CPBS

Center for Pedestrian and Bicyclist Safety

A Tier-1 University Transportation Center

An Unsupervised Learning Framework for Detecting Abnormal Driving Behavior via Utility-feature Sequences

August 2025

A Report From the
Center for Pedestrian and Bicyclist Safety

Xiao Liang
University of Wisconsin-Milwaukee

Xiaowei (Tom) Shi
University of Wisconsin-Milwaukee

Suzhou Huang
University of Wisconsin-Milwaukee

Xiao Qin
University of Wisconsin-Milwaukee

About the Center for Pedestrian and Bicyclist Safety (CPBS)

The Center for Pedestrian and Bicyclist Safety (CPBS) is a consortium of universities committed to eliminating pedestrian and bicyclist fatalities and injuries through cutting-edge research, workforce development, technology transfer, and education. Consortium members include: The University of New Mexico; San Diego State University; The University of California Berkeley; The University of Tennessee Knoxville; and The University of Wisconsin Milwaukee. More information can be found at: https://pedbikesafety.org

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

The U.S. Department of Transportation requires that all University Transportation Center reports be published publicly. To fulfill this requirement, the Center for Pedestrian and Bicyclist Safety provides reports publicly on its website, www.pedbikesafety.org The authors may copyright any books, publications, or other copyrightable materials developed in the course of, or under, or as a result of the funding grant; however, the U.S. Department of Transportation reserves a royalty-free, nonexclusive and irrevocable license to reproduce, publish, or otherwise use and to authorize others to use the work for government purposes.

Acknowledgments

This study was funded, partially or entirely, by a grant from the Center for Pedestrian and Bicyclist Safety (CPBS), supported by the U.S. Department of Transportation (USDOT) through the University Transportation Centers program. The authors would like to thank CPBS and the USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project.

TECHNICAL DOCUMENTATION

1. Project No.	2. Government Accession No.	3. Recipient's Catalog No.
24UWM02		

4. Title and Subtitle	5. Report Date
An Unsupervised Learning Framework for Detecting Abnormal Driving Behavior via Utility-feature Sequences	August 2025
	6. Performing Organization Code
	N/A
7. Author(s)	**8. Performing Organization Report No.**
Xiao Liang https://orcid.org/0009-0006-2187-8316 Xiaowei Shi https://orcid.org/0000-0002-7288-3186 Suzhou Huang https://orcid.org/0000-0002-7647-9375 Xiao Qin https://orcid.org/0000-0003-0073-3485.	N/A
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
Center for Pedestrian and Bicyclist Safety Centennial Engineering Center 3020 The University of New Mexico Albuquerque, NM 87131	**11. Contract or Grant No.** 69A3552348336
12. Sponsoring Agency Name and Address	**13. Type of Report and Period Covered**
United States of America Department of Transportation Office of Research, Development, and Technology (RD&T)	Final Report (June 2025)
	14. Sponsoring Agency Code
	USDOT OST-R

15. Supplementary Notes

Report available at the CPBS website [https://pedbikesafety.org] and DOI https://doi.org/10.21949/z2qe-6e59

16. Abstract

Abnormal driving poses a significant threat to all road users, especially VRUs such as pedestrians and bicyclists. Although traffic data collected from roadside cameras is widely available, it lacks ground-truth anomaly labels, making supervised learning impractical. Existing detection methods rely on individual vehicle features, such as speed, acceleration, which ignores interactions between vehicles. Moreover, as different vehicles have diverse driving behavior. Fixed thresholds fail to adapt in dynamic traffic environments. To address these challenges, we propose an unsupervised anomaly detection framework that integrates utility-based modeling and deep learning. A set of utility functions is defined to quantify driving safety risks. A GRU-based autoencoder learns temporal patterns in driving behavior, and latent features are clustered using a Gaussian Mixture Model (GMM). Meanwhile, the source of abnormality can be identified by analyzing reconstruction error from the GRU-AE. Experimental results on the HighSim dataset, demonstrate that the proposed framework can effectively detect abnormal behavior. Visualization confirms that detected anomalies correspond to spikes in interpretable utility features, validating both the effectiveness and the interpretability of the approach.

17. Key Words	18. Distribution Statement
[Anomaly detection; Driving behavior; Gated recurrent unit; Autoencoder; Gaussian mixture model]	No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	27	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized.**

An Unsupervised Learning Framework for Detecting Abnormal Driving Behavior via Utility-feature Sequences

A Center for Pedestrian and Bicyclist Safety Research Report

August 2025

Xiao Liang
University of Wisconsin-Milwaukee

Xiaowei Shi
University of Wisconsin-Milwaukee

Suzhou Huang
University of Wisconsin-Milwaukee

Xiao Qin
University of Wisconsin-Milwaukee

TABLE OF CONTENTS

List of Figures

List of Tables

Acronyms, Abbreviations, and Symbols

GRU	Gated Recurrent Units
AE	Autoencoder
GMM	Gaussian Mixture Model
SPE	Squared Prediction Error
VRUs	Vulnerable Road Users

Abstract

Abnormal driving poses a significant threat to all road users, especially vulnerable road users (VRUs) such as pedestrians and bicyclists. Although video-based traffic data collected from roadside cameras is widely available, it typically lacks ground-truth anomaly labels, making supervised learning approaches impractical. Existing anomaly detection methods rely on individual vehicle features, such as speed, acceleration, and steering angle, which ignore interactions between vehicles. Moreover, as different vehicles have diverse driving behavior, the use of fixed thresholds fails to account for variability in driving behavior and may perform poorly in dynamic traffic environments. To address these challenges, we propose an unsupervised anomaly detection framework that integrates utility-based modeling and deep learning to identify abnormal driving behavior from observable data. Specifically, we define a set of utility functions to represent driving safety risks and use a gated recurrent unit autoencoder (GRU-AE) to learn temporal patterns in vehicle behavior. Latent variables extracted by the GRU encoder are clustered using a Gaussian mixture model (GMM) to characterize driving behavior. Based on the GMM's probabilistic output, monitoring thresholds can be adaptively determined for each input. Meanwhile, the source of abnormality can be identified by analyzing reconstruction error from the GRU-AE. Experimental results on the HighSim dataset, a naturalistic driving dataset, demonstrate that the proposed framework can effectively detect abnormal behavior. Visualization confirms that detected anomalies correspond to spikes in interpretable utility features, validating both the effectiveness and the interpretability of the approach.

Executive Summary

With the increasing number of vehicles and vulnerable road users (VRUs), such as pedestrians and bicyclists, ensuring road safety has become a growing concern. Abnormal driving behaviors can have many manifestations for different drivers, such as hard acceleration, short headway, speeding, hard braking, lane departure, collision risk, and lane changing frequency, which are confirmed as the major contributors to traffic accidents.

We propose an unsupervised anomaly detection framework that integrates utility-based modeling and deep learning to identify abnormal driving behavior from observable data. Specifically, a set of utility functions is first defined to capture key features of driving behaviors. These interpretable features, unlike raw inputs in black box models, offer semantic insight into abnormal driving behavior. These utility functions are used as time-series inputs to a GRU-AE, which can capture latent variables of vehicle behavior. And then a GMM model is applied to latent variables to form clusters representing different behavior patterns. New vehicle sequences will be flagged as anomalous if their reconstruction errors exceed a dynamic threshold derived from the GMM components. This method achieves anomaly detection without requiring labeled anomalies in the dataset.

The proposed framework is evaluated using the HighSim dataset, which is a naturalistic driving dataset and shows real-world driving behavior, consisting of more than 4 million records from 1,200 vehicles. The experimental results demonstrate that the framework successfully detects abnormal driving behavior. Results show a strong alignment between detected anomalies and spikes in utility features, which verifies the model's effectiveness.

In conclusion, this report offers an interpretable and unsupervised solution for detecting abnormal vehicle behavior using traffic data. It provides an effective anomaly detection framework that does not require manually labeled data.

Introduction

Identifying the abnormal behavior of vehicles and protecting the safety of vulnerable road users (VRUs) is crucial. According to the National Highway Traffic Safety Administration (NHTSA), 39,345 people died in motor vehicle traffic crashes in 2024. Existing literature (*Wu et al. 2018)* has shown that early detection and intervention in cases of abnormal driving can avoid traffic accidents or at least reduce their severity.

Abnormal driving behavior refers to actions that deviate from expected or safe driving patterns. Such behavior often results from stress, perception errors, poor decision-making, or distraction and results in speeding, tailgating, improper lane changes, sharp turns, or sudden braking (*Ge et al. 2014).* Although video-based traffic data collected from roadside cameras is widely available, it typically lacks ground-truth anomaly labels, making supervised learning approaches impractical. In addition, existing methods for detecting abnormal behaviors face several limitations. Most current anomaly detection methods rely on individual vehicle features, such as speed ,acceleration and steering angle, which ignores interactions between vehicles. Moreover, due to the diversity in driving behavior among different vehicles, fixed-threshold methods lack adaptability and tend to perform poorly in dynamic traffic environments. To address these shortcomings, this project proposes a framework designed to improve the safety of pedestrians and bicyclists by reliably identifying abnormal driving behaviors.

We set a series of utility functions to model various aspects of driving behavior, integrating them into a GRU-AE framework for anomaly detection. By simulating realistic interactions and giving different behaviors different penalties, the utility-based functions enhance the fidelity of driving behavior and support more reliable detection of anomalies. At the same time, the rapid development of deep learning techniques offers new opportunities for improving anomaly detection. Deep learning models, particularly those for temporal sequences, such as autoencoders (AE) and gated recurrent units (GRU), can effectively capture temporal patterns in driving behavior. Moreover, by combining Gaussian mixture models (GMM) with deep learning methods, the system can cluster behavior patterns and determine adaptive anomaly thresholds for different behavior patterns. In this report, we propose an anomaly detection framework that integrates utility-based behavioral modeling with a deep learning model. Our method aims to verify the effectiveness and accuracy of abnormal driving detection methods.

Literature Review

Detecting abnormal driving behavior is a popular topic in the transportation safety field. Many researchers have developed various methods for anomaly detection. Broadly, these can be categorized into two main groups: rule-based approaches and model-based approaches. Rule-based methods rely on manually defined thresholds derived from expert knowledge or statistical summaries of characteristic indicators(*Mohamad et al. 2011, Bergasa et al. 2014, Levia-Bliech et al. 2019)*. Although interpretable and easy to understand, they are often inflexible and poorly generalize to diverse traffic scenarios.

Model-based methods learn behavioral patterns from data and can be further divided into statistical methods and machine learning-based methods. Statistical approaches such as principal component analysis (PCA), dynamic PCA (DPCA) (*Ku et al. 1995),* independent component analysis (ICA), and GMM (*Becker et al. 2006*) aim to capture underlying data structures without relying on labeled data. While PCA and DPCA are effective for linear patterns and dimensionality reduction, they are limited in capturing complex temporal dependencies and non-Gaussian distributions.

Machine learning-based methods include traditional models such as support vector machines (SVM) (*Li et al. 2013*), hidden Markov models (HMM) (*Zhang et al. 2017*), and Bayesian classifiers (*Eren et al. 2012*) and deep learning networks (*Hu et al. 2020*). Recent advances in deep learning have made it possible to model nonlinear, high-dimensional, and time-dependent behaviors. In particular, recurrent neural networks (RNN) (*Nanduri et al. 2016*), long short-term memory networks (LSTM) (*Tüfekci et al. 2022*), and GRU (*Tang et al. 2023*) are well-suited for modeling sequential driving data. Compared to RNN and LSTM, GRU has fewer parameters and requires less computation, making it more efficient for large-scale anomaly detection tasks.

A major challenge in this domain lies in data availability. While trajectory data collected from roadside cameras is abundant, abnormal behaviors are rare and usually unlabeled, making supervised approaches impractical. This has led to a surge in unsupervised learning approaches. AE has been widely used to reconstruct normal patterns and detect deviations via reconstruction error (*Taylor et al. 2016; Niu et al. 2020*). When combined with GRU networks, forming a GRU-AE architecture, the model can effectively learn temporal dependencies. For instance, Zhan proposes an anomaly detection and failure identification method based on GRU-AE, aimed at achieving anomaly detection in hydraulic support pressure data and equipment failure early warning. (*Zhan et al. 2025).*

Given that abnormal behaviors often exhibit non-Gaussian, irregular patterns, combining GRU-AE with probabilistic models such as GMM enables clustering of latent features and identification of low-probability anomalies. This also allows the system to determine adaptive thresholds dynamically, rather than relying on static rules.

In summary, combining interpretable utility functions with GRU-AE for temporal feature extraction and GMM for probabilistic clustering provides a promising, unsupervised, and interpretable framework for robust abnormal driving detection.

Data and Methodology

Data:

The project uses the High-Granularity Highway Simulation (HighSim)dataset (*Shi et al.2021),* which has vehicle trajectories from several high-resolution aerial videos recorded from helicopters in the I-75 highway segment, as shown in Figure 1. The data were collected at a sampling rate of 30 Hz, ensuring high temporal resolution. A summary of the dataset's contents is provided in **Error! Reference source not found.**.

Figure 1.The study area of HighSim dataset (*Shi et al. 2021*).

Table 1. Contents of HighSim dataset (*Shi et al. 2021*).

Column Name	Explanation
Vehicle ID	ID number for each vehicle
Global Time	Time in seconds from 12:00:00am of the day
Frame ID	Frame number in the corresponding video
Local X(ft)	Position in the direction perpendicular to road
Local Y(ft)	Position in the direction along the road
Global X	Vehicle's GPS longitude location
Global Y	Vehicle's GPS latitude location
Width	Vehicle width
Length	Vehicle length
Class (1 motor; 2 auto; 3 truck)	Vehicle class
Speed (ft/s)	Vehicle speed
Acceleration (ft/s^2)	Vehicle acceleration
Lane Num	Lane number
Space Highway (ft)	Distance between vehicle's front bumper to its following vehicle's front bumper

Methodology:

Utility function:

When vehicles operate on the road, drivers are often influenced by the surrounding traffic environment. Factors such as sudden traffic accidents, congestion, erratic behavior of nearby vehicles, or even road closures can significantly affect a driver's control over the vehicle. These external disturbances may lead to abnormal driving behaviors, including but not limited to: speeding, consistent deviation from the lane center, short headway and lateral distance, sudden acceleration or sharp deceleration, and frequent lane changes. Therefore, we set the following utility functions to capture the potential driving behavior risk:

- **Speeding penalty $\boldsymbol{\phi_1}$:**

This term penalizes vehicles whose speed deviates from the ideal speed.

$$\phi_1 = w_1 \left(\frac{v - v_0}{v_0}\right)^2,$$

where $w_1 = 1$, v_0 is the ideal speed, which is the speed of most vehicles in the current lane.

- **Hard acceleration or braking penalty $\boldsymbol{\phi_2}$:**

This term penalizes vehicles that have extreme speed changes.

$$\phi_2 = w_2 \left(log\left(1 + e^{k_1(a - a_{max})}\right) + log\left(1 + e^{-k_1(a - a_{min})}\right)\right),$$

where $w_2 = 1$, $k_1 = 15$, and $a_{max}, a_{min}$are determined based on the 3σ principle, assuming acceleration follows a Gaussian distribution.

- **Lane deviation penalty $\boldsymbol{\phi_3}$:**

This term penalizes deviation from the center of the lane for vehicles.

$$\phi_3 = w_3 min\left(\frac{(d^2 - d_0^2)^2}{\frac{3 \times W_l^4}{4}}, 1\right),$$

where $w_3 = 1, d_0$ is center of the lane and W_l is width of the lane equaling to 12ft.

- **Longitudinal and Lateral collision penalty $\boldsymbol{\phi_4}$:**

This term penalizes collisions between vehicles.

ϕ_{4_1} represent the longitudinal collision penalty for vehicles in the same lane.

$$\phi_{4_1} = \frac{1}{1 + e^{\left(k_2(d_x + l_x)\right)}} - 0.5 + \frac{1}{1 + e^{\left(k_2(l_x - d_x)\right)}} - 0.5,$$

ϕ_{4_2} represent the lateral collision penalty for vehicles in the different lanes.

$$\phi_{4_2} = \frac{1}{1 + e^{\left(k_3(d_y + l_y)\right)}} - 0.5 + \frac{1}{1 + e^{\left(k_3(l_y - d_y)\right)}} - 0.5,$$

$$\phi_4 = w_4 \times \phi_{4_1} \times \phi_{4_2},$$

where $w_4 = 1$, $k_2 = 0.5$, $k_3 = 9$, $l_x = 10$, $l_y = 2$, and dx, d_y denote the distance between vehicle edges.

$$d_x = \Delta x - \frac{(L_c - L_p)}{2}, \Delta x = L_p - x,$$

$$d_y = \Delta y - \frac{(W_c - W_p)}{2}, \Delta y = L_p - y,$$

where L_c, L_p are the length of the current vehicle and the previous vehicle, respectively. W_c, W_p are the width of the current vehicle and the previous vehicle, respectively. $\Delta x, \Delta y$ are distance between the length and width of adjacent vehicles.

- **Frequent lane changes penalty $\boldsymbol{\phi_5}$:**

This term penalizes frequent lane changes of vehicles.

$$\phi_5 = w_5 \begin{cases} 0, n = 0 \\ 0.5, n = 1, \\ 1, n \geq 2 \end{cases}$$

where $w_5 = 1$ and n is frequency of vehicle lane changes.

When a vehicle driver behaves following common driving patterns, the penalties for all utility functions are zero. Deviations from these typical patterns result in penalties of one or higher. These utility functions help determine which driving features are associated with higher levels of risk. The next step involves inputting these features into a GRU-AE to extract latent variables and anomaly detection.

