Artificial Intelligence and Road Safety

George Yannis
NTUA Professor

Together with:
Stella Roussou and Apostolos Ziakopoulos

Department of Transportation Planning and Engineering
National Technical University of Athens

World Conference on Transport Research

Toulouse, 7 July 2026

Objectives

- **Provide a structured overview of AI and digitalization solutions for road safety,** mapping the data, methods, and ethical considerations that can help break the road-fatality plateau on the way to Vision Zero.
- **Survey diagnostic, proactive and data-pipeline** roles that AI plays across the road transport system.
- **Summarize machine learning paradigms** and surrogate safety measures used in road safety research.
- **Highlight** privacy, fairness, transparency and human-oversight **principles needed for safety-critical AI**.

Outline

The AI Landscape in Road Safety

Drawing from:
Ziakopoulos, A., & Yannis, G. (2024). Key artificial intelligence and digitalization solutions towards vision zero in road safety. In Using Artificial Intelligence to Solve Transportation Problems (pp. 1-26). Cham: Springer Nature Switzerland.

A Global, Persistent Crisis

1.19M
road deaths/year worldwide (2021)

22,800
fatalities/year in EU-27 (2019)

70%
of urban fatalities are vulnerable road users

- Road crashes is a major societal problem worldwide, with 1,19 million road fatalities per year and more than 50 million of road injuries.

- Road crashes are preventable, humans are fallible, so road systems must forgive error (Vision Zero Network, 2024).

- Crash numbers have quite plateaued, new tools are needed to break the curve.

Several faces of AI

- Any technology, software and supporting hardware, belonging to the family of advanced computing techniques performing tasks traditionally requiring human intelligence and discretion, or newly invented tasks enabled by computational power.

Diagnostic (Reactive) — Causal modelling of crash occurrence and injury severity from historical data

Proactive (Anticipatory) — Surrogate Safety Measures: Time-to-Collision (TTC), Post-Encroachment Time (PET), harsh events

Data Pipeline Roles — Collection, transmission, cleaning, processing, storage, and conclusion formulation

- AI can be found in Road Transport:
  - Vehicle driving & navigation
  - Traffic management
  - Route optimization
  - Public transport planning
  - Driver Monitoring
  - Infrastructure Maintenance

The Promise and Limits of Big Data

Known for the 3 V's:

- **Volume:** Massive data points exceeding traditional database capacity
- **Velocity:** High generation & processing rate, enabling real-time analysis
- **Variety:** Structured, unstructured, semi-structured, JSON, XML, CSV and more

Indicative sources:

- Smartphone telematics (OSeven, ZenDrive)
- GPS & cellular signal data
- On-board diagnostics
- Roadside & dashcams
- OpenStreetMap
- Shared mobility (Uber, bike-share)
- Weather & census data

- Utility is bounded by:
  - the insight users can extract, and their domain expertise,
  - not by data volume alone.
- Some researchers argue big data's potential is overhyped and overstated

Necessary Caution

Representativeness : smartphone/app-based samples skew demographic and trip-type distributions.

Privacy & consent : GDPR compliance is mandatory, not optional, for telematics and video data.

Noise and missingness : raw sensor streams require extensive cleaning before any modelling begins.

Three Key Pillars of AI in Road Safety

01

Road User Behaviour

Modelling individual driving patterns:

- harsh events,
- speeding,
- distraction from naturalistic experiments
- telematics

02

Infrastructure Performance

Aggregating behaviour and infrastructure data across road networks to identify systemic risk

03

In-Depth Crash Analysis

Causal modelling:

- crash occurrence
- probability and
- injury severity factors

Machine Learning in Road Safety

- ML is the most visible pillar of AI in road safety, fast-growing, used even in industrial applications like object detection for automated vehicles.

- Algorithms and statistical models that learn from data, rather than being explicitly programmed for each task.

- Learns patterns, makes decisions, forecasts outcomes, must guard against bias and overfitting.

- Three learning paradigms:
  - Supervised -> labeled data, classification or regression
  - Semi-supervised -> small labeled + large unlabeled set
  - Unsupervised -> no labels, finds patterns (clustering, dimensionality reduction)

SSMs and Model Methods

- Surrogate Safety of Measures can act as **Proactive Indicators**, helping us going beyond crash counts:
  - Time-to-Collision (TTC)
  - Post-Encroachment Time (PET)
  - Harsh Events (acceleration, braking, cornering)

Classical Machine Learning

Logistic Regression

Decision Trees

Random Forest

SVM

k-NN

Deep Learning

Artificial Neural Networks

Convolutional Neural Networks (vision)

Long-Short Term Memory (sequences)

Graph & Self-Supervised

Graph Neural Networks

Contrastive Learning

GANs

Vision-Language Models

Classical model families to state-of-art

Ethical AI for Road Safety

- **Privacy & Consent:** Telematics and video data require GDPR-compliant, anonymized handling at scale.
- **Algorithmic Fairness:** Models trained on biased samples can systematically under-serve vulnerable road users.
- **Transparency:** Explainable AI (e.g. SHAP values) needed, so safety decisions can be audited and trusted.
- **Human Oversight:** AI augments, not replaces, expert judgement in safety-critical interventions is necessary.

Promising Future Directions

- **High-Impact Feature Engineering:** Designing richer, more predictive variables from raw sensor and map data.
- **Crash & Injury Causality Analysis:** Moving from correlation to causal inference for targeted interventions.
- **Ethical AI Applications:** Building fairness, privacy and transparency into deployment from day one.
- Multilevel and more frequent **dialogue channels** must be established between industrial and technical data holders and policymakers for **standardization and openness of data.**

A Dual Graph Framework for Road Safety Analysis

Paradiso, S., Ziakopoulos, A. & Yannis, G. (2026). “A Dual Graph Framework for Edge Embedding and Clustering in Road Safety Analysis”. Proceedings of the 11th Transport Research Arena TRA 2026, May 18-21, 2026, Budapest, Hungary.

