World Conference on Transport Research (WCTR)
6-10 July 2026, Toulouse, France

The Enhancing Risky Driving Behavior Classification And Using Conditional GANs (cGANs): A Data Augmentation Approach

Eleni Maria Theodoraki, Paraskevi Koliou, George Yannis

National Technical University of Athens, Department of Transportation Planning and Engineering, Athens, