Predicting pedestrian illegal crossing using vehicle dynamics with LightGBM

George Yannis
Professor

Together with:
Stella Roussou and Apostolos Ziakopoulos

Department of Transportation Planning and Engineering
National Technical University of Athens

IRCOBI
International Research Council on Biomechanics of Injury
Vilnius, Lithuania, 10–12 September 2025

Background

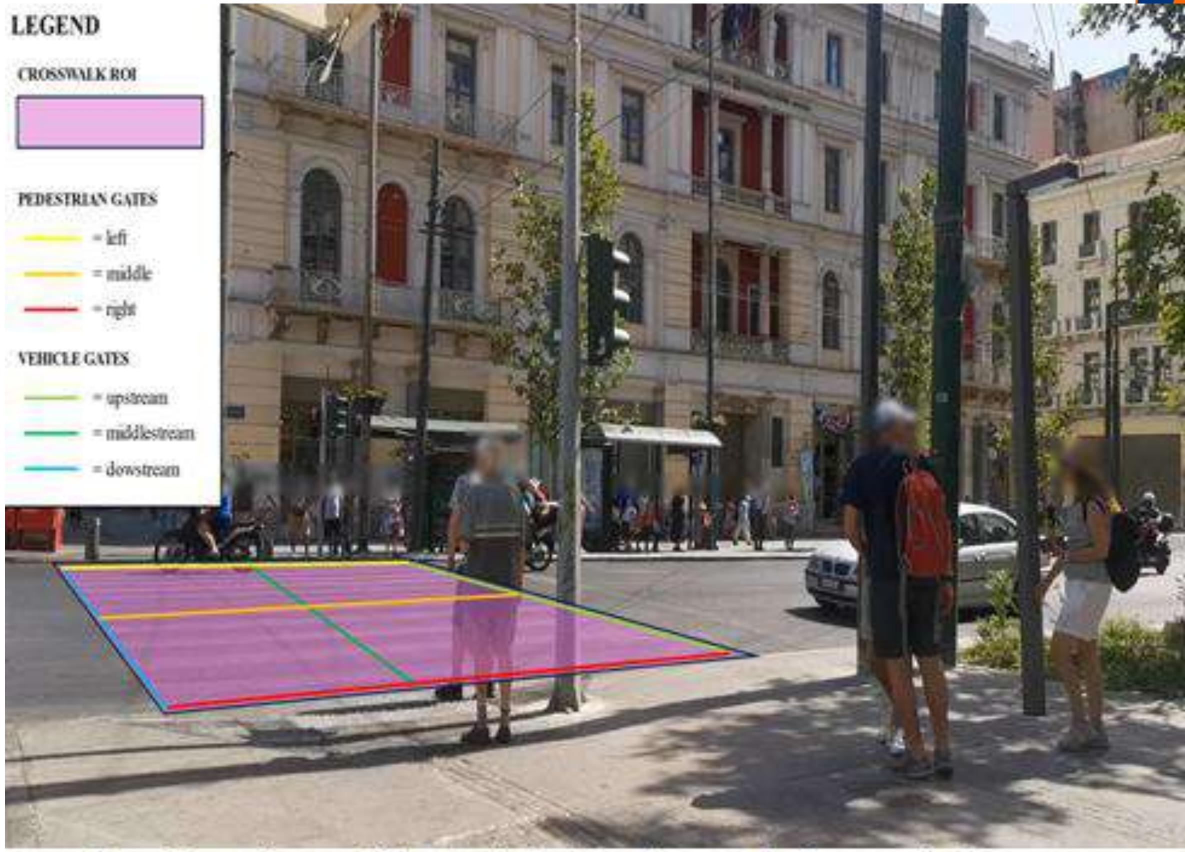
- Infrastructure, environmental, and driver-related features are determinants of pedestrian crash patterns, which can be uncovered through machine learning and econometric modelling.
- A dynamic but often hazardous urban environment is created due to the co-existence of vehicles, pedestrians, and vulnerable road users (cyclists, scooters, children, etc.).
- In Athens, Greece, there is high pedestrian activity combined with limited traffic monitoring infrastructure.
- Illegal crossings at signalized intersections increase the risk of crashes.
- To confront this issue, computer vision and video recognition technologies are providing tools to monitor and analyze traffic, becoming fundamental road safety techniques.
- The present research was carried out within the research project "PHOEBE - Predictive Approaches for Safer Urban Environment".

