UMEC
URBAN MOBILITY & EQUITY CENTER

Final Report

Two-Dimensional Modeling of Bicycle Behavior

Hesham Rakha
Virginia Tech Transportation Institute
Charles E. Via, Jr. Department of Civil and Environmental Engineering
Virginia Polytechnic Institute and State University
Phone: (540) 231-1505; Email: hrakha@vt.edu

Karim Fadhloun
Virginia Tech Transportation Institute
Phone: (540) 231-1500; Email: karim198@vt.edu

Date
January 27, 2023

Prepared for the Urban Mobility & Equity Center, Morgan State University, CBEIS 327, 1700 E. Cold Spring Lane, Baltimore, MD 21251

MORGAN STATE UNIVERSITY
GROWING THE FUTURE • LEADING THE WORLD

UNIVERSITY OF MARYLAND 1856

VirginiaTech 1872
Invent the Future

ACKNOWLEDGMENT

This research was supported by the Urban Mobility & Equity Center at Morgan State University and the University Transportation Center(s) Program of the U.S. Department of Transportation.

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 under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

1. Report No. UMEC-048	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Two-Dimensional Modeling of Bicycle Behavior		**5. Report Date** January 2023
		6. Performing Organization Code
7. Author(s) Karim Fadhloun (ORCID# 0000-0002-0442-3727) Hesham Rakha (ORCID # 0000-0002-5845-2929)		**8. Performing Organization Report No.**
9. Performing Organization Name and Address Virginia Tech Transportation Institute 3500 Transportation Research Road Blacksburg, VA 24060		**10. Work Unit No.**
		11. Contract or Grant No. 69A43551747123
12. Sponsoring Agency Name and Address US Department of Transportation Office of the Secretary-Research UTC Program, RDT-30 1200 New Jersey Ave., SE Washington, DC 20590		**13. Type of Report and Period Covered** Final January 2022 - December 2022
		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This research first develops a descriptive model that is capable of capturing the inherent non-lane-based traffic behavior characteristics of bicycles. To that end, the research team expands upon the existing Fadhloun-Rakha bicycle-following longitudinal motion model by complementing it with a lateral motion strategy, thus allowing for overtaking maneuvers and lateral bicycle movements. For the most part, the following strategy of the FR model remains valid for modeling the longitudinal motion of bicycles except under the conditions of the collision avoidance strategy, which are modified in order to allow for overtaking when possible. The proposed methodology is innovative in that it makes use of the intersection of certain pre-defined regions around the bicycles to decide on the feasibility of angular motion as well as its direction and magnitude. The resulting model is the first point-mass dynamics-based model to describe the longitudinal and lateral behavior of bicycles in both constrained and unconstrained conditions. In fact, by having the FR bicycle-following model as both the governing module of longitudinal behavior and a dynamic lateral module, the proposed model is able to model bicyclist behavior variability. Furthermore, given that the longitudinal logic used in the model was previously validated against experimental cycling data, it is the only existing model that is sensitive to the condition of the bicycle, the roadway surface, and the bicyclists' physical characteristics.

Next, the research team expanded this study by collecting a new naturalistic cycling dataset. Given that the collection of naturalistic cycling data is not achievable in the traditional vehicle approach, machine learning and computer vision techniques were used to construct the naturalistic dataset from existing video feeds. The videos used in the study come from a dataset collected in a previous Virginia Tech Transportation Institute study conducted in collaboration with SPIN in which continuous video data was recorded at a non-signalized intersection on the Virginia Tech campus. The research team applied existing computer vision and machine learning techniques to develop a comprehensive framework for the extraction of naturalistic cycling trajectories. In total, the proposed methodology resulted in the collection of 619 bicycle trajectories at a high level of precision with respect to the location, speed, and accelerations of the bicycles.

17. Key Words:	18. Distribution Statement
Bicycle lateral motion, Bicycle traffic flow dynamics, Naturalistic cycling data, Traffic efficiency and safety	No restrictions.

19. Security Classif. (of this report):	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	33	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This research first develops a descriptive model that is capable of capturing the inherent non-lane-based traffic behavior characteristics of bicycles. To that end, the research team expands upon the existing Fadhloun-Rakha bicycle-following longitudinal motion model by complementing it with a lateral motion strategy, thus allowing for overtaking maneuvers and lateral bicycle movements. For the most part, the following strategy of the FR model remains valid for modeling the longitudinal motion of bicycles except under the conditions of the collision avoidance strategy, which are modified in order to allow for overtaking when possible. The proposed methodology is innovative in that it makes use of the intersection of certain pre-defined regions around the bicycles to decide on the feasibility of angular motion as well as its direction and magnitude. The resulting model is the first point-mass dynamics-based model to describe the longitudinal and lateral behavior of bicycles in both constrained and unconstrained conditions. In fact, by having the FR bicycle-following model as both the governing module of longitudinal behavior and a dynamic lateral module, the proposed model is able to model bicyclist behavior variability. Furthermore, given that the longitudinal logic used in the model was previously validated against experimental cycling data, it is the only existing model that is sensitive to the condition of the bicycle, the roadway surface, and the bicyclists' physical characteristics.

Next, the research team expanded this study by collecting a new naturalistic cycling dataset. Given that the collection of naturalistic cycling data is not achievable in the traditional vehicle approach, machine learning and computer vision techniques were used to construct the naturalistic dataset from existing video feeds. The videos used in the study come from a dataset collected in a previous Virginia Tech Transportation Institute study conducted in collaboration with SPIN in which continuous video data was recorded at a non-signalized intersection on the Virginia Tech campus. The research team applied existing computer vision and machine learning techniques to develop a comprehensive framework for the extraction of naturalistic cycling trajectories. In total, the proposed methodology resulted in the collection of 619 bicycle trajectories at a high level of precision with respect to the location, speed, and accelerations of the bicycles.

Key words: Bicycle lateral motion, Bicycle traffic flow dynamics, Naturalistic cycling data, Traffic efficiency and safety

1. INTRODUCTION

As a result of the ever-increasing number of vehicles on the road and the infeasibility of further road enlargement, congestion is becoming a major problem facing modern cities around the world. A popular solution adopted by policymakers to lessen traffic congestion in central downtown areas is to advocate cycling as a sustainable commuting mode. This is because short-distance bike commuting often takes less time when accounting for congestion and delays in public transportation and presents the most efficient way to increase the road capacity while maintaining existing infrastructure. Several major cities are promoting the use of bikes between public transportation hubs and private transportation through the implementation of bike sharing systems. In fact, 119 US cities had a bike sharing system in 2017. Besides being beneficial to human health, cycling has significant positive impacts on the environment as well, as it significantly reduces fossil fuel consumption and vehicle emissions.

Despite the growing interest in bicycle use over the past decade and the urgent need to develop models and planning techniques for bicycle traffic operations, traffic researchers have barely scratched the surface of bicycle traffic flow dynamics. In fact, while vehicular traffic flow dynamics have been studied extensively, investigations into the use of bicycles as a mode of transportation remain relatively scarce in general [1-7]. The observed literature gap between vehicular and bicycle traffic research can be explained by the scarcity of naturalistic and experimental cycling data. While the physics of bicycles were investigated as early as the 1960s, it is only quite recently that cycling was explored from a traffic engineering perspective. It is noteworthy that most of those research studies investigated the interactions between bicycles, cars and other entities. Technically speaking, a significant portion of existing cycling research falls under one of two main categories. The first category involves research based on the Cellular Automata (CA) model that involves discretizing the time and space domain using a non-continuous cell grid. As an example, one can mention the work of Jiang et al. [8] who investigated the dispersion effect of bicycle traffic at a signalized intersection using the Cellular Automata approach. Another example of cycling research that used the Cellular Automata scheme is the work of Ren et al. [9] who proposed to simulate the lateral interactions of heterogeneous traffic using the bicycle spilling maneuver phenomenon. Jia et al. [2] have also proposed a CA model for mixed traffic involving bicycles and cars. Furthermore, Gould and Karner [6] used the CA model for the modeling of bicycle facility operations. Besides the CA approach, researchers have made use of social force models to simulate bicycle longitudinal and lateral traffic behavior. For instance, Qu et al. [10] proposed a social force model for the simulation of the interactions of electric bicycles and cars based on five different forces. Several other models [5, 11] were also developed using the SFM approach because of its advantages in terms of simulating dynamic lateral dispersion characteristics of mixed traffic.

Regardless of the approach used, it is quite noticeable that existing cycling research is mostly oriented towards capturing the effect of bicycles in a mixed traffic environment rather than the fundamental concepts behind bicycle motion. In fact, research that deals exclusively with capturing bicycle traffic flow behavior is quite limited [1-3]. To illustrate the extent to which modeling the longitudinal motion of bicycles has been ignored historically, there is no better

argument than noting that the first model specifically designed to simulate the following behavior of bicyclists, the Necessary Deceleration Model (NDM), was only developed in 2012 [2]. Based on the observation that there are no major differences between the dynamics of single-file bicycle traffic and vehicular traffic, another approach used by researchers to model the longitudinal motion of bicycles consisted of capturing cyclists' behavior by revamping certain aspects of existing car-following models. That is the case, for example, in the Intelligent Driver Model (IDM) [4], which, after a simple re-parameterization, was proven to be a good descriptor of bicycle-following behavior [1]. In a similar fashion, the research team proposed a longitudinal motion model for bicycles [3] derived from the Fadhloun-Rakha (FR) car-following model to address the gap in the literature regarding effects that the cyclist and the road environment have on bicycle motion behavior [5].

This research effort was mainly initiated to develop a descriptive model that is capable of capturing the inherent non-lane-based traffic behavior characteristics of bicycles. Specifically, bicycle traffic is relatively unconstrained in relation to lateral motion and overtaking when compared to vehicular traffic, which is bounded by lane limits. With this in mind, it is quite noticeable that previous attempts at developing bicycle-following models were limited to capturing the longitudinal motion behavior of bicyclists. Moreover, all of the proposed models [1-3] are only valid as descriptors of bicycle motion dynamics under very specific assumptions and conditions. In fact, as mentioned earlier, those models were developed by applying vehicular traffic techniques based on the assumption that there are significant similarities between the traffic flow dynamics of bicycles and cars when single-file motion is considered. While those assumptions are clearly stated by the researchers, they result in significant limitations for the proposed models.