GRU-AE:

In the framework, the utility functions are inputted as the encoder of a GRU-based autoencoder. Specifically, we adopt a GRU-based autoencoder, where the encoder leverages GRUs to capture the temporal patterns of vehicle behavior. GRU-AE is a neural network model that combines GRUs and an Autoencoder architecture, primarily used for feature learning and dimensionality reduction in sequential data.

As illustrated in Figure 2, GRU-AE consists of two components: GRU and Autoencoder. The GRU, a variant of RNNs, has a simpler structure than LSTM but efficiently captures long-term dependencies in sequences using update and reset gates. The Autoencoder includes an encoder and a decoder: the encoder, which compresses the input sequence into a low-dimensional latent representation, and the decoder, which reconstructs the original sequence from this compressed form.

The detailed computation within a GRU unit is described as follows:

$$r_t = \sigma(W_r[h_{t-1}, x_t]),$$
$$c_t = \sigma(W_c[h_{t-1}, x_t]),$$
$$\widetilde{h_t} = tanh(W_{\widetilde{h}}[r_t * h_{t-1}, x_t]),$$
$$h_t = c_t * h_{t-1} + (1 - c_t) * \widetilde{h_t},$$

where x_t is the input at time t. h_{t-1}t is he hidden state from the previous time step. r_t is the reset gate. c_t is the update gate. $\widetilde{h_t}$ is the candidate hidden state. h_t: the updated hidden state.

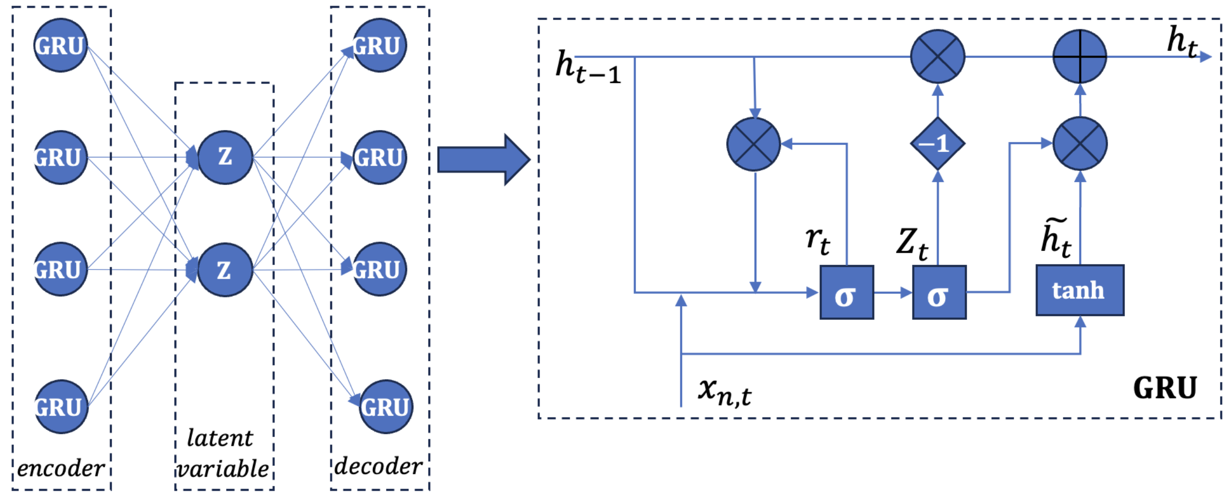

Figure 2. The structure of GRU-AE.

The encoder processes the input sequence step-by-step using GRU layers, ultimately generating a low-dimensional latent variable that has essential features of the sequence. The decoder, also based on GRU, reconstructs the original sequence from this latent variable by minimizing the difference between the reconstructed output and the original input.

Gaussian mixture model:

GMM is a probabilistic model that represents data as a combination of multiple Gaussian (normal) distributions, allowing for flexible modeling of complex datasets. Unlike a single Gaussian distribution, which assumes data is clustered around one mean, a GMM captures scenarios where data arises from several distinct subpopulations, each with its mean μ and variance σ, as shown in Figure 3.

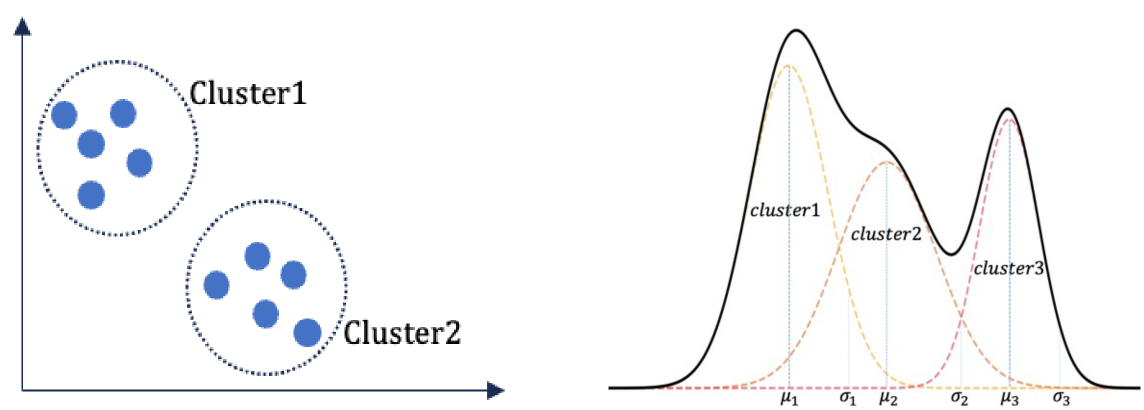

Figure 3. The structure of GMM.

