Network Performance in an Urban Context

Predicting Pedestrian Violations Using Object Detection and Deep Learning: A Comparative Study of LSTM & GRU Models

Host – Symposium Sponsor

International Partners

National Supporters

HELLENIC REPUBLIC
MINISTRY OF INFRASTRUCTURE AND TRANSPORT

HELLENIC REPUBLIC
Ministry of Digital Governance

Predicting Pedestrian Violations Using Object Detection and Deep Learning: A Comparative Study of LSTM & GRU Models

Stella Roussou

Ph.D Candidate, Research Associate

Together with:

Apostolos Ziakopoulos, Roberto Ventura, and George Yannis

Urban Road Safety

- **Human behavior**, road characteristics (design and condition of roads), vehicle safety standards, environmental factors, and socioeconomic differences are some of the complex factors that contribute to road incidents.
- 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 and methods to monitor and analyze traffic, and in this way, becoming fundamental techniques for road safety.
- The present research was carried out within the research project "**PHOEBE - Predictive Approaches for Safer Urban Environment**".

Study 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**, like the city of Athens.

- **Predicting Pedestrian Violations using Deep Learning Models (LSTM & GRU):**
  - Capture sequential, time-dependent behavior.
  - **Forecast illegal pedestrian crossings** during traffic light phases.

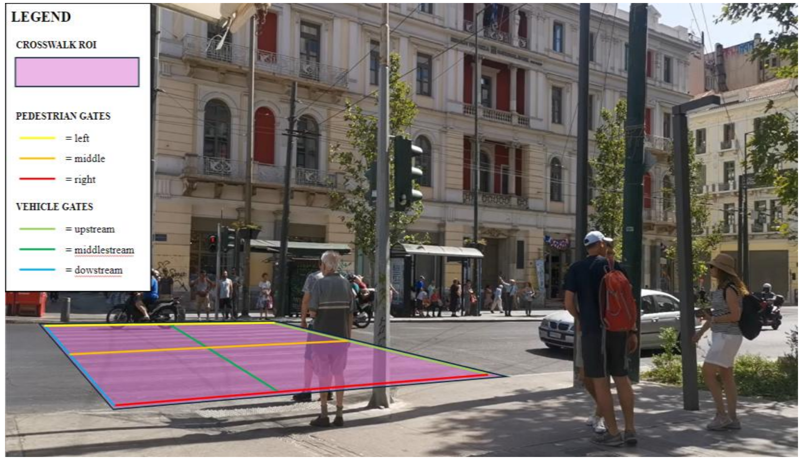
- **Study Objectives**
  - Evaluate and compare LSTM and GRU performance for pedestrian violation prediction.
  - **Enhance urban traffic safety** by integrating AI with object detection and predictive analytics.

Data Collection and Algorithmic Logic

- Video footage captured manually at the end of **Panepistimiou Street**, Athens - a critical urban corridor connecting Syntagma and Omonoia Squares.
- Smartphone cameras with tripods are used to overcome the **lack of fixed infrastructure.**

The system uses a **multi-step tracking logic including**:

- Object detection **YOLOv8 detector**, which is a neural network-based object detection framework, detecting objects including pedestrians and vehicles.
- **ResNet-50 and Re-Identification** network in order to maintain consistent tracking of objects across multiple frames.
- The **Hungarian Algorithm** for matching the detected objects to the tracked ones, which is crucial for maintaining consistent identities.
- **Kalman filter-based prediction**, which tracks continuous movement, as objects can become occluded or temporarily disappear from the scene (e.g., a pedestrian walks behind a vehicle).
- 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.



LSTM and GRU Model Results

➢Why LSTM and GRU?

- Recurrent Neural Networks (RNNs) designed to capture sequential, time-dependent behavior.
- Ideal for modeling dynamic pedestrian actions at intersections.

➢LSTM (Long Short-Term Memory):

- Two bidirectional layers (64 and 32 units).
- ReLU activations, dropout 0.5 to prevent overfitting.
- Optimized with Adam, learning rate: 0.0002
- Captures long-term dependencies in pedestrian behavior.

Confusion Matrix (LSTM Results)		
	Positive	Negative
Positive	449	161
Negative	502	57

Classification Metrics (LSTM Results)	
Accuracy	0.812
Precision	0.736
Recall	0.887
F1-Score	0.805

➢GRU (Gated Recurrent Unit):

- Two bidirectional GRU layers, optimized for fast learning.
- ReLU activation, dropout for regularization
- Simpler structure, fewer parameters, reduced computational cost.
- Strong recall for detecting illegal crossings.

