

COMPARISON OF MANUAL AND AUTOMATED INFRASTRUCTURE LABEL EXTRACTION FOR HARSH EVENT PREDICTION

Julia Alves Porto*, Apostolos Ziakopoulos, George Yannis
National Technical University of Athens, *julia_porto@mail.ntua.gr
Daniel Felipe Lopez
FRED Engineering
John Avgeros
OSeven Telematics

ABSTRACT

Despite a lot of progress in recent years towards increasing road safety levels, road crashes remain a leading cause of death worldwide, especially amongst youngster individuals. This scenario incentivizes the continuous development of fast and proactive methods for safety assessments. To engage on the proactive aspect, many researchers, analysts and policymakers rely on surrogate safety measures, such as harsh events (HE) and near-misses, and road infrastructure data to estimate safety levels without relying on historical crash data, such as the International Road Assessment Programme (iRAP) star rating method. Moreover, computational advancements have increased the speed in which data collection is made, allowing for even faster analyses. Infrastructural assessment and behavior assessment, however, rarely draw on the same data source, limiting integrated analysis connecting both. This study addresses that gap through a proof-of-concept comparison of two parallel pipelines for predicting HE occurrence along approximately 64 km of road in northern Italy: one built on manually coded iRAP infrastructure attributes, the other on infrastructure attributes extracted automatically through computer vision (CV) and complementary open and commercial geospatial datasets. HE are defined by the presence of an abrupt disturbance in driver's behavior, flagged by either harsh acceleration or harsh braking, and derived from OSeven Telematics data, together with trip counts and mean driving speed. The data was aggregated over the same 100-meter segment used for iRAP coding. Corresponding video footage was processed with a semantic segmentation model (DeepLabV3+, trained on Cityscapes), a lane-delineation model, and a YOLO11m object detector to generate CV-based counterparts to the iRAP variables retained after correlation- and performance-based screening, alongside KML linestring manipulations and OSM attribute extraction. Two XGBoost regression models, sharing an identical repeated stratified 5-fold cross-validation scheme (10 repetitions, 50 splits), were trained to predict the number of HE per segment from the manually coded and the CV-derived attribute sets, respectively. The manual-based model achieved a mean R^2 of 0.26 and Spearman's rho of 0.33; the CV-based model achieved an R^2 of 0.20 and rho of 0.29, both outperforming a naive mean-prediction baseline despite strong zero-inflation in the target (92% zero-HE segments). For both models, telematics exposure variables, namely mean speed and trip counts, were by far the most influential predictors. The two models diverged, however, in which infrastructure attributes carried explanatory weight: the manual-based model relied most on guardrail presence, traffic volume and speed limit, while the CV-based model relied on traffic-control objects, wall presence and curvature. This divergence is attributed largely to differences in encoding logic, namely iRAP's forced single-category coding versus the CV pipeline's independent, continuous attribute outputs, rather than to a gap in overall predictive power. The results indicate that automatically extracted infrastructure attributes can approximate manually coded iRAP data for HE prediction at comparable, while also showing that the two approaches are not interchangeable at the level of individual attribute importance. These findings support the use of CV to help scale proactive safety assessments, and motivate further work on attribute fusion, larger samples, and directional driver-gaze effects.

Keywords: road safety, telematics, computer vision, harsh events, infrastructure, segmentation

1. INTRODUCTION

Road fatalities account for 1.2 million deaths per year worldwide, thus remaining an important issue despite recent advancements in road safety rates (WHO, 2023). Since the early 2000s, the road safety paradigm has shifted from reactive approaches, relying on crash counts to prevent future crashes, to proactive approaches, assessing infrastructure and behaviour data independently of crash occurrences, commonly relying on surrogate measures of safety (SMoS) (Singh et al., 2024).

Infrastructure-based proactive safety assessments, such as Road Safety Inspection (RSI), survey the road network and roadside elements and their quality, providing a diagnosis of road quality for each user type (Singh et al., 2024). Most approaches rely on video inspection and manual annotators to classify the road's visual elements. Among other methods, the International Road Assessment Programme (iRAP) star rating method can be highlighted for its great detail level (iRAP - International Road Assessment Programme, 2025; Paliotto et al., 2024) and global adoption (Brkić et al., 2022; Murozi et al., 2022; Sunil and Sharma, 2021).

Among behavior-based proactive measures, conflict (near-miss) is the most widely used surrogate (Tarko, 2018), defined as either an evasive action taken to avoid collision or as spatial or temporal approximation between two road users (Saunier and Lareshyn, 2021). Alternatively, harsh events (HE), such as harsh braking and harsh acceleration, provide a single-vehicle evasive-action metric (Hossain et al., 2026) and have likewise been used as a SMoS in several recent studies (Gupta et al., 2024; Hossain et al., 2026; Nikolaou et al., 2025; Ziakopoulos, 2024; Ziakopoulos et al., 2022; Ziakopoulos and Yannis, 2024).

These proactive methods rely on different data sources: network-wide assessments primarily use continuous street-level imagery (González-Hernández et al., 2024), conflicts use often facility-level observations over time (Arun et al., 2021), and HE use either connected vehicle (Gupta et al., 2024; Hossain et al., 2026) or telematics data (Nikolaou et al., 2025; Ziakopoulos, 2024; Ziakopoulos et al., 2022; Ziakopoulos and Yannis, 2024). Bridging those different surrogate approaches have hardly been attempted in the literature: their data sources do not connect, precluding integrated analyses that use network-wide assessment data to predict HE occurrence. This study addresses that gap through a proof-of-concept methodology that combines video-based infrastructure coding with telematics-derived HE occurrence over the same road segments, and further tests whether the infrastructure side of this pipeline can be automated through computer vision without materially compromising predictive quality.

2. DATA DESCRIPTION

The base dataset used in this study covers approximately 64 km of road across three routes in northern Italy: SS54, SS55 and SS13. This dataset is part of a road safety inspection conducted by Fred Engineering following the iRAP methodology. For each road, a CSV file containing manually coded iRAP variables at 100-m intervals was provided, along with the corresponding video footage used during the inspection. In total, nine videos were supplied at 29.97 frames per second and a 2304x1296 pixel

resolution, each accompanied by a KML file recording the vehicle's geolocation at approximately 1-second intervals. The surveyed roads cross predominantly rural areas and are classified as tertiary roads by OpenStreetMap (OSM). Data collection took place on 17-18 September 2024.

For the same study area, telematics data were provided by OSeven Telematics, a Greece-based telematics company that processes data collected from mobile phones. All records were fully anonymized prior to sharing, in compliance with General Data Protection Regulation (GDPR). The telematics dataset spans trips recorded between 2021 and 2025 and comprises 322 trips unevenly distributed across the study area. For each recorded telematics point, driving speed, harsh-event classification (harsh braking, harsh acceleration, or no event) and harsh-event intensity (in a 1-3 scale) were provided, sampled at 1-second resolution per trip.

To integrate the iRAP and telematics datasets, each telematics point was assigned to its nearest iRAP segment (represented as a point geometry) using a spatial nearest-neighbor join, subject to a 100-m maximum distance threshold. Every iRAP segment had at least one telematics point within this threshold, so no segments were excluded from the joined dataset. For each segment, the number of unique trips and the number of telematics points were obtained by counting the corresponding records, while the number of HE was obtained by summing event occurrences. Mean driving speed was computed as the average of all telematics points' speeds within the segment. The resulting distribution of HE across segments is markedly imbalances, as shown in Table 1. Figure 1 shows the spatial distribution of the surveyed roads (blue) and the segments where HE were recorded (red).

Table 1: Data summary

Road	Coverage (km)	Number of Trips	Number of Data Points	Number of Harsh Events (HE)	Number of HE Segments
SS13	14.3	75	14,766	15	10
SS54	34.8	30	20,035	217	17
SS55	15.6	217	16,060	48	24

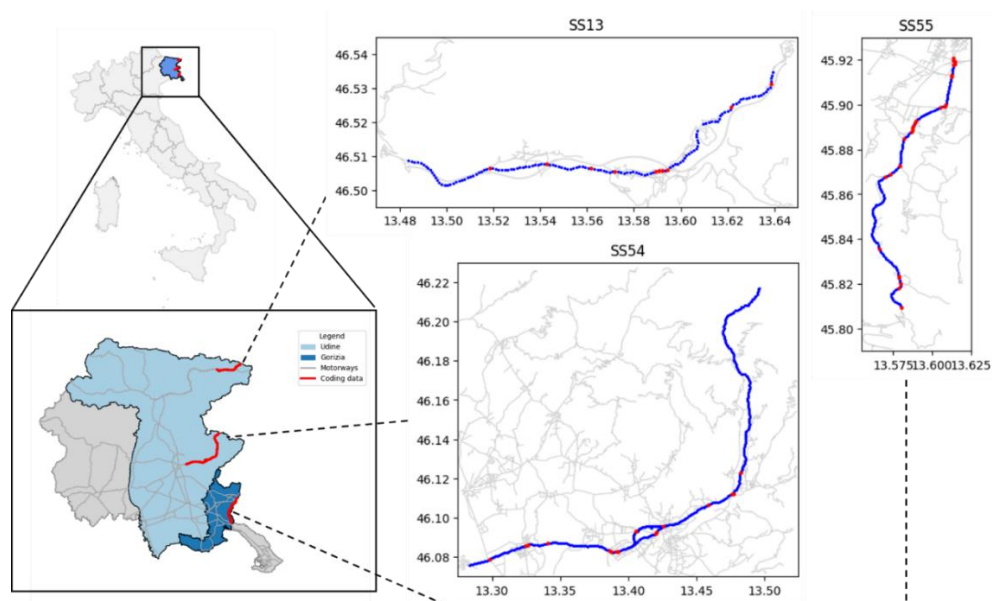


Figure 1: geolocation of survey data

For the automatic-based assessment, the KML datafile was similarly split into approximately 100-m segments aligned with the iRAP segmentation, with the end and start points of consecutive segments bridged for continuity and a voting procedure used to keep points originating from the same source video together. The corresponding video frame ranges were retained for each segment, allowing frame-level attributes to be extracted from the same footage used for the iRAP-aligned segments.

3. METHODS

In this Section, we describe the target variable and baseline model used, the attributes retained and the metrics used to evaluate and compare the models.

3.1 Target variable and baseline model

After excluding variables without variance, highly correlated variables (> 0.9 Pearson correlation), we compared three target variables (presence of HE, number of HE and ratio between mean speed and speed limit) and four model types (econometric, CatBoost, XGBoost and Random Forest). We converted categorical variables into dummy (one-hot-encoded) variables and eliminated those whose correlation was too high after this process to allow the econometric models to converge.

After this process, we kept 20 iRAP variables, of which a considerable percentage was subdivided into dummy variables for model processing. Of the tested models, number of HE with XGBoost showed the greatest correct values and metrics, and were defined as respectively our target variable and baseline model for further step.

3.2 Retained attributes

For each of the 20 retained iRAP attributes, we sought a computational counterpart from open or freely accessible commercial sources, such as the TomTom API. A proxy was retained if it achieved either over 50% correctness against its own validation or test data or, where that was unavailable, an absolute Pearson correlation above 0.20 with the manually coded variable (taken in absolute value because some proxies, such as lane-delineation quality, are inversely coded relative to iRAP). Table 2 summarizes the retained and excluded variables and the criterion used to judge proxy quality.

Table 2: Retained and excluded variables

iRAP Variable	Automated proxy	Quality	Retained
Building / rigid structure – passenger side (obj. 13)	Semantic segm.: building (right)	89% IoU	Yes
Sidewalk presence	Semantic segm.: sidewalk	76% IoU	Yes
Building / rigid structure – driver side (obj. 13)	Semantic segm.: building (left)	89% IoU	Yes
Tree ≥ 10 cm – passenger side (obj. 11)	Semantic segm.: vegetation (right)	89% IoU	Yes
Street lighting (poles)	Semantic segm.: poles	44% IoU	No
Tree ≥ 10 cm – driver side (obj. 11)	Semantic segm.: vegetation (left)	89% IoU	Yes
Sign / post / pole ≥ 10 cm – driver side (obj. 12)	Semantic segm.: sign, light, pole (left)	51% IoU	Yes
Sign / post / pole ≥ 10 cm – passenger side (obj. 12)	Semantic segm.: sign, light, pole (right)	51% IoU	Yes
Vertical face / slope – driver side (obj. 6 and 7)	Semantic segm.: wall (left)	54% IoU	Yes

Vertical face / slope – passenger side (obj. 6 and 7)	Semantic segm.: wall (right)	54% IoU	Yes
Vehicle flow (AADT)	TomTom: sample size – same road	$r = 0.52$	Yes
Intersecting road volume	TomTom: sample size – intersecting road	$r = -0.04$	No
Speed limit	TomTom: speed limit	$r = 0.53$	Yes
Curvature class	Spline second derivative, 10-knot fit	$F1 = 0.71$	Yes
Lane delineation quality	Mean lane-model confidence	$r = -0.41$	Yes
Land use by side (commercial / industrial / residential)	OSM buildings by category, 50m buffer	$r = 0.02-0.30$	No
Roundabout	OSM roundabouts, 50m buffer	$r = 0.44$	Yes
Merge lane	OSM “turn lane”, 50m buffer	$r = -0.01$	No
Unified land use - commercial	OSM commercial buildings, 50m buffer	$r = 0.31$	Yes
Unified land use - industrial	OSM industrial buildings, 50m buffer	$r = 0.30$	Yes
Unified land use - residential	OSM residential buildings, 50m buffer	$r = 0.04$	No
No intersection	OSM graph, degree ≤ 2	$r = 0.34$	Yes
3-leg / 4-leg intersection	OSM graph, degrees 3 / 4	$r = 0.18 / 0.08$	No
Guardrail (metal barrier) – passenger side (obj. 1)	Object detection: guardrail (left)	97% mAP-50	Yes
Guardrail (metal barrier) – driver side (obj. 1)	Object detection: guardrail (right)	97% mAP-50	Yes
Low rigid object $\geq 20\text{cm}$ – passenger side (obj. 16)	Object detection: bollard (left)	32% mAP-50	No
Low rigid object $\geq 20\text{cm}$ – driver side (obj. 16)	Object detection: bollard (right)	32% mAP-50	No

Semantic segmentation used DeepLabV3+ (Chen et al., 2018) trained on Cityscapes (Cordts et al., 2016); lane delineation used a dual-branch multiclass lane segmentation model (Porto et al., 2026); and object detection used YOLO11m (Jocher et al., 2023) trained on the DORIE dataset (Katsamenis et al., 2025). Telematics exposure variables, mean speed and trip count, were retained in both the manual- and CV-based models given their high contribution to HE occurrence

Several attributes considered early on were ultimately excluded for lack of a reliable proxy. Property access points were dropped because the underlying OSM data was too unreliable to attach confidently to road pixels or the road network. Paved shoulder and median type both required road-edge or centerline delineation beyond the reliability of the lane model; median type would additionally require a training set specifically covering central-line hatching, which was considered out of scope for the current work. For area type (rural or urban), a Eurostat LUCAS classification mapped to a 2-km grid correlated only -0.05 with the manually coded iRAP variable and was therefore not considered a reliable proxy.

Roadside object distance proved particularly difficult to approximate. In the iRAP scheme, distance refers to a single, most-hazardous object at each location; an equivalent computational measure would need to combine object identity and distance into one value, discarding information about other objects present in the same image. This attribute was therefore excluded from the current study.

