applied sciences

MDPI

Article

Predictive Safety Analytics of Double-Lane Roundabouts: An Explainable AI Approach to Geometric Compatibility and Opportunistic Driving

Salvatore Leonardi * and Natalia Distefano

Department of Civil Engineering and Architecture (DICAR), University of Catania, via Santa Sofia 64, 95123 Catania, Italy; natalia.distefano@unict.it
* Correspondence: salvatore.leonardi@unict.it

Featured Application

The analytical framework and the specific geometric critical intervals identified in this study provide road safety engineers and infrastructure planners with a reliable, data-driven tool to evaluate and mitigate operational risks at double-lane roundabouts. By objectively quantifying the behavioral anomalies associated with permissive double-lane configurations, these empirical findings can be directly applied within road safety audits to inform the physical retrofitting of existing uncanalized roundabouts into strictly channelized layouts, such as turbo-roundabouts. Furthermore, the extracted dimensional constraints for entry widths, exit widths, and connection radii offer targeted design parameters to mitigate opportunistic driving trajectories, actively guiding the proactive transition toward Self-Explanatory Roads in modern infrastructure planning.

Abstract

While two-lane roundabouts increase operational capacity, their excessive lateral space often induces opportunistic driving and safety-critical events. This study presents a data-driven framework to evaluate the non-linear impact of geometric constraints and vehicle kinematics on operational adequacy. Naturalistic driving data (5622 observations) were collected at four urban two-lane roundabouts. To address severe class imbalance without altering empirical distributions, a Cost-Sensitive Random Forest classifier was implemented. The algorithmic logic was decoded using Permutation Feature Importance and SHapley Additive exPlanations (SHAP) to evaluate risk associations across predefined geometric ranges. Results identify the executed maneuver as the dominant predictor, making geometric criticality context-dependent. SHAP dependence analysis revealed that intermediate two-lane exit widths (6–8 m) and exit radii (10–30 m) neutralize traffic calming, significantly increasing inappropriate trajectories and inappropriate speeds. Furthermore, wide two-lane entry widths (>8 m) are positively associated with yielding violations, a pattern hypothesized to stem from attentional dispersion and opportunistic gap acceptance. Overall, the observed behavioral classifications highlight the operational vulnerabilities of uncanalized two-lane roundabouts. By quantifying the geometric incompatibility between permissive layouts and angular maneuvers, these behavioral findings provide empirical motivation to consider physically channelized designs (e.g., turbo-roundabouts) as a design countermeasure.

Keywords: double-lane roundabouts; explainable AI; road safety; geometric compatibility; machine learning; naturalistic driving; opportunistic driving; turbo-roundabouts

Check for updates

Academic Editor: Luís Picado Santos

Received: 9 October 2026
Revised: 26 October 2026
Accepted: 27 October 2026
Published: 30 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Roundabouts now represent a globally consolidated design standard for managing road intersections, thanks to their proven ability to reduce severe accidents and streamline vehicular flows. Over the past two decades, the transition from signalized intersections to roundabouts has produced a drastic reduction in head-on and orthogonal collisions, due to the physical deflection imposed at the entrance and the one-way counterclockwise flow. However, while the effectiveness of single-lane roundabouts is widely documented and universally recognized, the geometry of double-lane roundabouts continues to raise heated debates within the international scientific and technical community. The increase in operational performance offered by the latter is systematically accompanied by a proliferation of conflict points, primarily linked to lane-changing maneuvers, side-by-side driving, and cutting trajectories during the crossing phases of the circulatory roadway [1,2].

The international academic literature has extensively investigated the causal link between the spatial architecture of roundabouts and user behavioral dynamics. Numerous studies have highlighted how the geometric width of the circulatory roadway and the high curvature radii at entry and exit tend to disregard the fundamental principle of Self-Explanatory Roads. In the absence of physical constraints and strict channelizations, drivers perceive excessive lateral space, increasing approach speeds and assuming opportunistic trajectories [3,4]. Empirical safety evaluations have further revealed a marked discrepancy across crash severity levels when comparing single-lane and double-lane configurations [2,4–6]. Specifically, both single-lane and double-lane roundabouts remain highly effective in curbing fatal and severe injury crashes compared to conventional intersections, as their curvilinear geometry physically eliminates high-energy head-on and right-angle impacts [2,6]. However, whereas single-lane layouts consistently reduce crashes across all severity tiers, uncanalized two-lane roundabouts exhibit significantly higher rates of property-damage-only (PDO) crashes and less favorable reductions in minor-injury collisions [5,6]. This divergence across severity levels stems from the absence of physical lane separation, which promotes frequent low-to-moderate impact angle, sideswipe, and exiting-circulating conflicts triggered by improper lane keeping and ambiguous yielding interactions [5,6]. Precisely in response to these operational criticalities, road engineering has developed alternative layouts, among which turbo-roundabouts stand out. The latter, through the use of mountable curbs and mandatory lane pre-selection, ensures the physical separation of flows, reducing the probability of sideswipe collisions and simultaneously optimizing the intersection's capacity [7,8].

From an analytical perspective, the modeling of driver behavior and the prediction of accidents have undergone a profound evolution. The first macroscopic approaches, based on safety performance functions (SPF) and generalized linear models [9–11], have been progressively complemented by microscopic and naturalistic investigation methodologies. Responding to the growing need to integrate data-driven mobility analytics into infrastructural engineering, the use of advanced machine learning techniques has made it possible to overcome rigid parametric assumptions, capturing the intrinsic non-linearities and the unobserved heterogeneity typical of the human factor [12–16]. Recent investigations conducted by combining tree-based classification algorithms and Explainable AI have explored the impact of isolated geometric parameters, such as the deflection angle and the width of the central island, confirming their ability to alter risk perception [17–21].

Despite such literary richness, a significant knowledge gap remains regarding the study of synergistic interactions between spatial variables and the user's cognitive workload during maneuvering phases. Most previous studies evaluated infrastructural parameters independently of one another, failing to quantify how the combined effect, for example, of

the circulatory roadway width and the entry radius, directly influences the aggressiveness of the entry maneuver [22–25].

The present study fits into this research path by proposing an advanced analytical framework, based on supervised machine learning and oriented towards predictive risk analysis. The investigation aims to mathematically formalize and quantify the weight of geometric configurations and their interactions in triggering behavioral deviations within double-lane roundabouts. Through the implementation of a Cost-Sensitive Random Forest classifier, integrated with the non-linear interpretability of SHAP (SHapley Additive exPlanations) values, the methodological framework extracts objective predictors from observational dynamics to identify the critical dimensional intervals responsible for human error. This empirical foundation also contextualizes the regulatory provisions of the Italian Ministerial Decree of 19 April 2006 [26], which explicitly prohibits multi-lane configurations on the circulatory roadway and at exit legs. In real-world traffic practice, this regulatory ban has introduced significant operational and management challenges. On the one hand, restricting new intersections to single-lane exits and rings often creates capacity bottlenecks and queue spillbacks along high-volume urban and suburban arterials, especially since channelized multi-lane alternatives (such as turbo-roundabouts) are not yet standardized within the national decree. On the other hand, hundreds of pre-existing uncanalized double-lane roundabouts built prior to 2006 remain in daily operation across the Italian road network as legacy, non-compliant nodes. Road agencies are thus faced with the practical dilemma of either restriping these wide platforms into single-lane layouts at the expense of severe congestion or diagnosing their specific geometric vulnerabilities to prioritize targeted physical retrofitting.

2. Materials and Methods

To translate the complex network of spatial interactions into a quantitative form and overcome the purely linear assumptions of traditional road safety models, this study implements a multi-stage analytical framework based on supervised machine learning. In dynamic environments such as 2-lane roundabouts, the onset of an anomalous maneuver rarely stems from a single isolated infrastructural deficiency but rather emerges from simultaneous interference among geometric constraints, vehicle kinematics, and human perception. To model these interactions, the analytical workflow is organized into three sequential steps (Figure 1). First, a Cost-Sensitive Random Forest classifier is trained to map driving behaviors while accounting for class imbalance without altering the empirical data distribution. Second, Permutation Feature Importance is computed on the test set to rank the global predictive weight of individual geometric features. Third, SHapley Additive exPlanations (SHAP) are applied to interpret local non-linear relationships.

This framework exploits the non-parametric versatility of tree-based algorithms to identify complex spatial interactions without imposing a priori functional forms. In turn, the SHAP analysis quantifies the direction and magnitude of feature contributions across predefined geometric categories, while remaining bounded by the categorical resolution of the input variables and the observational scope of the dataset.

2.1. Preliminary Phase: Data Acquisition and Pre-Processing

The empirical foundation of this study is based on a naturalistic driving investigation including continuous kinematic and behavioral observations, collected along a route of approximately 6 km located in a peripheral urban context of the city of Catania. In particular, the sample used in the survey consisted of four urban double-lane roundabouts (both on the circulatory roadway and at the entries/exits), chosen for the variety of their geometric layouts and for their mutual spatial proximity along the same itinerary. This route continuity

was strategically identified to facilitate survey operations and ensure continuous observation, without alterations or interruptions, of the behaviors adopted by the test drivers for the entire duration of the experimental investigation.

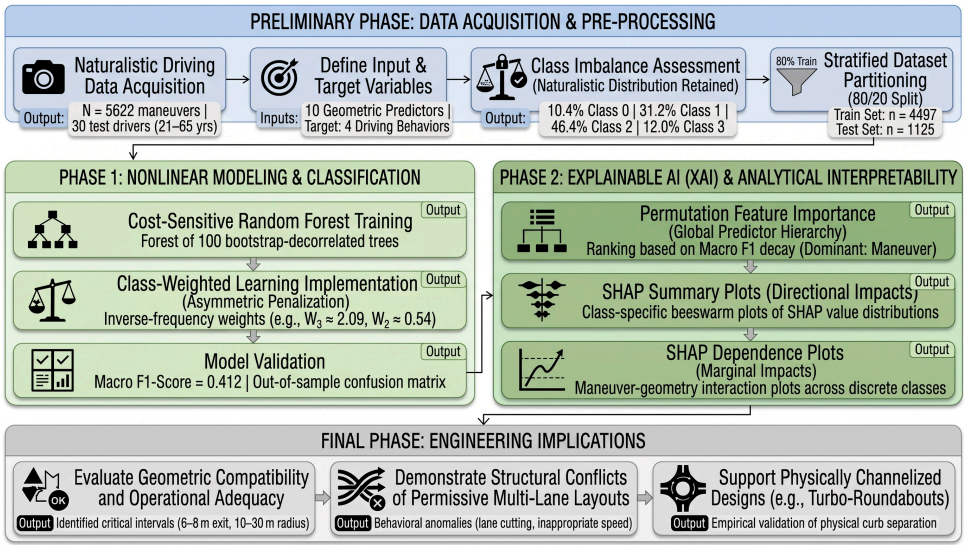


Figure 1. Hybrid analytical framework for double-lane roundabout assessment.