Scope

- **Develop a graph neural network framework** that generates edge embeddings encoding road-network structure and driving behaviour, enabling data-driven clustering of segments by risk profile.
- Combine telematics data with OpenStreetMap topology into a single road-network graph, so **we can have a network representation**.
- Apply Graph Attention Networks (GAT) to **encode structural and behavioural information** for every edge.
- Use K-Means directly on the numerical embeddings to **partition the network into distinct risk groups**.

Why Graph Neural Networks?

- A road network is naturally a graph but road safety modelling usually ignores that structure, treating segments as independent rows in a table.

Graph Neural Networks	Extend neural nets to graph-structured data, generating node embeddings that capture structure + neighbourhood context
The Dual-Graph Trick	Road safety analysis needs edge-level insight (segments), not just nodes (intersections), so edges become nodes.
Self-Supervised Training	A Graph Attention Network (GAT) learns via contrastive loss, no manual labels required.

Athens, Mapped and Measured

➤ Sources combined

1. **OSeven Telematics ->** Smartphone sensor data: speed, harsh acceleration/braking, speeding, mobile use, GDPR compliant.
2. **OpenStreetMap (OSM) ->** Graph extraction: road geometry, type, connectivity, matched to nearest edge per trip.

13,000+ anonymized trips	**1 Hz** sampling rate
4 months of 2024 telematics data	**43,378** road edges analyzed

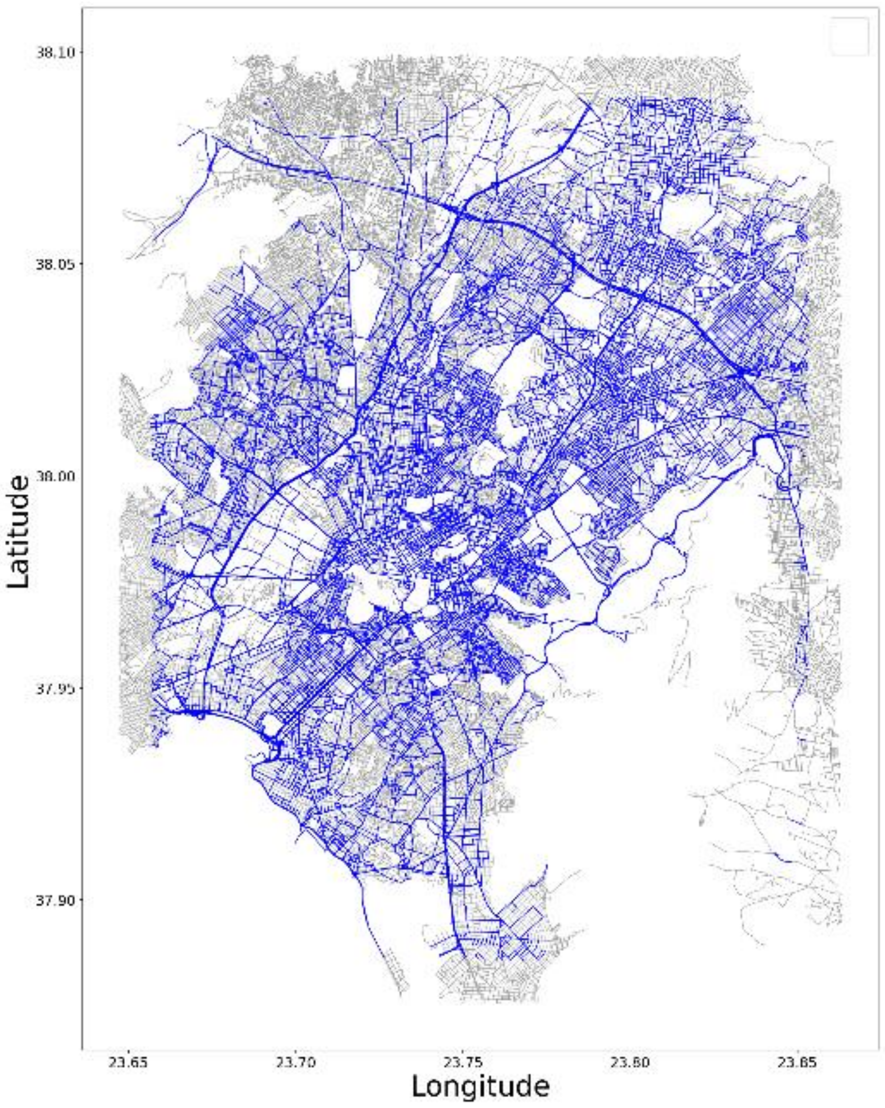


Figure: Road Network with Telematics-Characterized Nodes and Edges

From Raw Edges to Embeddings

1. **Build dual graph**: each road edge becomes a node; shared-node edges become connections.
2. **Train a Graph Attention Network** (2 GAT layers, 3-head attention) with self-supervised contrastive loss.
3. Generate **numeric embeddings for every edge**, encoding both its features and neighbourhood context.
4. **Cluster embeddings with K-Means**, compared against K-Prototypes clustering on raw features.

GAT update rule: $hi\ ' = f\emptyset(hi, AGGREGATE(\{hj \mid j \in \mathcal{N}i\}))$

GNN Embeddings Win Decisively

0.92

Peak silhouette score (K=2) with GNN embeddings -> a clear, well-defined partition

2

Distinct road-network clusters discovered: **high-speeding rural roads** vs. moderate urban roads

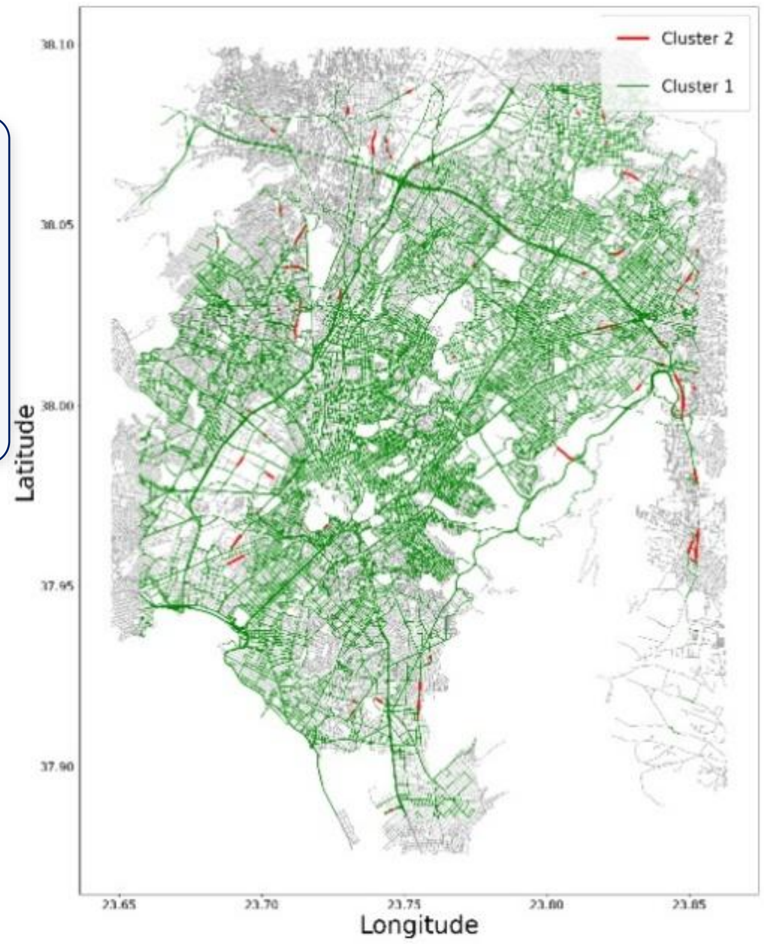


Figure: Road Network Partitioning Based on GNN-**Generated Edge Embeddings**

The discussed novel approach:

- improves clustering quality
- yields more accurate and meaningful network partitions
- enables the use of K-Means by working exclusively with numerical data
- doesn't rely on mixed-type clustering methods like K-Prototypes
- leverages the well-established framework of node embedding tasks to generate edge embeddings.

Getting There Wasn't Easy

Edges aren't points

Most GNN research targets nodes, representing road segments required the dual-graph workaround itself

No crash ground truth

Clusters reflect behavioural risk proxies, not validated crash outcomes, a correlation, not yet a proof

Interpretability loss

Embeddings are abstract vectors, cluster meaning had to be reverse-engineered by mapping back to raw features

Spatial confounds

Urban vs. rural and one-way vs. bi-directional roads were not yet separated, potentially blending distinct risk patterns

Can Your Phone Detect a Dangerous Corner?

Tsoutsanis, A., Ziakopoulos, A. & Yannis, G. (2026). “Self-Supervised Detection of Harsh Cornering Events at Scale using Smartphone Sensor Data”. Proceedings of the 11th Transport Research Arena TRA 2026, May 18-21, 2026, Budapest, Hungary.

Scope

- **Design a self-supervised framework** that detects harsh cornering events directly from smartphone sensor data, without on manually labelled ground truth.
- **Feature extraction:** Derive orientation-invariant motion features from GPS, accelerometer and gyroscope streams.
- **Turn validation:** Cross-check candidate turns against OpenStreetMap intersection geometry.
- **Classifier benchmarking:** Train and compare multiple classifiers to separate harsh from normal cornering in latent space.

Learning Without Labels

- Harsh cornering, linked to nearly a third of passenger fatalities via rollover, remains underexplored, largely because nobody has ground-truth labels for it.

Self-Supervised Learning	No manual annotation. The model generates its own training signal from the structure of the data itself
Anomaly Detection as Labels	An ensemble of anomaly detectors flags statistical outliers as 'pseudo-labels' for harsh events
Orientation Invariance	Phones sit anywhere in a car, so features must be mathematically blind to device orientation