Objectives

- Real-Time Detection & Tracking with YOLOv8 + ResNet-50:
  - Accurate, real-time detection of pedestrians and vehicles.
  - Enables traffic monitoring even in cities lacking integrated surveillance infrastructure, (e.g. Athens)
- Develop a LightGBM classifier to predict pedestrian illegal crossings using only vehicle-dynamics.
- Explore potential for lightweight, scalable solutions deployable in:
  - Smart infrastructure
  - On-board vehicle systems, integrating AI with object detection and predictive analytics.

Data Collection

- Video footage captured in Panepistimiou Str., Athens: a critical urban corridor connecting Syntagma and Omonoia Squares.
- Smartphone cameras with tripods are used to overcome the lack of fixed infrastructure.
- Main features concerned vehicle position (x,y), velocity (x,y), speed magnitude, and direction.
- The algorithm checks their positions relative to Regions of Interest (ROI), which define critical areas, i.e. crosswalks or traffic light positions.
- If a pedestrian is found inside the crosswalk during a pedestrian red traffic light, their behavior is flagged as illegal.

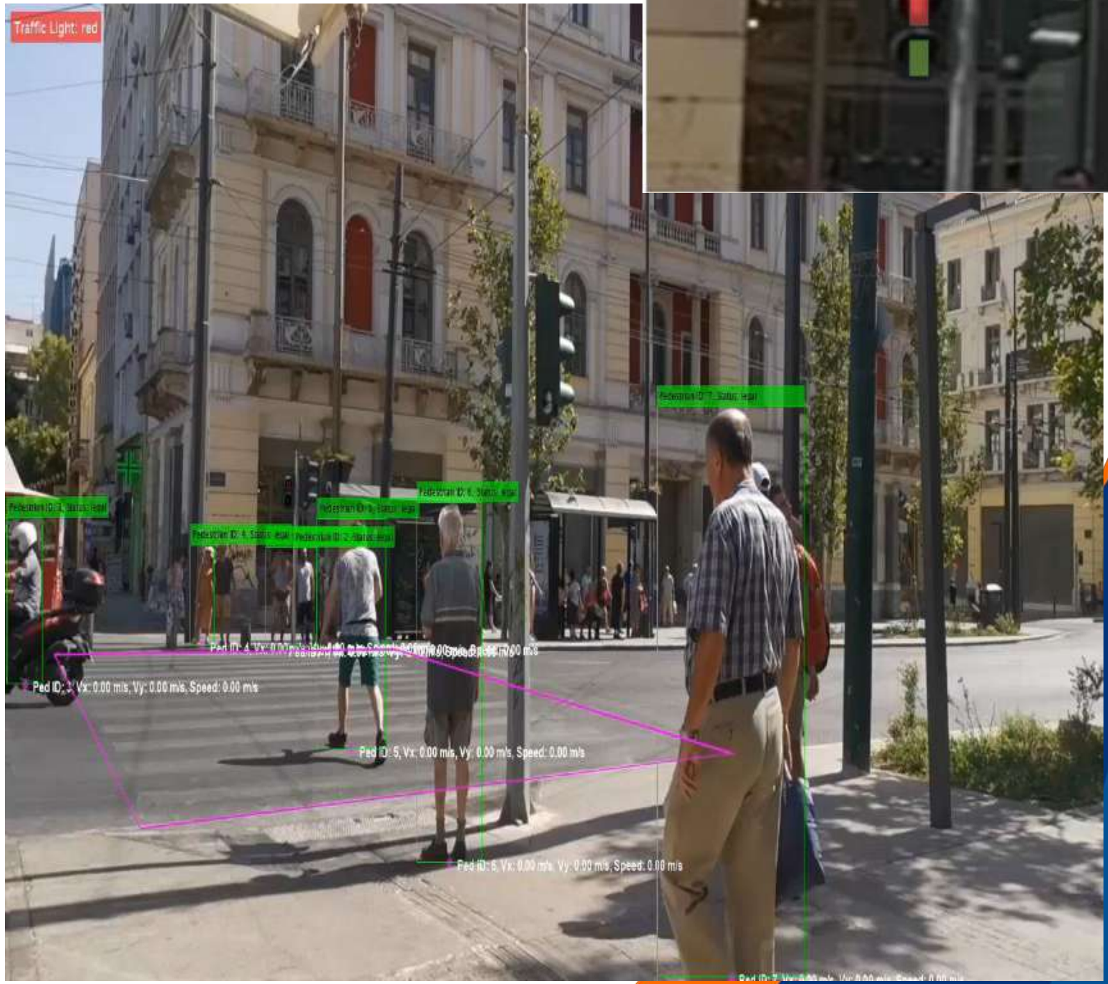


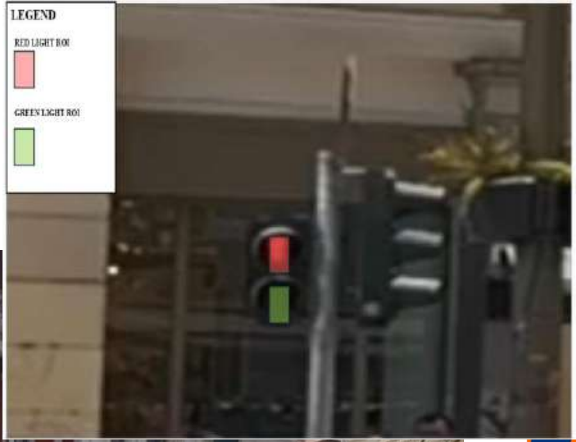
Tracking Area of Panepistimiou Street & Omonoia Square

Video Detection Framework

The system uses a multi-step tracking logic:

- **Object detection YOLOv8** detector, a neural network, detecting objects including pedestrians and vehicles.
- **ResNet-50 and Re-Identification** network to maintain consistent tracking of objects across multiple frames.