Specifically, a GMM describes the overall data distribution as a weighted sum of K Gaussian components. Each component models a subset of the data and contributes to the total probability density based on its associated weight. The GMM captures complex, multimodal distributions more effectively than a single Gaussian distribution by combining multiple Gaussian components. The probability density function of GMM is given by:

$$p(x|\theta) = \sum_{k=1}^{K} \pi_k p(x|\theta_k),$$

where π_k is the mixing coefficient for the kth Gaussian component, $\sum_{k=1}^{K} \pi_k = 1$ and $0 \leq \pi_k \leq 1$.θ_k is the parameters involving mean and covariance, $\theta_k = \mu_k, \Sigma_k$.

Each component $p(x|\theta_k)$ is a multivariate Gaussian distribution, the mentioned parameters can be estimated by the expectation maximization (EM) method:

$$p(x|\theta_k) = \frac{1}{(2\pi)^{(m/2)}|\Sigma_k|^{(1/2)}} e^{\left(-0.5(x-\mu_k)^T \Sigma_k^{(-1)}(x-\mu_k)\right)},$$

where $x \epsilon R^m$ is input vector. $\mu_k \epsilon R^m$ is mean of kth component. $\Sigma^k \epsilon R^{mxm}$ is the covariance matrix. $|\Sigma_k|$ is the determinant of the covariance matrix.

Overall, within the framework, the GMM enables soft clustering, allowing each data point to belong to multiple clusters with varying probabilities. This probability method is a key advantage and the primary reason we adopt GMM for defining a dynamic threshold in anomaly detection.

GMM-GRU-AE:

Figure 4 shows the framework of the whole process. High-dimensional time-series data are first encoded using a GRU-based autoencoder. The GRU encoder captures temporal dependencies and compresses the sequence into a low-dimensional latent variable Z, which summarizes the essential features of the input. This latent vector is then modeled using a GMM, which estimates the underlying distribution through a set of parameters including the π_k, μ_k, and Σ_k of each Gaussian component. Meanwhile, the decoder reconstructs the original input sequence from the latent vector, and the reconstruction quality is measured using the Squared Prediction Error (SPE), computed as the squared difference between the original and reconstructed sequences. A common measure of this reconstruction error is the Mean Squared Error (MSE), defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \widetilde{x_i})^2.$$

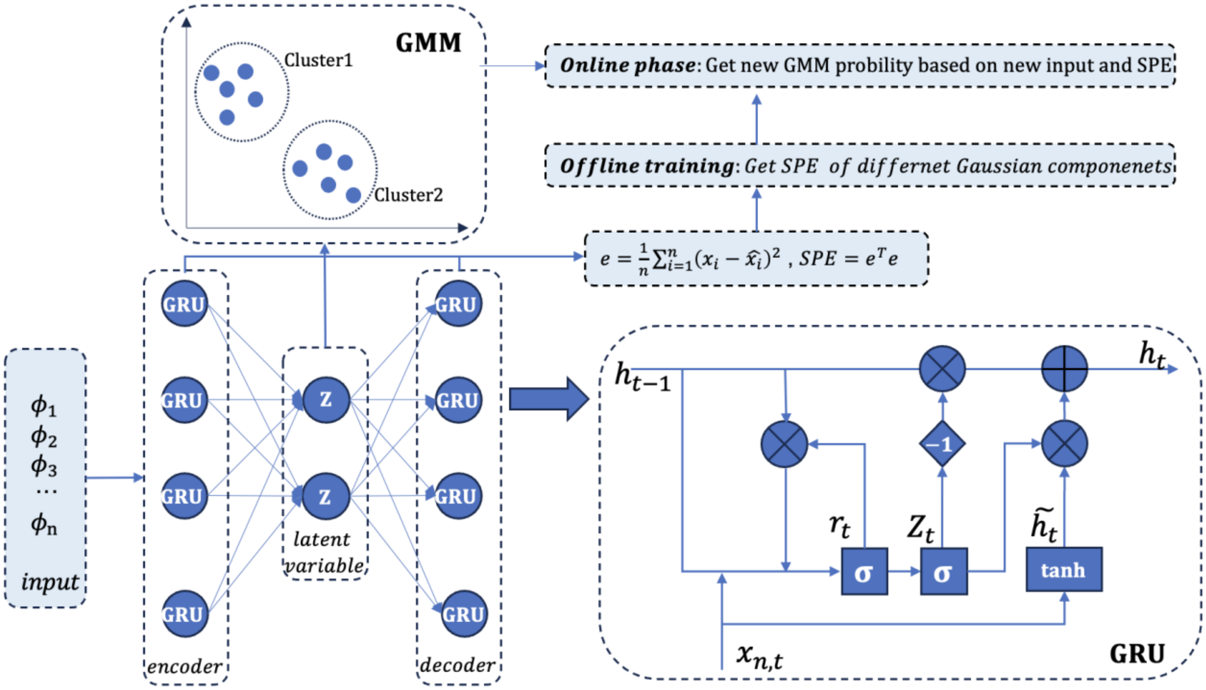

Figure 4. The structure of GMM-GRU-AE.

SPE can be used to evaluate the reconstruction error at each individual timestep. Specifically, for the ith input, the SPE is computed as:

$$SPE_i = e_i^T e = (x_i - \hat{x}_i)^2 = \left(x_i - f(W^T h_t)\right).^2$$

MSE can be expressed as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} SPE_i = E(SPE).$$

Offline training and online phase are two important parts of our framework. During offline training, the model learns latent behavior representations and reconstructs input data to compute SPE values. These values are clustered using GMM, and the component SPE thresholds are determined via cumulative distribution functions to support adaptive anomaly detection during the online phase. In the online phase, each new input is first encoded into a latent representation, which is then evaluated using the trained GMM to obtain its individual probabilities across the Gaussian components. These probabilities are used to compute an adaptive, weighted SPE threshold tailored to the current input. The actual SPE of the input is then compared with this threshold to determine whether the observed behavior is normal or abnormal. The specific steps of offline training and the online phase are shown in Table 2 and

Table 3.

Table 2. Offline training.

