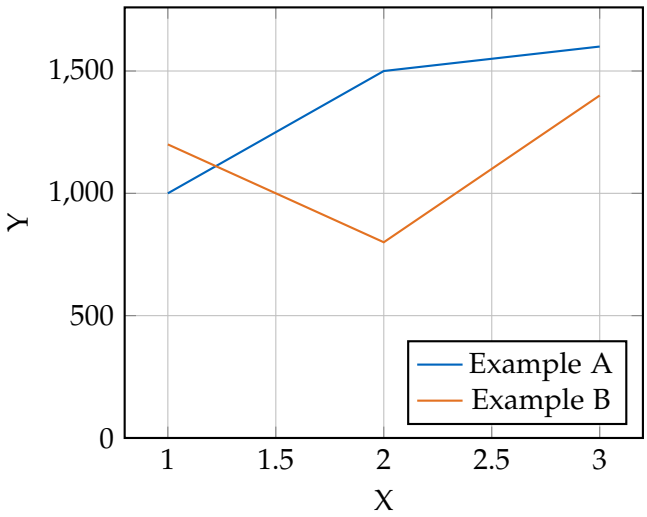

Figure 1.2.: An example for a simple plot.

```
SELECT * FROM tbl WHERE tbl.str = "str"
```

Figure 1.3.: An example for a source code listing.

Figure 1.4.: Includegraphics searches for the filename without extension first in logos, then in figures.

Figure 1.5.: For pictures with the same name, the direct folder needs to be chosen.

(a) The logo.

(b) The famous slide.

Figure 1.6.: Two TUM pictures side by side.

This will be a few tough weeks.)]
[(Done: Nevertheless, celebrate it when it is done!)]

2. Background

2.1. Scene graph generation

2.2. large language model

2.3. behavior prediction

2.4. graph

2.4.1. Dynamic graph

2.4.2. graph neural networks

3. Methodology

This work aims to construct a graph neural network-based architecture for predicting, analyzing, and detecting any potentially abnormal behavior regarding the driver during the whole driving process. In particular, The model extracts a description graph, the so-called scene graph, of the driver from the video filmed inside the vehicle and trains itself with these data to learn for future behavior prediction. The result will be used to compare and detect any abnormal behavior. Here we would lay most emphasis on the construction of the training model. To make precious anomaly detection we aim to predict not only if there is a behavior between humans and a specific kind of object but the type of behavior as well, which will cause several adaptions based on existing model *JODIE*.

3.1. scene graph generation

3.1.1. video data extracting

3.1.2. graph generating

3.2. model architecture

After comparing all the training results of the below models we would find that **JODIE** is one coming up with the best prediction. However, the model jodie still fail to predict the state of the predicted edge. In my masterwork I would like to rewrite the embedding function and the loss function of **JODIE to make the state prediction possible.

- function from **JODIE**:

embedding function

$$\mathbf{u}(\mathbf{t}) = \sigma(W_1^u \mathbf{u}(\mathbf{t}^-) + W_2^u \mathbf{i}(\mathbf{t}^-) + W_3^u f + W_4^u \Delta_u)$$

$$\mathbf{i}(\mathbf{t}) = \sigma(W_1^i \mathbf{i}(\mathbf{t}^-) + W_2^i \mathbf{u}(\mathbf{t}^-) + W_3^i f + W_4^i \Delta_i)$$

loss function(BCE)

$$L = -(j_{pos} \log \tilde{j} + j_{neg} log(1 - \tilde{j}))$$

where

$$\tilde{j}(t + \Delta) = W_1 \hat{u}(t + \delta) + W_2 \bar{u} + W_3 i(t + \Delta^-) + W_4 \bar{i} + B$$

- functions adapted in my work:

embedding function

$$\mathbf{u}(\mathbf{t}) = \sigma(W_1^u \mathbf{u}(\mathbf{t}^-) + W_2^u \mathbf{i}(\mathbf{t}^-) + W_3^u f + W_4^u s + W_5^u \Delta_u)$$

$$\mathbf{i}(\mathbf{t}) = \sigma(W_1^i \mathbf{i}(\mathbf{t}^-) + W_2^i \mathbf{u}(\mathbf{t}^-) + W_3^i f + W_4^i s + W_5^u \Delta_i)$$

we will change it from BCE to CE for predictiing state.

$$\tilde{j}(t + \Delta) = W_1 \hat{u}(t + \delta) + W_2 \bar{u} + W_3 i(t + \Delta^-) + W_4 \bar{i} + W_5 s + B$$

A. General Addenda

If there are several additions you want to add, but they do not fit into the thesis itself, they belong here.

A.1. Detailed Addition

Even sections are possible, but usually only used for several elements in, e.g. tables, images, etc.

B. Figures

B.1. Example 1

✓

B.2. Example 2

✗

List of Figures

List of Tables