- The **Hungarian Algorithm** for matching the detected objects to the tracked ones (crucial for maintaining consistent identities).
- **Kalman filter-based prediction**, tracking continuous movement, as objects can become occluded or temporarily disappear from the scene.
- The **traffic light detection** component isolates the regions corresponding to the traffic lights using Regions of Interest (ROIs) and applies **color segmentation** to determine whether the light is red or green.





Modeling Approach

- LightGBM classifier, which is a gradient boosting framework based on decision trees, known for its efficiency in classification tasks.
- It was trained on pedestrian behavior data to detect illegal crossings.
- Features: pedestrian/vehicle coordinates and speeds, traffic light status, and TTC.
- Class imbalance (66% legal / 34% illegal) → balanced with SMOTE.
- Train/test split: 80/20
- Metrics: accuracy, precision, F1-score, and recall.

Results – Omonoia Square

- Accuracy achieved was at 82%.
- The precision for the illegal class was 0.76.
- Key predictor: Vehicle speed (absolute velocity is the most important).
- Precision–Recall curve demonstrates robust performance despite class imbalance. Trade-off between detecting more illegal crossings (recall) and avoiding false alarms (precision).

	Precision	Recall	F1-Score	Support
0 (Legal)	0.84	0.89	0.86	3705
1 (Illegal)	0.76	0.66	0.71	1910
Accuracy		0.82		5615
Macro avg	0.80	0.78	0.79	5615
Weighted avg	0.81	0.82	0.81	5615