Data Collection and Distribution of Turn Events

Sensors Used

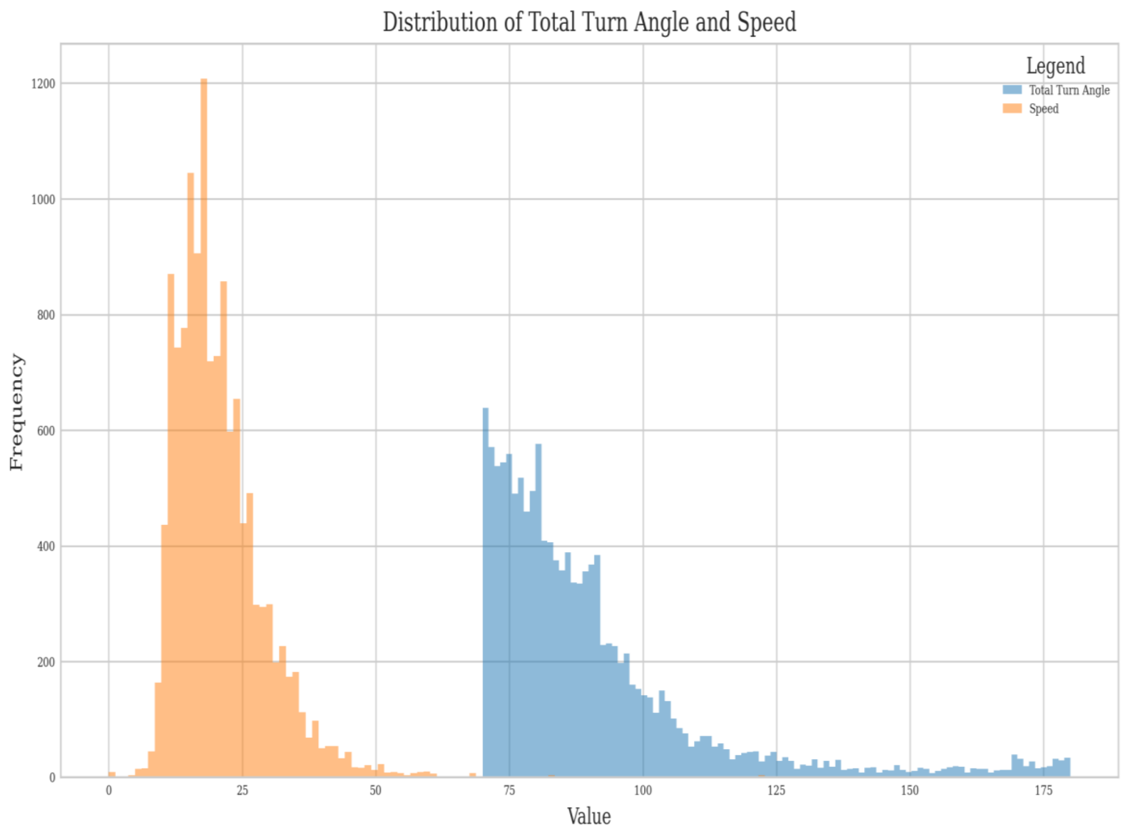
- **GPS:** Location, speed, heading, used to derive yaw rate and turn angle over 1–4s horizons
- **Accelerometer + Gyroscope:** 3-axis readings, converted to orientation-invariant magnitude via Euclidean norm
- **OpenStreetMap:** Validates candidate turns: must occur within 15m of a real intersection node

4,017 trips

1,900 validated turns

1.76M data points

1 Hz sampling rate



Distribution of Turn Events

- The histogram shows the distribution of total turn angles and vehicle speeds across all turning events in the dataset
- Predominantly involving 70°–180° turns at speeds below 30 km/h

From Raw Sensors to Classifier

1. **Detect**
   Flag candidate turns where heading change > 60° in a 4-second window

2. **Validate**
   Confirm turns sit within 15m of a real OSM intersection, filters out lane changes & noise

3. **Seed Labels**
   Ensemble anomaly detection (Isolation Forest + LOF + One-Class SVM); top 5% = positive seeds

4. **Classify**
   Train 4 supervised models on pseudo-labels, Logistic Regression, SVM, Random Forest, MLP

GAT update rule: $hi' = f\emptyset(hi, AGGREGATE(\{hj \mid j \in \mathcal{N}i\}))$

Strong Separation, No Labels Needed

0.95

SVM achieves the best ROC-AUC, all 4 models exceed 0.87

0.76

MLP latent-space silhouette score, harsh vs. normal corners cleanly separate

- Self-supervised framework detects **harsh cornering without manual annotation**
- Orientation-invariant features and OSM validation
- All classifiers **AUC >0.87**
- Applications: **insurance telematics**, fleet management, urban safety planning

The Obstacles

No ground truth, anywhere

Manual labelling of cornering events is prohibitively expensive at this scale, forcing reliance on unsupervised proxies

Arbitrary phone placement

Devices sit in cup holders, pockets, dashboards — raw x/y/z axes are meaningless without orientation correction

Pseudo-labels are noisy

Anomaly-detector seeds are a best guess, not ground truth — errors propagate into downstream classifiers

Validation still pending

Cross-region testing and OBD-based validation remain future work before real-world deployment

Lane Segmentation from Street-Level Imagery via Noisy Label Generation and Contrastive Self-Supervision

Porto, J. A., Ziakopoulos, A., Lopez, D.F. & Yannis, G. (2026). “Lane Segmentation from Street-Level Imagery via Noisy Label Generation and Contrastive Self-Supervision”. Proceedings of the 11th Transport Research Arena TRA 2026, May 18-21, 2026, Budapest, Hungary.

Scope

- Test how far lane detection can be pushed without costly, pixel-perfect manual annotation, comparing noisy-label supervision against contrastive self-supervised pretraining.
- Automatically generate training labels for real-world street-level imagery.
- Train a LinkNet model on the noisy labels using a ResNet backbone.
- Self-supervised alternative: Pre-train with BYOL contrastive learning and validate generalisation on the TuSimple benchmark.

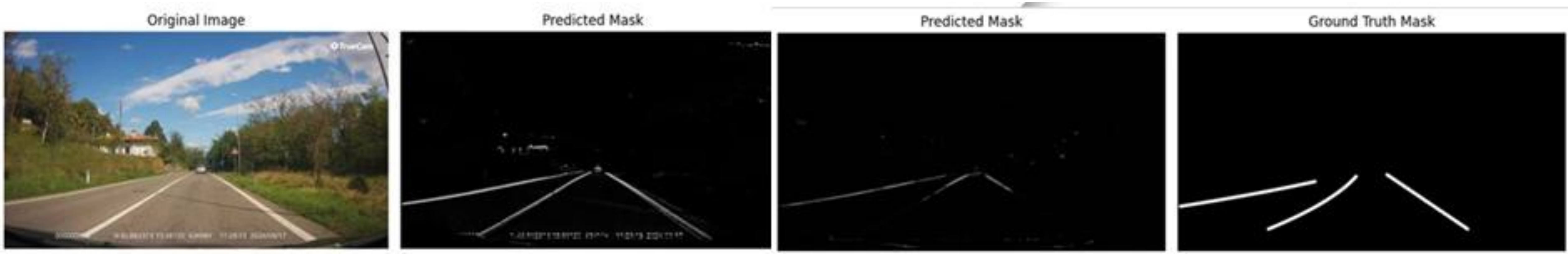
Teaching a Model to See Lanes

- Lane detection underpins automated navigation and behaviour monitoring.
- But pixel-perfect manual annotation of lane markings is slow and expensive.
- So how far can we get without it?