3.3 Model comparison

To reduce bias due to train/test partition choice, we employed a stratified k-fold cross-validation strategy, stratified on the binary presence or absence of HE, with 10 repetitions of 5 folds each, totalling of 50 train/test splits. The same fold assignments and random seed were used for both the manually coded and computationally derived feature sets, ensuring that per-fold results are paired. For each fold, we tracked R^2 , MAE, RMSE, Spearman correlation on the test set, the fraction of predictions rounding to zero, R^2 restricted to the subset of test cases with a nonzero actual target, and Spearman correlation restricted to that same nonzero subset. Mean SHAP-based attribute importance was computed per model, together with each attribute's frequency among the five most important features per split.

4. RESULTS

Table 3 reports the monitored metrics, averaged across folds per model and Figure 2 presents the corresponding SHAP values distributions. As a reference point, a naive baseline predicting the training-fold mean achieved a mean absolute error of 0.264 (manual) and 0.271 (automated). This slight difference is credited for two segments been excluded during KML-segment matching.

Table 3: Metric outputs per model

Metric	Manual-based model		Automatic-based model	
	Value	Std. Dev.	Value	Std. Dev.
R^2	0.2623	0.1345	0.1982	0.1300
MAE	0.1728	0.0326	0.1735	0.0268
RMSE	0.5088	0.1700	0.5537	0.1594
Spearman rho	0.3256	0.0692	0.2920	0.0680
Zero-prediction ratio	0.9677		0.9844	
Nonzero-subset R^2	-2.1703	2.7048	-1.7620	1.7440
Nonzero-subset Spearman rho	0.4565	0.2593	0.4321	0.2386

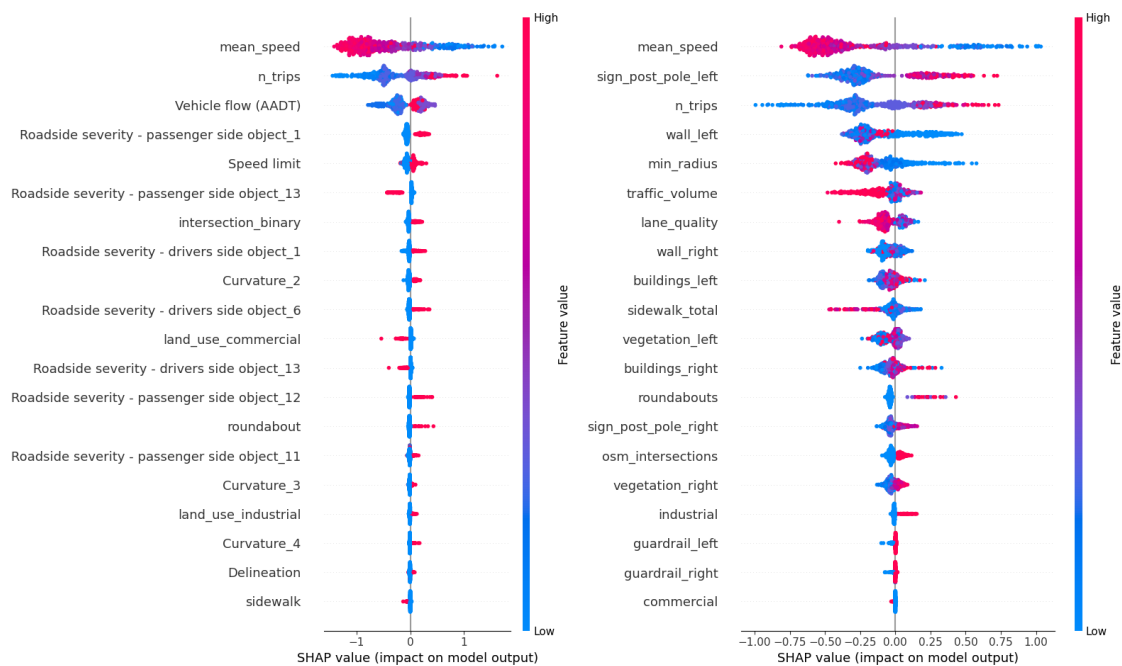


Figure 2: SHAP-values for manual- (left) and automatic-based model (right)

The automatic-based model achieved lower metrics compared to the manual-based one, but both surpassed the naïve MAE. They both tended to over-estimate the null-events, but that is partially attributable to the original dataset being very imbalanced itself, with 92% of 0-event occurrences. For both models, the telematics-based exposure metrics, mean speed and number of trips, were the most important attributes to predict HE occurrence. However, the importance of infrastructure-based metrics was substantially different between two models. Table 4 presents the mean SHAP-value and attribute frequency among the top-5 most important ones.