As illustrated in detail in Figure 2, the sample includes three three-leg roundabouts and one four-leg roundabout.

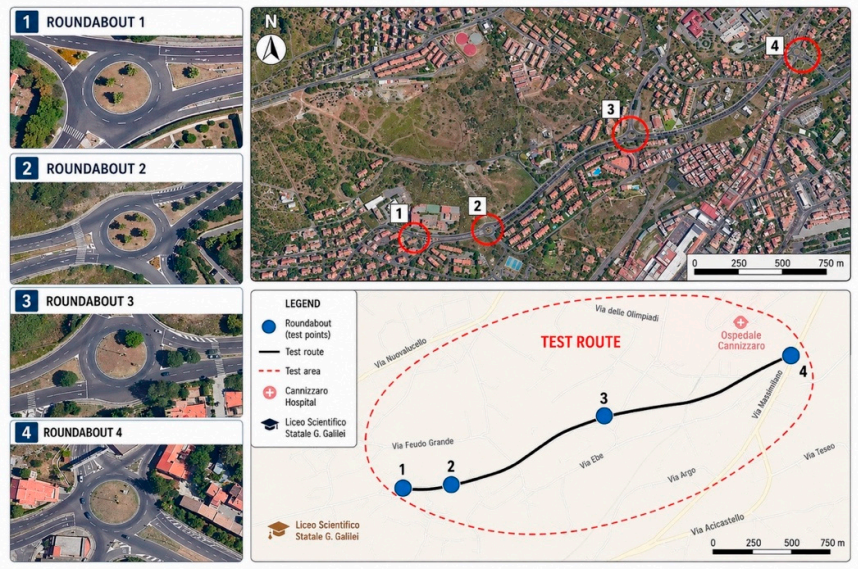


Figure 2. Geographical distribution and aerial details of the 4 roundabouts analyzed.

To ensure adequate demographic and behavioral representativeness, the testing panel consisted of 30 test drivers, strictly balanced by gender (15 men and 15 women). The participants were recruited from non-professional passenger car drivers aged between 21 and 65 years old, all holding a valid driving license with at least 3 years of active driving experience and regular annual mileage. Specific exclusion criteria were applied during recruitment: novice drivers (holding a license for less than 3 years), elderly drivers over 65 years old, and professional drivers (such as taxi, bus, delivery, or commercial heavy-vehicle operators) were explicitly excluded from the panel, alongside individuals with uncorrected visual impairments or medical conditions affecting psychomotor performance.

Excluding professional drivers prevented operational bias linked to specialized driving skills and superior familiarity with complex road geometries, ensuring that the sample reflected ordinary road users. Furthermore, all selected participants were kept unaware of the specific safety purposes of the test. This methodological precaution was necessary to prevent alterations or overcompensations of their natural driving style.

Data acquisition took place in naturalistic conditions, not vitiated by the typical conditioning of simulated driving environments, and was conducted during off-peak hours to minimize queue formation and severe vehicular congestion, thereby allowing drivers to adopt free-flow speed and trajectory choices primarily guided by the intersection layout (although instantaneous circulating traffic volumes and exact gap availability were not directly logged as continuous exposure covariates). Consistent with consolidated on-road investigation protocols [27], data collection involved the presence of two trained observers on board the vehicle. The first observer, seated in the front passenger seat, provided instructions on the maneuvers to be executed, while the second, positioned in the rear seats, recorded behavioral anomalies in real-time on a tablet. The recording utilized a predefined list of unsafe driving behaviors, derived from the sector's scientific literature, ensuring objectivity and uniformity in the survey [27]. This methodology guarantees that the recorded reactions reflect the users' real interaction with the complex double-lane geometry.

To generate a data volume suitable for training machine learning models, the protocol required each driver to navigate the circuit performing multiple cycles. The itinerary was planned so that each driver would execute a complete matrix of maneuvers (right turn, through crossing, left turn, and U-turn, where topologically permitted) approaching the four roundabouts from different entry sections. Considering practical scheduling constraints and the need to test drivers exclusively in time windows compatible with optimal flow regimes, the entire data collection campaign was distributed over a four-week time horizon. Against an experimental plan of 200 scheduled transits for each participant, corresponding to an overall theoretical volume of 6000 maneuvers, the subsequent data cleaning phase resulted in the exclusion of 6.3% of the recordings. This normal data attrition, attributable to slight instrumental anomalies or aborted transits due to unexpected traffic interactions or external traffic perturbations, made it possible to consolidate the final database into 5622 valid surveys.

To formalize the analytical framework, the objectively observed driving behaviors constitute the dependent variable (target), which was classified into four discrete and mutually exclusive clusters, representative of the most critical risk phenomenologies: 0 (absence of errors, i.e., the safety baseline), 1 (inappropriate speed), 2 (inappropriate trajectory, which includes lane invasion phenomena), and 3 (failure to yield). Following the consolidated on-board observation protocol [12,27], each behavioral class was identified through strict operational criteria directly verifiable by the on-board observers. Specifically, 'Inappropriate speed' (Class 1) was recorded when the driver maintained an operating speed (directly monitored by the observers via the vehicle's speedometer) exceeding the local regulatory and design threshold for urban roundabouts (>40 km/h along the circulatory roadway or >50 km/h at the entry/exit approaches) or executed abrupt braking or acceleration maneuvers incompatible with smooth traffic calming progression while maintaining lane discipline. 'Inappropriate trajectory' (Class 2) was assigned whenever the vehicle's wheels physically crossed the longitudinal lane markings separating the two adjacent lanes during entry, circulation, or exit, including path straightening (lane cutting) across the circulatory ring, or encroached upon the central apron or splitter islands. 'Failure to yield' (Class 3) was documented when an approaching test driver forced entry into the circulatory roadway in the immediate proximity of an oncoming circulating vehicle, visibly compelling the circulating user to brake or alter their trajectory to avoid a collision. Finally, 'Absence of

errors' (Class 0) applied to transits that fully complied with all speed, lane-keeping, and yielding criteria. Whenever multiple anomalies occurred during the same transit, mutual exclusivity was enforced by applying a chronological primary-trigger rule: the passage was labeled according to the initial behavioral deviation that triggered the unsafe dynamic along the approach-to-exit sequence, ensuring that the model predicts the root operational failure rather than secondary compensatory actions. Furthermore, to ensure label reliability, both trained observers monitored the execution of each maneuver; any ambiguous transits where immediate inter-observer consensus was not reached, or where external traffic perturbations interfered with the driver's natural choice, were systematically excluded during the 6.3% data cleaning phase.

In parallel, the independent variables encode the entire ecosystem of spatial and operational constraints experienced by the driver during the various intersection usage phases. These include continuous geometric parameters measured in the field, specifically the entry width, exit width, and circulatory roadway width (where all three width parameters consistently denote the total two-lane carriageway width across the manuscript), as well as the roundabout radius, entry radius, exit radius, path radius, deflection radius, and deflection angle, alongside categorical parameters defining the exact maneuver executed (right turn, through crossing, left turn, U-turn) and the presence of pedestrian crossings.

All physical and kinematic variables were accurately recorded and discretized into distinct categorical levels. This discretization procedure is strategic for multiple reasons: not only does it facilitate a much more robust and interpretable probabilistic modeling, but it aggregates the spatial configurations into engineering tolerance bands, preventing estimation distortions linked to local micro-variations or extreme anomalous values (outliers) in the survey. The complete coding scheme for all considered variables, which serves as the basis for subsequent logical modeling, is detailed in Table 1, where the final column specifies both the absolute sample sizes (n) and the percentage distributions across the investigated roundabouts, maneuvers, and geometric categories within the entire acquired database (N = 5622). Specifically, the boundaries defining the Low, Medium, and High classes were established a priori based on functional geometric ranges consolidated in the international technical literature [1,12], grouping the approach and exit legs of the four intersections into homogeneous operational intervals. Because the model evaluates these predefined categorical bands, the subsequent XAI analysis characterizes behavioral risk associations across established engineering intervals rather than estimating continuous point-wise transitions.

Table 1. Categorization, coding scheme, and sample distribution of the considered variables (N = 5622).

Variable	Description and Classes	Coding Values	Sample Size (n) and Distribution
Roundabout Site (Descriptive)	Roundabout 1 (3-leg, large radius)	—	1719 (30.58%)
	Roundabout 2 (3-leg, compact radius)	—	483 (8.59%)
	Roundabout 3 (3-leg, medium radius)	—	615 (10.94%)
	Roundabout 4 (4-leg, large radius)	—	2805 (49.89%)
Driving Behavior	Absence of errors	0	588 (10.46%)
	Inappropriate speed	1	1752 (31.16%)
	Inappropriate trajectory	2	2610 (46.42%)
	Failure to yield	3	672 (11.95%)

Table 1. *Cont.*

Variable	Description and Classes	Coding Values	Sample Size (n) and Distribution
Maneuver	Right turn	0	390 (6.94%)
	Through	1	2322 (41.30%)
	Left turn	2	1680 (29.88%)
	U-turn	3	1230 (21.88%)
Roundabout Radius	Low (≤14 m)	0	483 (8.59%)
	Medium (Between 14 and 20 m)	1	615 (10.94%)
	High (>20 m)	2	4524 (80.47%)
Circulatory roadway width (total for 2 lanes)	Low (≤6 m)	0	483 (8.59%)
	Medium (Between 6 and 8 m)	1	3255 (57.90%)
	High (>8 m)	2	1884 (33.51%)
Entry width (total for 2 lanes)	Low (≤6 m)	0	615 (10.94%)
	Medium (Between 6 and 8 m)	1	819 (14.57%)
	High (>8 m)	2	4188 (74.49%)
Exit width (total for 2 lanes)	Low (≤6 m)	0	483 (8.59%)
	Medium (Between 6 and 8 m)	1	408 (7.26%)
	High (>8 m)	2	4731 (84.15%)
Entry Radius	Low (≤10 m)	0	747 (13.29%)
	Medium (Between 10 and 30 m)	1	615 (10.94%)
	High (>30 m)	2	4260 (75.77%)
Exit Radius	Low (≤10 m)	0	510 (9.07%)
	Medium (Between 10 and 30 m)	1	276 (4.91%)
	High (>30 m)	2	4836 (86.02%)
Path Radius	Low (≤50 m)	0	1368 (24.33%)
	Medium (Between 50 and 100 m)	1	1005 (17.88%)
	High (>100 m)	2	3249 (57.79%)
Deflection Radius	Low (<50 m)	0	747 (13.29%)
	Medium (Between 50 and 100 m)	1	3744 (66.60%)
	High (≥100 m)	2	1131 (20.12%)
Deflection Angle	Low (≤30°)	0	1614 (28.71%)
	Medium (Between 30° and 45°)	1	3261 (58.00%)
	High (>45°)	2	747 (13.29%)
Pedestrian crossing	Absent	0	1719 (30.58%)
	Present	1	3903 (69.42%)

The preliminary exploration of the target variable highlighted a natural and predictable class imbalance within the overall dataset, comprising 5622 observations. This condition is intrinsic and structural to real-world traffic safety scenarios and naturalistic observation. In the examined sample, inappropriate trajectories constitute the absolute majority (n_2 = 2610, 46.42%), followed by inappropriate speeds (n_1 = 1752, 31.16%). Behaviors linked to severe violations of gap acceptance, such as failures to yield (n_3 = 672, 11.95%), and operational regularity or absence of errors (n_0 = 588, 10.46%) are instead configured as minority classes.