Confusion Matrix (GRU Results)		
	Positive	Negative
Positive	471	278
Negative	384	36

Classification Metrics (GRU Results)	
Accuracy	0.731
Precision	0.629
Recall	0.929
F1-Score	0.749

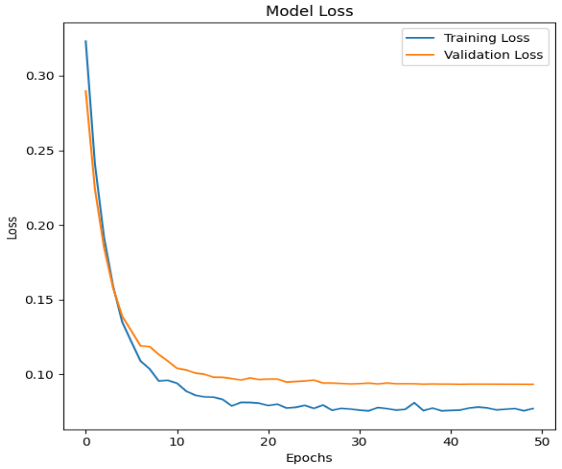
Deep Learning for Predicting Pedestrian Violations

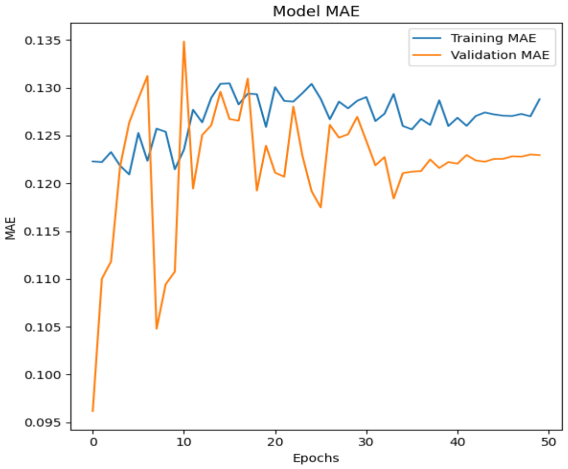
➢LSTM (Long Short-Term Memory):

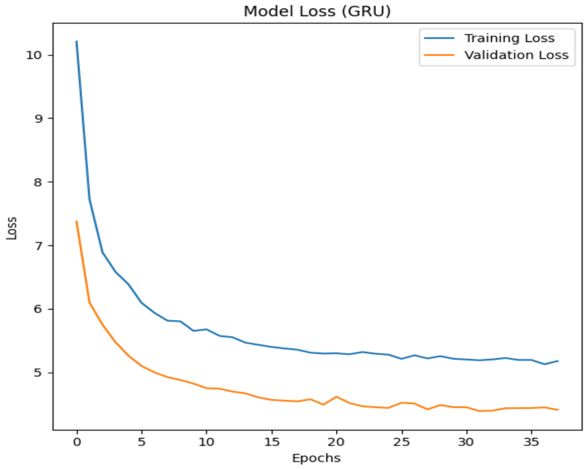
- Both the training and validation losses decrease over time, while accuracy improves.
- The model is learning effectively from the data **and is not overfitting**, as there is little divergence between the training and validation curves.
- The steady decrease in both loss curves and the increase in accuracy indicate that the LSTM model is fitting the data well and generalizing effectively.
- The slight fluctuations in the validation loss indicate occasional overfitting, but overall, the model is improving with each epoch.

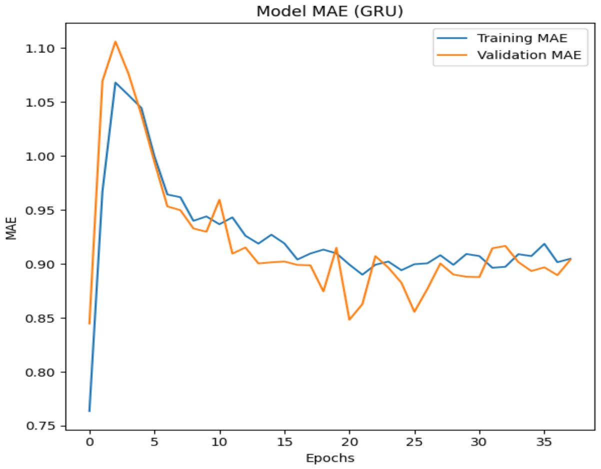
➢ GRU (Gated Recurrent Unit):

- The GRU model shows a steady decrease in loss and an increase in accuracy over time, suggesting effective learning.
- The validation loss follows a similar trend to the training loss, indicating that the GRU model is generalizing well without significant overfitting.
- The GRU model shows faster convergence with both training and validation loss decreasing rapidly within the first few epochs. This is indicative of the GRU's ability to learn more efficiently from sequential data, especially in tasks like pedestrian behavior prediction.
- The smooth convergence and consistent performance suggest that the GRU model is stable and less prone to overfitting compared to the LSTM.