As opposed to multi-lane vehicular traffic, which can be modeled by combining a car-following model and a lane-changing model, the naturalistic behavior of congregated bicycle traffic is characterized by flock-like behavior. Bicycles demonstrate significantly more flexibility than cars with respect to executing lateral and overtaking maneuvers as well as lane striping. For that reason, a simplistic analogy with vehicular behavior cannot capture the lateral motion behavior of bicycles. Subsequently, for a model to be successful in terms of replicating lateral bicycle motion, it would need to demonstrate its ability to capture angular motion and directional changes with a high level of detail. Furthermore, unlike lane-changing models for vehicles, a lateral motion control strategy for bicycles would need to be unified with a longitudinal motion model in a single comprehensive model in order to be able to work in harmony together.

In the first part of this study, the research team thus proposes to build on the aforementioned FR bicycle-following model by complementing it with a lateral motion strategy to allow for overtaking maneuvers. For the most part, the following strategy of the FR model remains valid for modeling the longitudinal motion of bicycles except under the conditions of the collision avoidance strategy in order to allow for overtaking when possible. The biggest challenge that faced the research team in this phase mostly related to choosing an adequate strategy that would result in a gradual change of the angular motion while overtaking. A successful strategy would also need to be sensitive to several parameters in order to induce a certain degree of randomization in its outputs.

Next, given that most existing cycling datasets were collected in an experimental setting in which the naturalistic characteristics of bicyclists were not fully captured, the research team worked on expanding this study by collecting a new naturalistic cycling dataset. Specifically, the second part of this report describes the collection of naturalistic cycling data from video feeds for use in different mobility applications. To achieve this, the research team first applied computer vision, machine learning, and data reduction techniques to a video dataset in order to identify and extract bicycle trips from the videos. The selected video dataset is derived from a previous Virginia Tech Transportation Institute study conducted in collaboration with SPIN in which continuous video data at a non-signalized intersection at the Virginia Tech campus was collected. Next, using the results of a high-precision surveying campaign of the observed area, the collected trajectories were projected in the Northing-Easting coordinate system allowing for the determination of the actual locations, speeds, and accelerations of the bicycles at a high level of precision. This effort resulted in the collection of 619 bicycle trajectories.

The contribution of this study is significant for the research community as it is the first of its kind to propose a dynamics-based longitudinal and lateral motion model for bicycles that works for both constrained and unconstrained conditions. Not only that, given the choice of the FR-model as the longitudinal module, the proposed model captures bicyclist behavior variability. That makes it the only existing model to be sensitive to factors related to the bicyclists' physical characteristics, such as age, gender, size and stamina as well as roadway conditions and bicycle characteristics. The research team acknowledges the limitations of this study, as bicycle behavior is more complicated than assumed herein. Despite that, the contribution of this study remains significant, as it constitutes a necessary first step towards the development of a comprehensive two-dimensional model that is descriptive of empirical bicycle motion behavior. Another noteworthy contribution of this study relates to the collected naturalistic cycling dataset, which we expect to be beneficial to the research community for use in different mobility applications.

The report will begin by presenting a comprehensive overview of the Fadhloun-Rakha model variations for both vehicles and bicycles, which will serve as the foundation of the proposed model. After that, the methodology used to achieve the formulation of the proposed lateral model module is described. The research team will study the possibility of controlling the angular motion of bicycles by investigating the interactions between certain deterministic and dynamic elliptic areas that are defined around the bicycles. The formulation of the proposed lateral module could be cast as a multi-step algorithmic process rather than a mathematical expression as observed in typical motion models. The performance and adequacy of the resulting model as a descriptor of bicycle longitudinal and lateral motion is then assessed using some basic scenarios. Thereafter, the report provides a detailed description of the different methodologies and techniques involved in the extraction of the naturalistic cycling trajectories from the video feeds. Finally, the conclusions of the project are presented.

2. THE FADHLOUN-RAKHA MODEL

Given that the FR-model was initially developed as a car-following model before being adapted to the motion of bicycles, we will provide a description that covers the two variations of the model.

2.1. The Fadhloun-Rakha Car-Following Model

One of the simplest car-following strategies entails following the lead vehicle at a constant headway, which is typically taken equal to the driver perception-reaction time T, as illustrated in Equation 1. This model is also known as the Pipes or GM-1 model [6-8]. This time headway ensures that the subject vehicle n follows its leader at a safe distance in order to avoid a collision under steady-state conditions (i.e. when both vehicles are traveling at the same constant velocity and assuming that the subject vehicle's deceleration maneuver starts T seconds after the lead vehicle decelerates).

$$\tilde{s}_n = s_j + T v_n \tag{1}$$

In the context of car-following modeling, Van Aerde [9] and Van Aerde and Rakha [10] proposed a more general formulation that reflects empirical driver behavior better than other models. This formulation combines the Pipes (Equation 1) and the Greenshields models to generate a more general formulation [11-13], presented in Equation 2.

$$\tilde{s}_n = c_1 + \frac{c_2}{(v_f - v_n)} + c_3 v_n \tag{2}$$

Here, $c_1, c_2,$ and c_3 are model coefficients that can be computed using key roadway traffic stream parameters (Equation 3) [11], namely: the free-flow speed, v_f; the speed-at-capacity, v_c; the roadway capacity, q_c; and the roadway jam density, k_j (the inverse of the jam density spacing, s_j).

$$\begin{aligned} c_1 &= \frac{v_f}{k_j v_c^2}(2v_c - v_f); \\ c_2 &= \frac{v_f}{k_j v_c^2}(v_f - v_c)^2; \\ c_3 &= \frac{1}{q_c} - \frac{v_f}{k_j v_c^2} \end{aligned} \tag{3}$$

A slight modification is applied to the expression of the steady-state model presented in Equation 2 when the lead vehicle is traveling at a lower velocity than the following vehicle (non-steady-state conditions). In those specific scenarios, a safety distance margin is considered in addition to the steady-state spacing as shown in Equation 4. The additional term is computed as the braking distance needed for the follower to decrease its speed to that of the leader at a deceleration level *d* (*d* positive in the equation). The purpose of the desired safe following spacing is to allow the following driver to drive at a distance greater than the steady-state spacing when the vehicle ahead of it is driving at a lower speed. It is noteworthy that the additional term is formulated such that it is only active when the following vehicle is approaching the leading vehicle. Otherwise, it is equal to zero which leads back to Equation 2.

$$\tilde{s}_n = \max\left(c_1 + \frac{c_2}{(v_f - v_n)} + c_3 v_n + \frac{v_n^2 - v_{n-1}^2 + \sqrt{(v_n^2 - v_{n-1}^2)^2}}{4d}, s_j\right) \tag{4}$$

Finally, the vehicle acceleration behavior is governed by vehicle dynamics to ensure that vehicle accelerations are realistic, as demonstrated in Equation 5.

$$a_{max} = \frac{\min\left(\frac{\beta\eta_d P_n}{v_n}, m_{ta} g\mu\right) - \frac{\rho C_d C_h A_f v_n^2}{2} - mgC_{r0}(C_{r1}v_n + C_{r2}) - mgG}{m} \quad (5)$$

Rakha and Lucic [14] introduced the β factor in order to account for the gearshift impacts when trucks are accelerating at low speeds. This factor is set to 1.0 for light-duty vehicles [15]. Other parameter definitions are: η_d is the driveline efficiency (unitless); P is the vehicle power (W); m_{ta} is the mass of the vehicle on the tractive axle (kg); g is the gravitational acceleration (9.8067 m/s^2); μ is the coefficient of road adhesion or the coefficient of friction (unitless); ρ is the air density at sea level and a temperature of 15°C (1.2256 kg/m^3); C_d is the vehicle drag coefficient (unitless), typically 0.30; C_h is the altitude correction factor equal to 1-0.000085h, where h is the altitude in meters (unitless); A_f is the vehicle frontal area (m^2), typically 0.85 multiplied by the height and width of the vehicle; C_{r0} is a rolling resistance constant that varies as a function of the pavement type and condition (unitless); C_{r1} is the second rolling resistance constant (h/km); C_{r2} is the third rolling resistance constant (unitless); m is the total vehicle mass (kg); and G is the roadway grade (unitless).

To capture the driver input, the f_p factor is introduced, which ranges between 0.0 and 1.0. The final FR model formulation, which considers deceleration to avoid a collision with a slower traveling leader, is demonstrated in Equation 6. This equation includes two terms. The first term is the acceleration term while the second term is the deceleration term. Both terms ensure that the following vehicle does not collide with its leader.

$$a_n = f_p a_{max} - \frac{\left[v_n^2 - v_{n-1}^2 + \sqrt{(v_n^2 - v_{n-1}^2)^2}\right]^2}{16(d_{des} - gG)\left(s_n - s_j\right)^2} \quad (6)$$

Here f_p is computed using Equation 7 where X_n is calculated using Equation 8.

$$f_p = e^{-g_1 X_n}\left(1 - X_n^{g_2} e^{g_2(1-X_n)}\right)^{g_3} \quad (7)$$

$$X_n = \frac{\min\left(\tilde{s}_n, \tilde{s}_n((1-\alpha)v_f)\right)}{\min\left(s_n, \tilde{s}_n((1-\alpha)v_f)\right)} \cdot \frac{v_n}{\tilde{v}_n} \quad (8)$$

Here $\tilde{s}_n$ is the desired spacing for the current speed (computed using Equation 4); $\tilde{v}_n$ is the desired speed for the current spacing (which is computed by solving for the driver's desired speed based on its current spacing using Equation 4); α is the percentile of v_f (suggested to be 2.5%); d_{des} is the desired deceleration level; g_1, g_2, and g_3 are model parameters that are calibrated to a specific driver, and model the driver power input through the application of the gas pedal.