Offline Training
1. Normalize the input sequence X, then feed it into the GRU encoder to obtain the latent representation Z 2. Apply GMM clustering to Z, and extract the parameters of each Gaussian component: the weight π, mean μ and covariance matrix Σ for each component i 3. Reconstruct Z using a single fully connected layer and compute the reconstruction error e. 4. Calculate the SPE for each sample using $SPE = e^T e$. 5. Based on the GMM clustering results, use the cumulative distribution function (CDF) to determine the SPE threshold SPE_{UCL} for each Gaussian component i.

Table 3. Online phase.

Online phase
1. Normalize the new input x_{new} and feed it into the GRU encoder to obtain the latent representation z_{new}. 2. Use the GMM to compute the probability z_{new} that belongs to the j_{th} Gaussian component, calculated as: $\mathrm{p_j} = \frac{\pi_j\, p(z_{new} \mid \theta_j)}{\sum_{i=1}^{k} \pi_i\, p(z_{new} \mid \theta_i)}$ 3. Reconstruct z_{new} using a single fully connected layer, and compute the reconstruction error e_{new}. 4. Calculate the Squared Prediction Error (SPE) for the new sample using: SPE new $= e_{new}^T e_{new}$ 5. Determine the control threshold for SPE new by computing a weighted sum of the component-wise thresholds SPE_{UCL} ,weighted by p_j. $$SPE_{UCLnew} = \sum j = 1^k p_j \ \mathrm{SPE}_{\mathrm{UCL}_j}$$ 6. If SPE new $\geq SPE_{UCLnew}$ at time step t, calculate the increment of the reconstruction error as: $\Delta e = \lvert e_t - e_{t-1} \rvert$

Results

Environment setting

Dataset

To evaluate the performance of the proposed anomaly detection framework, we use HighSim dataset, considering the massive volume of data. A total of 1,200 vehicles were selected, comprising 4,068,625 rows of time-series data. The dataset was split into two parts: 1,000 vehicles were used for training, and the remaining 200 vehicles were reserved for testing and evaluation.

GRU-AE-GMM framework

To model temporal sequences in driving behavior, we first applied a GRU-AE to encode and reconstruct the vehicle time-series data. The GRU encoder extracted the latent variable of driving patterns. These latent variables were then modeled using a GMM in the latent space to represent clusters of normal driving behaviors. After encoding, the GRU decoder reconstructed the input sequences. SPE was computed between each original input and its reconstruction to measure how well the model captured the temporal dynamics. Higher SPE values indicated greater reconstruction error, suggesting potential anomalies.

Online detection

For real-time anomaly detection, we adopt a dynamic thresholding mechanism based on GMM probabilities. To further enhance the temporal resolution and responsiveness of the detection process, we implement an adaptive sliding window strategy, in which both the window size and step size vary dynamically according to the likelihood of abnormal behavior.

During periods of stable driving, a larger window size and longer step (e.g., 90 frames with a stride of 30) are used to capture long-term behavioral patterns while minimizing computational cost. However, when the system detects a rising trend in reconstruction error, indicating a potential deviation from normal behavior, the window size is progressively reduced (e.g., 60 or 30 frames), enabling the model to focus on shorter and more recent temporal segments. At the same time, the step size is decreased to increase the overlap between windows, thereby improving temporal granularity in the region of suspected anomalies.

For each time window, an adaptive threshold is calculated as a weighted upper control limit (UCL), where the weights are based on the GMM-derived probabilities of the input belonging to each Gaussian component. A window is flagged as anomalous if its SPE exceeds this adaptive threshold.

Figure 5 illustrates the anomaly detection results for a subset of testing vehicles (Vehicle IDs 80 to 90). The blue line represents the SPE of each window, the red dashed line indicates the dynamic threshold, and the black crosses denote the windows where the SPE exceeds the threshold, and an anomaly is detected. The model demonstrates strong temporal sensitivity, effectively identifying points in the sequence where reconstruction quality deteriorates sharply. This dynamic and probabilistic approach enhances both the responsiveness and robustness of the anomaly detection system in real-world conditions.

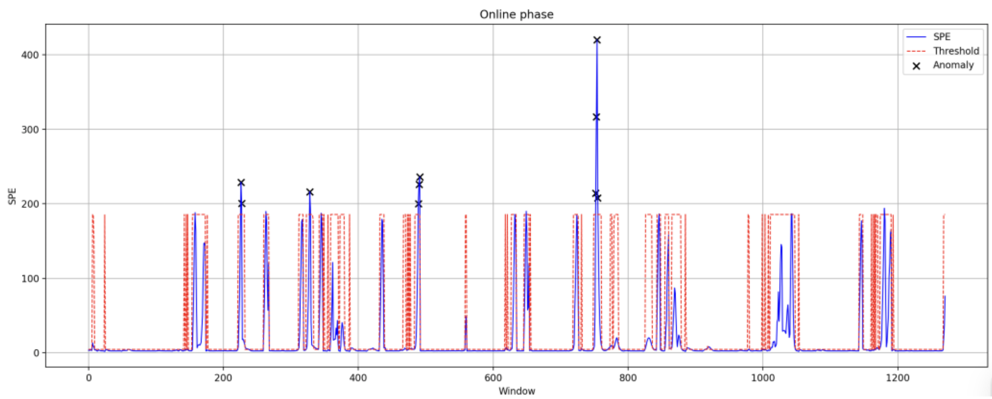

Figure 5. Anomaly points of testing vehicles.

Figure 5 also illustrates the distribution of SPE statistics across different vehicles, showing variability in reconstruction quality and potential abnormality over time. To demonstrate this variability, we selected a subset of vehicle IDs from 80 to 90 for a detailed view shown in Figure 6.