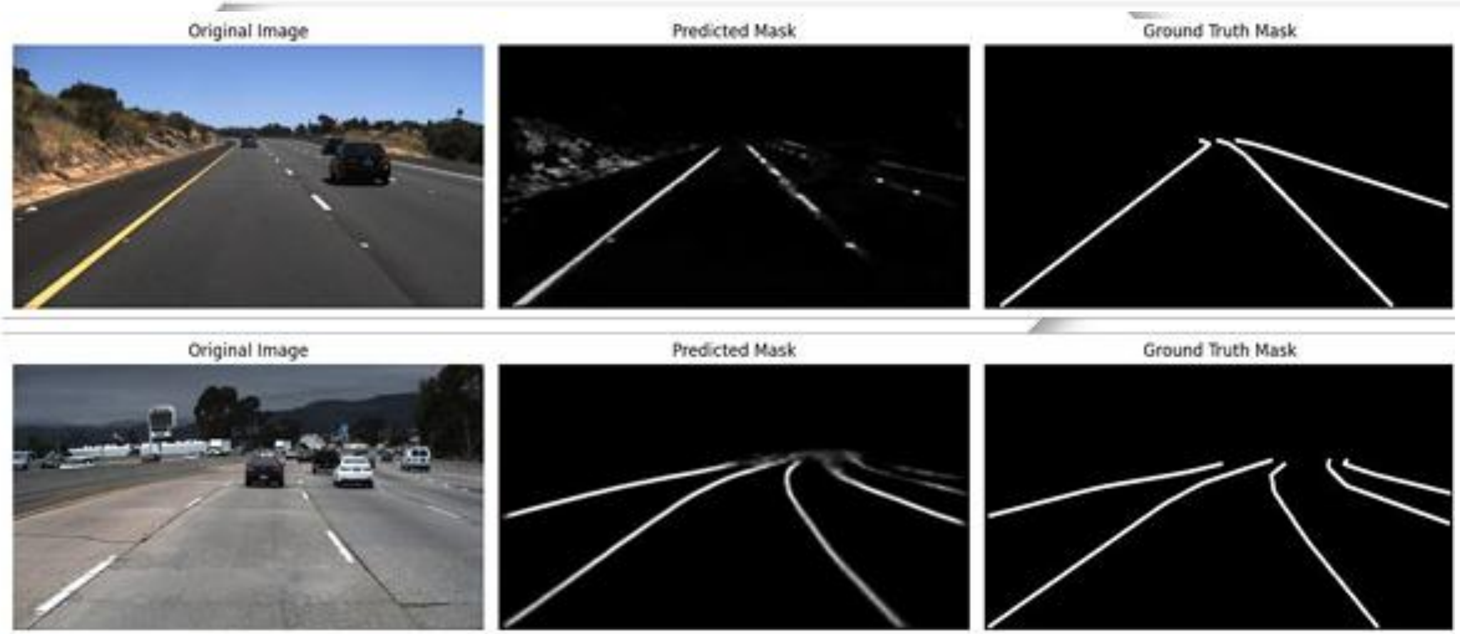
Noisy Label Generation	Colour thresholding, edge detection, and DBSCAN clustering auto-generate approximate lane labels
Weak Supervision	Training proceeds on these imperfect, automatically-generated labels rather than hand-annotated ground truth
Contrastive Self-Supervision	BYOL learns from two augmented views of the same image, no negative examples required

Two Datasets, Two Worlds

- **Proprietary Dataset** : Real-world street-level imagery with labels generated via the noisy-label pipeline, reflects messier, in-the-wild conditions.



- **TuSimple Dataset**:



Public benchmark dataset for lane detection, used to validate methods against a well-established standard.

Two Training Regimes, Compared

- **Label generation pipeline:**

Colour thresholding → Edge & contour detection → DBSCAN clustering → Polyline fitting → Lane merging