In much of the machine learning-based literature, class imbalance is addressed ex ante by modifying the dataset using heuristic over-sampling techniques (e.g., SMOTE) or generative adversarial networks (e.g., CTGAN). In this study, however, the original data distribution was retained without synthetic resampling. The generation of synthetic instances risks irretrievably distorting the true joint probability distribution $P(X, Y)$ [28,29]. Since driving behaviors are inextricably linked to physical and kinematic constraints dictated by parameters X (radii, widths, deflection angles), artificial interpolation would generate spurious covariance patterns, creating combinations of maneuver and geometry that are physically implausible or dynamically unrealistic [30–32]. Consequently, the original topology of the space-behavioral domain was strictly preserved.

At the end of the pre-processing phase, to validate the real generalization capacity of the predictive model and avoid overfitting, the entire dataset was partitioned into a training set (80%) and a test set (20%). This operation was performed using stratified sampling, which preserves the empirical class proportions (n_i/N) across both subsets.

Within this partitioning scheme, the predictor matrix relies exclusively on geometric and kinematic parameters, intentionally omitting any driver-specific identifiers, demographic covariates, or site-level labels. Because all participants executed the same balanced protocol of maneuvers across the same infrastructure elements, identical input vectors naturally map to varying behavioral outcomes across individuals. This structural formulation ensures that the classifier learns generalizable relationships between spatial configurations and driving behavior across the sampled geometric domain, treating inter-individual variability as natural stochastic variance and establishing a robust diagnostic baseline for the subsequent XAI threshold extraction.

In this context, it is also important to consider the behavioral distinction between drivers navigating a roundabout for the first time and everyday familiar users, as well as its relationship with the probabilistic formulation of the model. Unfamiliar drivers approaching a double-lane roundabout for the first time typically experience higher cognitive and navigational workload, which tends to manifest as more cautious approach speeds, hesitation during gap acceptance, or late lane corrections near the exit when attempting to identify the correct outbound leg. Conversely, everyday commuters (as well as test drivers after completing initial circuit loops) possess established spatial familiarity; relieved of wayfinding uncertainty, habitual users tend to exploit permissive, uncanalized geometries more aggressively by straightening trajectories and maintaining higher operating speeds to minimize lateral acceleration. Despite these different cognitive triggers, the underlying geometric vulnerability remains shared: an uncanalized multi-lane platform fails to provide clear visual guidance for first-time users while simultaneously lacking physical confinement to restrain habitual drivers. Regarding the impact on the empirical results, because the experimental protocol involved regionally recruited drivers performing repeated cycles along the same corridor, the dataset predominantly reflects familiar driving dynamics, meaning that the recorded prevalence of lane cutting (46.42%) and inappropriate speed (31.16%) represents habitual operational adaptation rather than initial navigational confusion. Within the probabilistic architecture of the Random Forest classifier, where node splitting evaluates the empirical class probability distribution $p(i|t)$ across bootstrap samples, the absence of explicit familiarity covariates means that behavioral shifts between initial exploratory transits and subsequent familiarized runs are absorbed as intra-node stochastic variance. Consequently, the ensemble probabilities and resulting SHAP attributions capture the baseline geometric propensity of each infrastructural configuration to induce unsafe behavior across repeated exposures.

2.2. Phase 1: Non-Linear Modeling and Classification

To model the complex behavioral relationship, markedly non-linear and characterized by high latent multicollinearity among geometric variables, an Ensemble Learning algorithm based on decision trees was implemented: the Random Forest classifier (RF). Within the specific scope of this investigation, the RF architecture offers distinct methodological advantages over traditional parametric discrete-choice models (such as multinomial logit formulations) and single decision trees. First, in double-lane roundabouts, geometric design elements are physically interdependent (for instance, deflection and path radii are geometrically constrained by the roundabout radius, circulatory roadway width, and entry or exit widths). Whereas parametric models suffer from variance inflation and unstable parameter estimates under severe multicollinearity, the tree-based ensemble handles correlated predictors without imposing predetermined functional forms. Second, driver adaptation at multi-lane intersections is governed by higher-order conditional interactions between the intended turning maneuver and local cross-sectional dimensions, which recursive tree partitioning captures automatically without requiring the manual specification of multiplicative interaction terms. Third, aggregating multiple decorrelated trees reduces predictive variance compared to single-tree methods (such as CHAID), providing higher robustness against the inter-individual behavioral variability inherent in open-road naturalistic observations.

To structurally mitigate the inductive bias towards majority classes (inappropriate trajectories and speeds), the RF algorithm was regularized by applying an asymmetrically weighted cost function (class-weighted learning) [33–35]. During the training phase, the weight W_i assigned to each i-th class was not imposed arbitrarily, but was computed in inverse proportion to its empirical frequency, according to the following mathematical formulation:

$$W_i = \frac{N}{k \cdot n_i}$$

where $N = 5622$ is the cardinality of the dataset, $k = 4$ is the total number of behavioral classes, and ni is the absolute frequency of empirical observations belonging to the i-th class. Through this asymmetric penalization, a classification error committed by the loss function on the minority class "failure to yield" ($W3 \approx 2.09$) receives a proportionally higher penalty during node splitting compared to an error on the majority class "inappropriate trajectory" ($W2 \approx 0.54$).

The model's architecture was implemented by generating a forest of 100 independent decision trees. Each tree is trained on a sample generated via bootstrap (random sampling with replacement). To minimize the correlation among the ensemble's trees, the search for the best split parameter in each node occurs by exclusively evaluating a random subset of predictors of dimension $\sqrt{p}$, where p is the total number of features. In the context of roundabout safety analysis, this random subspace selection is particularly advantageous: by preventing dominant predictors (such as the executed maneuver or exit width) from monopolizing the primary splits of every tree, it allows secondary yet safety-critical geometric features (such as entry width in yielding conflicts) to be evaluated across the ensemble and subsequently quantified via TreeExplainer. The final classification for each observation x is then determined through majority voting:

$$\hat{y} = \arg \max_{c \in \{0,1,2,3\}} \sum_{m=1}^{M} I(h_m(x) = c)$$

where $M = 100$ is the number of trees, $h_m(x)$ is the prediction of a single tree, and $I(.)$ is the indicator function.

The optimization criterion for splitting the decision nodes is based on the minimization of the Gini impurity. For a given node t, containing a fraction of the training data, the impurity is defined as:

$$\text{Gini}(t) = 1 - \sum_{i=0}^{k-1} [p(i|t)]^2$$

where $p(i|t)$ represents the empirical proportion of samples belonging to class i present in node t. At each node, the algorithm evaluates the candidate predictors to identify the split that maximizes the reduction in Gini impurity (Gini Gain).

The predictive effectiveness of the entire cost-sensitive framework was validated on the independent test set using the Macro F1-Score. In multi-class classification problems governed by strongly asymmetric target distributions, the adoption of global accuracy (Overall Accuracy) is methodologically inadequate. Accuracy, in fact, returns excessively optimistic and misleading metrics, being almost entirely dominated by the model's performance on majority classes. This statistical indicator assigns equal importance to each behavioral class, ensuring that less frequent but critically relevant dynamics (such as failures to yield and absence of infractions) receive equal weight in the evaluation compared to ordinary infractions.

Starting from the multi-class confusion matrix, whose diagonal and non-diagonal elements define True Positives (TP_i), False Positives (FP_i), and False Negatives (FN_i) for the i-th class, Precision and Recall are analytically defined:

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i}$$

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i}$$

The local F1-Score (relative to the single class) is calculated as the harmonic mean between Precision and Recall, a function that severely penalizes models that unilaterally maximize a single metric to the detriment of global generalization capacity:

$$F1_i = 2 \cdot \frac{\text{Precision}_i \cdot \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}$$

Consequently, the aggregate Macro F1-Score is calculated as the unweighted arithmetic mean of the local F1-Scores relative to the entire domain of the k classes:

$$\text{Macro F1} = \frac{1}{k} \sum_{i=0}^{k-1} F1_i$$

2.3. Phase 2: Analytical Interpretability and Explainable AI (XAI)

The primary inferential objective requires quantifying the extent to which design choices induce specific behaviors. To this end, the global importance of the predictors was calculated using Permutation Feature Importance. This technique quantifies the critical impact of a geometric parameter by measuring the degradation of predictive performance (specifically, the Macro F1-Score) when the values of this variable are randomized exclusively within the independent test set.

The importance I_j for the j-th variable is measured analytically as the average difference between the reference score and the permuted score calculated over K iterations:

$$I_j = s_{base} - \frac{1}{K} \sum_{k=1}^{K} s_{k,j}$$

where s_{base} represents the Macro F1-Score calculated on the intact test set, K is the number of independent repetitions of the random permutation, and $s_{k,j}$ is the score obtained in the k-th iteration with the variable Xj permuted.

Evaluating permutation importance on the held-out test set measures how much the model relies on each feature for out-of-sample classification. However, Permutation Feature Importance provides only an aggregated global ranking: it does not indicate the direction of the effect, cannot separate class-specific impacts, and may be influenced by correlated predictors. To address these limitations, SHapley Additive exPlanations (SHAP) were computed using the TreeExplainer algorithm. Based on cooperative game theory, SHAP values quantify the marginal contribution of each feature to the prediction for each individual observation across all possible feature coalitions. Class-specific Summary Plots (beeswarm plots) were generated to display the distribution of SHAP values alongside the coded feature levels. Positive SHAP values indicate a positive marginal contribution toward predicting a given behavioral class relative to the model's base expectation, whereas negative values indicate a shift away from that class. These attributions serve as comparative diagnostic indicators of feature influence within the model rather than calibrated absolute risk probabilities.

Finally, SHAP Dependence Plots were used to examine how feature contributions varied across geometric categories and maneuver types. By plotting the SHAP values of a given geometric variable against its discrete categories and coloring observations by the executed maneuver, these plots illustrate how local interaction effects modify behavioral risk associations within predefined engineering intervals.

3. Results

This section presents the quantitative results derived from the proposed analytical framework. The exposition is divided into three sequential levels of investigation:

- In the first instance, the predictive effectiveness of the Cost-Sensitive Random Forest classifier is validated through the extraction of global performance metrics, those disaggregated by single class, and the topological observation of the confusion matrix;
- Subsequently, the impact hierarchy of geometric predictors is defined through the calculation of Permutation Feature Importance on the independent test set;
- Finally, the model's decision-making logic is decoded in depth through the Explainable AI (SHAP) framework.

This final step extends the investigation from an aggregate statistical evaluation to non-linear local interpretability, using Summary Plots and Dependence Plots to quantify the asymmetric directional impact of single variables and identify the critical intervals of geometric compatibility between the road platform and vehicular maneuvers.

3.1. Global Metrics and Classification Report

The evaluation of the model on the test set, consisting of N = 1125 independent observations (equal to exactly 20% of the original partitioned database), produced the metrics reported in Table 2. The model registered an overall Macro F1-Score equal to 0.412, derived from a Macro Precision of 0.400 and a Macro Recall of 0.446. The achieved Macro F1-Score (0.412), which substantially outperforms the 0.250 uniform random baseline in a four-class setting, reflects the strictly geometric and kinematic nature of the feature space. By evaluating infrastructure configurations shared across multiple drivers, the model captures the structural influence of roundabout geometry on operational behavior, while the residual variance naturally accounts for the unobserved human heterogeneity inherent to naturalistic open-road driving.

Disaggregating the results by single behavioral class, the maximum Recall (0.661) is observed for Class 0 (absence of errors). Class 2 (inappropriate trajectory), which represents the relative majority of the test sample (n = 522), obtained the highest values for Precision (0.585) and F1-Score (0.554). Conversely, Class 3 (failure to yield) records the lowest

performance values, with an F1-Score equal to 0.234. From a methodological perspective, achieving a 0.661 Recall on the smallest minority class (Class 0, 10.46% of the dataset) demonstrates the effectiveness of the cost-sensitive weighting scheme in preventing the ensemble from collapsing onto the two dominant behavioral classes (Classes 1 and 2, which jointly account for 77.58% of the sample). At the same time, the low Recall (0.239) and F1-Score (0.234) recorded for Class 3 indicate that approximately three-quarters of yielding violations are not captured by static geometric and maneuver predictors alone. Consequently, while the classifier is not intended for standalone real-time prediction of individual yielding infractions, it provides a statistically meaningful diagnostic foundation for isolating the marginal directional influence of planimetric features via XAI.

Table 2. Classification report of the Cost-Sensitive Random Forest model (test set).

Driving Behavior (Target)	Precision	Recall	F1-Score	Support
Absence of errors (0)	0.382	0.661	0.484	118
Inappropriate speed (1)	0.401	0.356	0.377	351
Inappropriate trajectory (2)	0.585	0.527	0.554	522
Failure to yield (3)	0.230	0.239	0.234	134

3.2. Observations on the Confusion Matrix

The analysis of the confusion matrix (Figure 3) calculated on the independent test set allows for isolating precise interaction dynamics among behavioral deviances, extracting transport interpretations directly from the algorithm's error rates.

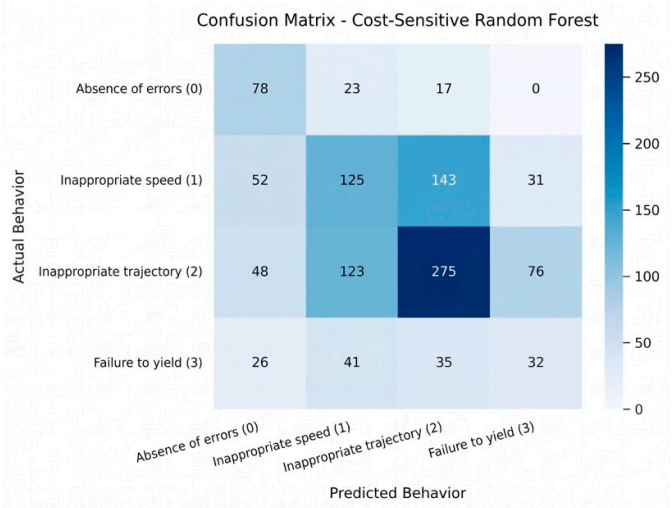


Figure 3. Confusion matrix of the Cost-Sensitive Random Forest model evaluated on the independent test set.

The most significant engineering data lies in the strong predictive overlap between inappropriate speed (Class 1) and inappropriate trajectory (Class 2). The model manifests a systematic difficulty in clearly discriminating these two categories: specifically, 143 real instances of inappropriate speed are classified as trajectory deviances, while conversely 123 anomalous trajectories are labeled by the model as speed infractions. From an analytical standpoint, this algorithmic permeability highlights a clear classification ambiguity within the purely geometric feature space. Rather than proving direct kinematic coupling by itself, this mutual misclassification demonstrates that both behavioral deviations emerge under virtually identical spatial boundary conditions: permissive geometric configurations

simultaneously lower the threshold for both speed infractions and lane-cutting maneuvers, causing the two phenomena to share overlapping domains within the predictor matrix.

Conversely, the model establishes a clear logical separation between regular driving dynamics (Class 0) and failure to yield (Class 3), recording no predictive overlap strictly speaking (0 cases of correct maneuvers mistaken for failure to yield). However, it is precisely Class 3 that registers the greatest algorithmic interception difficulty: out of a total of 134 observed events, against only 32 True Positives, the false negatives are dispersed almost uniformly across all other categories (41 instances assigned to speed, 35 to trajectory, and 26 to absence of errors). This behavior reflects the intrinsically different nature of this infraction: unlike purely kinematic errors, failure to yield is linked to inter-vehicular interaction and gap acceptance, dynamics driven by cognitive ergonomics and visual stress that the model struggles to uniquely frame using exclusively planimetric variables. Specifically, three underlying factors explain this lower classification performance (F1-Score = 0.234). First, from a traffic flow perspective, a yielding violation can physically materialize only when conflicting vehicles are simultaneously circulating within the ring at critical headways during the exact instant of entry; because the predictor matrix includes only static geometric variables and omits instantaneous circulating traffic volumes or gap availability, identical geometric inputs frequently result in unconflicted entries when no circulating vehicle is present. Second, gap-acceptance thresholds are heavily influenced by unobserved inter-individual human factors (such as risk propensity, visual scanning, and reaction time), which introduce substantial stochastic variance across identical approach geometries. Third, within permissive intersection legs, forced entries often occur under the same wide spatial boundary conditions that facilitate high-speed approaches and lane-cutting maneuvers, explaining why the classifier frequently assigns false negatives of Class 3 to inappropriate speed (41 cases) and inappropriate trajectory (35 cases). Consequently, static infrastructure features can only capture the baseline environmental predisposition to yielding conflicts rather than deterministic event occurrence.

3.3. Extraction of Variable Importance (Permutation Feature Importance)

Figure 4 illustrates the geometric and operational predictors in decreasing order of importance calculated by permutation on the independent test set, represented through boxplots that highlight the median and the interquartile range (IQR).

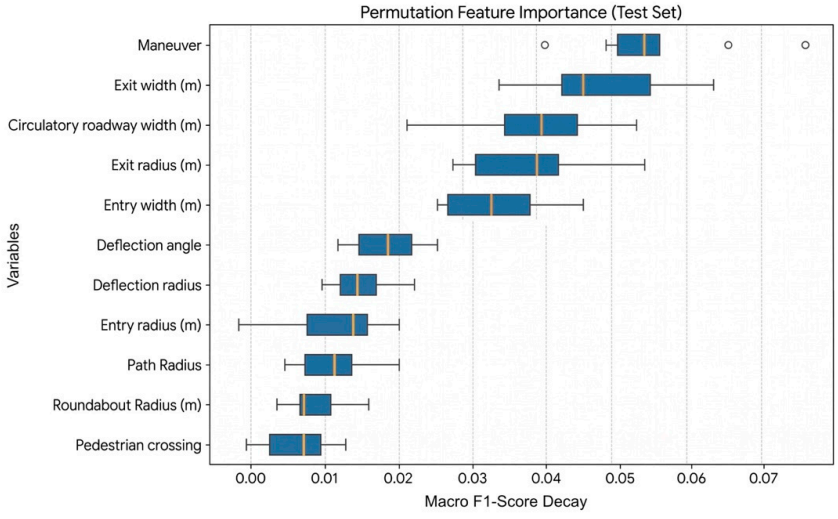


Figure 4. Feature Importance evaluated via Permutation Feature Importance (Macro F1-Score degradation on the test set).

The maneuver variable emerges as the dominant predictor, recording the most marked median decay in predictive performance following spatial randomization, while manifesting a relatively wide dispersion (IQR) due to the strong contextual dependence of human behavior.

Following this is the block of parameters relating to the outbound vehicular flow and the circulatory roadway: exit width, circulatory roadway width, and exit radius. The analysis of their respective boxplots highlights not only median importances significantly higher than the parameters associated with roundabout entry, but also whiskers of significant extension, proving how the perturbation of these specific geometric dimensions upon exit causes a consistent and systematic degradation of the algorithm's ability to correctly classify maneuvers on unseen data.

The parameters associated with roundabout entry (entry width and radius), geometric deflection (radius and angle), and the pedestrian crossing variable are all located in the lower portion of the ranking. In this segment, the boxplots show marginal median decreases (close to zero) and markedly narrower interquartile ranges compared to the exit parameters, indicating that the Random Forest classifier is structurally less sensitive to the alteration of such planimetric configurations in predicting overall behavioral error.

3.4. Global and Local Interpretability via SHAP Analysis

The analytical core of this investigation is based on the application of the SHAP (SHapley Additive exPlanations) metric to decode the decision-making logic of the Random Forest model. Overcoming the limits of traditional global importance measures, the use of Summary Plots (or Beeswarm plots) allowed for quantifying the marginal impact of each geometric parameter, exploring the distribution of spatial impact (SHAP values on the *x*-axis) as a function of the magnitude assumed by the variable (chromatic coding from blue for low values to red for high values).

The examination of the Summary Plot relative to inappropriate trajectory (Class 2) highlights how the exit planimetric configuration likely acts as a driving factor (Figure 5). A clear polarization trend is observed: high dimensional values (red dots) of the exit width and circulatory roadway width tend to shift SHAP values into positive territory. This trend suggests that wide transverse geometries can increase the probability that the model classifies the maneuver as a trajectory error.