In order to ensure that the parameters (g_1, g_2, and g_3) result in a minimal maximum value of f_p in the deceleration domain, an iterative procedure was developed. The iterative procedure, presented in Equation 9, is only approximate and converges relatively fast (within four to five

iterations) to the location of the maximum of f_p, which is then verified to be below a threshold ε (for instance, $\varepsilon = 0.1$). By doing so, it is decided whether the chosen values for the g_1, g_2, and g_3 parameters are accepted or rejected. Of course, this procedure was only adopted after ensuring that the number of the different combinations of g_1, g_2, and g_3 that would result in $f_p(X_{k\to inf}) < \varepsilon$ is significant. It is worth clarifying why this additional step was needed given that a simple binary function would have been sufficient. In fact, this step is necessary in order to make the model computationally friendly as it significantly reduces the feasible region for the model parameters without major repercussions on the model performance. In fact, the defined feasible region for the variables is big enough to allow the model's flexibility and adaptability under different scenarios.

$$\begin{cases} X_0 = 3\left(-1+\sqrt{2\ln(3)}\right) \\ X_{k+1} = 3\left[-1+\sqrt{2\ln\left(3\left[1+\frac{g_2 g_3}{g_1}\left(1-\frac{1}{X_k}\right)\right]^{1/g_2}\right)}\right] \end{cases} \tag{9}$$

Finally, three noise variables were added to the model formulation in order to capture the perception and control inaccuracies of the drivers. The first two signals attempt to model the perception errors in estimating the leader's speed and the gap distance separating the two vehicles. They consist of two Wiener processes that are incorporated in the model formulation as presented in Equations 10 and 11. On the one hand, Equation 10 emulates the driver's inability to have an exact estimation of the speed of the leading vehicle. On the other hand, Equation 11 simulates the error committed while estimating the spacing separating them. Additionally, the white noise signal presented in Equation 12 is added to capture the control errors during the acceleration and deceleration maneuvers. The compounding effect of these three signals makes the model output more representative of human behavior. The model output is computed as the sum of Equation 12 and Equation 6 in which $\widetilde{u_n}(t)$ and $\widetilde{s_{n+1}}(t)$ are used instead of u_n and s_{n+1}.

$$\begin{cases} \widetilde{u_{n-1}}(t) = u_{n-1}(t-\Delta t) - 0.01\left(s_n - s_j\right)\left(e^{-0.01}.W_l(t-\Delta t) + \sqrt{0.02}.N(0,1)\right) \\ W_l(1) = N(0,1) \end{cases} \tag{10}$$

$$\begin{cases} \widetilde{s_n}(t) = s_n(t-\Delta t) \times e^{0.1\left(e^{-0.01}.W_s(t-\Delta t)+\sqrt{0.02}.N(0,1)\right)} \\ W_s(1) = N(0,1) \end{cases} \tag{11}$$

$$\breve{a}_n(t) = N(0, 0.25) \tag{12}$$

2.2. Fadhloun-Rakha Bicycle-Following Model

For the most part, the car-following strategy of the FR-model remains valid for modeling the longitudinal single-file motion of bicycles. Specifically, the functions governing collision avoidance, steady state behavior, and human behavior modeling would have the same functional forms. For the aforementioned functions, the differences between vehicular traffic and bicycle

traffic would be expressed in the adopted values of their different parameters. The latter is not the case for the vehicle dynamics model [16], which requires the implementation of structural modifications in order to make it descriptive of the maximum acceleration behavior of bicycles. The needed modifications [3] implemented at the level of the vehicle-dynamics model to make it representative of longitudinal bicycle traffic are summarized in TABLE 1.

TABLE 1 Summary of the dynamics model for bicycles and cars[17]

	Bicycles	**Cars**
Tractive Force	$\min\left(\eta_{eff}\eta_{gears}\frac{m_{cyclist}P_{ftp}}{v_n}, m_{ta}g\mu\right)$ η_{eff} depends on bicycle chain η_{gears} depends on gears and bike geometry m_{ta} depends on center of gravity position	$\min\left(\frac{\beta\eta_d P_n}{v_n}, m_{ta}g\mu\right)$
Rolling Resistance	mgC_{rr} C_{rr} depends on road type	$mgC_{r0}(C_{r1}v_n + C_{r2})$
Aerodynamic Resistance	Same formulation except: C_dA_f depends on cyclist physique and posture on bike	C_dA_f depends on car shape
Grade Resistance	Same formulation	

In the above table, the power produced by the bicyclist as a result of pedaling is estimated as the product of the cyclist weight and the highest average power that can be sustained over a certain period of time, commonly known as the functional threshold power (FTP factor in W/kg). Understandably, the FTP factor depends on several variables such as gender, stamina, and the time interval as shown in TABLE 2. For instance, a male cyclist in good shape is able to generate an average of 3.91 W/kg over a 1-hour period and a higher average of 8.28 W/kg over a 1-minute period. For a female cyclist in the same shape, these values decrease slightly to 3.39 W/kg and 6.75 W/kg over the same time periods. Finally, in order to account for the losses incurred while the pedaling power is transmitted to the rear wheel, several efficiency factors are applied. These factors attempt to model the effect of the bicycle gears and the friction at the level of the bicycle chain.

TABLE 2 Maximal power outputs for different cyclist categories [18]

Bicyclist Condition	Male			Female		
	1 min	5 min	1 hour	1 min	5 min	1 hour
World Class	11.50	7.60	6.40	9.29	6.61	5.69
Exceptional	10.35	6.57	5.51	8.38	5.68	4.87
Excellent	9.66	5.95	4.98	7.84	5.13	4.38
Very good	8.97	5.33	4.44	7.3	4.57	3.88
Good	8.28	4.70	3.91	6.75	4.02	3.39
Moderate	7.48	3.98	3.29	6.12	3.37	2.82
Fair	6.79	3.36	2.75	5.57	2.82	2.32
Untrained	5.87	2.53	2.04	4.85	2.07	1.67

Through the reconfiguration of vehicle-related input variables and the integration of new parameters that replicate the characteristics of bicycle/bicyclist system, the suitability of the resulting bicycle following model was validated in both constrained and unconstrained conditions using experimental data. The main benefit of the model lies in its robustness and its ability to model bicyclist behavior variability. In fact, the proposed model is the only existing model that is sensitive to bicyclists' physical characteristics and roadway surface conditions.

3. LATERAL MODEL FORMULATION

In the first part of this project, the FR bicycle-following model was used as the foundation to develop a two-dimensional motion model that allows lateral movements and overtaking. The following section provides a step-by-step process used to implement the proposed lateral module. After that, the performance of the model is illustrated using some basic scenarios.

3.1. Model Formulation

Step 1

The first step of the proposed methodology defines three elliptic regions around the bicycle that will be used in the following stages of the algorithm for the determination of the angular motion of the bicycles. FIGURE 1 presents a schematic view of the defined ellipses, which have the following characteristics:

- *View zone region*: As its name connotes, the biggest half ellipse is used to define the view zone of the bicyclist. Only objects/bicycles that are present within this zone have an influence on the subject entity for which the calculations are being made. The constructed list of neighboring bicycles is then used to investigate the bike's surroundings and find the most likely longitudinal leader along with any potential lateral co-riders that would

obstruct desired overtaking maneuvers by the subject bicycle. Alternatively, if no bikes are present within that region, the subject bicycle will be considered a leader and behave as such. One might argue here that a simpler optimization technique, such as the k-d tree algorithm, could have easily achieved that objective. While that is mostly true, the research team opted to use the proposed methodology because of its adaptability and its flexibility in relation to changing the view zone area by controlling the major and minor radii of the ellipse. For the remainder of this study, the ellipse radii R_v and r_v will be set to 15m and 2m respectively.

- *Steady-state zone region*: The definition of this semi-elliptic region plays a fundamental role in the proposed model and its stability. The area of this ellipse varies at every time step as its major radius R_{ss} is a function of the steady-state spacing of the Van Aerde steady-state traffic stream model, making it sensitive to both the speed v_n and the roadway characteristics. The value of the parameter R_{ss} is calculated as presented in Equation 13. In the equation, m_{ss} designates a steady-state margin factor that is greater than 1. The main purpose of this parameter is to allow the cyclist to look at the area beyond the steady-state spacing in order to investigate any need for lateral motion before getting too close to the leader. In other words, the margin induced by m_{ss} aims at ensuring the initiation of any necessary overtaking maneuver before the collision avoidance strategy kicks in. In what follows, m_{ss} will be set to 1.25, and the minor radius of the ellipse r_{ss} to 0.75m.

$$R_{ss} = m_{ss}\left(c_1 + \frac{c_2}{(v_f - v_n)} + c_3 v_n\right) \tag{13}$$

- *Safety zone region*: Put simply, this area plays the same role as the spacing at jam-density in longitudinal motion models. The ellipse represents the area occupied by the bicycle and the rider that is supplemented by a minimum safety margin in the two-dimensional space. In that regard, the major radius R_s is set to be consistent with longitudinal following models as expressed in Equation 14.

$$R_s = s_j - \frac{l_n}{2} \tag{14}$$

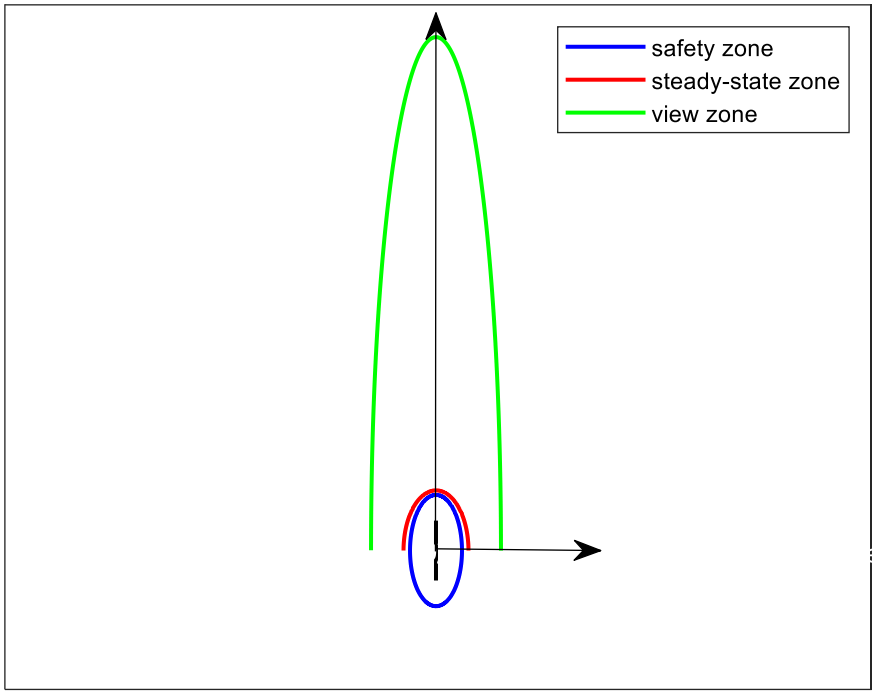

FIGURE 1 A schematic view of the different regions used in the proposed model

Step 2

Here, the core component of the lateral control strategy is developed. Specifically, the main purpose of this step is to determine the angular motion behavior of the subject bicycle at the next time step based on the current system status. The output consists of the angle θ_n, which reflects the bicycle heading. The biggest challenge at this stage was ensuring that the determination of the angle θ_n is the result of a dynamic and stochastic procedure. The randomness induced by the model stochasticity would make it more adaptable and tunable to naturalistic bicycle behavior.

Technically speaking, the proposed methodology is innovative in that it makes use of the intersection of the steady-state zone region of the following bicycle and the safety zone region of its leader to determine the angular motion parameter θ_n. The mathematical expression of θ_n is shown in Equation 15-17.

$$\theta_n = max(\theta_1, \theta_2) \tag{15}$$

$$\theta_1 = atan\left(\frac{y_{P1} - y_n}{x_{P1} - x_n - \frac{l_n}{2}}\right) \tag{16}$$

$$\theta_2 = atan\left(\frac{y_{P2} - y_n}{x_{P2} - x_n - \frac{l_n}{2}}\right) \quad (17)$$

Where x_n and y_n denote the longitudinal and lateral positions of the following bicycle respectively, and x_{P1}, y_{P1}, x_{P2}, and y_{P2} are the coordinates of the intersection points between the two aforementioned ellipses.

Step 3

The factor θ_n is not the only parameter responsible for the angular motion in the proposed lateral control strategy. In fact, it needs to be complemented by another variable that is representative of the feasibility of lateral motion at the specified time step. In that regard, we introduce λ_n as a trinary variable with states represented by the numbers -1, 0, or 1. The product $\lambda_n.\theta_n$ is then used as the global indicator of the rate of angular motion and its direction. The associated three-level logic to determine λ_n is as follows:

- $\boldsymbol{\lambda_n}$ **= 1**: This value indicates the feasibility of a lateral maneuver to the left. To ascertain that, we verify that there are no obstacles, whether in relation to the road geometry or other bicycles surrounding the subject bicycle that would stop it from moving laterally to the left.
- $\boldsymbol{\lambda_n}$ **= -1**: If left movements are deemed impossible, the possibility of initiating a right lateral maneuver is investigated in the same manner described above.
- $\boldsymbol{\lambda_n}$ **= 0**: If it is determined that lateral motion cannot be achieved in both directions, then setting λ_n to zero would ensure that the following bicycle continues its longitudinal motion. That would result in it continuously closing the gap separating it from its leader until the collision avoidance strategy is activated. The adopted strategy is fundamentally similar to that used in social force models where certain repulsive forces are used to achieve a similar objective.

Step 4

Once all the parameters responsible for lateral motion are determined, we proceed with calculating the acceleration using the expression of the FR bicycle-following model described in the background section. For illustration purposes in what follows, we consider a typical scenario in which a bicyclist is traveling on a dry and flat asphalt road. This would result in the values of the grade G, the rolling coefficient C_{rr}, and the friction coefficient μ to be equal to 0, 0.004, and 0.8 respectively. For the remaining variables, the following assumptions are used again:

- The cyclist is an untrained male (P_{ftp}=2.04 W/kg),
- The bicycle weighs 8 kg,
- The proportion of the total mass on the rear axle equals 0.60,
- The aerodynamics coefficients are such that $C_d A_f$ = 0.4,
- The desired deceleration level d_{des} is equal to 1.5 m/s^2,
- The total efficiency factor $\eta_{eff}\eta_{gears}$ = 0.62.

Considering the above along with a cyclist weight of 75 kg, the computation of the acceleration using the FR model would only require the values of the road characteristics (u_f, u_c, q_c, k_j) along with those of the three calibration parameters (g_1, g_2, g_3).

Step 5

At this level, the acceleration vector of the following bicycle is completely defined in terms of direction and magnitude. Subsequently, we can proceed with updating the speed vector along with both the lateral and longitudinal speeds and positions as presented in Equations 18-20.

$$v_n(t) = v_n(t - \Delta t) + a_n(t).\Delta t \tag{18}$$

$$x_n(t) = x_n(t - \Delta t) + v_n(t)cos\big(\lambda_n(t).\theta_n(t)\big).\Delta t \tag{19}$$

$$y_n(t) = y_n(t - \Delta t) + v_n(t)sin\big(\lambda_n(t).\theta_n(t)\big).\Delta t \tag{20}$$

Step 6

Lastly, the final step of the algorithm deals with the way in which the subject bicycle returns to its original lane (or around its initial lateral position) once it overtakes its original leader. At this stage, the research team opted to use a behavior symmetrical to that used by the subject bicycle at the start of the passing maneuver.

Mathematically, that is achieved through memorizing the opposite of the angular motion parameters θ_n, specific to a certain overtaking maneuver from start to end, in a vector variable Γ_n in a reversed order as defined in Equation 21.

$$\Gamma_n(\theta_n) = -\begin{bmatrix} \theta_n(t_{end}) \\ \theta_n(t_{end} - \Delta t) \\ \theta_n(t_{end} - 2\Delta t) \\ \vdots \\ \theta_n(t_{ini} + \Delta t) \\ \theta_n(t_{ini}) \end{bmatrix} \tag{21}$$

Where t_{ini} is the time step at which the concerned overtaking maneuver started and t_{end} presents the final time step after which the initiated lateral motion is no longer needed as the leading bicycle is no longer obstructing the longitudinal motion of the follower.

The collected values in Γ_n are used in a sequential order to define the angular motion of the overtaking bicycle once the condition presented in Equation 22 is met. This condition verifies that the rear of the subject bicycle is past the safety zone region of the bicycle it is overtaking. Finally, we note that given the possibility of overtaking, the algorithm needs to be updated at every time step in order to identify the current ranking of the different bicycles in the queue.

$$x_n - l_n \; > \; x_{n-1} + R_s \tag{22}$$

A final remark regarding the above equation relates to the indexing choice of x_n and x_{n-1}. Even though at the time Equation 22 becomes valid bicycle *n* is actually the leader of bicycle *n-1*, we chose the above indexing to be consistent with the other equations and to avoid any confusion.

3.2. Model Evaluation

To demonstrate the model's performance, the study implements the proposed algorithm for different basic case scenarios. In each of the scenarios, all the model parameters except for the geometry of the road were set to a fixed value. Moreover, the speed of the leading bicycle was set to a constant low value in order to allow overtaking conditions to arise. The chosen scenarios involve only a single follower to the leading slow bicycle. That would allow for a perfect visualization of the algorithm's behavior and its generated trajectories. In fact, a higher number of followers would result in significant interactions between the bicycles, making it harder to understand the actual model logic. The following case scenarios are considered:

- *Scenario 1*: The road geometry would only allow overtaking to the left side as shown in FIGURE 2.a. This scenario is equivalent to the case in which overtaking is feasible in both directions. This is because the left side was prioritized during the computation of the directional parameter λ_n.
- *Scenario 2*: The road has borders on both sides, making overtaking impossible even if desired by the following bicycle as illustrated in FIGURE 2.b.

FIGURE 2.a demonstrates the ability of the proposed lateral control strategy to overtake a leading vehicle traveling at a low speed while responding to the constraint imposed by the road border. While it might be argued that the proposed model is only tentative and does not necessarily represent empirical bicycle behavior, this is partially negated by the information presented in FIGURE 2.b. In fact, the model output for the second scenario is merely an extension of the original FR longitudinal motion model, which was already validated using longitudinal bicycle data and was proven to replicate observed empirical behavior effectively. While lateral motion was triggered in the beginning, it was quickly stopped, and the following bicycle continued following its leader longitudinally. Furthermore, FIGURE 2.b demonstrates that even though the lateral and longitudinal modules interact with each other, they are implemented separately from one another. That opens the door to other researchers to implement any longitudinal motion logic they see fit for their purposes.

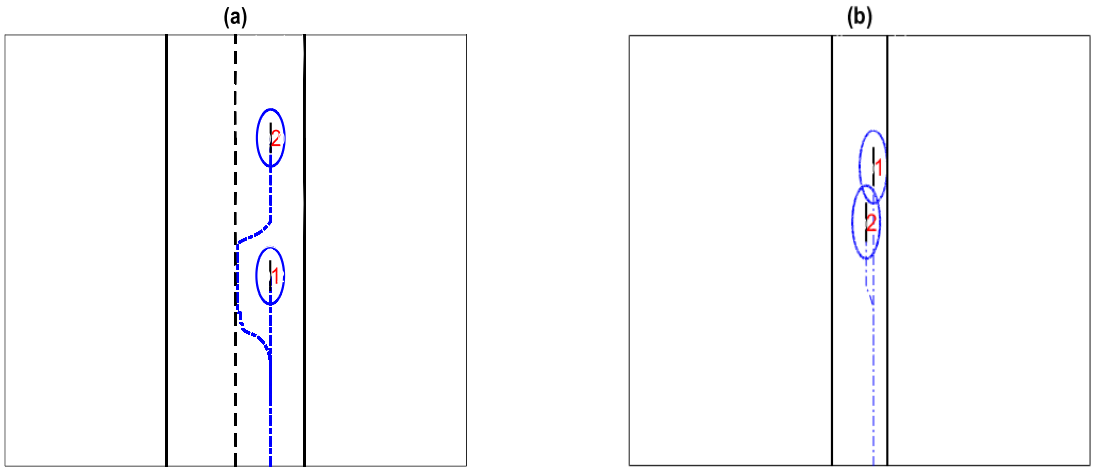

FIGURE 2 Projection of the simulated trajectories on the two-dimensional domain for: a) Scenario 1; b) Scenario 2

4. NATURALISTIC DATA COLLECTION AND MODEL VALIDATION

The second part of this project describes the different methodologies and techniques involved in extracting naturalistic cycling trajectories from existing video feeds. This is significant because most existing cycling datasets are experimental and were collected in a controlled environment under very specific conditions (single-file motion, no overtaking).

Due to the proliferation of machine learning and computer vision techniques, it is becoming feasible to acquire reliable naturalistic traffic data in a cheap and efficient manner from video datasets. This is especially true in the case of bicycles as they are not as instrumented as cars, which would not allow the capture of their full surroundings in the context of a naturalistic data collection study. In the case of this study, the complete video dataset is the result of a previous Virginia Tech Transportation Institute study conducted in collaboration with SPIN in which continuous video data at several fixed locations at the Virginia Tech campus was collected over a seven-month period. For the purpose of this research, only a portion of the above dataset from a single location is used. The selected dataset was collected over 55 days between the months of September and December 2019 using a roof-mounted high-definition camera facing a non-signalized three-way stop intersection. The selected dataset includes approximately 810 hours of 3720 x 1728 pixels videos recorded at a frequency of 30 Hz. It is worth mentioning that the ultimate purpose of this task is to collect the naturalistic data at a high level of detail. From a machine learning and computer vision perspective, we are using existing tools to achieve our objective in the simplest way possible.

4.1. Video Processing

The first step of this research effort involves the identification of bicycle events from the different videos. Given the large size of the video feed, manual data reduction was considered impractical as it would be a costly and lengthy process. Instead, the research team opted for an automated

method that makes use of existing object detection techniques. Specifically, a two-step object-detection algorithm was developed.

The first step of the proposed algorithm uses a cascade detector based on the histograms of oriented gradients (HOG) with 11 stages to detect regions of interest in the video frames that might be bicycles. The number of stages used to train the detector is not random. In fact, the research team initially used a database composed of 400 positive images and 900 negative images to train detectors with different numbers of stages (5, 7, 9, 11, and 13) and a false alarm rate fixed at 2.5%. The number of training images and stages were purposefully set relatively low in order to ensure a quick training process. The focus of the research team at this point was to ensure that the number of stages of the detector is high enough to detect a significant percentage of true positives regardless of the number of false positives as these would be addressed and eliminated later. Next, the trained detectors were run on a one-hour video from the database at 5-second intervals to quantify their performance. The outputs from this step consisted of bounded areas that highlight regions that might be inclusive of bicycles in the examined video frame, as illustrated in FIGURE 3.

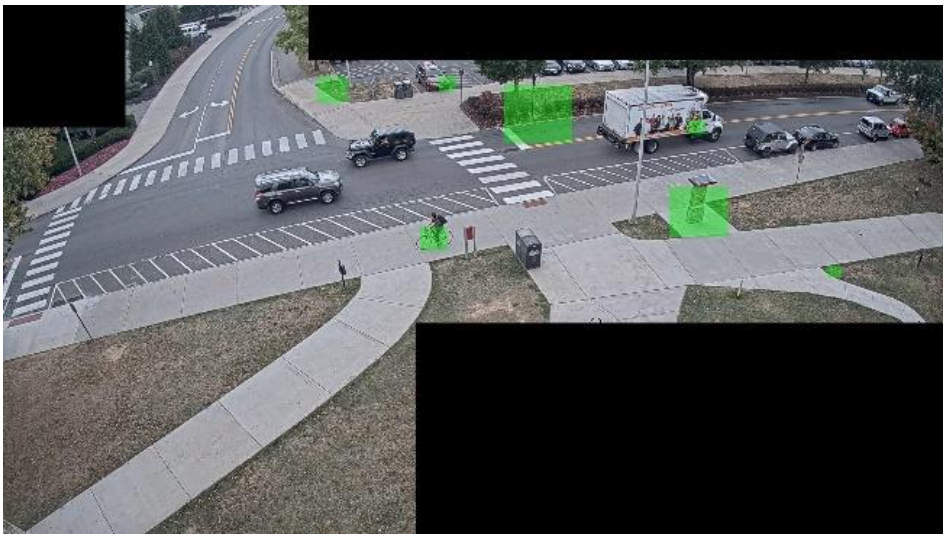

FIGURE 3 Sample output of HOG detectors

The following metrics were used to evaluate each of the detectors:

- The number of true positives: These refer to the bounded areas identified correctly by the detectors in that they contain a bicycle.
- The number of false positives: These correspond to the bounded areas identified wrongfully by the detectors.
- The number of false negatives: These account for the cases in which a bicycle was present in the video frame without being detected.

It is noteworthy that all the detectors, regardless of the number of stages, were able to identify 42 out of the 44 bicycle trips. However, a deeper look into the results highlighted the significant differences between the metrics listed above. FIGURE 4 plots the variation of the true positives (FIGURE 4.a), false positives (FIGURE 4.b), and false negatives (FIGURE 4.c) against

the number of stages used to train the detector. The main revelation from the figures is that the total number of false positives significantly decreases as the number of training stages increases. However, the observed decrease is also accompanied by a decrease in the number of true positives and an increase of the number of false negatives. Based on the observed patterns, it is evident that the detector with 11 stages is the best among those investigated albeit with a relatively high number of false positives. To address that issue, the bicycle-detection algorithm was complemented with another technique to reduce false detections.

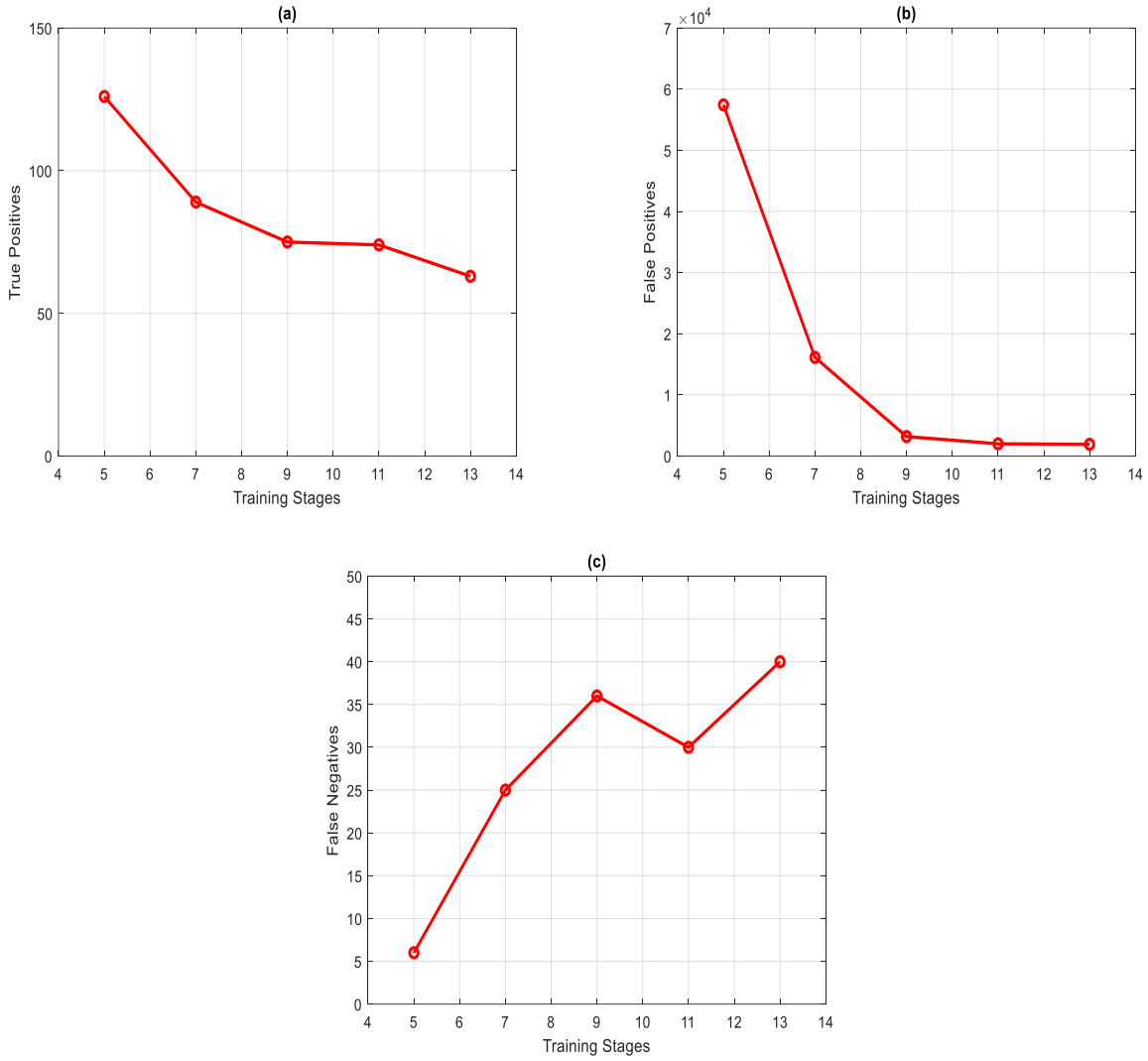

FIGURE 4 Variation of the detector metrics as a function of the number of training stages a) True positives; b) False positives; c) False negatives

In the second stage of the algorithm, the highlighted areas of interest are selected for further examination using a semantic segmentation network that attempts to classify every pixel in them and assign them to different classes. For that purpose, the research team selected an existing pre-trained DeepLabv3+ network [19], which is a convolutional neural network (CNN) designed for semantic image segmentation. The network is available for download at the Mathworks website

and was trained using the CamVid dataset [20] from the University of Cambridge. The dataset consists of a collection of street-level images that are segmented at the pixel-level using 32 semantic classes (such as bicyclist, pedestrian, and car) as shown in FIGURE 5.

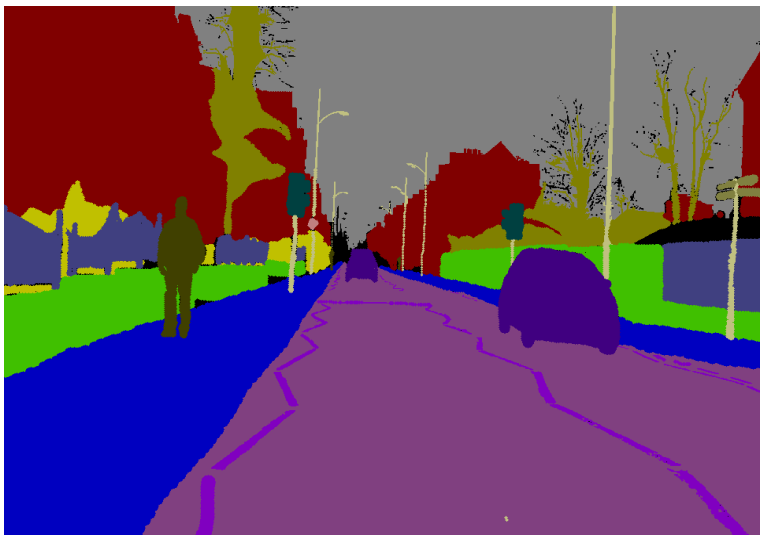

FIGURE 5 Sample image from the CamVid dataset

As mentioned earlier, the main reason behind the semantic segmentation phase is to eliminate the false positives that were detected by the HOG detector in the previous step. This was achieved by comparing the number of pixels that were classified as bicyclist and the total number of pixels in the investigated area. If the ratio between the two values is greater than a set threshold of 5% in at least one of the highlighted regions, the examined video frame was saved for manual confirmation. Otherwise, it was rejected (FIGURE 6). The application of the semantic segmentation algorithm over the areas identified by the HOG detector proved to be quite successful. Using the two algorithms together, the number of frames selected for further investigation decreased from 683 frames to 89 frames without any decrease in the number of bicycles detected. The algorithm was able to accurately detect 42 out of the existing 44 bicycle events (95.5%).