Table 4: Manual- and computer-based models SHAP frequency comparison

Manual-based variable	Mean SHAP value	Top-5 Freq.	Computer-based variable	Mean SHAP value	Top-5 Freq.
mean_speed	0.703	1.00	mean_speed	0.459	1.00
n_trips	0.464	1.00	n_trips	0.288	1.00
Vehicle flow (AADT)	0.248	1.00	traffic volume	0.106	0.18
Guardrail, passenger side	0.090	0.70	guardrail_right	0.006	0.00
Speed limit	0.075	0.56	speed_limit_tomtom	0.000	0.00
Building, passenger side	0.067	0.40	buildings_right	0.063	0.00
Intersection (binary)	0.044	0.08	osm_intersections	0.036	0.00
Guardrail, driver side	0.040	0.06	guardrail_left	0.008	0.00
Curvature (class 2)	0.039	0.16	min_radius	0.166	0.76
Wall / slope, driver side	0.031	0.00	wall_left	0.207	0.96
Building, driver side	0.024	0.02	buildings_left	0.067	0.00
land_use_commercial	0.023	0.00	commercial	0.001	0.00
Sign / light / pole, passenger side	0.022	0.00	sign_post_pole_right	0.048	0.00
Tree >= 10cm, passenger side	0.021	0.02	vegetation_right	0.041	0.00
Roundabout	0.019	0.00	roundabouts	0.046	0.00
Curvature (class 3)	0.015	0.00			
Lane delineation quality	0.012	0.00	lane_quality	0.084	0.06
Tree >= 10cm, driver side	0.011	0.00	vegetation_left	0.061	0.02
land_use_industrial	0.011	0.00	industrial	0.015	0.00
Sidewalk	0.011	0.00	sidewalk_total	0.072	0.00
Curvature (class 4)	0.010	0.00			
Wall / slope, passenger side	0.009	0.00	wall_right	0.068	0.04
Sign / light / pole, driver side	0.008	0.00	sign_post_pole_left	0.289	0.98

Specifically, guardrail presence on the passenger side, segment traffic volume and speed limit rank among the top five most influential variables in the manual-based model but not in the computer-based model. Conversely, the presence of traffic signs, traffic lights, poles or wall on the driver's side and segment curvature are influential in the computer-based model but considerably less so in the manual-based model. This divergence in variable ranking is most plausibly attributable to a structural difference in encoding logic: iRAP's manual coding selects a single classification from among 14 possible roadside attributes (here condensed into seven category types), whereas the computer-based assessment independently records the frequency of each object type across the road network. The latter approach permits the model to capture co-occurrence and interaction effects among roadside objects that are precluded by the mutually exclusive, forced-choice structure of the iRAP coding scheme.

A related but distinct factor is the difference in encoding granularity between the two approaches. The elevated significance of traffic sign, traffic light, and pole variables in the computer-based model, relative to their null significance under manual coding, is consistent with the model deriving more signal from continuous, granular outputs than from binary presence indicators. The same pattern is evident for curvature: the continuous “min_radius” variable ranks as highly significant in the computer-based model, whereas the corresponding categorical curvature classes derived from iRAP coding do not.

For both models, despite the significant variables being different, one variable side appears as much more significant than the other. For the manual-based model, guardrail (object 1) on passenger’s side has a much larger value than its counterpart on driver’s side. The same happens with buildings. For the computer-based model, both wall pixels and the aggregated traffic light, traffic sign or poles are much more significant on the left than on the right. Since there was no distinction when aggregating the telematics data regarding the travel direction, it cannot be attributed to driver gaze. Actually, it suggest that the left/right distinction is not meaningful for model prediction. When one has to decide a single category of object to assign, though, removing the left/right distinction means to halve the possible detected objects. However, using continuous model output, the simultaneous presence of objects can be more easily retained, especially if combined with distance measurements (not used in the presence study). Nevertheless, investigating whether driver gaze does linger on one side of the road or another should provide fruitful information about how to implement, for example, signalization without disrupting traffic and is a recommended line of work for future research.