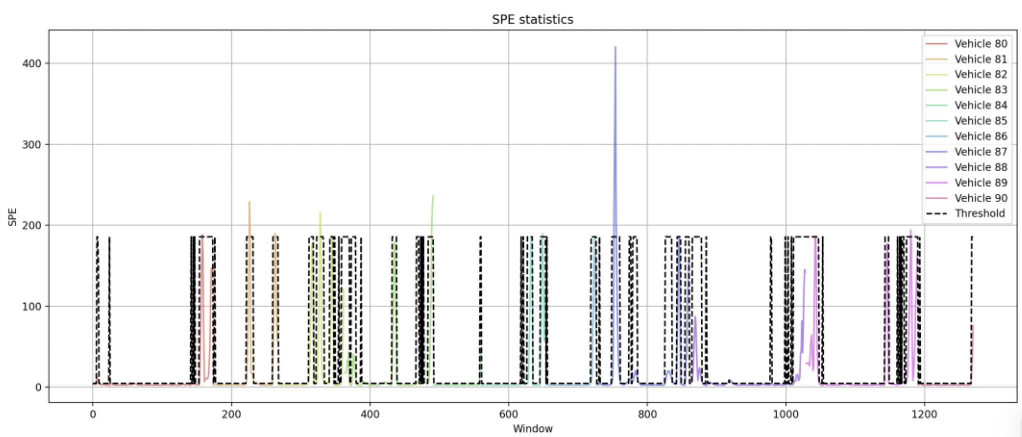

Figure 6. SPE statistics of different vehicles.

Utility function Comparison:

As shown in Figure 7, multiple utility dimensions such as ϕ_2(hard acceleration/deceleration penalty) and ϕ_5(frequent lane changing penalty) exhibit significant spikes in short periods. These sudden increases suggest aggressive acceleration and lateral movements that deviate from normal driving behavior.

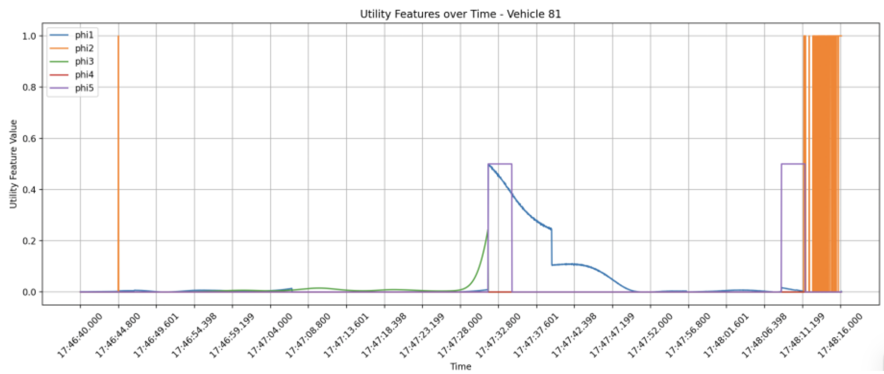

Figure 7. Utility value of vehicle 81.

Figure 8 further illustrates the relationship between ϕ_1 (speeding penalty) and ϕ_5 (frequent lane-change penalty). The observed abrupt transitions, such as simultaneous spikes in speed violations and lane-change penalties, indicate that drivers tend to accelerate sharply or exceed speed limits during lane changes. These sudden behavioral shifts align closely with the anomaly windows detected by the model, which is consistent with real-world driving experiences. This empirical consistency provides indirect evidence supporting the effectiveness of the proposed anomaly detection framework.

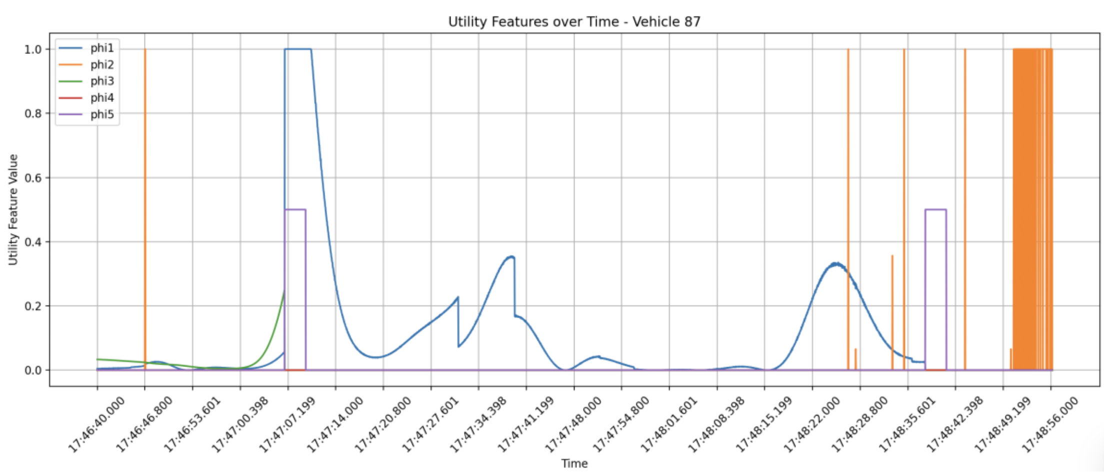

Figure 8. Utility value of vehicle 87.

Discussion

This project's results confirm that the GRU-AE-GMM model effectively detects anomalies in driving behavior using HighSim trajectory data. Traditional approaches often rely on pre-defined rules (e.g., speed limits, headway thresholds) or supervised learning with labeled anomalies, which are expensive and often incomplete. In contrast, our GRU-AE-GMM model learns behavioral patterns directly from the data, eliminating the need for hand-crafted features or ground-truth labels. This makes it highly scalable and adaptable to different road scenarios.

By capturing the time-based relations among vehicle utility features and applying probabilistic clustering in the latent space, the system autonomously learns to identify normal and abnormal behaviors without the need for labeled training examples. The use of SPE with adaptive thresholds is context sensitive, allowing for situational adjustment, which is crucial in heterogeneous traffic conditions due to road layout, surrounding vehicles, or driving habits. Employing GMM within the latent space enhances interpretability and flexibility as each component signifies a typical mode of operation. The dynamic thresholding based on GMM components' probabilities aligns with risk factors and thereby enhances robustness. The fact that many detected anomalies correspond to commonly recognized forms of dangerous driving behavior, like sharp acceleration, frequent lane changes, and inadequate headway.