With the algorithms ready, the different videos of the database were processed using the HOG detector in conjunction with the semantic segmentation at 5-second intervals. This is mainly due to the heavy computational toll of those algorithms. However, that did not have much effect on the accuracy of the algorithm in bicycle detection as demonstrated earlier. Furthermore, to further illustrate the performance of the algorithm in relation to false positives, it was run on the 4-hour video between 6AM and 10AM on Christmas day, which is a period in which no bicycles were present. The algorithm saved only 21 frames for further investigation out of the 2880 frames examined (< 1%).

FIGURE 6 Sample output after semantic segmentation

This step resulted in an image database in which the video frames selected for further investigation were saved separately with information pertaining to the date, time, and timestamps in their respective videos. Through a manual data reduction process, the resulting database was investigated to identify the different cycling trips and note their start and end times. The result of this process was the identification of 2259 cycling events. Out of those, a total of 619 cycling events were selected to be processed in the next step. Most of the selected trips have a duration between 10 and 20 seconds with an average of 16.1 seconds and a median of 14.0 seconds.

4.2. Trajectory Extraction

With the start and end times of the different trips known, the corresponding video sections were isolated and prepared for the next step, which relates to the extraction of the trajectories in the video pixel domain.

There are two approaches that can be used to achieve the latter. The first approach is quite straightforward but is only possible for a relatively low number of trajectories. For each of the trajectories, a simple script is used to extract the frames from the video at 0.2 seconds allowing for the user to manually click on the position of the bicycle and the vehicles interacting with it. Two moving perpendicular lines are implemented to help a data reductionist detect the intersection of the front of the bicycle wheel and the pavement as shown in FIGURE 7. In the background, the script saves the location of the clicks in the (x, y) domain of the video frames (a 3720×1728 pixel grid), thus collecting the trajectories for further processing. It is necessary to note here that if any obstructions (such as a car, a tree, or a structure) interfere with the collection of precise bicycle location data, the bicycle coordinates will not be collected for that specific timeframe. An interpolation algorithm will be used to estimate the location of the bicycle at those time steps at a later step.

FIGURE 7 Sample screenshot from the trajectory collection process

Before moving on to the description of the next step, we would like to note that our final objective is to extend this work by extracting bicycle trip events that occurred over the entire seven-month period at all 14 locations on campus. However, this would require additional automation to complete the trajectory extraction process. In fact, only 1.2 out of the available 49.5 terabytes of available videos have been used so far. Assuming, hypothetically, that a perfect proportionality exists between the number of bicycle trips and the size of the video database, the expected number of trips expected to be found in the entire video dataset would be in excess of 90,000. Even more, once the tasks requiring manual labor are removed, the research community would have access to a comprehensive automated trajectory extraction framework that can be applied to similar videos.

In that regard, the research team is currently developing an automated tool that can extract the trajectories and replace the data reduction process. Without going into much detail, the algorithm uses the Hough transform for the detection of bicycle wheels to determine their contact point with the road surface. To achieve that, edge detection techniques are first used to isolate the bicycle trip on a black and white background as shown in FIGURE 8. After that, the Hough transform is used to detect the wheels as shown in FIGURE 9. However, the research team is still working on solving the most challenging part of this process, which deals with the fine-tuning of the algorithm to assign detected points to their corresponding trajectories and automatically exclude false positives.

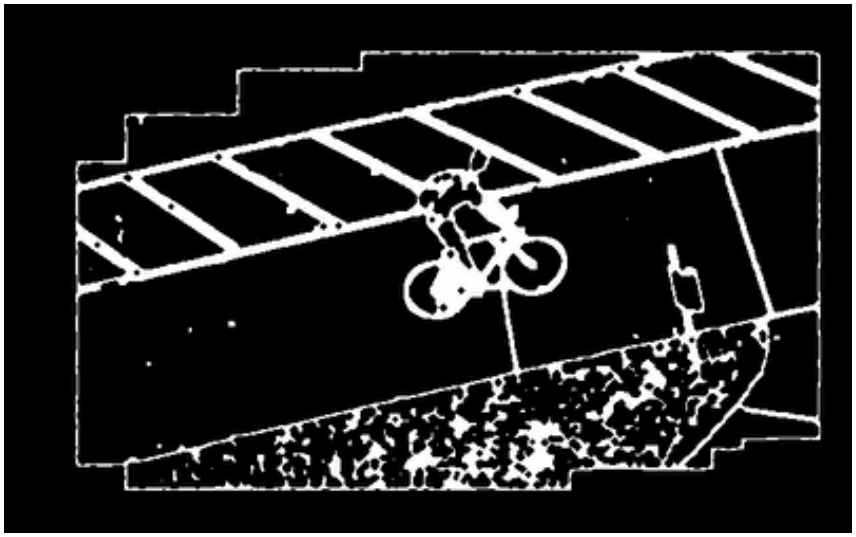

FIGURE 8 Image filtering using edge detection techniques

FIGURE 9 Detection of bicycle wheels using Hough transform

4.3. Intersection Surveying

In order to convert the extracted trajectories into naturalistic trajectories, a grid map overlay of GPS coordinates at specific locations is needed. For this, the aforementioned map allows the conversion of the pixel-based trajectories into distance-based trajectories using the multi-step algorithm described thereafter.

To achieve the stated objective, the research team started by creating a mesh of approximately 400 points as shown in FIGURE 10. As the figure shows, the points are heavily concentrated around the edges of sidewalks and road crossings because they are the easiest to identify in the videos and in the field. This is quite useful for the next step as both sidewalks and road crossings are frequently used by the bicyclists to complete their trips. Initial attempts to collect the GPS coordinates at the specified locations were made using accessible tools such as Google

Earth and existing GPS mobile applications. However, these attempts proved unsuccessful due to the small distances involved and the relative low accuracy of those tools when used in this context. As a result, a surveying campaign using professional high-precision tools was conducted to acquire the required coordinates, which are expressed in the Northing-Easting-Elevation coordinates system. Since the investigated area is relatively flat, the elevation data can be ignored without major repercussions on the results. In what follows, we will refer to the data collected in this step by the transform matrix.

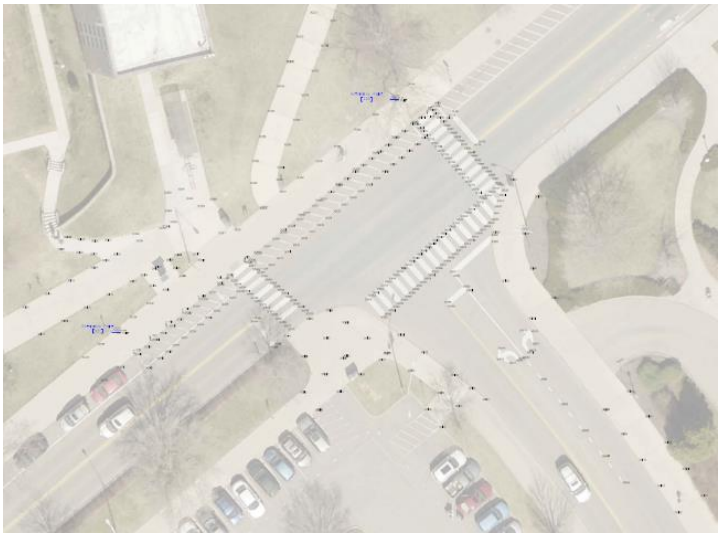

FIGURE 10 An aerial view of the surveyed area and the collection points

4.4. Trajectory Transformation and Results

The final step of this methodology deals with the conversion of the extracted trajectories expressed in the video pixel domain to actual naturalistic trajectories and the determination of speed, acceleration, and distances traveled. This constitutes the final product of this study and allows us to validate the proposed model against the resulting naturalistic bicyclist dataset. The trajectory transformation process is achieved using the following multi-step algorithm:

1. A linear interpolation algorithm is initially used to complement the extracted trajectories with estimated values at the time steps for which the determination of the bicycle location was impossible due to the presence of visual obstructions.
2. Next, the trajectories are exponentially smoothed using a smoothing factor of 0.5. The purpose of the exponential smoothing operation is to address the noise and the zigzag-like features that might be present as a result of the manual trajectory extraction process. At this level, the trajectories will look similar to the two sample trajectories presented in FIGURE 11.
3. For each of the observations composing a trajectory, one of the closest convex hulls containing the observation and delimited by three points from the transform matrix is identified.
4. Since we have access to the coordinates of the points defining the convex hull in both coordinate systems, the coordinates of the trajectory observation could be approximated in the Northing-Easting coordinate system using a triangulation algorithm.

5. Once Step 4 is completed for all the observations, the speed profile associated with the obtained trajectory is determined and smoothed through the application of a third order Savitzky–Golay filter.
6. In a similar fashion to Step 5, the acceleration profile is obtained from the smoothed speed profile and smoothed using a similar Savitzky–Golay filter.
7. Finally, the speed profile, the distance traveled, and the coordinates of the trajectory in the Northing-Easting coordinate system are updated backwards to account for the effect of the two-layer filtering that was applied.