Fully Supervised

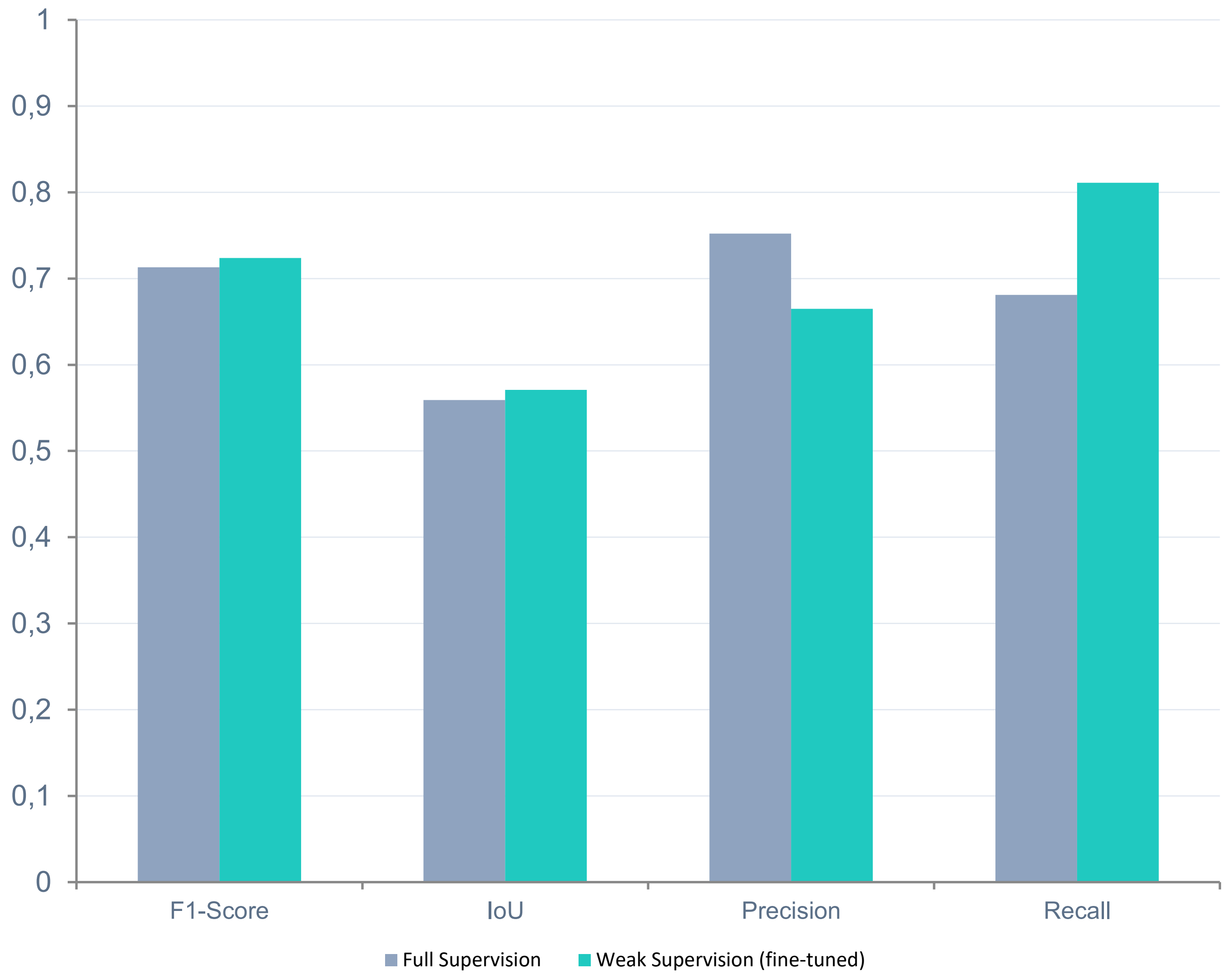
- LinkNet model, ResNet backbone.
- Trained on noisy generated labels.
- Binary Cross-Entropy loss, Adam optimizer.
- High precision, lower recall

Weak + Self-Supervised

- BYOL contrastive pre-training.
- No negative examples needed.
- Loss: 2 − 2×cosine similarity.
- Better generalization, prevents overfitting.

A Precision-Recall Trade-off

- Weak supervision recovers more true lane pixels (higher recall) but is less precise.
- A useful trade-off where missing a lane matters more than a false positive.



TuSimple dataset, 50-epoch comparison at IoU threshold 0.3 (best-performing threshold)

Getting There Wasn't Easy

Noisy labels are still noisy

Even with careful geometric heuristics, auto-generated labels introduce systematic errors the model must absorb

Ambiguous convergence

The weak-supervision branch converged to a higher loss — unclear if the model learned different features or simply hadn't finished training

Small proprietary dataset

Limited training data constrained how far the label generator and segmentation model could be pushed

Reviewer-driven rework

Required re-running at reduced resolution, new early-stopping criteria, and transfer-learning tests across both datasets

Comparing Machine Learning Techniques for Predictions of Motorway Segment Crash Risk Level

Nikolaou, D., Ziakopoulos, A., Dragomanovits, A., Roussou, J. & Yannis, G. (2023). "Comparing machine learning techniques for predictions of motorway segment crash risk level." Safety, 9(2), 32

Scope

- Identify which of five common machine learning techniques most accurately classifies crash risk level on motorway segments, and test whether its logic can be trusted.
- Have a Model comparison of Benchmark Logistic Regression, Decision Tree, Random Forest, SVM and K-Nearest Neighbours on the same dataset.
- Combine road geometry with naturalistic driving-behaviour metrics per segment.
- Apply SHAP values to confirm the best-performing model's logic can support proactive screening of hazardous segments.

Which Model Should You Trust?

- Motorways are statistically the safest roads, but when crashes happen there, severity is high.
- This study asks a deceptively simple question: of five common ML techniques, which best classifies a segment's crash risk?

Head-to-Head Benchmarking	Logistic Regression, Decision Tree, Random Forest, SVM, and k-Nearest Neighbours , trained and tested on the same data.
Classification, Not Regression	The target is a discrete crash risk level per segment, not a continuous number, a categorization task.
Explainability via SHAP	Shapley Additive Explanations quantify which features actually drove each prediction, opening the 'black box'.

Data Collection

668 motorway segments

75% / 25% train / test split

5 ML techniques compared

Feature categories used:

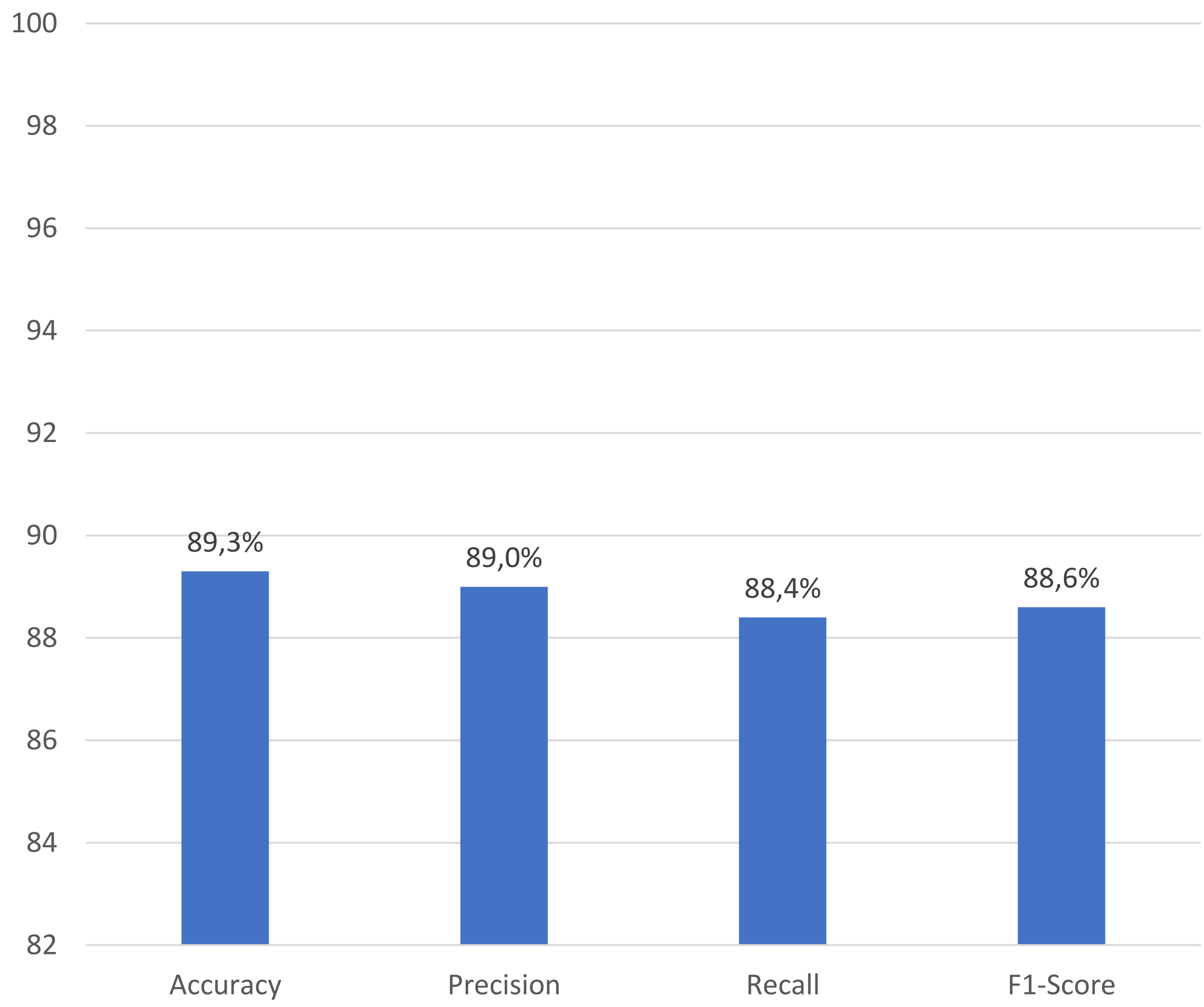
- **Road Geometry:** Lane width, curvature, gradient, and other static infrastructure characteristics.
- **Driving Behaviour Metrics:** Naturalistic driving data, speed profiles, harsh events, aggregated per segment
- **Crash Risk Level (target):** The response variable: a categorical risk classification per motorway segment.

Five Contenders, One Test Set

- Logistic Regression
- Decision Tree
- Random Forest ★ **Best performer**
- Support Vector Machine
- K-Nearest Neighbours

Random Forest Takes the Lead

- **Evaluation metrics:** overall accuracy, macro-averaged precision, recall, and F1-score, with SHAP values layered on top for interpretability,
- **89.3% overall classification** accuracy, best among all five techniques tested,
- **SHAP analysis** confirmed the model's logic was interpretable, supporting its use as a proactive screening tool for hazardous segments.



The barriers

Small sample for ML

668 segments is modest by deep-learning standards, favoring simpler, more data-efficient models like Random Forest

Class imbalance risk

Truly hazardous motorway segments are rare by definition, risking models that default to predicting 'low risk'

Black-box trade-off

The best-performing model (Random Forest) is inherently less transparent than Logistic Regression, requiring SHAP as a fix

Static + dynamic data fusion

Combining fixed road geometry with time-varying driving behaviour metrics required careful feature engineering

Using computer vision and street-level videos for pedestrian-vehicle tracking and behaviour analysis

Ventura, R., Roussou, S., Ziakopoulos, A., Barabino, B. & Yannis, G. (2025). “Using Computer Vision and Street-Level Videos for Pedestrian-Vehicle Tracking and Behaviour Analysis”. Transportation Research Interdisciplinary Perspectives, 30, 101366.

Scope

- **Build and test a low-cost computer vision pipeline** that extracts pedestrian and vehicle behaviour metrics from ordinary street-level video.
- **Identify pedestrians, vehicles and traffic-light status** from smartphone-recorded intersection footage.
- **Derive surrogate safety measures** such as illegal crossings and near-miss interactions.
- Assess real-time, **low-cost monitoring** for intersection design and traffic-safety planning.

The AI concept

- ➢ Surrogate Safety Measures (SSMs) let researchers study near-misses, not only crashes, but computing them from raw video requires three capabilities working together.