Table 1: *Classification Report for LightGMB Classifier*

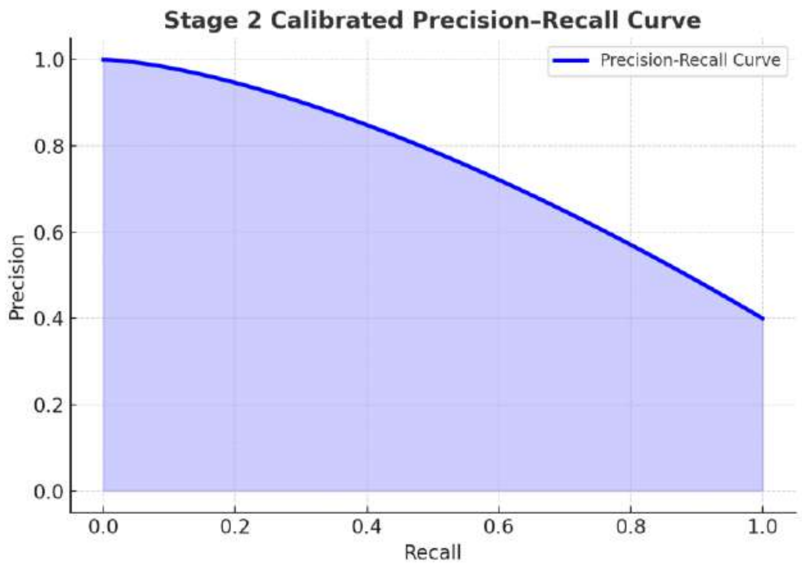


Figure 1: *Precision-Recall Curve*

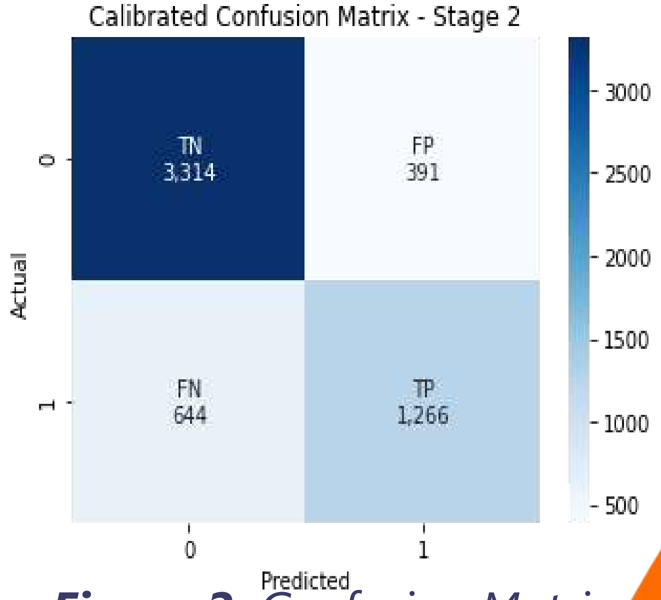


Figure 2: *Confusion Matrix*

Discussion

- The environment of Omonia Square allowed LightGBM classifier to perform well, achieving 82% accuracy.
- It confirms the potential of vehicle-based prediction of pedestrian behavior.
- It showcases the importance of vehicle data, meaning no absolute need for pedestrian data, which is difficult to find and based mostly on unmanned aerial devices.
- The results suggest the scalability of the study to smart traffic systems.
- One of the limitations of the current study is the lack of a larger number of intersections and their sensitivity to the temporal alignment.

Conclusions & Future Work

- Vehicle dynamics alone can predict illegal crossings with promising accuracy.
- Useful for integration into urban safety systems and contribution
  to intelligent crosswalk monitoring systems.
- Findings support the use of machine learning for monitoring pedestrian violations.
- Future research directions:
    - Apply to more intersections and more diverse datasets.
    - Incorporate contextual data (weather, road layout, and pedestrian demographics).
    - Enhance trajectory fusion techniques.
    - Include surrogate safety measures (time-to-collision, post encroachment time)

Predicting pedestrian illegal crossing using vehicle dynamics with LightGBM

George Yannis
Professor

Together with:
Stella Roussou and Apostolos Ziakopoulos

Department of Transportation Planning and Engineering
National Technical University of Athens

IRCOBI
International Research Council on Biomechanics of Injury
Vilnius, Lithuania, 10–12 September 2025