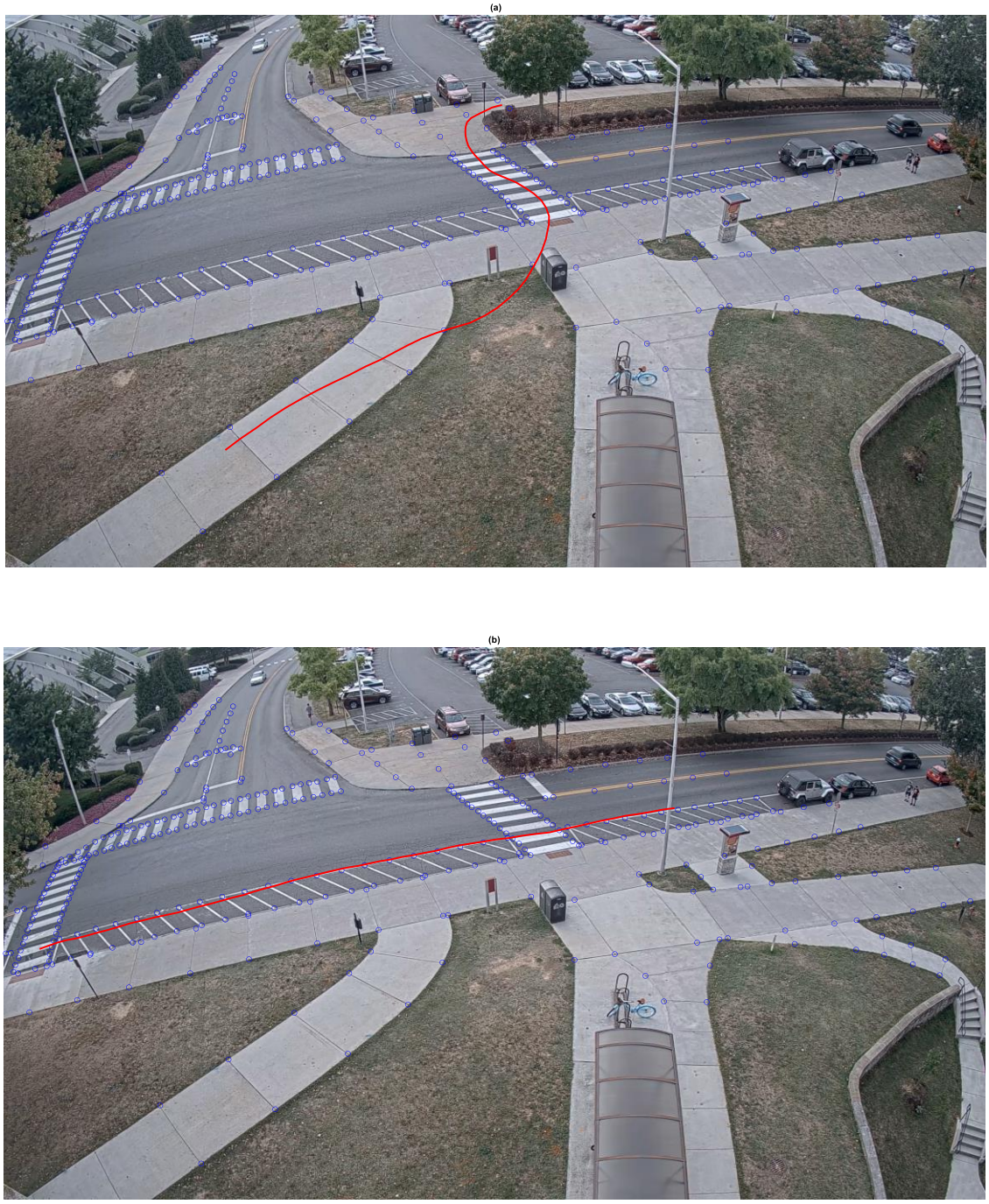

FIGURE 11 Sample trajectories in the video pixel domain

Sample results from this step are presented in FIGURE 12 and FIGURE 13. FIGURE 12 shows the resulting trajectories in the Northing-Easting coordinate system corresponding to the two trajectories presented in FIGURE 11. The figure demonstrates the success of the proposed multi-step algorithm in conserving the shape and main features of the extracted trajectory.

Meanwhile, FIGURE 13 illustrates the distance traveled, speed, and acceleration profiles corresponding to the trajectory presented in FIGURE 11.a and FIGURE 12.a.