Real-time object detection	YOLOv8 detects pedestrians & vehicles frame-by-frame from ordinary roadside video
Re-identification & tracking	ResNet-50 features + Kalman filtering keep track of the same person/vehicle across frames
Geometric grounding	Homography transforms pixel coordinates into real-world positions, enabling TTC computation

From Street Video to Behaviour Metrics

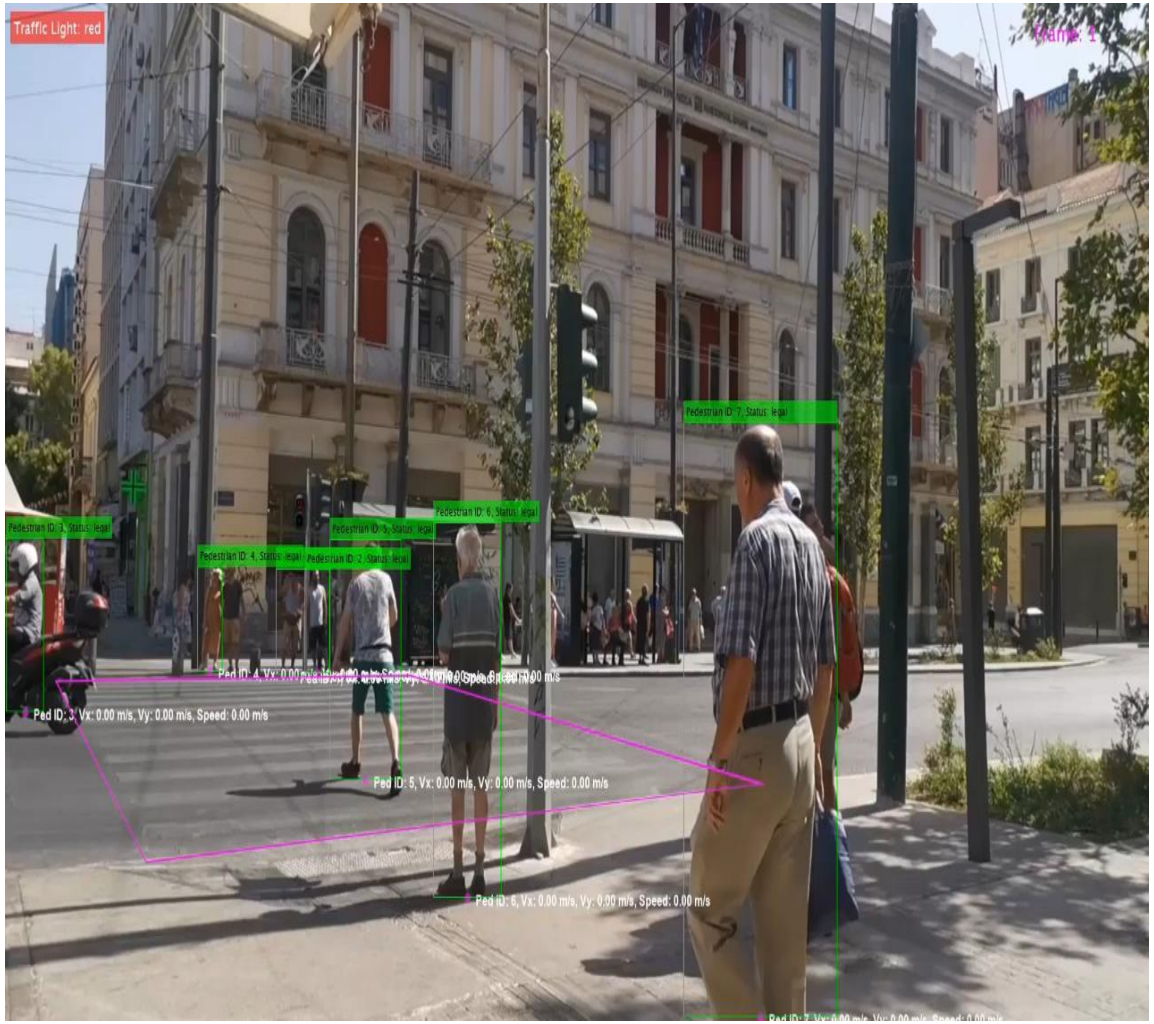
- A single smartphone camera, aimed at 8 busy intersections in Athens and it was enough to feed the entire pipeline below.

1. Detect pedestrians and vehicles per frame using YOLOv8 object detection
2. Track identities across frames with Kalman filtering and ResNet-50 re-identification
3. Apply homography transformation to convert pixel positions to real-world coordinates
4. Compute TTC and PET cross-reference behaviour with traffic light status

- Data source: smartphone-recorded roadside video from intersections in Athens, Greece, chosen for low-cost, easily transferable data collection.

Promising, But Imperfect Vision

- Approximately **70% accuracy detecting traffic light status from video**, which is workable but inconsistent under real-world lighting and occlusion.
- 23% average discrepancy between manual and automated counts of illegal pedestrian crossings.
- Despite imperfect accuracy, the **framework demonstrates real-time, low-cost monitoring** is feasible, providing actionable insight for intersection design and traffic safety planning.



Facing Obstacles

Occlusion & clutter

Dense urban intersections mean pedestrians and vehicles frequently block each other from camera view

Traffic light ambiguity

Detecting signal status from variable camera angles and lighting conditions caps accuracy at 50–70%

Homography drift

Converting 2D pixels to real-world coordinates accumulates geometric error, especially at frame edges

Manual vs. automated gap

A 23% discrepancy with human counts shows automated detection still needs human-in-the-loop validation

The common threads

Five Concepts, One Common Thread

- **The values of the variables (labels) are the bottleneck**
  Self-supervised and weakly-supervised methods recur across all analyses, manual annotation simply doesn't scale
- **Interpretability is non-negotiable**
  SHAP values, cluster profiling, and latent-space analysis are needed everywhere in order that AI meets safety-critical decisions
- **Real-world data is messy by default**
  Orientation-invariant sensors, noisy labels, occlusion and every project must be engineered around imperfect & in-the-wild data
- **Validation remains the frontier**
  More regions, more ground truth, more real-world testing are needed

AI won't solve road safety alone, but paired with data, transparency, and human oversight, it moves us measurably closer to Vision Zero.

Artificial Intelligence and Road Safety

George Yannis
NTUA Professor

Together with:
Stella Roussou and Apostolos Ziakopoulos

Department of Transportation Planning and Engineering
National Technical University of Athens

World Conference on Transport Research

Toulouse, 7 July 2026