Integrating Data and Rules: A Hybrid Approach for Robust Lane Change Intention Prediction under Distribution Shift

Yao Long Teng
School of Computer Science and Engineering, Nanyang Technological University
Singapore
teng0144@e.ntu.edu.sg

Htet Naing
School of Computer Science and Engineering, Nanyang Technological University
Singapore
htetnain001@e.ntu.edu.sg

Wentong Cai
School of Computer Science and Engineering, Nanyang Technological University
Singapore
aswtcai@ntu.edu.sg

ABSTRACT

Lane change intention prediction is critical in improving road safety. In general, this task is achieved via two distinct approaches: The data-based approach and the rule-based approach. Generally, the former outperforms the latter under in-distribution scenarios. However, under out-distribution scenarios, data-based methods tend to perform extremely poorly. Rule-based methods are robust to such problems since they are formulated based on theory. Instead of relying exclusively on each approach, it is more advantageous to combine them to exploit both of their strengths. Thus, in this paper a physics guided hybrid model is proposed for lane change intention prediction to improve model robustness under distribution shift. The effectiveness of the proposed model is shown in the experimental results where the proposed model outperforms rule-based models under in-distribution scenarios by 4%, and outperforms both rule-based models and data-based models in out-distribution scenarios by 3% and 60% respectively.

CCS CONCEPTS

• **Computing methodologies → Modeling methodologies**.

KEYWORDS

Lane Change, Distribution Shift, Physics-guided Machine Learning, Rule-based, Data-based

ACM Reference Format:
Yao Long Teng, Htet Naing, and Wentong Cai. 2024. Integrating Data and Rules: A Hybrid Approach for Robust Lane Change Intention Prediction under Distribution Shift. In *Proceedings of In ACM SIGSIM Conference on Principles of Advanced Discrete Simulation (ACM SIGSIM PADS '24)*. ACM, New York, NY, USA, 10 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Predicting lane change intentions accurately is a cornerstone in the quest to enhance road safety. Official data suggests that a minimum of 33% of all road accidents occur when vehicles change lanes or exit the road [10]. As such, predicting lane change intentions accurately can significantly reduce the risk of collisions. Additionally in the domain of autonomous driving, it is especially crucial for intelligent driving vehicles to comprehend and anticipate shifts in the behavior of human-operated vehicles, as this significantly impacts driving decisions [1]. Furthermore in the simulation context, more accurate simulations can contribute greatly in safety analysis and risk assessment studies. By understanding when and why drivers are likely to change lanes, simulations can identify high-risk scenarios and contribute to the development of safety protocols and policies, advancing in the testing and development of advanced driver assistance systems (ADAS) [14, 30].

In literature, the task of lane change intention prediction can be achieved mainly with two distinct methods [16]: 1) rule-based and 2) data-based models. Rule-based approaches, such as the *Minimising Overall Braking Induced by Lane Change* (MOBIL) model by Kesting et al [15], are grounded in theoretical principles and strive to pinpoint the basic underlying mechanisms that prompt a driver to initiate a lane change. On the other hand, data-driven methodologies, such as neural networks, operate without any modeling assumptions, relying instead on learning complex patterns from extensive datasets [26].

Data-based methods can be refined and improved over time with more data. They are also better at dealing with complex relationships, and can automatically derive features from data, reducing the need for domain expertise. However, many of such models are not interpretable, making it hard to understand decision-making processes due to their black box nature. Additionally, these models are sensitive to the distribution of the data they are trained on, and performance degrades heavily when faced with data from different distributions [5]. This is especially problematic as autonomous vehicles continually encounter shifts in data distribution when navigating different cities and experience varying weather conditions [4]. Conversely, rule-based methods rely on a set of explicit rules defined by experts. These rules are based on domain knowledge and are often deterministic. Such models are interpretable since they are formulated based on theory, and have stable performance unless the rules are modified. However, these models often underperform compared to data-driven methods in complex tasks under in-distribution scenarios.

To address the aforementioned challenges, it is therefore desirable to formulate a new approach that has competitive and stable performance regardless of being in or out distribution, while retaining most of its analytical explainability. To the best of our knowledge, none of the existing works have attempted:

- To investigate robustness of lane change intention prediction models under distribution shift.
- To improve robustness of lane change intention prediction models under distribution shift.
- To combine rule-based and data-based methods to model lane change intention prediction.

Thus, our main contributions can be summarised as follows:

(1) MOBIL and a physics-guided neural network are integrated to create a hybrid model.
(2) Performance of various lane change intention prediction models under distribution shift were investigated.
(3) Experiments were conducted based on a real-world dataset to show the effectiveness of our model.

The structure of the paper is as follows: Section 2 provides an overview of related works found in existing literature. Section 3 introduces our proposed model. In Section 4, the details and analysis of our experimental results are presented . The paper is then concluded in Section 5.

2 RELATED WORKS

2.1 Rule-based models

Rule-based lane change intention prediction models offer valuable insights into driver behavior by predicting whether a driver intends to change lanes in the near future. This crucial data enhances traffic flow modeling and supports the development of intelligent transportation systems. As mentioned before, rule-based models rely on predefined rules and logic derived from traffic theory, empirical data analysis, and expert knowledge. These rules consider various factors influencing lane change decisions, including gap acceptance, lane selection and driver heterogeneity [7]. Early models like Gipps [12] laid the foundation for rule-based lane change prediction. The model uses a set of predefined rules to determine whether a driver will change lanes. However, Gipps' model primarily focuses on the behavior of an individual driver and the immediate factors influencing their decision to change lanes.

In 2007, Kesting et al introduced the MOBIL model [15], a notable development in rule-based prediction. MOBIL combined the advantages and risks associated with lane changes into an acceleration utility function derived from car-following models. Its design facilitated seamless incorporation into current traffic simulation systems and provided a more inclusive method for making lane change decisions. MOBIL is an improvement to Gipps in several ways as MOBIL considers both the benefits (e.g., reaching desired speed) and risks (e.g., potential braking) of lane changes, providing a more balanced decision-making process compared to Gipps' deterministic rules. Additionally, MOBIL uses an acceleration function to represent lane change decisions, making it more flexible and adaptable to different traffic situations. Moreover, MOBIL allows for incorporating driver heterogeneity by adjusting the politeness parameter in the model to capture different driving styles and preferences.

2.2 Data-based models

The recent surge in data availability and advancements in machine learning have ushered in a new era of data-driven lane change intention prediction. Such models learn directly from large amounts of data, allowing the model to learn complex patterns and relationships from the data itself. However, akin to rule-based models, the inputs typically include distance, speed difference, and other spatial-temporal variables related to the ego vehicle and its neighbours [16]. Notable examples of machine learning algorithms for predicting lane change intention include Random Forests, Support Vector Machines and feed-forward Neural Networks [6, 9].

In general, deep-learning models outperform other models by a fair margin due to their ability to learn from the data well [31, 32]. However, this is a double-edged sword as relying on patterns learnt from data coming from a particular distribution would lead to the inability to generalise to data from another distribution even though both sets of data are representations of the same problem. Current research on lane change intention prediction models focuses on in-distribution scenarios only, which is a research gap we address in this paper.

2.3 Physics-guided machine learning

Physics-guided machine learning (PGML) is an emerging field that integrates principles from physics with machine learning techniques. This interdisciplinary approach aims to enhance the performance and interpretability of machine learning models by incorporating domain-specific knowledge from physics. Such techniques have been employed successfully in various fields such as that of blast protection engineering [25], structural dynamics simulation and many others [11, 29].

Physics-guidance allows for the reduced dependence on datasets, while making the models more interpretable, as a layer of understandable and verifiable logic is added. Common ways to integrate physics include constraint-based integration [19], where physical constraints are directly imposed in the learning algorithm and physics-informed feature engineering [20], where physical insights are used to guide the selection or creation of features for machine learning models. Despite its adoption in car-following [22], no such method is used for lane changing. Hence we fill this gap.

3 PROPOSED LANE CHANGE INTENTION PREDICTION MODEL

Since we aim to create a model that is not only robust to distribution shift but is also analytically explainable, the foundation of our model needs to be rule-based. For this paper, the lane changing model MOBIL will be our area of focus. For this section, we first introduce the various models and its related components to be used in our hybrid model, followed by how we combine these models to form our proposed hybrid model.

3.1 Minimising Overall Braking Induced by Lane Change

The MOBIL model [15] is essential in traffic simulation, analyzing vehicular lane-changing behaviors. It optimizes lane changes to minimize braking, enhancing traffic flow and ensure safety. The model has three key components: (i) the safety criterion, (ii) incentive criterion and (iii) politeness factor. To fulfil the safety criterion, MOBIL ensures that a lane change does not result in significant deceleration for the following vehicle in the target lane. This is quantified using a safety threshold parameter:

$$\tilde{a} \geq -b_{safe} \tag{1}$$

where $\tilde{a}$ is the deceleration of the follower in the target lane and b_{safe} is the safe limit . The incentive criterion and politeness factor is quantified by the symmetric lane-changing rule (Refer to the original work [15] for more details):

$$\tilde{a_{ego}} - a_{ego} + p\left(\tilde{a_{new}} - a_{new} + \tilde{a_{old}} - a_{old}\right) \geq a_{thr} \tag{2}$$

where:

a_{ego}: current acceleration of the ego vehicle
$\tilde{a_{ego}}$: new acceleration of the ego vehicle after lane change
$\tilde{a_{new}}$: acceleration of the ego's new follower after lane change
a_{new}: acceleration of the ego's new follower in the target lane before lane change
$\tilde{a_{old}}$: acceleration of the ego's old follower after lane change
a_{old}: acceleration of the ego's old follower before lane change
a_{thr}: acceleration threshold

The MOBIL lane change model uses the Intelligent Driver Model (IDM) as its car following model [27] to determine vehicle accelerations on the current and target lane. The incentive criterion formulated by the acceleration functions, generally assesses whether a lane change can enhance a driver's traffic conditions. Equation (2) can be understood as such: if the personal incentive of the ego vehicle exceeds the weighted sum of the disincentives of both the new and old following vehicles by a predefined threshold, a lane change occurs.

3.2 Intelligent Driver Model

The IDM [27] characterizes how a driver adjusts their vehicle's acceleration based on the front gap distance and relative velocity to the vehicle ahead. The IDM can be quantified by the following equations:

$$acc_{t+1} = acc_{max}\left[1 - \left(\frac{v_t}{v_0}\right)^4 - \left(\frac{s^*(\Delta v_t, v_t)}{s_t}\right)^2\right] \tag{3}$$

where:

$$s^*(\Delta v_t, v_t) = s_0 + \max\left(0, v_t T + \frac{v_t \Delta v_t}{2\sqrt{acc_{\max} dacc_{\max}}}\right) \tag{4}$$

The IDM outputs the acceleration in the next time step acc_{t+1} of a follower vehicle by accepting the follower's current velocity v_t, the front gap distance and relative velocity with respect to the leader s_t and Δv_t respectively, and the minimum desired gap distance $s^*(\Delta v_t, v_t)$ given by Equation (4) as input.

Additionally, the set of model parameters to be calibrated are as follows:

- v_0:the maximum desired velocity
- T: the desired time headway
- s_0: the jam spacing distance
- acc_{max}: the maximum desired acceleration
- $dacc_{max}$: the maximum desired deceleration

3.3 Physics-guided acceleration prediction

In Microscopic Traffic Simulation, physics-guided machine learning is commonly used to inject theory into data-based models [21, 22]. This is commonly done with a physics-guided loss function [23] that consists of two components:

(1) Data loss - the mathematically quantified difference between predicted and actual values in a data-based model.
(2) Physics loss - the mathematically quantified difference between predicted values of the data-based model and the predicted values of a physics-based model .

As mentioned previously, it is generally established that data-based models outperform other types of models due to their ability to leverage large amounts of data in their learning process. However, a major weakness of such data-based models is its inability to perform satisfactorily under out-distribution scenarios. Physics models on the other hand are relatively robust to different data distributions due to the fact that they are formulated based on theory, which is invariant under different distributions. Injecting theory into data-based models with physics loss helps to find a middle ground by regularising the data-based model from over-fitting to the training distribution.

The physics-guided loss function is defined as:

$$\mathcal{L}(\theta) = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \left(a_i^{\text{data}} - a_i\right)^2 + \lambda \left(a_i^{\text{physics}} - a_i\right)^2 \tag{5}$$

where a_i^{physics} is defined by Equation (3), a_i^{data} is the true label from the dataset, a_i the prediction from the data-based model of choice and $\lambda \in [0, 1]$ controls the amount of weight to put on the physics loss. λ can be interpreted as the amount of theory we wish to inject into our data-based model, where a larger λ means a larger contribution of the physics loss towards the total loss which leads to greater regularisation of the data-based model. The physics guided loss function is then the sum of squared error between the predictions of the data-based model with respect to the actual values and the predictions of the physics model, averaged over the number of samples $\mathcal{N}$.

In our paper, a basic feed-forward neural network with input features similar to the IDM is used as the data-based model for acceleration prediction. The loss function used is the physics-guided loss function as explained in Equation (5) to increase robustness of the model to distribution shift.

3.4 Physics-guided rule-data hybrid model

As mentioned before, the foundation of our model needs to be rule-based since we aim to create a model that is not only robust to distribution shift but also retains most of its analytical explainability. Additionally, we also aim to leverage the effectiveness of data-based methods in learning complex patterns from data. Thus, it is beneficial to combine the MOBIL model with a neural network by incorporating the neural network as its acceleration prediction component. However, neural networks are prone to overfitting. To mitigate this, we utilise concepts from physics-guided machine

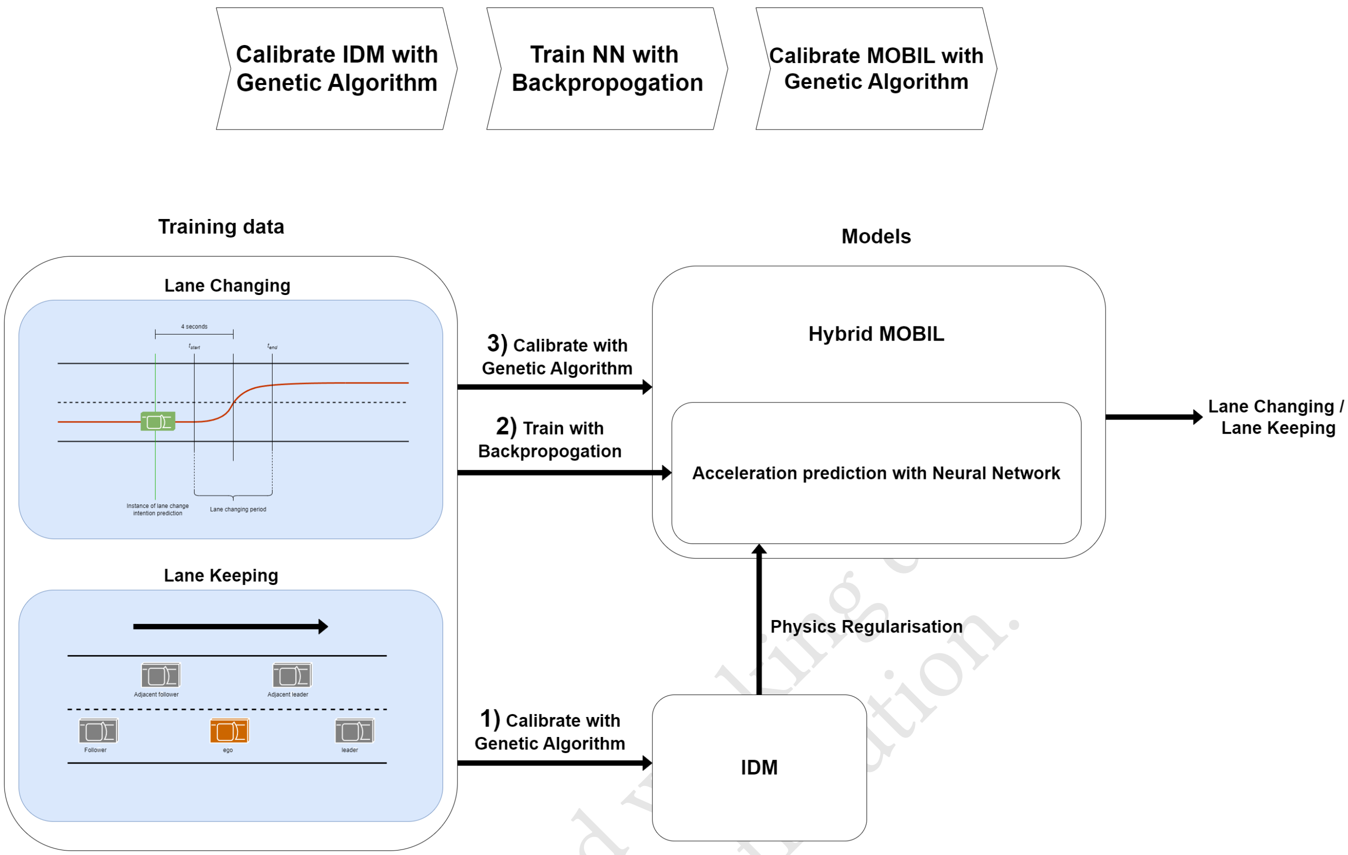

Figure 1: Training of physics-guided rule-data hybrid model

learning for physics-guided acceleration prediction explained in Section 3.3.

Putting it all together, the overview of our physics-guided hybrid model (PG-hMOBIL) is shown in Figure 1. The framework of the model can be grouped mainly into three steps:

(1) Pre-calibration of IDM with Genetic Algorithm
(2) Training the physics-guided neural network for acceleration prediction
(3) Post-calibration of PG-hMOBIL

The Genetic Algorithm has been widely used for the calibration of various traffic simulation models [8, 17]. Without loss of generality, we use the Genetic Algorithm to calibrate our model for similar purposes. For simplicity, the models are trained in a disjoint matter. Joint training is left for future work. The IDM is first calibrated using Genetic Algorithm on the training data. The parameters to be calibrated are: $\{v_0, T, s_0, acc_{max}, dacc_{max}\}$, with parameters and bounds shown in Section 4. Subsequently, the feed-forward neural network is trained with physics-guided loss function (5) for acceleration prediction of the next time step acc_{t+1}. Next, this neural network is used as the acceleration prediction component of MOBIL, giving us the hybrid model. Finally, the hybrid model is calibrated with Genetic Algorithm. More in-depth pseudocode for this process is shown in Algorithm 1.

In the Genetic Algorithm described in Algorithm 1, the fitness function used is defined in Equation (6) with reasons explained in Section 4.3. The calculation of the fitness function is done by the local fitting of input-output pairs with the proposed model. This fitness function is then used in the selection process. The rest of the function follows the standard Genetic Algorithm operations.

The pseudocode before describes the calibration of our hybrid model with Genetic Algorithm. However, before calibrating the hybrid model we require the physics-guided neural network for acceleration prediction as the acceleration component of the original MOBIL model. The pseudocode for training of the physics-guided neural network is given in Algorithm 2.

The flow of steps in using the calibrated model for inference is given by Figure 2. The accelerations of the various vehicles mentioned in Equation (2) are first estimated with the physics-guided neural network and then used for lane change intention prediction with MOBIL.

4 EXPERIMENTS

4.1 Experiment Setup

The experiments were conducted using a real world traffic dataset, namely the highD dataset [18]. The highD dataset presents a collection of naturalistic vehicle trajectories from German highways, captured through drone technology which promises a positioning

Algorithm 1: Calibration of Hybrid model with Genetic Algorithm

Result: Hybrid model f_2 with parameters $\theta^* = \{P, a_{thr}\}$;
Initialization:
Initialize training data $Data_{train}$;
Initialize population size N;
Initialize number of generations G;
Initialize crossover rate c_r;
Initialize mutation rate m_r;
Function `InitializePopulation`(N)**:**
  population ← random set of parameters;
  return *population*;
Function `CalculateFitness`(*individual*, *data*)**:**
  Local fitting of input-output pairs with hybrid model and its *individual* parameters;
  Compare model output to *data*;
  return fitness value;
Function `SelectParents`(*population*,*fitness*)**:**
  Select individuals based on fitness to reproduce;
  return parents;
Function `Crossover`(*parent*1, *parent*2, c_r)**:**
  Produce offspring by combining parameters of parents;
  return offspring;
Function `Mutate`(*individual*, m_r)**:**
  Mutate parameters of *individual* with probability m_r;
  return mutated individual;
Procedure `Main`()**:**
  population ← `InitializePopulation`(N);
  fitness ← `CalculateFitness`(*population*, *data*);
  for $g \leftarrow 1$ **to** G **do**
    newPopulation ← $\emptyset$;
    for $i \leftarrow 1$ **to** N **do**
      parents ← `SelectParents`(*population*, *fitness*);
      offspring ← `Crossover`(*parents*[0], *parents*[1], c_r);
      offspring ← `Mutate`(*offspring*, m_r);
      newPopulation ← *newPopulation* + *offspring*;
    end
    population ← *newPopulation*;
    fitness ← `CalculateFitness`(*population*, *data*);
    bestIndividual ← individual with highest fitness in *population*;
  end
  θ^* ← parameters of *bestIndividual*;
return

accuracy within ten centimeters. The dataset is constructed from 60 recordings from six locations at a collection range of 420m and a frame frequency of 25 Hz. The dataset was also already postprocessed to retrieve various features such as smooth positions, speeds, and accelerations of the vehicles.

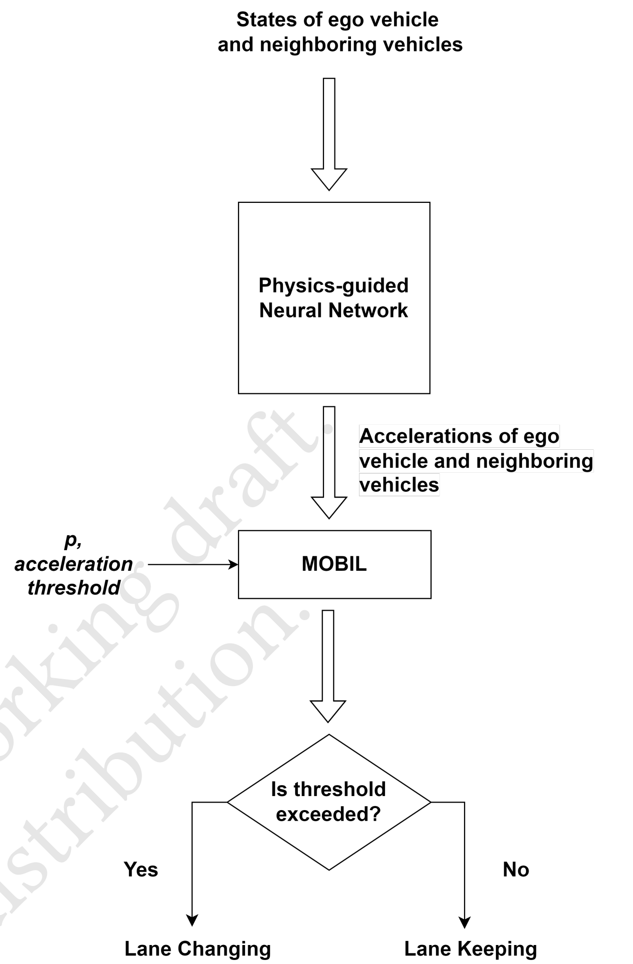

Figure 2: Prediction using physics-guided rule-data hybrid model

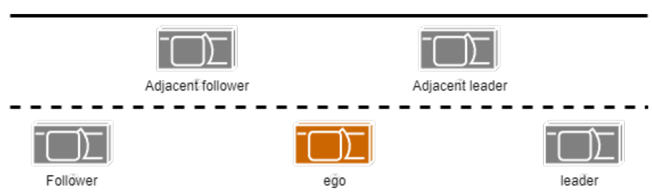

Figure 3: Two-lane scenario extracted

As mentioned previously, MOBIL relies on the ego vehicle and its immediate neighbours to make predictions on lane changes. Since the focus of this paper is to investigate lane change intention prediction under distribution shift, we restrict our experiment setup to the two-lane scenario context shown in Figure 3 to reduce inconsistencies.

Additionally, it is recently highlighted by Ali et al [2] that, lane-changing models are often improperly calibrated due to the fact that the time instant where a vehicle crosses the lane marking is

Algorithm 2: Acceleration prediction model training

Result: Physics-guided NN f_1 with parameters Θ;
Initialization:
Initialize training data $Data_{train}$;
Initialize network weights Θ;
Initialize IDM calibrated with Genetic Algorithm;
Initialize desired value of $\lambda \in [0, 1]$;
Initialize batch size N;
for *epoch = 1,..., E* **do**
  Shuffle D_{train} to create mini-batches B_k such that $D_{train} = \cup_k B_k$;
  forall *batches* B_k **do**
    Initialize gradient accumulation: $\Delta\Theta = 0$;
    forall $(x^{(i)}, y^{(i)}) \in B_k$ **do**
      Perform forward propagation to compute prediction $\hat{y}^{(i)}$;
      Compute output $\hat{y}^{(i)}_{phy}$ with pre-calibrated IDM;
      Compute gradient of the cost function (5) w.r.t. parameters and $\nabla_\Theta J(\Theta; x^{(i)}, y^{(i)}, \hat{y}^{(i)}_{phy})$;
      Accumulate gradients: $\Delta\Theta \leftarrow \Delta\Theta + \nabla_\Theta J(\Theta; x^{(i)}, y^{(i)}, \hat{y}^{(i)}_{phy})$;
    end
    Update parameters: $\Theta \leftarrow \Theta - \alpha \cdot \frac{\Delta\Theta}{|B_k|}$
  end
end

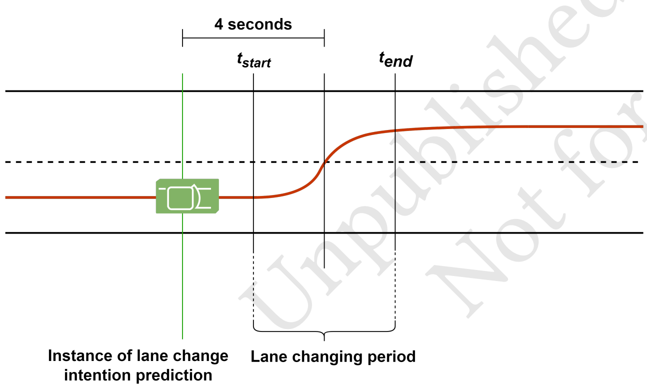

Figure 4: Lane changing trajectory

oftentimes incorrectly considered as the lane-changing decision point, when in fact the lane-changing decision is made earlier.

With reference to Figure 4, it is erroneous to use data between t_{start} and t_{end} for lane change intention prediction since lane change behaviour is already occurring, meaning the predictions are using data from an already occurring lane change maneuver. Lane change maneuver duration $t_{end} - t_{start}$ typically last 3-5 seconds [28]. Thus, we extract data from the instance 4 seconds before the ego vehicle crosses the lane marking for our lane change intention prediction task so as to minimise the possibility of using data of an already occurring maneuver.

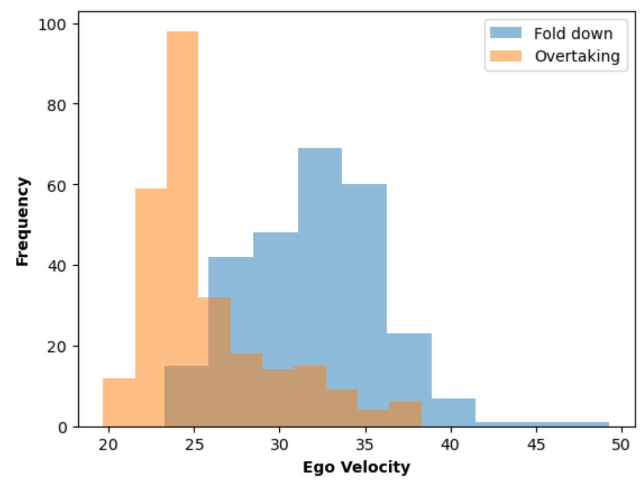

Figure 5: Data distributions of ego velocity under in and out-distribution scenarios

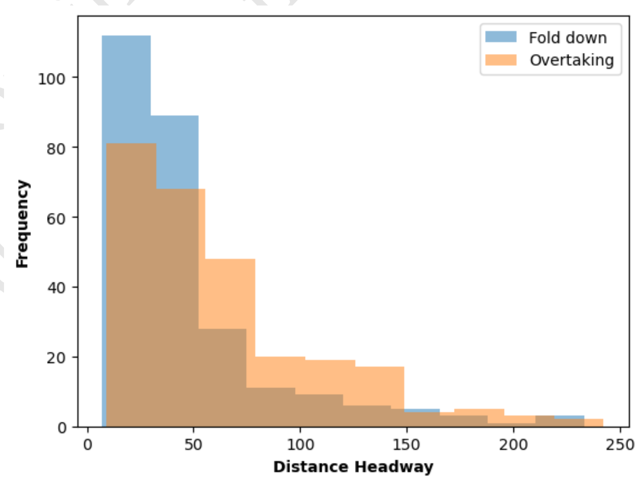

Figure 6: Data distributions of distance headway under in and out-distribution scenarios

To obtain data from two different distributions, data extracted was divided into fold-down and overtaking lane change maneuvers. Fold-down maneuvers consists of the vehicle changing lanes from the fast to slow lane (left to right lane) while overtaking maneuvers consists of the vehicle changing from the slow to fast lane (right to left lane). As observed in Figure 5 and Figure 6, these two scenarios give us data from different distributions.

4.2 Parameters and bounds

In this section, we present the parameters and bounds used in the experiments. Bounds used for calibration of IDM and MOBIL were derived from existing studies [15, 16, 21, 24, 27] and presented in Table 1 and Table 2 respectively.

Table 1: IDM parameter bounds

Parameter	Bound
Maximum desired velocity, v_0	[20,80] m/s
Desired time headway, T	[0,5] s
Jam spacing distance, s_0	[1,5] m
Maximum acceleration, acc_{max}	[0,5] m/s^2
Maximum deceleration, $dacc_{max}$	[0,5] m/s^2

Table 2: MOBIL parameter bounds

Parameter	Bound
Politeness factor, p	[0,1]
Acceleration threshold, a_{thr}	[-4,4] m/s^2

Table 3: Neural network parameters

Parameter	Value
Batch Size	32
Learning Rate	0.001
Initialization type	Xavier
λ used in loss function	0.5

Table 4: Genetic Algorithm Parameters

Parameter	Value
Population Size	60
Cross Over Rate	0.5
Mutation Probability	0.3
Mutation Noise	0.10
Mutation Method	Uniform noise addition
Proportion of Parents in Population	0.5
Maximum number of generations	100
Crossover Method	Uniform
Crossover Point	1
Selection Method	2-way tournament

For acceleration prediction of the ego vehicle using a neural network, the input features used are similar to that of the inputs of the IDM. More concretely, current velocity v_t, front gap distance s_t and relative velocity with respect to the leader Δv_t are used as inputs for the neural network. The loss function used in the training of this neural network is the physics-guided loss function described in Equation (5). The model used is a simple three-layer neural network with hidden layers of sizes 64-128-64 . The exact architecture and its related parameters are shown in Figure 7 and defined in Table 3.

The parameters used in the calibration with Genetic Algorithm are as shown in Table 4.

4.3 Experiment Results

Similar to existing literature [31], non-lane-changing samples extracted far outnumbers lane-changing samples. As such, we use balanced accuracy as defined in Equation (6) as our performance assessment of this imbalanced dataset. Balanced accuracy ensures that the performance of both majority and minority classes are considered in the evaluation, making it a fairer measure of a model's performance and addressing the impact of class imbalance.

$$\text{Balanced Accuracy} = \frac{1}{2}\left(\frac{\text{TP}}{\text{TP}+\text{FN}} + \frac{\text{TN}}{\text{TN}+\text{FP}}\right) \tag{6}$$

where:

- TP: True Positives
- TN: True Negatives
- FP: False Positives
- FN: False Negatives

Table 5: In-distribution experimental results

Model	Balanced Accuracy
Binomial choice	0.487
Pure MOBIL	0.660
Neural Network	0.789
Non-physics guided hybrid model	0.669
Physics-guided hybrid model	**0.689**

Table 6: Out-distribution experimental results

Model	Balanced Accuracy
Binomial choice	0.482
Pure MOBIL	0.584
Neural Network	0.372
Non-physics guided hybrid model	0.572
Physics-guided hybrid model	**0.601**

The following models were considered for performance comparison:

(1) **Binomial Choice**: Lane changes are predicted with the probability derived from the proportion of lane changing to lane keeping instances in the training data. That is, if the proportion of lane changing instances is p and lane-keeping instances is $1 - p$, the predictions on the test set will follow this probability distribution.
(2) **Pure MOBIL**: the MOBIL model described in Section 3.1 calibrated with Genetic Algorithm.
(3) **Neural Network**: A simple feed-forward neural network with input features {ego velocity, front gap distance, preceding neighbor velocity, proceeding neighbor, follower velocity, leader velocity }. The neural network has 4 hidden layers, each of 8 neurons. ReLU activation is used for each hidden layer and a sigmoid function is used in the last layer for lane-changing/ lane-keeping prediction.

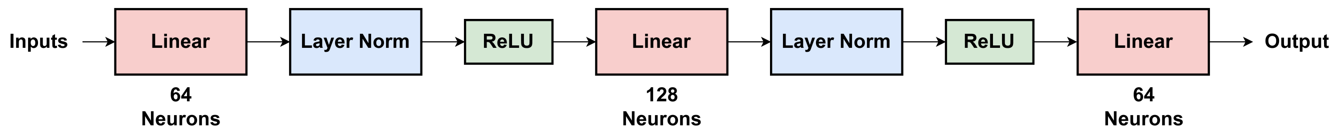

Figure 7: Network architecture of neural network used in acceleration prediction

(4) **Non-physics guided hybrid model**: Same model architecture as described in Figure 1, but the acceleration prediction component is trained with the standard mean squared error loss instead of the physics-guided loss function.
(5) **Physics-guided hybrid model**: The model described in Section 3.4 and Figure 1, trained with physics-guided loss function given by Equation (5).

The baseline chosen is a simple binomial model where the probability of lane-changing and lane-keeping prediction is given by the proportion of lane-changing and lane-keeping instances in the training data. A visualisation of the balanced accuracy of the tested models w.r.t the baseline model is given in Figure 8.

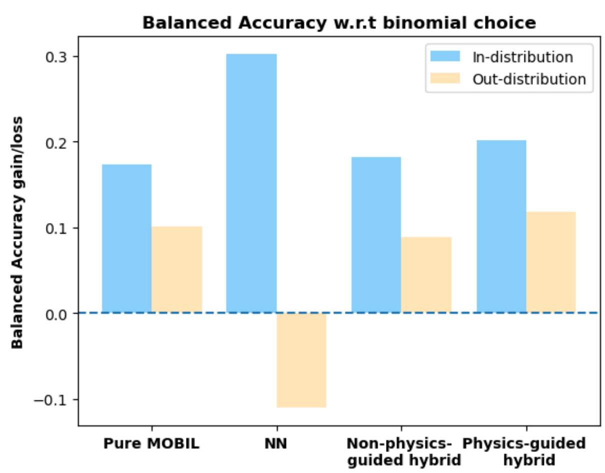

Figure 8: Performance of models w.r.t baseline

As observed from the above experiment results in Table 5 and Table 6, the physics-guided hybrid model is able to produce competitive performance under in-distribution scenarios, while producing the best results amongst all the models under out-distribution scenarios. This is in contrast to the data-based neural network, which is only able to deliver good performance under in-distribution scenarios but breaks down under out-distribution scenarios. More concretely, the proposed model outperforms rule-based models under in-distribution scenarios by 4%, and outperforms both rule-based models and data-based models in out-distribution scenarios by 3% and 60% respectively.

4.3.1 In-distribution performance assessment.

From Table 5, the neural network model achieves the best performance under in-distribution scenarios unsurprisingly. This is due to the fact that it is a parametric model having high flexibility and strong ability to learn the patterns in the dataset. The non-physics guided hybrid model which is the MOBIL model integrated with a neural network for acceleration prediction, performs slightly better than that of the pure MOBIL model. This can be attributed to the neural network producing better estimates in the task of acceleration prediction as compared to that of the IDM. In the domain of supervised learning, the total error from a model can be defined as the sum of reducible and irreducible error [13]. The reducible error arises from the fact that no model can perfectly estimate f, which is the function that maps any input variable to its correct output variable. On the other hand, the irreducible error arises from the stochastic nature of the system. Reducible errors can be decreased by improving our estimates. The neural network fills this role by providing better estimates as compared to the IDM. Better acceleration estimation then leads to more accurate lane-change predictions. Subsequently, a larger improvement occurs when a physics-guided loss function is used for acceleration prediction. The incorporation of the physics-guided loss function brings about a regularisation effect, as the flexibility of neural networks tends to lead to the overfitting on training data. As mentioned before, the incorporation of the physics-guided loss function also introduces theory into our otherwise black box model. This can be considered as instilling physical consistency into the model's parameter space through gradient-based techniques [23].

4.3.2 Out-distribution performance assessment.

From Table 6, the neural network achieves results worse than the baseline model. The neural network's flexibility stems from the large number of parameters it estimates. The complex nature of the neural network allows for the model to fit the patterns in the data closely. However this is a double edged sword as the model also fits to the noise closely, which is overfitting. This means that the model has learnt patterns and relationships specific to the training data distribution, but fails to generalise from the training data to unseen examples. In contrast, rule-based models are designed to be more robust to distribution shifts because they are explicitly defined by rules which are formulated based on theory and not based on training data. When a rule-based model encounters data from a different distribution, it can gracefully handle such cases by applying appropriate rules. As such, MOBIL is able to maintain its performance even under an out-distribution scenario. Following, the non-physics guided hybrid model is also able to outperform the baseline model as well as the neural network although its performance is slightly worse than that of MOBIL. This is expected as the acceleration prediction component of this model is a neural network, which can easily overfit to the training data as well. The less accurate estimates for accuracy under the out-distribution scenario increases the error of the overall model, leading to a slight decrease in performance as compared to the MOBIL model. Finally, the PG-hMOBIL achieves the best performance with the incorporation of the physics-guided loss function. The neural network used for acceleration prediction is now less prone to overfitting due to the regularisation effect provided by the physics-guided loss function. From the experiment results and their accompanying analysis, while it can be observed that the MOBIL backbone contributes the most to making our model robust under distribution shift; it can also be deemed that physics-guidance is of great importance when incorporating a data-based method with a rule-based method to create a hybrid model that is robust to distribution shift. This is because injecting theory into the model allows the model to learn the true underlying function better, rather than simply fitting on the data.

5 CONCLUSIONS

In this paper, a new physics-guided hybrid model combining rule-based and data-based methods is proposed to offer robust lane change intention prediction under distribution shift. Since the model is built upon a rule-based algorithm as its foundation, it also retains most of its theoretical explainability unlike data-based black box models. Additionally, results have shown that such a model produces competitive performance under in-distribution scenarios, while maintaining good performance under out-distribution scenarios. For future work, it might be worthwhile to test the acceleration prediction component with a more sophisticated neural network architecture. Additionally, it might be advantageous to transform the whole framework into a single differentiable model [3]. This might lead to better calibration results, with a unified approach which could potentially result in a shorter training time. Other ways of joint calibration methods can also be explored.

ACKNOWLEDGMENTS

This work was supported by

REFERENCES

[1] Mohamed Abdel-Aty, Ou Zheng, et al. 2023. Lane Change Intention Recognition and Vehicle Status Prediction for Autonomous Vehicles. (2023).
[2] Yasir Ali, Zuduo Zheng, and Michiel CJ Bliemer. 2023. Calibrating lane-changing models: Two data-related issues and a general method to extract appropriate data. *Transportation Research Part C: Emerging Technologies* 152 (2023), 104182.
[3] Philipp Andelfinger. 2023. Towards differentiable agent-based simulation. *ACM Transactions on Modeling and Computer Simulation* 32, 4 (2023), 1–26.
[4] Christina Baek, Yiding Jiang, Aditi Raghunathan, and J Zico Kolter. 2022. Agreement-on-the-line: Predicting the performance of neural networks under distribution shift. *Advances in Neural Information Processing Systems* 35 (2022), 19274–19289.
[5] Tiffany Tianhui Cai, Hongseok Namkoong, Steve Yadlowsky, et al. 2023. Diagnosing model performance under distribution shift. *arXiv preprint arXiv:2303.02011* (2023).
[6] Dongho Choi and Sangsun Lee. 2021. Comparison of machine learning algorithms for predicting lane changing intent. *International journal of automotive technology* 22 (2021), 507–518.
[7] Charisma Farheen Choudhury. 2007. *Modeling driving decisions with latent plans.* Ph. D. Dissertation. Massachusetts Institute of Technology.
[8] André Luiz Cunha, José Elievam Bessa, and José Reynaldo Setti. 2009. Genetic algorithm for the calibration of vehicle performance models of microscopic traffic simulators. In *Progress in Artificial Intelligence: 14th Portuguese Conference on Artificial Intelligence, EPIA 2009, Aveiro, Portugal, October 12-15, 2009. Proceedings 14*. Springer, 3–14.
[9] Ürün Dogan, Johann Edelbrunner, and Ioannis Iossifidis. 2011. Autonomous driving: A comparison of machine learning techniques by means of the prediction of lane change behavior. In *2011 IEEE International Conference on Robotics and Biomimetics*. IEEE, 1837–1843.
[10] Danish Farooq, Sarbast Moslem, Arshad Jamal, Farhan Muhammad Butt, Yahya Almarhabi, Rana Faisal Tufail, and Meshal Almoshaogeh. 2021. Assessment of significant factors affecting frequent lane-changing related to road safety: An integrated approach of the AHP–BWM model. *International journal of environmental research and public health* 18, 20 (2021), 10628.
[11] Salah A Faroughi, Nikhil Pawar, Celio Fernandes, Maziar Raissi, Subasish Das, Nima K Kalantari, and Seyed Kourosh Mahjour. 2022. Physics-guided, physics-informed, and physics-encoded neural networks in scientific computing. *arXiv preprint arXiv:2211.07377* (2022).
[12] Peter G Gipps. 1986. A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological* 20, 5 (1986), 403–414.
[13] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor. 2023. *An introduction to statistical learning: With applications in python.* Springer Nature.
[14] Felipe Jiménez, José Eugenio Naranjo, José Javier Anaya, Fernando García, Aurelio Ponz, and José María Armingol. 2016. Advanced driver assistance system for road environments to improve safety and efficiency. *Transportation research procedia* 14 (2016), 2245–2254.
[15] Arne Kesting, Martin Treiber, and Dirk Helbing. 2007. General lane-changing model MOBIL for car-following models. *Transportation Research Record* 1999, 1 (2007), 86–94.
[16] Basma Khelfa, Ibrahima Ba, and Antoine Tordeux. 2023. Predicting highway lane-changing maneuvers: A benchmark analysis of machine and ensemble learning algorithms. *Physica A: Statistical Mechanics and its Applications* (2023), 128471.
[17] Kyu-Ok Kim and Laurence R Rilett. 2001. Genetic-algorithm based approach for calibrating microscopic simulation models. In *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 01TH8585)*. IEEE, 698–704.
[18] Robert Krajewski, Julian Bock, Laurent Kloeker, and Lutz Eckstein. 2018. The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 2118–2125. https://doi.org/10.1109/ITSC.2018.8569552
[19] Dehao Liu and Yan Wang. 2023. Physics-constrained neural networks with minimax architecture for multiphysics dendritic growth problems in additive manufacturing. *Manufacturing Letters* 35 (2023), 1060–1071.
[20] Sen Liu, Branden B Kappes, Behnam Amin-ahmadi, Othmane Benafan, Xiaoli Zhang, and Aaron P Stebner. 2021. Physics-informed machine learning for composition–process–property design: Shape memory alloy demonstration. *Applied Materials Today* 22 (2021), 100898.
[21] Zhaobin Mo, Rongye Shi, and Xuan Di. 2021. A physics-informed deep learning paradigm for car-following models. *Transportation research part C: emerging technologies* 130 (2021), 103240.
[22] Htet Naing, Wentong Cai, Hu Nan, Wu Tiantian, and Yu Liang. 2022. Dynamic Data-driven Microscopic Traffic Simulation using Jointly Trained Physics-guided Long Short-Term Memory. *ACM Transactions on Modeling and Computer Simulation* 32, 4 (2022), 1–27.
[23] Htet Naing, Wentong Cai, Hu Nan, Wu Tiantian, and Yu Liang. 2023. Physics-guided Graph Convolutional Network for Modeling Car-following Behaviors under Distributional Shift. (2023).
[24] Htet Naing, Wentong Cai, Tiantian Wu, and Liang Yu. 2022. Dynamic car-following model calibration with deep reinforcement learning. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 959–966.
[25] Jordan J Pannell, Sam E Rigby, and George Panoutsos. 2022. Physics-informed regularisation procedure in neural networks: An application in blast protection engineering. *International Journal of Protective Structures* 13, 3 (2022), 555–578.
[26] Mohammad Mustafa Taye. 2023. Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers* 12, 5 (2023), 91.
[27] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. 2000. Congested traffic states in empirical observations and microscopic simulations. *Physical review E* 62, 2 (2000), 1805.
[28] Wu Xiaorui and Yang Hongxu. 2013. A lane change model with the consideration of car following behavior. *Procedia-Social and Behavioral Sciences* 96 (2013), 2354–2361.
[29] Yang Yu, Houpu Yao, and Yongming Liu. 2020. Structural dynamics simulation using a novel physics-guided machine learning method. *Engineering Applications of Artificial Intelligence* 96 (2020), 103947.
[30] Nanbin Zhao, Bohui Wang, Yiping Xiong, and Rong Su. 2022. Safety Prompt Advanced Driver-Assistance System with Lane-Change Prediction and Free Space Detection. In *2022 IEEE 17th International Conference on Control & Automation (ICCA)*. IEEE, 734–739.
[31] Jian Zheng, Koji Suzuki, and Motohiro Fujita. 2014. Predicting driver's lane-changing decisions using a neural network model. *Simulation Modelling Practice and Theory* 42 (2014), 73–83.
[32] Qiiie Zou, Yingli Hou, and Zumin Wang. 2019. Predicting vehicle lane-changing behavior with awareness of surrounding vehicles using LSTM network. In *2019 IEEE 6th International Conference on Cloud Computing and Intelligence Systems*

(CCIS). IEEE, 79–83.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009