Finally, land use (either commercial or industrial) and sidewalk presence were non-significant in both models. Contrary to our expectation that these variables would serve as reasonable proxies for an urbanized environment (commercial land use more so than industrial), these results suggest either that urban-area presence alone is not associated with increased harsh-event occurrence, or that these particular characteristics do not meaningfully influence driving behaviour.

5. CONCLUSION

This work presents a side-by-side comparison of HE prediction models, built on (i) iRAP-style manual infrastructure attribute labeling and (ii) computer-based labeling derived from multiple data sources, over approximately 64 km of road in northern Italy. Using XGBoost and repeated cross-validation strategy, both models exhibited comparable predictive performance and identified telematics trip-related variables (mean speed and number of trips) as the most influential predictors of HE occurrence. The infrastructural elements contributing most to each model’s predictability, however, varied substantially between models: the computer-based model placed high importance on the presence of traffic lights, traffic signs, poles, walls and curvature; while the manual-based model placed high importance on the presence of buildings, guardrails and the speed limit.

This work is subject to several limitations. First and foremost, the analysis was restricted to a limited road network spanning a few kilometres in northern Italy. Second, the dataset is limited not only by the number of road segments but also by

the number of telematics trips which compose the target variable. Consequently, the results presented here are not generalizable. Third, driving direction was not incorporated when aggregating telematics events. The purpose of this research was to provide a proof-of-concept comparison between manual and computer-based labelling approaches for safety-related assessment, rather than to establish a decisive causal account of HE occurrence. Driving direction was not incorporated when aggregating telematics events.

Future work will pursue additional data sources to cover attributes not yet addressed, particularly intersection presence and volume and object-distance proxies; a larger dataset with more trips per segment, to move from proof-of-concept toward a full case study; incorporation of driver gaze into HE prediction; and a dedicated investigation of whether the influence of roadside elements on driving behaviour differs between the left and right sides of the road from the driver's point of view. We also plan on including data balancing techniques on future work on larger datasets.

Acknowledgements

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101119590.

Bibliography

- Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. *Accid. Anal. Prev.* 153, 106016. <https://doi.org/10.1016/j.aap.2021.106016>
- Brkić, I., Miler, M., Ševrović, M., Medak, D., 2022. Automatic Roadside Feature Detection Based on Lidar Road Cross Section Images. *Sensors* 22, 5510. <https://doi.org/10.3390/s22155510>
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. <https://doi.org/10.48550/arXiv.1802.02611>
- Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3213–3223. <https://doi.org/10.1109/CVPR.2016.350>
- Global Status Report on Road Safety 2023, 1st ed. ed, 2023. . World Health Organization, Geneva.
- González-Hernández, B., Ngwah, E.C., Usami, D.S., Persia, L., 2024. In-built network-wide road safety assessment methodologies for rural roads. *Traffic Inj. Prev.* 25, 1048–1054. <https://doi.org/10.1080/15389588.2024.2361461>
- Gupta, N., Jashami, H., Savolainen, P.T., Gates, T.J., Barrette, T., Powell, W., 2024. Examining the Relationship between Connected Vehicle Driving Event Data and Police-Reported Traffic Crash Data at the Segment- and Event Level. *Transp. Res. Rec.* 2678, 1378–1394. <https://doi.org/10.1177/03611981241243329>
- Hossain, M.T., Lee, J., Besenski, D., Spasovic, L., 2026. Evaluating Harsh Braking Events as a Surrogate Measure of Crash Risk Using Connected-Vehicle Telematics. *Vehicles* 8. <https://doi.org/10.3390/vehicles8030068>
- IRAP - International Road Assessment Programme, 2025. AiRAP project. URL <https://irap.org/project/ai-rap/> (accessed 1.5.26).
- Jocher, G., Qiu, J., Chaurasia, A., 2023. Ultralytics YOLO.
- Katsamenis, I., Bakalos, N., Lappas, A., Protopapadakis, E., Martín-Portugués Montoliu, C., Doulamis, A., Doulamis, N., Rallis, I., Kalogeras, D., 2025. DORIE: Dataset of Road Infrastructure Elements—A Benchmark of YOLO Architectures for Real-Time Patrol Vehicle Monitoring. *Sensors* 25, 6653. <https://doi.org/10.3390/s25216653>
- Murozi, A.-F.M., Ishak, S.Z., Nusa, F.N.M., Wai Hoong, A.P., Sulistyono, S., 2022. The Application of International Road Assessment Programme (iRAP) as a Road Infrastructure Risk Assessment Tool, in: 2022 IEEE 13th Control and System Graduate Research Colloquium (ICSGRC). Presented at the 2022 IEEE 13th Control and System Graduate Research Colloquium