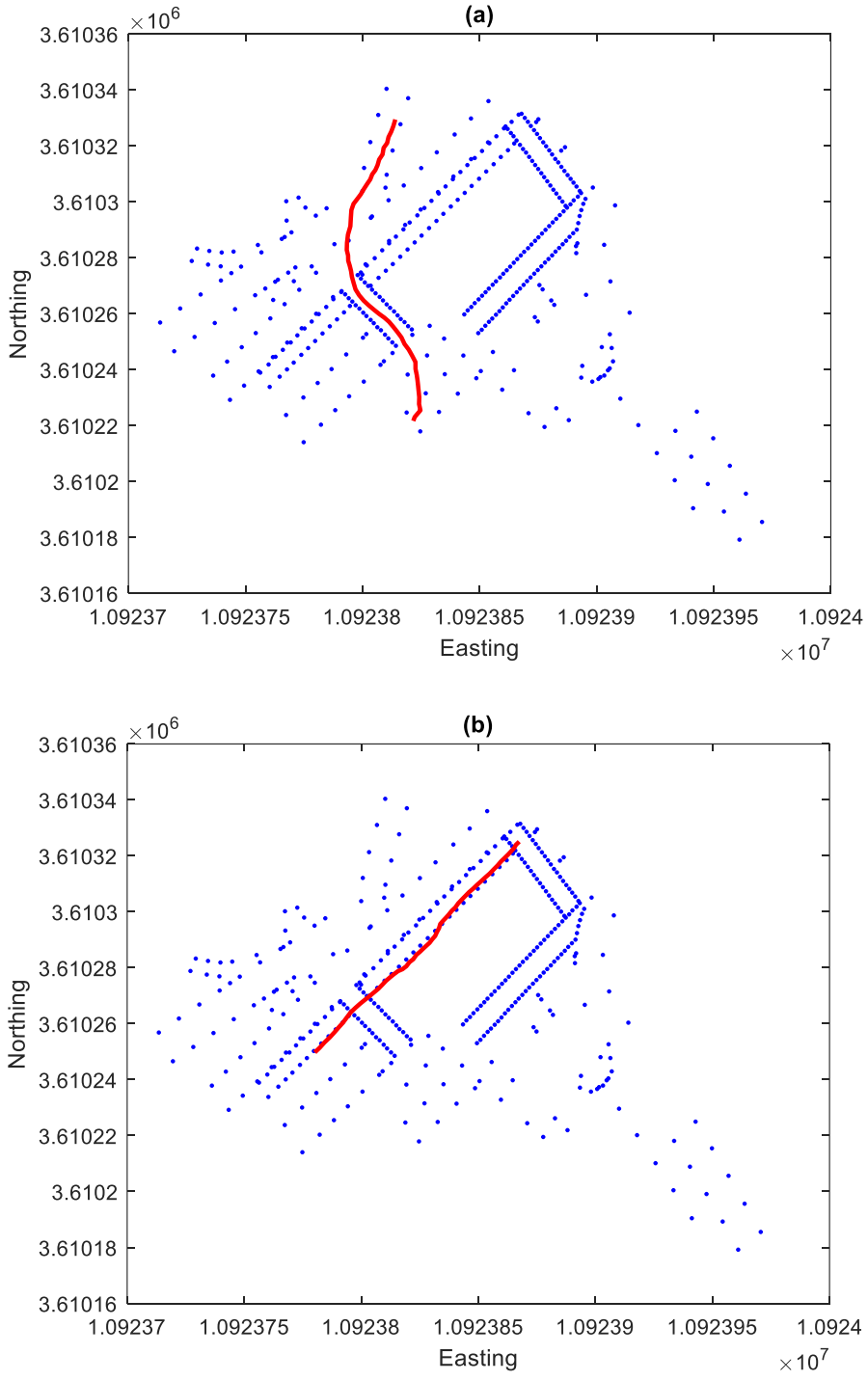

FIGURE 12 Sample naturalistic trajectories after the triangulation procedure

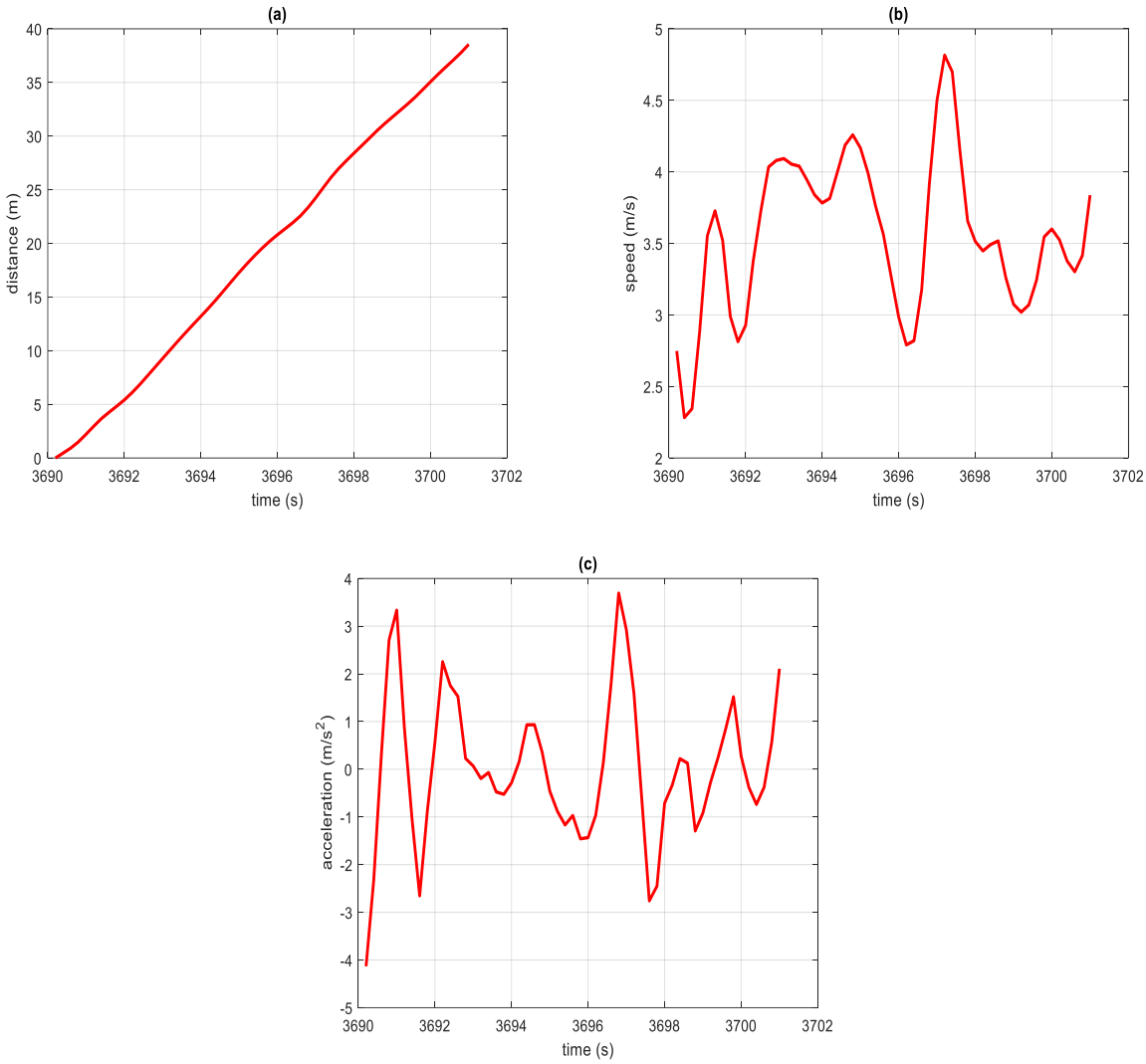

FIGURE 13 Extraction of the distance traveled, speed, and acceleration profiles for a naturalistic trajectory a) Distance traveled; b) Speed profile; c) Acceleration profile

Finally, the histograms of the instantaneous accelerations and speeds from all the 619 trajectories is investigated to confirm the consistency of the obtained values with bicycle behavior. The results, which are plotted in FIGURE 14, show that the results are concentrated around low acceleration levels and speeds that are quite typical for bicyclists. Furthermore, the range of the observed values are physically feasible for a bicycle. With the completion of this step, the collection of the naturalistic cycling dataset is ready to validate the proposed model in what follows.

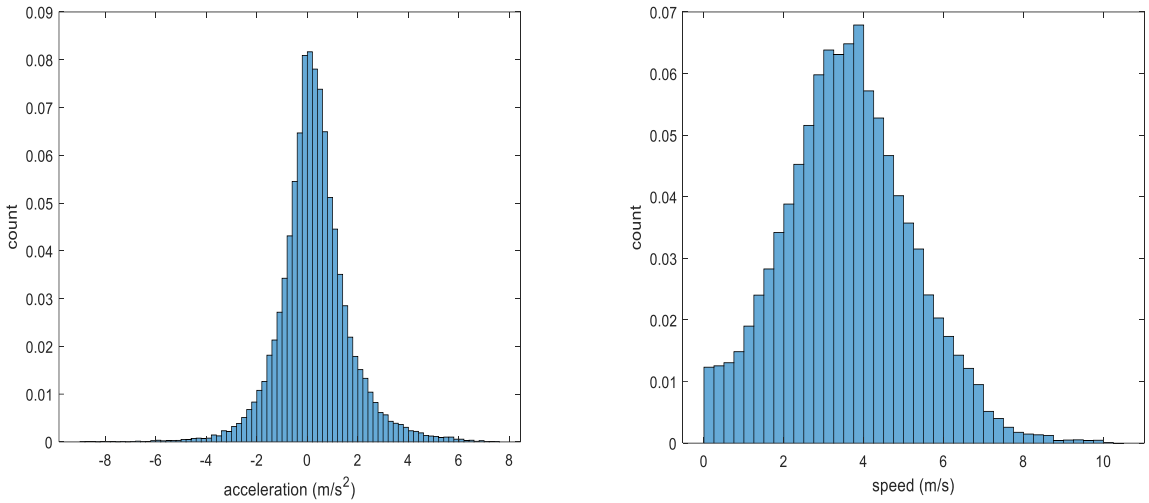

FIGURE 14 Histogram of the instantaneous accelerations and speeds of the aggregated extracted trajectories

5. CONCLUSION

In this project, the research team proposed a two-dimensional comprehensive model for bicycle motion behavior modeling. The formulation of the model is achieved by integrating a lateral control strategy with the Fadhloun-Rakha longitudinal motion model formulation. The developed lateral module makes use of new parameters that control the angular and directional motion of the bicycle, allowing for overtaking to occur while accounting for the bicycle/bicyclist system surroundings. The proposed model is the first point-mass dynamics-based model for the description of the longitudinal and lateral behavior of bicycles in both constrained and unconstrained conditions. By having the FR bicycle-following model as both the governing module of longitudinal behavior and a dynamic lateral module, the proposed model is able to model bicyclist behavior variability. Furthermore, it is the only existing model that is sensitive to roadway surface conditions and the bicyclist's physical characteristics.

Preliminary investigations of the proposed lateral module show that it is successful in terms of allowing lateral movements and overtaking. However, the research team is aware that the lateral model outputs may need further tuning to be representative of naturalistic bicycle behavior. In that regard, the research team would consider extending the collected naturalistic dataset by extracting the trajectories of the different entities interacting with the bicycles. That would give us access to a two-dimensional comprehensive dataset that is both inclusive of the bicycles' trajectories as well as those of their surrounding entities.

In relation to the acquisition of the naturalistic dataset, this paper described the development of a comprehensive framework that would allow for the collection of naturalistic cycling trajectories from video feeds. Even though the current naturalistic dataset is composed of only 619 trajectories, it will be useful to traffic researchers in several mobility applications such as the validation of studies investigating bicycle motion behavior. Furthermore, once this dataset is complemented by the trajectories of the surrounding cars, the resulting dataset will contribute to

a better understanding of bicyclists' behavior around cars, leading to a better understanding of the interactions between bicycles and other modes of transportation.

The research team faced two main challenges during this study. The first challenge deals with automating the process of extracting the bicycle trajectories from the videos through the detection of bicycle wheels. In fact, the number of trajectories in the resulting dataset is limited due to the problems encountered while trying to complete that process. Once these problems are addressed and the process is entirely automated, the size of the trajectory database will increase significantly. More importantly, the proposed methodology will be completely transferable for use by other researchers at different locations. The second challenge relates to the collection of the transform matrix needed to transform the video trajectories into actual trajectories. Due to the small distances involved, typical tools such as Google Maps and existing GPS applications cannot be used; instead, a professional surveying campaign of the observed area is needed.

6. REFERENCES

1. Kurtc, V. and M. Treiber, *Simulating bicycle traffic by the intelligent-driver model-Reproducing the traffic-wave characteristics observed in a bicycle-following experiment.* Journal of Traffic and Transportation Engineering (English Edition), 2020.
2. Andresen, E., et al. *Basic Driving Dynamics of Cyclists*. 2014. Berlin, Heidelberg: Springer Berlin Heidelberg.
3. Fadhloun, K.R., Hesham; Mittal, Archak, *Bicycle Longitudinal Motion Modeling.* 100th Annual Meeting Transportation Research Board, 2021.
4. Treiber, M., A. Hennecke, and D. Helbing, *Congested traffic states in empirical observations and microscopic simulations.* Physical review E, 2000. **62**(2): p. 1805.
5. Fadhloun, K. and H. Rakha, *A novel vehicle dynamics and human behavior car-following model: model development and preliminary testing.* International journal of transportation science and technology, 2020. **9**(1): p. 14-28.
6. Gazis, D., R. Herman, and R. Rothery, *Nonlinear follow-the-lead models of traffic flow.* Operations Research, 1961. **9(4)**: p. 545-567.
7. Pipes, L.A., *Car-following models and the fundamental diagram of road traffic.* Transportation Research, 1967. **1**: p. 21-29.
8. Pipes, L.A., *An operational analysis of traffic dynamics.* Journal of Applied Physics, 1953. **24**: p. 274-287.
9. Van Aerde, M. *Single regime speed-flow-density relationship for congested and uncongested highways*. in *74th TRB Annual Conference*. 1995. Washington DC.
10. Van Aerde, M. and H. Rakha. *Multivariate calibration of single regime speed-flow-density relationships*. in *Proceedings of the 6th 1995 Vehicle Navigation and Information Systems Conference*. 1995. Seattle, WA, USA: Vehicle Navigation and Information Systems Conference (VNIS) 1995. IEEE, Piscataway, NJ, USA,95CH35776..
11. Rakha, H., *Validation of Van Aerde's simplified steadystate car-following and traffic stream model.* Transportation letters, 2009. **1**(3): p. 227.
12. Wu, N. and H. Rakha, *Derivation of Van Aerde Traffic Stream Model from Tandem-Queuing Theory.* Transportation Research Record: Journal of the Transportation Research Board, 2009. **2124**(1): p. 18-27.

13. Rakha, H., P. Pasumarthy, and S. Adjerid, *A simplified behavioral vehicle longitudinal motion model.* Transportation Letters: The International Journal of Transportation Research, 2009. **1**(2): p. 95-110.
14. Rakha, H., et al., *Vehicle dynamics model for predicting maximum truck acceleration levels.* Journal of Transportation Engineering, 2001. **127**(5): p. 418-425.
15. Rakha, H., M. Snare, and F. Dion, *Vehicle dynamics model for estimating maximum light-duty vehicle acceleration levels.* Transportation Research Record: Journal of the Transportation Research Board, 2004. **1883**(1): p. 40-49.
16. Rakha, H., Pasumarthy, P., and Adjerid, S. , *A Simplified Behavioral Vehicle Longitudinal Motion Model.* Transportation Letters: The International Journal of Transportation Research, Vol. 1(2), pp. 95-110., 2009.
17. Soden, P. and B.J.J.o.B. Adeyefa, *Forces applied to a bicycle during normal cycling.* 1979. **12**(7): p. 527-541.
18. https://zwiftinsider.com/rider-categorization-based-on-ftp-how-do-you-rank/. 2016 [cited 2020 07/16].
19. Chen, L.-C., et al. *Encoder-decoder with atrous separable convolution for semantic image segmentation*. in *Proceedings of the European conference on computer vision (ECCV)*. 2018.
20. Brostow, G.J., J. Fauqueur, and R.J.P.R.L. Cipolla, *Semantic object classes in video: A high-definition ground truth database.* 2009. **30**(2): p. 88-97.