In general, the method offers a robust and unsupervised framework for identifying abnormal driving behavior. From the perspective of VRUs such as pedestrians and cyclists, the ability to detect and anticipate abnormal vehicular behavior is crucial. By identifying vehicles that demonstrate erratic or aggressive patterns early, this model could be integrated with vehicle-to-everything (V2X) systems to issue real-time warnings to nearby road users or adjust traffic signals to improve safety.

Conclusions and Recommendations

This research develops an unsupervised anomaly detection framework employing GRU-AE and GMM for abnormal vehicle behavior recognition using time-series data from the HighSim dataset. The sequential driving data were modeled with a GRU-based autoencoder to capture the underlying latent behavioral patterns, which were then clustered using a Gaussian Mixture Model. The reconstruction errors, specifically the SPE, computed, provided an adequate measure for quantifying deviations from normative patterns. Anomalies were detected adaptively across varying driving behaviors by applying dynamic, probabilistically weighted thresholds based on GMM. Findings illustrate that such approaches can detect abnormalities in driving patterns, including sudden acceleration, significant lane deviation, close headway with other vehicles, and excessive tailgating and frequently changing lanes without depending on labeled training data. This characteristic makes such methods particularly useful in large-scale real-world applications where labeled anomalies are rare. This unsupervised characteristic increases the method's applicability to real-world large-scale scenarios, where labeled anomalies are scarce. Detection robustness is further enhanced using the adaptive thresholding approach, which accounts for variability among different vehicles.

Looking ahead, there may be an opportunity to validate the detected anomalies against some ground truth data—perhaps manually assigned or generated through controlled test scenarios—to assess how accurately these methods label discrepancies through false positives and false negatives. It would, however, be useful to assess how this approach performs in comparison with other statistical methods or supervised neural networks so that its strengths and weaknesses can be neatly quantified relative to other techniques. This also assists us in enabling the possible adaptation of the approach to the crosswalk situation.

References

National Highway Traffic Safety Administration (NHTSA).(2025). https://www.nhtsa.gov/press-releases/nhtsa-2023-traffic-fatalities-2024-estimates.

Wu, X., Boyle, L.N., Marshall, D. and O'Brien, W., 2018. The effectiveness of auditory forward collision warning alerts. Transportation research part F: traffic psychology and behaviour, 59, pp.164-178.

Ge, Y., Qu, W., Jiang, C., Du, F., Sun, X., & Zhang, K. (2014). The effect of stress and personality on dangerous driving behavior among Chinese drivers. *Accident Analysis & Prevention*, *73*, 34-40.

Mohamad, I., Ali, M. A. M., & Ismail, M. (2011, July). Abnormal driving detection using real time global positioning system data. In *Proceeding of the 2011 IEEE international conference on space science and communication (IconSpace)* (pp. 1-6). IEEE.

Bergasa, L. M., Almería, D., Almazán, J., Yebes, J. J., & Arroyo, R. (2014, June). Drivesafe: An app for alerting inattentive drivers and scoring driving behaviors. In *2014 IEEE Intelligent Vehicles symposium proceedings* (pp. 240-245). IEEE.

Levi-Bliech, M., Kurtser, P., Pliskin, N., & Fink, L. (2019). Mobile apps and employee behavior: An empirical investigation of the implementation of a fleet-management app. *International Journal of Information Management*, *49*, 355-365.

Ku, W., Storer, R. H., & Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and intelligent laboratory systems*, *30*(1), 179-196.

Becker, T., & White, P. R. (2006). Independent Component Analysis using Gaussian Mixture Models.

Li, N., Jain, J. J., & Busso, C. (2013). Modeling of driver behavior in real world scenarios using multiple noninvasive sensors. *IEEE transactions on multimedia*, *15*(5), 1213-1225.

Zhang, M., Chen, C., Wo, T., Xie, T., Bhuiyan, M. Z. A., & Lin, X. (2017). SafeDrive: Online driving anomaly detection from large-scale vehicle data. *IEEE Transactions on Industrial Informatics*, *13*(4), 2087-2096.

Eren, H., Makinist, S., Akin, E., & Yilmaz, A. (2012, June). Estimating driving behavior by a smartphone. In *2012 IEEE Intelligent Vehicles Symposium* (pp. 234-239). IEEE.

Hu, J., Zhang, X., & Maybank, S. (2020). Abnormal driving detection with normalized driving behavior data: A deep learning approach. *IEEE transactions on vehicular technology*, *69*(7), 6943-6951.

Nanduri, A., & Sherry, L. (2016, April). Anomaly detection in aircraft data using Recurrent Neural Networks (RNN). In *2016 Integrated Communications Navigation and Surveillance (ICNS)* (pp. 5C2-1). Ieee.

Tüfekci, G., Kayabaşı, A., Akagündüz, E., & Ulusoy, İ. (2022, October). Detecting driver drowsiness as an anomaly using lstm autoencoders. In *European Conference on Computer Vision* (pp. 549-559). Cham: Springer Nature Switzerland.

Tang, C., Xu, L., Yang, B., Tang, Y., & Zhao, D. (2023). GRU-based interpretable multivariate time series anomaly detection in industrial control system. *Computers & Security*, *127*, 103094.

Taylor, A., Leblanc, S., & Japkowicz, N. (2016, October). Anomaly detection in automobile control network data with long short-term memory networks. In 2016 IEEE international conference on data science and advanced analytics (DSAA) (pp. 130-139). IEEE.

Niu, Z., Yu, K., & Wu, X. (2020). LSTM-based VAE-GAN for time-series anomaly detection. Sensors, 20(13), 3738.

Zhan, K., Wang, C., Zheng, X., Kong, C., Li, G., Xin, W., & Liu, L. (2025). Seq2Seq-based GRU autoencoder for anomaly detection and failure identification in coal mining hydraulic support systems. *Scientific Reports*, *15*(1), 542.

Shi, X., Zhao, D., Yao, H., Li, X., Hale, D. K., & Ghiasi, A. (2021). Video-based trajectory extraction with deep learning for High-Granularity Highway Simulation (HIGH-SIM). *Communications in transportation research*, *1*, 1000