Conclusions

➢ LSTM and GRU results:

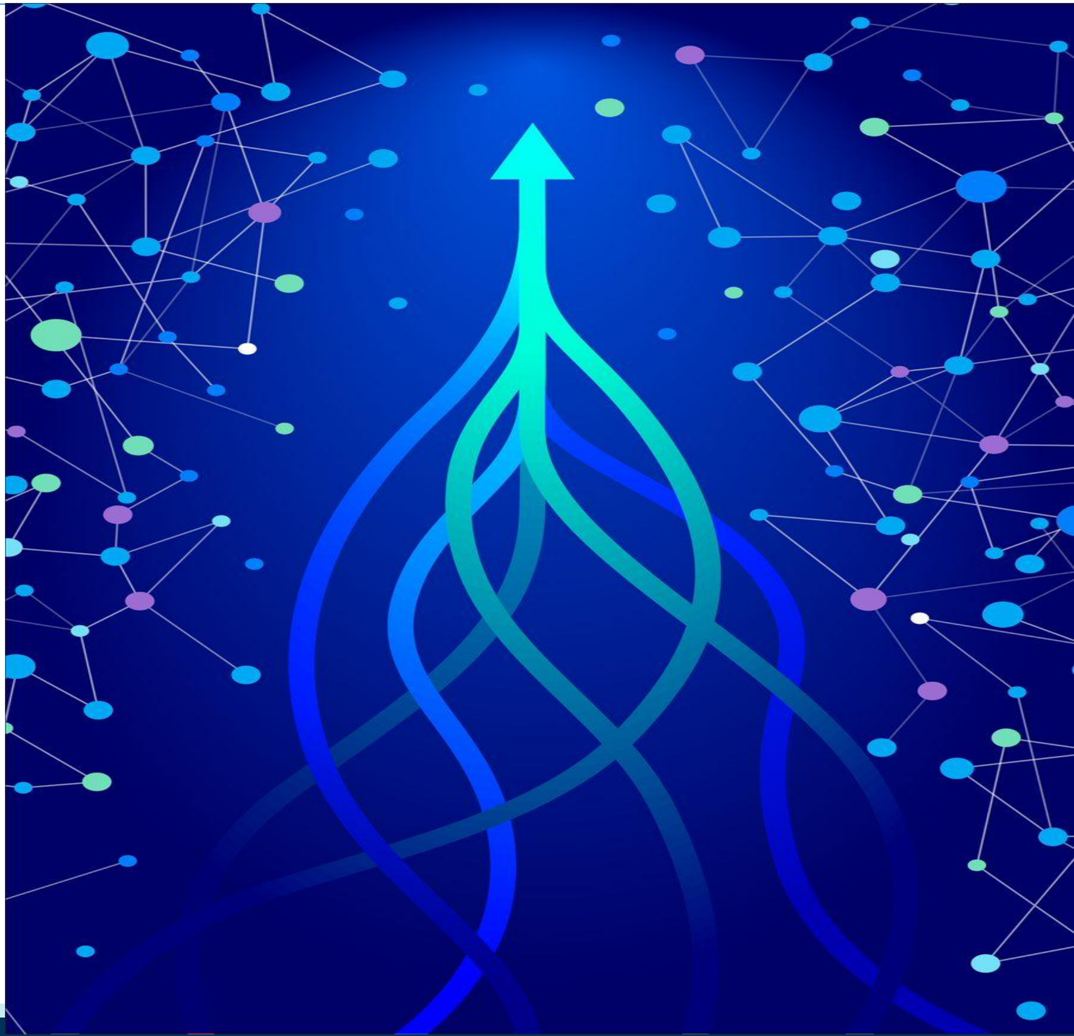
- **Both LSTM and GRU models** demonstrated a strong ability to predict illegal pedestrian crossings during different traffic light phases.
- **LSTM**: Best balance between precision and recall; minimizes false alarms; suited for **real-time traffic safety systems requiring reliability**.
- **GRU**: Maximizes detection rate; suitable for applications **focused on capturing as many violations as possible**, even with increased false positives.
- YOLOv8 + ResNet-50 provided accurate real-time detection, forming a reliable dataset for model training.

➢ Study Contributions:

- Demonstrates the potential of LSTM and GRU models to predict illegal pedestrian crossings in complex urban environments.
- Integrates advanced object detection (YOLOv8) with deep learning for real-time violation detection.
- Addresses challenges in cities with limited monitoring infrastructure, such as Athens.
- **Practical Implications:**
- LSTM recommended for traffic safety systems where minimizing false alarms is critical.
- GRU preferred when maximum detection is prioritized, accepting more false positives.

Future Work & Broader Impact

- **Broader Impact**
  - Supports the **global target of 50% reduction in traffic fatalities** by 2030 (EU & WHO goals).
  - Enables cost-effective, AI-powered monitoring in cities with limited infrastructure.
- **Next Steps for Research & Application**
  - Enhanced Feature Integration by **including road conditions, pedestrian density, and weather data** to refine prediction accuracy.
  - Multi-Modal Data Fusion by combining video, sensor, and connected vehicle data for more robust traffic monitoring systems.
  - Real-Time Deployment by integrating models into live traffic management platforms for proactive safety interventions.
- **Future Directions**
  - Integrate additional environmental factors (e.g., pedestrian density, weather).
  - Expand to multi-sensor data sources for enhanced system robustness.
  - Supports **PHOEBE project goals**: Smarter urban traffic management, safer mobility for all road users.



THANK YOU!

Stella Roussou

s_roussou@mail.ntua.gr

Host – Symposium Sponsor

International Partners

National Supporters