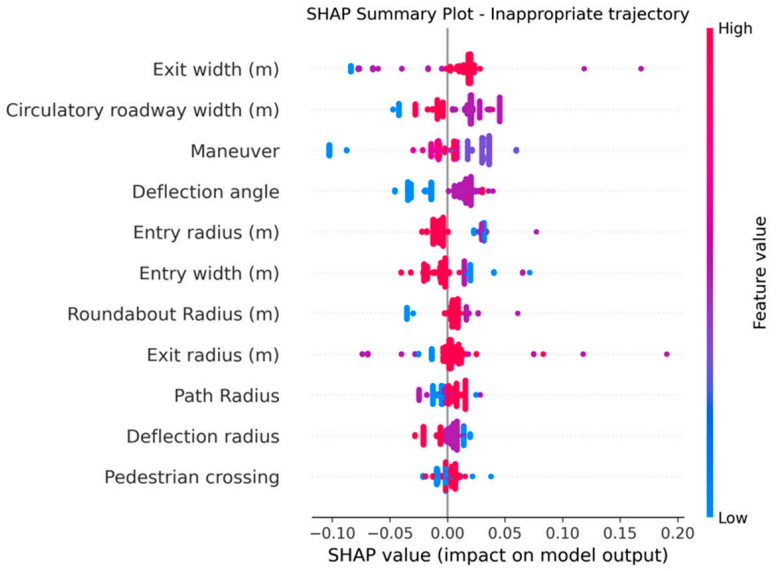


Figure 5. SHAP Summary Plot for inappropriate trajectory.

Analyzing the diagram for inappropriate speed (Class 1), the predictive hierarchy is reshuffled, placing the exit radius at the top (Figure 6). Here, too, the color distribution outlines a consistent pattern: high values of the exit radius (red dots) predominantly extend into the positive semi-axis of SHAP values, indicating that particularly wide connecting radii are frequently associated with a kinematic increase.

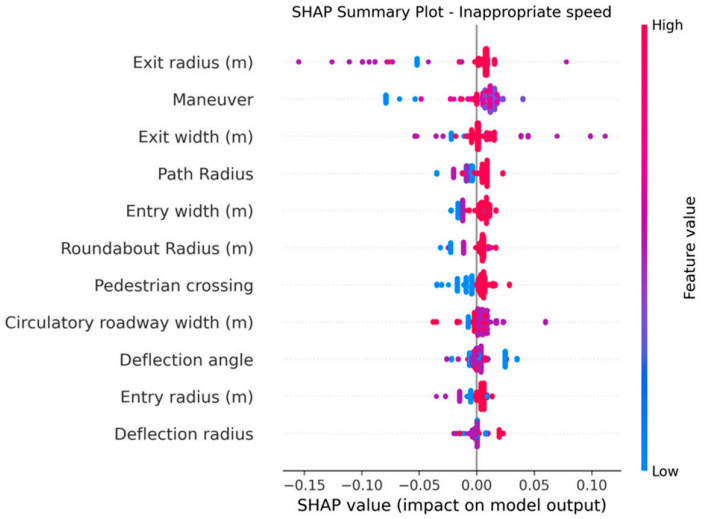


Figure 6. SHAP Summary Plot for inappropriate speed.

The diagram for failure to yield (Class 3) confirms a greater dispersion of impact on planimetric variables (Figure 7). Unlike kinematic errors, where the color-SHAP gradient appears clearer, entry conflicts show a marked chromatic overlap around zero. This configuration indicates that neither wide nor narrow configurations seem, on their own, to uniquely govern the phenomenon of incorrect gap acceptance, assigning a decidedly greater interpretive weight to the type of maneuver undertaken.

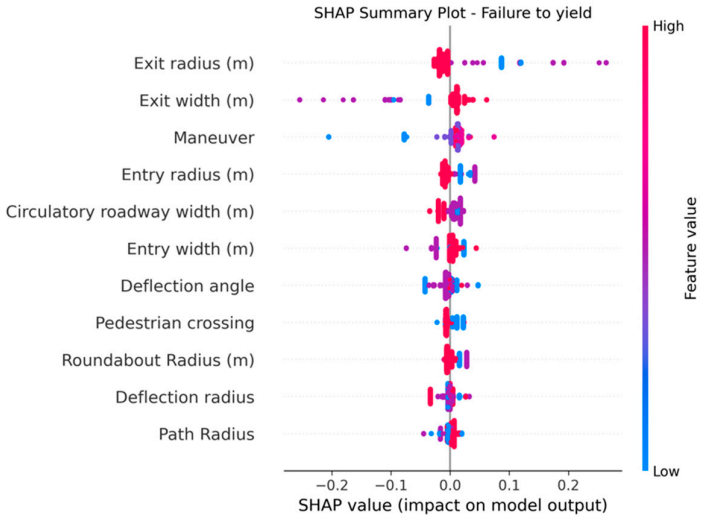


Figure 7. SHAP Summary Plot for failure to yield.

Finally, the Summary Plot for the absence of errors provides a complementary consistency check (Figure 8). The predictive probability that a driver maintains correct driving conduct experiences a generally positive impact (SHAP values above zero) in correspondence with kinematically direct maneuvers and more constrained exit sections. This statistical trend supports the hypothesis that the adoption of restrictive planimetric geometries

favors the emergence of operational behaviors that are adequate and compatible with the intersection's configuration, helping to radically mitigate the spatial degrees of freedom that trigger trajectory drifts.

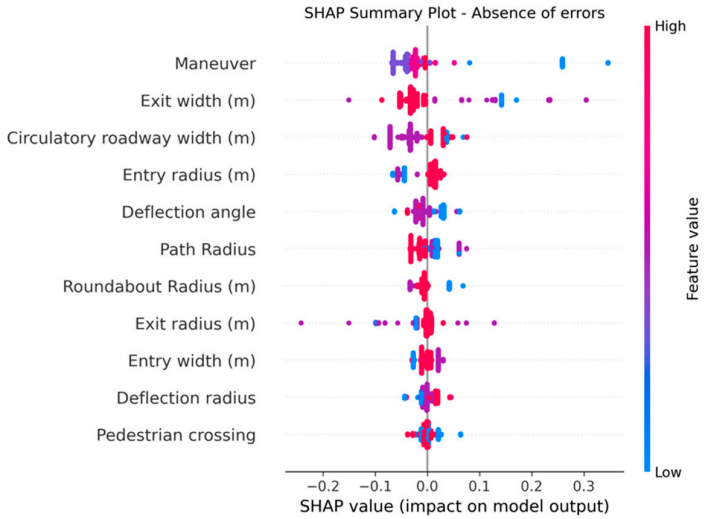


Figure 8. SHAP Summary Plot for absence of errors.

3.5. Marginal Dependence Analysis and Critical Dimensional Intervals

To quantify the non-linear marginal effects of infrastructural parameters and identify behavioral risk associations across the predefined geometric classes, the analytical framework was extended through the implementation of SHAP Dependence Plots. Within these two-dimensional representations, the impact of each planimetric variable was systematically evaluated in interaction with the maneuver type (maneuver). This structural configuration analytically derives from the Permutation Feature Importance results (Section 3.3), which demonstrated how maneuver kinematics represent a predictor of primary importance for the model; consequently, the degree of criticality of a geometric element should not be evaluated in isolation, but it is appropriate to correlate it with the specific type of trajectory executed by the vehicle within the circulatory roadway.

The architecture based on dependency diagrams is formalized here as an intrinsically scalable and expandable analytical criterion, theoretically capable of mapping and decoding spatial interactions for the entire domain of geometric covariates. However, to meet necessary synthesis requirements and to maximize the effectiveness of the scientific discussion, the present contribution limits itself to exposing a narrow but highly representative sample of interactions.

The delimitation of the investigation context is not random but responds to the need to test the model on conditions of greatest infrastructural vulnerability. The visual architecture was therefore exclusively focused on exploring behavioral deviances (speed, trajectory, and failure to yield), omitting the regular driving baseline to narrow the analytical focus solely to risk-triggering factors and relative spatial criticality conditions. Furthermore, the isolation of these specific priority setups did not merely trace the global hierarchy of Feature Importance, which is intrinsically polarized by the preponderant statistical weight of exit geometries in overall model performance. Instead, the choice was guided by a spatial-functional mapping, associating each specific error with the planimetric parameter that physically characterizes the intersection sector where the conflict materializes. Under this lens, for anomalies manifesting upon exit, the analysis of inappropriate trajectories and inappropriate speed focuses respectively on exit width and exit radius, confirmed as significant predictors for such dynamics. Conversely, failure to yield dynamics, which

by engineering definition are triggered at the entry section, are explored through entry width, a globally secondary variable but locally crucial for defining the visual field and gap acceptance.

The detailed examination of the diagrams shown in Figures 9–11 returns a coherent phenomenological reading, suggesting plausible triggering dimensional intervals for each deviation. It should be noted that, because both the geometric categories and the maneuver types are discrete variables (Table 1), multiple test-set observations sharing the exact same leg geometry and turning movement overlap at identical coordinates in the scatter plots.

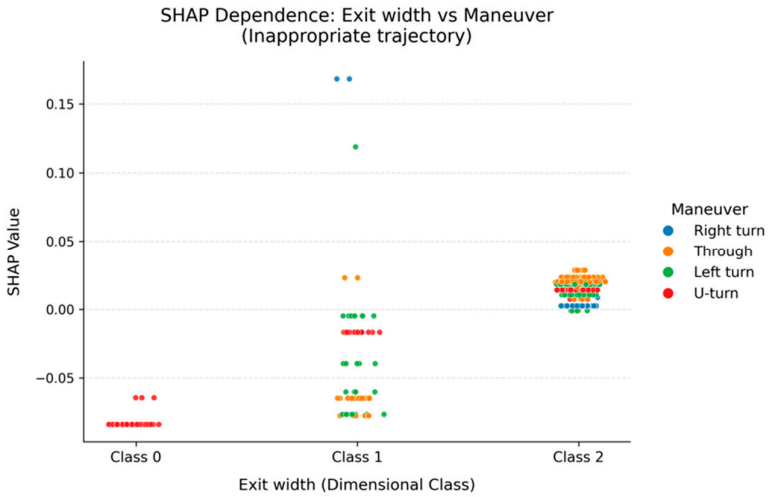


Figure 9. SHAP Dependence Plot: exit width vs. inappropriate trajectory.