- (ICSGRC), IEEE, Shah Alam, Malaysia, pp. 237–242.
<https://doi.org/10.1109/ICSGRC55096.2022.9845149>
- Nikolaou, D., Ziakopoulos, A., Kontaxi, A., Theofilatos, A., Yannis, G., 2025. Spatial analysis of telematics-based surrogate safety measures. *J. Safety Res.* 92, 98–108.
<https://doi.org/10.1016/j.jsr.2024.09.012>
- Paliotto, A., Meocci, M., Terrosi, A., La Torre, F., 2024. Systematic review, evaluation and comparison of different approaches for the implementation of road network safety analysis. *Heliyon* 10, e28391. <https://doi.org/10.1016/j.heliyon.2024.e28391>
- Porto, J.A., Ziakopoulos, A., Yannis, G. (2026). Weak supervision and fine-tuning with contrastive learning for multiclass lane marking segmentation [Paper presentation]. World Conference on Transportation Research, Toulouse, France.
- Saunier, N., Laureshyn, A., 2021. Surrogate Measures of Safety, in: *International Encyclopedia of Transportation: Volume 1-7*. pp. 662–667. <https://doi.org/10.1016/B978-0-08-102671-7.10197-6>
- Singh, D., Das, P., Ghosh, I., 2024. Bridging conventional and proactive approaches for road safety analytic modeling and future perspectives. *Innov. Infrastruct. Solut.* 9, 128.
<https://doi.org/10.1007/s41062-024-01426-4>
- Sunil, Sharma, A., 2021. Roadway Safety Assessment and Star Rating using iRAP along SH-11A (Jind-Kaithal) in Haryana. *IOP Conf. Ser. Earth Environ. Sci.* 889, 012036.
<https://doi.org/10.1088/1755-1315/889/1/012036>
- Tarko, A.P., 2018. Surrogate Measures of Safety, in: Lord, D., Washington, S. (Eds.), *Safe Mobility: Challenges, Methodology and Solutions*. Emerald Publishing Limited, p. 0.
<https://doi.org/10.1108/S2044-994120180000011019>
- Ziakopoulos, A., 2024. Analysis of harsh braking and harsh acceleration occurrence via explainable imbalanced machine learning using high-resolution smartphone telematics and traffic data. *Accid. Anal. Prev.* 207. <https://doi.org/10.1016/j.aap.2024.107743>
- Ziakopoulos, A., Vlahogianni, E., Antoniou, C., Yannis, G., 2022. Spatial predictions of harsh driving events using statistical and machine learning methods. *Saf. Sci.* 150.
<https://doi.org/10.1016/j.ssci.2022.105722>
- Ziakopoulos, A., Yannis, G., 2024. Key Artificial Intelligence and Digitalization Solutions Towards Vision Zero in Road Safety, in: Śladkowski, A. (Ed.), *Using Artificial Intelligence to Solve Transportation Problems*. Springer Nature Switzerland, Cham, pp. 1–26.
https://doi.org/10.1007/978-3-031-69487-5_1