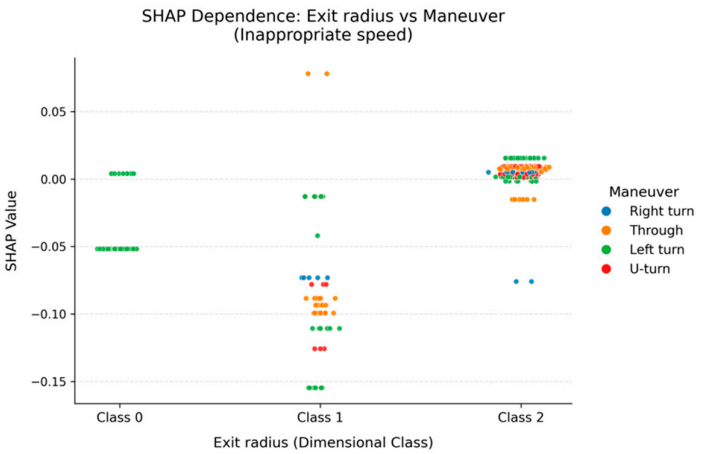


Figure 10. SHAP Dependence Plot: exit radius vs. inappropriate speed.

In the case of inappropriate trajectory (Figure 9), while the majority class of wide exits (>8 m, Class 2, n = 4731 in Table 1) exhibits a consistent, compact cluster of positive SHAP values across through, left-turn, and U-turn maneuvers, the model also highlights a pronounced maneuver-specific sensitivity within the predefined intermediate two-lane exit width category (6–8 m, Class 1, n = 408). This specific dimensional interval generates wide vertical dispersion, with negative SHAP values for left-turn and U-turn movements but projecting the SHAP value toward positive peaks (approaching 0.17) for the overlapping clusters of observations executing a right turn or specific through movements. From a transport perspective, a width of this magnitude can be ambiguous: it appears excessive for strictly channelizing a single vehicle, but insufficient to accommodate side-by-side exit flows in total safety. It is plausible that this excessive lateral space may induce drivers, especially

during insertion into the right exit arm, to straighten the curve and invade the adjacent lane to minimize kinematic effort.

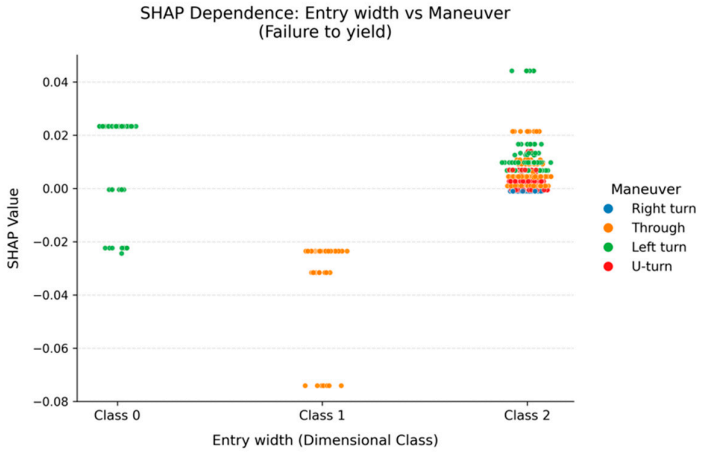


Figure 11. SHAP Dependence Plot: entry width vs. failure to yield.

Analyzing inappropriate speed (Figure 10), high exit radii (>30 m, Class 2, n = 4836 in Table 1) systematically yield positive SHAP values across almost all maneuvers, whereas the dependency diagram for the intermediate exit radius category (10–30 m, Class 1, n = 276) registers marked maneuver-dependent polarization. While most turning movements in this intermediate range exhibit negative SHAP values, a distinct cluster of direct through-crossing observations shifts into positive territory (SHAP value exceeding 0.08). Geometrically, an exit radius exceeding 10 m entails a progressive reduction in curvature, likely attenuating the geometric severity of the maneuver and the transverse acceleration perceived by the driver. When this permissive geometry pairs with a linear trajectory like a direct crossing, the central island's deflection effect tends to be mitigated; the driver perceives an optically and physically straighter path, which can encourage them to maintain or increase approach speed, potentially undermining the intersection's traffic calming function.

The most evident functional transition emerges, however, in the failure-to-yield scenario (Figure 11), explored through the entry width. In correspondence with wide two-lane entry widths exceeding 8 m, the impact on the model becomes generally positive, reaching SHAP values above 0.04. Unlike previous scenarios, this result denotes a transversal tendency to trigger more opportunistic behaviors, regardless of the maneuver the driver has planned to undertake. In double-lane configurations, where side-by-side flow constitutes the normal operational premise, it is hypothesized that a dimensionally excessive entry dilates the visual field and attenuates the perceptual constriction effect that naturally induces caution. Instead of rigorously channelizing approach trajectories, this spatial abundance seems to favor competition between parallel vehicles, potentially encouraging them to simultaneously vie for the same useful gap in the circulatory roadway and fostering, in fact, a potentially forced and dangerous risk acceptance upon entry.

4. Discussion

The structural integration between supervised learning (Cost-Sensitive Random Forest) and non-linear local interpretability (SHAP) has made it possible to deeply decode the complex interactions among geometric infrastructure, vehicular kinematics, and user cognitive ergonomics. Compared to earlier single-tree decision approaches such as Chi-squared Automatic Interaction Detection (CHAID) previously applied to roundabout behavior [12,27], the proposed ensemble architecture overcomes the structural instability and variance of single hierarchical trees, where lower-level splits are strictly conditioned by the top-node

chi-squared partitioning, by aggregating 100 bootstrap-decorrelated trees and extracting game-theoretic SHAP attributions across all possible feature coalitions. It is also instructive to compare the proposed Explainable AI framework against conventional crash-based safety performance functions (SPFs) and generalized linear models (such as Poisson and Negative Binomial regressions) widely reported in road safety literature [9–11,16,22–24]. Regarding its strengths, the RF-SHAP behavioral approach is inherently proactive and microscopic: whereas SPFs rely on multi-year historical crash records and suffer from severe under-reporting of property-damage-only (PDO) collisions (which predominate at double-lane roundabouts [5,6]), naturalistic behavioral modeling identifies high-frequency operational anomalies before crashes accumulate. Furthermore, while parametric SPFs impose predefined log-linear functional forms that struggle to isolate maneuver-specific geometric interactions at individual intersection legs, the non-parametric TreeExplainer architecture captures non-linear conditional dependencies without restrictive distributional assumptions. Conversely, the limitations of the proposed framework relative to traditional SPFs must also be acknowledged. Conventional SPFs are directly anchored to verified crash outcomes, explicitly account for macroscopic traffic exposure via Annual Average Daily Traffic (AADT) volumes, and yield parametric Crash Modification Factors (CMFs) suitable for network-wide economic appraisal [16,23]. In contrast, the proposed XAI framework models behavioral precursors rather than actual crash frequencies, requires intensive on-road data collection, and yields relative feature attributions (SHAP values) within a finite categorical domain without continuous traffic-volume covariates. Consequently, the proposed methodology should be viewed as a diagnostic microscopic complement to macroscopic SPF modeling rather than a standalone replacement. The analytical evidence emerging from cross-validation via Permutation Feature Importance and Dependence Plots converges to suggest a relevant transportation assumption: double-lane roundabouts present an excessive planimetric permissiveness that can compromise the systemic safety of the node. Such configurations, historically justified by the theoretical need to maximize traffic capacity and levels of service [36–38], delegate the maintenance of operational trajectories solely to horizontal markings within a permissive platform, radically disregarding the "Self-Explaining Road" principle. The absence of physical channelization transforms excessive lateral space into an operational margin that drivers tend to exploit to arbitrarily alter their operating speed and design trajectory, revealing intrinsic geometric incompatibilities, especially for turning maneuvers characterized by sharper angular deviations.

From a geometric design and vehicle dynamics perspective, the most critical quantitative evidence is linked to the exit maneuver. The hierarchical preeminence assumed by exit width and exit radius suggests how the infrastructure may fail in its traffic calming intent if transverse dimensions fall within permissive geometric ranges [12,23]. Specifically, the marginal dependency analysis isolated both wide exits (>8 m) and intermediate exit widths (6–8 m in total two-lane width) as critical dimensional bands associated with behavioral instability. Within this dimensional range, the exit width frequently generates positive risk peaks for lane cutting, particularly affecting right-turn and through movements [39,40]. Specifically, through movements are exposed to a greater risk of lateral path deviations, as they require the driver to negotiate a reverse-curve ("S") trajectory within a carriageway characterized by excessive lateral space; this configuration explains the driver's tendency to straighten the path to keep lateral acceleration within comfortable limits [12,19,41].

Similarly, adopting intermediate exit radii, between 10 and 30 m, is associated with a propensity for inappropriate speed, supporting kinematics in which the vehicle tends to rectify its path. A plausible phenomenological interpretation suggests that, in an optically permissive environment, the driver does not perceive the infrastructure as a moderating element, potentially tending to anticipate the radial acceleration phase even before physically

clearing the circulatory roadway. In order to minimize jerk (the derivative of acceleration over time) and counteract transverse acceleration, which is often exacerbated by negative cross slopes functional to hydraulic drainage, it is likely that the user is physiologically induced to increase the curvature radius of their path. From this perspective, systematic lane cutting (inappropriate trajectory) [42] and inappropriate speed can reasonably be interpreted not as trivial errors of inattention, but as probable kinematic responses of postural and dynamic adaptation, facilitated by an infrastructure granting excessive lateral degrees of freedom.

Shifting the predictive focus to vehicular interactions and the triggering of "Failure to yield", the analysis moves from rigid body dynamics to cognitive ergonomics (human factors). The dependency diagrams revealed that entry width seems to act as a relevant conflict catalyst, peaking in criticality at wide entry widths (>8 m in total two-lane width) [12,43]. Beyond this physical limit, the impact on the model becomes predominantly positive, triggering aggressive behavior regardless of the maneuver undertaken by the driver. A geometrically excessive entry can generate dangerous tactical ambiguity: by reducing the sense of optical constriction, it tends to induce an opportunistic approach and encourage poorly pondered insertions. While ocular and cognitive processes were not directly measured during the on-road runs, a theoretical behavioral hypothesis grounded in human factors literature can be formulated to interpret this empirical association. Specifically, the need to monitor multiple unchanneled conflicting flows within a wide entry area may overload the driver's attentional resources, a condition known in the literature to favor gaze concentration and peripheral perceptual narrowing (tunnel vision) [44,45]. Combined with potential visual occlusions from the vehicle's A-pillar during curvilinear approaches, such hypothesized attentional dispersion could contribute to "Looked-but-Failed-to-See" errors [46,47], impairing gap-acceptance judgments. Distinguishing this inferred cognitive-overload mechanism from directly measured psychological stress requires dedicated biometric monitoring; in this regard, recent real-world driving research has successfully utilized physiological indicators such as heart rate variability (HRV) to quantify how infrastructure demands and environmental conditions interactively elevate driver stress [48], offering a rigorous empirical paradigm to validate driver responses at complex roundabout entries alongside eye-tracking metrics.

Ultimately, maneuver kinematics is confirmed as a primary directional discriminant in the double-lane roundabout ecosystem. As amply documented by the dependency diagrams relative to exit maneuvers, the same permissive geometry produces markedly divergent effects depending on the trajectory type associated with the specific maneuver undertaken by the driver. While, on the one hand, excessive lateral space passively facilitates tangential or straight-crossing maneuvers by providing room to flatten the curve, albeit inducing dangerous lane cutting and inappropriate speed, on the other hand, it fails to provide the necessary physical guidance for maneuvers characterized by sharper angular deflections, such as left turns and U-turns. These latter, forced to develop within a road platform lacking impassable physical confinement, considerably dilate trajectory dispersion and increase exposure to transverse inter-vehicular conflicts. At the same time, because the empirical dataset was collected across four intersections within a single peripheral urban corridor in Italy, the generalizability of the derived geometric intervals to other road functional classes and international contexts requires careful consideration. While the underlying physical and perceptual mechanisms identified by the XAI framework (namely, path straightening to limit lateral acceleration across wide exits and attentional dispersion at wide uncanalized entries) reflect fundamental principles of vehicle dynamics and human factors consistent with international guidelines [1,4], the specific numerical boundaries (such as the 6–8 m and >8 m two-lane width classes or the 10–30 m exit radius range) are calibrated to urban

arterials with moderate operating speeds. To guide engineering practice, three distinct operational scenarios can be defined regarding the applicability of these findings. First, the derived geometric thresholds can be applied directly, without retraining the model, during road safety audits or retrofitting evaluations of uncanalized urban and suburban two-lane roundabouts characterized by moderate speed limits (40–50 km/h) and passenger-car-dominated traffic. In a common practical case, a municipal road agency inspecting an existing uncanalized two-lane roundabout can directly compare its leg dimensions against the identified critical bands: if an exit leg exhibits a two-lane width of 6–8 m (or >8 m) paired with an exit radius above 10 m, or if an entry leg exceeds 8 m in width, practitioners can immediately flag those specific approaches for physical narrowing via splitter-island extensions, curb realignments, or conversion into a turbo-roundabout without collecting new behavioral data. Second, retraining the Cost-Sensitive RF-SHAP model on locally collected trajectory or naturalistic data is required when evaluating uncanalized multi-lane roundabouts belonging to different functional classes or traffic environments, such as high-speed rural highways (70–90 km/h approach speeds), industrial corridors with a high share of articulated heavy vehicles requiring wider swept paths, or international jurisdictions with markedly different lane-marking conventions and gap-acceptance norms. Third, applying either the derived thresholds or the present model structure is impossible when the intersection topology or control regime violates the underlying operational assumptions: specifically, the framework cannot be applied to signalized multi-lane roundabouts, single-lane or mini-roundabouts, or already physically channelized layouts where inter-lane weaving is structurally prevented, nor can it be used for real-time deterministic crash prediction under gridlock congestion or for evaluating pedestrian and cyclist conflicts without expanding the feature space.

Although the present study did not directly evaluate physically channelized intersections or before-and-after retrofitting interventions, the behavioral vulnerabilities documented in the surveyed uncanalized layouts provide empirical motivation to consider structurally channelized design alternatives promoted by international guidelines. Advanced configurations like turbo-roundabouts, implementing physical mountable curbs right from the entry approach, are specifically designed in the literature to reduce excessive lateral space and address the trade-off between lateral acceleration control and trajectory straightening. By physically inhibiting lateral path deviations and guiding trajectories, such layouts are widely recognized in prior comparative studies as a promising engineering countermeasure to mitigate the behavioral anomalies observed herein while preserving traffic capacity [49,50].

5. Conclusions

This study developed a machine learning framework to quantitatively examine the interaction between double-lane roundabout geometry and driving behavior. The methodology combined multi-class classification with post hoc explainability: first, training a Cost-Sensitive Random Forest algorithm addressed the imbalance across behavioral classes. By assigning inverse-frequency weights to minority classes directly within the loss function rather than generating synthetic samples (such as via SMOTE or CTGAN), the approach preserved the empirical joint distribution and the natural variance of the observed driving behaviors. Subsequently, Permutation Feature Importance ranked the influence of geometric and maneuver predictors on out-of-sample Macro F1-Score decay, while SHAP Dependence Plots examined how feature contributions varied across predefined geometric intervals and maneuver types. Together, these analyses indicate that the operational adequacy of a geometric element cannot be evaluated in isolation, as its behavioral impact depends on the specific turning maneuver executed by the driver.

Through this approach, the research highlights significant structural and functional criticalities of double-lane roundabouts lacking physical channelization. It clearly emerged that excessive lateral space acts as a primary operational vulnerability factor, reducing traffic calming effectiveness. High geometric permissiveness, evidenced by the positive risk associations observed within the predefined categories for total two-lane exit widths (6–8 m and >8 m) and exit radii (10–30 m and >30 m), tends to neutralize the infrastructure's physical constriction. Drivers exploit wide cross-sections to opportunistically increase their operating speed, consequently tending to straighten their path and encroach into adjacent lanes to limit lateral acceleration. This suggests that lateral path deviation and inappropriate speed do not constitute independent anomalies but rather represent adaptive kinematic responses to the same permissive layout.

In parallel, the analysis of vehicular interactions and failures to yield highlighted how the impact of any planimetric element is inextricably linked to maneuver kinematics, which emerged as the dominant predictor. The double-lane infrastructure, while passively supporting tangential trajectories, shows evident limits in guiding maneuvers characterized by sharper angular deviations. In this context, the presence of wide two-lane entry widths exceeding 8 m is positively associated with yielding infractions. It is hypothesized that the need to monitor multiple unchanneled traffic streams within a wide entry zone may induce attentional dispersion and impair temporal gap evaluation, increasing the likelihood of opportunistic or forced entries.

Although the measured outcomes represent behavioral classifications rather than crash frequencies or conflict-severity indicators, and no direct comparison with channelized intersections was performed, the high prevalence of lane-cutting and speed anomalies observed in uncanalized layouts highlights the practical relevance of exploring physically separated design alternatives. Considering the established international literature, eliminating spatial ambiguity through physical channelization (such as spiral or turbo-roundabout schemes) represents a coherent engineering countermeasure worthy of pilot evaluation and before-and-after safety assessment, particularly in national contexts like Italy, where existing uncanalized multi-lane roundabouts require retrofitting.

The intrinsic limitations of this research outline useful directions for future developments. Although the predictive framework relies on 5622 observations that characterize behavioral adaptation within the surveyed intersections, the participant pool was limited to 30 non-professional drivers aged 21 to 65 from the same regional context. Consequently, future investigations should expand both the demographic and geographic sampling domains (incorporating novice drivers, older drivers above 65, and professional operators, as well as broader driving cultures) to improve sample representativeness and validate the model's out-of-site transferability across additional roundabout configurations, including direct comparative evaluations with turbo-roundabouts. Furthermore, future data collection should explicitly log instantaneous circulating traffic volumes, available temporal gaps, and yield opportunities, integrating surrogate safety measures (such as Post-Encroachment Time and Time-to-Collision) to separate geometric effects from traffic exposure and link behavioral anomalies to objective conflict severity.

In addition, future extensions of the algorithmic architecture should explicitly account for variations in traffic flow composition, particularly the presence of heavy vehicles, cyclists, and micromobility users, which can substantially alter operational dynamics at double-lane roundabouts. Specifically, a high share of heavy vehicles (such as trucks, buses, and articulated lorries) introduces severe physical and kinematic constraints: due to their large turning radii and wide swept paths, heavy vehicles frequently encroach upon adjacent lanes out of mechanical necessity rather than opportunistic speed optimization, while simultaneously constraining the lateral space available to surrounding passenger cars, lowering average

circulating speeds, and causing visual occlusion at entry approaches. Conversely, the presence of cyclists and e-scooter riders circulating along the outer edge of the ring introduces pronounced speed differentials and lateral clearance demands: passenger car drivers often execute lateral swerving maneuvers into the inner lane to overtake cyclists safely (generating interaction-induced trajectory shifts rather than purely geometric lane cutting), while multi-lane entries and exits amplify yielding ambiguities and visual scanning conflicts between exiting motor vehicles and circulating two-wheelers. Incorporating traffic composition percentages and user-class interactions into the predictor matrix will therefore be essential to distinguish geometry-driven opportunism from heterogeneous-flow adaptations.

Finally, testing the inferred cognitive-overload and tunnel-vision hypotheses upon entry will require dedicated empirical validation, combining virtual reality simulations, eye-tracking instrumentation, and real-world physiological monitoring (such as heart rate variability measurements) to objectively quantify psychological stress and visual ergonomics degradation across varying geometric constraints.

Author Contributions: Conceptualization, N.D. and S.L.; methodology, N.D. and S.L.; software, N.D. and S.L.; validation, S.L.; formal analysis, N.D.; investigation, N.D. and S.L.; resources, N.D. and S.L.; data curation, N.D. and S.L.; writing—original draft preparation, N.D. and S.L.; writing—review and editing, N.D. and S.L.; visualization, N.D.; supervision, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rodegerdts, L.A.; Bansen, J.; Tiesler, C.; Knudsen, J.; Myers, E.; Johnson, M.; Moule, M.; Persaud, B.; Lyon, C.; Hallmark, S.; et al. *Roundabouts: An Informational Guide*, 2nd ed.; NCHRP Report 672; Transportation Research Board: Washington, DC, USA, 2010; pp. 1–408. [CrossRef] [PubMed]
2. Elvik, R. Road safety effects of roundabouts: A meta-analysis. *Accid. Anal. Prev.* **2017**, *99*, 364–371. [CrossRef] [PubMed]
3. Distefano, N.; Leonardi, S.; Pulvirenti, G. Factors with the greatest influence on drivers' judgment of roundabouts safety. An analysis based on web survey in Italy. *IATSS Res.* **2018**, *42*, 265–273. [CrossRef]
4. Daniels, S.; Brijs, T.; Nuyts, E.; Wets, G. Explaining variation in safety performance of roundabouts. *Accid. Anal. Prev.* **2010**, *42*, 393–402. [CrossRef] [PubMed]
5. Hu, W.; McCartt, A.T.; Jermakian, J.S.; Mandavilli, S. Public opinion, traffic performance, the environment, and safety after construction of double-lane roundabouts. *Transp. Res. Rec.* **2014**, *2402*, 47–55. [CrossRef]
6. Wang, J.; Cicchino, J.B. Safety effects of roundabout conversions in Carmel, Indiana, the Roundabout City. *J. Saf. Res.* **2022**, *82*, 159–165. [CrossRef] [PubMed]
7. Tollazzi, T.; Rencelj, M. Modern and alternative types of roundabouts: State of the art. In *Proceedings of the 9th International Conference on Environmental Engineering (ICEE), Vilnius, Lithuania, 22–24 May 2014*; Vilnius Gediminas Technical University Press Technika: Vilnius, Lithuania, 2014; pp. 1–8. [CrossRef]
8. Fortuijn, L.G.H. Turbo roundabouts: Design principles and safety performance. *Transp. Res. Rec.* **2009**, *2096*, 16–24. [CrossRef]
9. Montella, A. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. *Accid. Anal. Prev.* **2011**, *43*, 1451–1463. [CrossRef] [PubMed]
10. Mauro, R.; Branco, F. Comparative analysis of compact multilane roundabouts and turbo-roundabouts. *J. Transp. Eng.* **2010**, *136*, 316–322. [CrossRef]
11. Macioszek, E. The road safety at turbo roundabouts in Poland. *Arch. Transp.* **2015**, *33*, 57–67. [CrossRef]
12. Pulvirenti, G.; Distefano, N.; Leonardi, S.; Tollazzi, T. Are double-lane roundabouts safe enough? A chaid analysis of unsafe driving behaviors. *Safety* **2021**, *7*, 20. [CrossRef]

13. Pulvirenti, G.; De Ceunynck, T.; Daniels, S.; Distefano, N.; Leonardi, S. Safety of bicyclists in roundabouts with mixed traffic: Video analyses of behavioural and surrogate safety indicators. *Transp. Res. Part F Traffic Psychol. Behav.* **2021**, *76*, 72–91. [CrossRef]
14. Zheng, L.; Sayed, T.; Mannering, F. Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions. *Anal. Methods Accid. Res.* **2021**, *29*, 100142. [CrossRef]
15. Giuffrè, O.; Granà, A.; Tumminello, M.L.; Giuffrè, T.; Trubia, S. Surrogate measures of safety at roundabouts in AIMSUN and VISSIM environment. In *Roundabouts as Safe and Modern Solutions in Transport Networks and Systems*; Macioszek, E., Akçelik, R., Sierpiński, G., Eds.; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2019; Volume 52, pp. 59–79. [CrossRef]
16. Washington, S.P.; Karlaftis, M.G.; Mannering, F. *Statistical and Econometric Methods for Transportation Data Analysis*; Chapman and Hall/CRC: New York, NY, USA, 2003; pp. 1–425. [CrossRef]
17. Yoshioka, K.; Nakamura, H.; Shimokawa, S.; Morita, H. Modeling of a novel risk index for evaluating the geometric designs of roundabouts. *Accid. Anal. Prev.* **2020**, *145*, 105702. [CrossRef] [PubMed]
18. Gkyrtis, K.; Kokkalis, A. An overview of the efficiency of roundabouts: Design aspects and contribution toward safer vehicle movement. *Vehicles* **2024**, *6*, 433–449. [CrossRef]
19. Maji, A.; Ghosh, I. Assessing roundabout safety perceptions under heterogeneous traffic: Socio-demographic and geometric influences in Indian urban contexts. *Int. J. Transp. Sci. Technol.* **2026**, *15*, 214–229. [CrossRef]
20. Klobučar, M.; Šurdonja, S.; Deluka-Tibljaš, A. Impact of geometric parameters and operational speed on traffic safety at roundabouts: A conflict analysis using microsimulation models. *J. Road Saf.* **2025**, *36*, 45–58. [CrossRef]
21. Dwekat, T.M.; Almsre, A.A.; Ashqar, H.I. Effect of roundabout design on the behavior of road users: A case study of roundabouts with application of Unsupervised Machine Learning. *arXiv* **2023**, arXiv:2309.14540. [CrossRef]
22. Zubaidi, H.A.; Anderson, J.C.; Hernandez, S. Understanding roundabout safety through the application of advanced econometric techniques. *Int. J. Transp. Sci. Technol.* **2020**, *9*, 309–321. [CrossRef]
23. Vinayaraj, V.S.; Perumal, V. Developing safety performance functions and crash modification factors for urban roundabouts in heterogeneous non-lane-based traffic conditions. *Transp. Res. Rec.* **2023**, *2677*, 644–661. [CrossRef]
24. Al-Nabulsi, D.; Hassouneh, A. Development and calibration of roundabout safety performance functions using machine learning: A case study from Amman, Jordan. *J. Eng. Appl. Sci.* **2025**, *72*, 104. [CrossRef]
25. Zhou, X.; Yuan, Z.; Han, Y.; Xu, T.; Yang, L.T. CARLA-Round: A multi-factor simulation dataset for roundabout trajectory prediction. *arXiv* **2026**, arXiv:2601.12119. [CrossRef]
26. Ministero delle Infrastrutture e dei Trasporti. *Norme Funzionali E Geometriche per la Costruzione Delle Intersezioni Stradali, D.M. 19 Aprile 2006*; Gazzetta Ufficiale della Repubblica Italiana, Serie Generale n. 170: Roma, Italy, 2006.
27. Distefano, N.; Pulvirenti, G.; Leonardi, S.; Tollazzi, T. Unsafe driving behaviours at single-lane roundabouts: Empirical evidence from CHAID method. In *New Metropolitan Perspectives*; Bevilacqua, C., Calabrò, F., Della Spina, L., Eds.; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2021; Volume 217, pp. 145–163. [CrossRef]
28. Elreedy, D.; Atiya, A.F.; Kamalov, F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Mach. Learn.* **2024**, *113*, 4903–4923. [CrossRef]
29. Akritidis, L.; Bozanis, P. A Conditional GAN for Tabular Data Generation with Probabilistic Sampling of Latent Subspaces. *arXiv* **2025**, arXiv:2508.00472. [CrossRef]
30. Zhao, M. Modeling Driving Behavior at Single-Lane Roundabouts. Ph.D. Thesis, Technische Universität Braunschweig, Braunschweig, Germany, 2019. [CrossRef]
31. Rahmani, S.; Calvert, S.C.; van Arem, B. Decentralized Modeling of Vehicular Maneuvers and Interactions at Urban Junctions. *arXiv* **2025**, arXiv:2507.21547. [CrossRef]
32. Lin, Z.; Tian, Z.; Lan, J.; Zhao, D.; Wei, C. Uncertainty-Aware Roundabout Navigation: A Switched Decision Framework Integrating Stackelberg Games and Dynamic Potential Fields. *IEEE Trans. Veh. Technol.* **2026**, *75*, 9027–9039. [CrossRef]
33. Mahmoodi, N.; Shirazi, H.; Fakhredanesh, M. Automatically weighted focal loss for imbalance learning. *Neural Comput. Appl.* **2025**, *37*, 4035–4052. [CrossRef]
34. Roy, D.; Pramanik, R.; Sarkar, R. Margin-Aware Adaptive-Weighted-Loss for Deep Learning Based Imbalanced Data Classification. *IEEE Trans. Artif. Intell.* **2024**, *5*, 776–785. [CrossRef]
35. Li, Y.; Jin, J.; Ma, J.; Zhu, F.; Jin, B.; Liang, J.; Chen, C.L.P. Imbalanced least squares regression with adaptive weight learning. *Inf. Sci.* **2023**, *648*, 119541. [CrossRef]
36. Ali, H.K.M.; Majid, H.M. Comparative Evaluation of Roundabout Capacities Methods for Single-lane and Multi-lane Roundabout. *J. Eng.* **2023**, *29*, 80–95. [CrossRef]
37. Bie, J.; Lo, H.K.; Wong, S.C. Capacity evaluation of multi-lane traffic roundabout. *J. Adv. Transp.* **2010**, *44*, 245–255. [CrossRef]
38. Šarić, A.; Lovrić, I. Multi-lane roundabout capacity evaluation. *Front. Built Environ.* **2017**, *3*, 42. [CrossRef]
39. Al-Mistarehi, B.; Imam, R.; Al-Shawabkah, M.M.; Shtayat, A.; Al-Omari, A. Assessing the effect of geometric design and land use on roundabouts using video camera. *Civ. Eng. J.* **2024**, *10*, 1805–1821. [CrossRef]

40. Zheng, Y.; Shyrokau, B.; Keviczky, T. Comfort and time efficiency: A roundabout case study. In *Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021*; IEEE: Piscataway, NJ, USA, 2021; pp. 1618–1623. [CrossRef]
41. Mandavilli, S.; McCartt, A.T.; Retting, R.A. Crash patterns and potential engineering countermeasures at Maryland roundabouts. *Traffic Inj. Prev.* **2009**, *10*, 44–50. [CrossRef] [PubMed]
42. Fernandes, P.; Ferreira, E.; Macedo, E.; Coelho, M.C. Unraveling roundabout dynamics: Analysis of driving behavior, vehicle performance, and exhaust emissions. *Transp. Res. Part D Transp. Environ.* **2024**, *133*, 104308. [CrossRef]
43. Salamati, K.; Schroeder, B.J.; Geruschat, D.R.; Rouphail, N.M. Event-based modeling of driver yielding behavior to pedestrians at two-lane roundabout approaches. *Transp. Res. Rec.* **2013**, *2389*, 1–11. [CrossRef] [PubMed]
44. Arias-Portela, C.Y.; Mora-Vargas, J.; Caro, M. Situational awareness assessment of drivers boosted by eye-tracking metrics: A literature review. *Appl. Sci.* **2024**, *14*, 1611. [CrossRef]
45. Abbasi, J.A.; Mullins, D.; Ringelstein, N.; Reilhac, P.; Jones, E.; Glavin, M. An analysis of driver gaze behaviour at roundabouts. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 8715–8724. [CrossRef]
46. Herslund, M.-B.; Jørgensen, N.O. Looked-but-failed-to-see-errors in traffic. *Accid. Anal. Prev.* **2003**, *35*, 885–891. [CrossRef] [PubMed]
47. Koustanaï, A.; Boloix, E.; Van Elslande, P.; Bastien, C. Statistical analysis of "looked-but-failed-to-see" accidents: Highlighting the involvement of two distinct mechanisms. *Accid. Anal. Prev.* **2008**, *40*, 461–469. [CrossRef] [PubMed]
48. Wu, J.; Wang, Z.; Wang, J.; Huang, S.; Pervez, A. Effects of day-night conditions and tunnel length on drivers' psychological stress: Insights from a real-world driving study. *Transp. Saf. Environ.* **2026**, *8*, tdag055. [CrossRef]
49. Mauro, R.; Cattani, M.; Guerrieri, M. Evaluation of the safety performance of turbo roundabouts by means of a potential accident rate model. *Balt. J. Road Bridge Eng.* **2015**, *10*, 28–38. [CrossRef]
50. Tesoriere, G.; Campisi, T.; Canale, A.; Zgrablić, T. The surrogate safety appraisal of the unconventional elliptical and turbo roundabouts. *J. Adv. Transp.* **2018**, *2018*, 2952074. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.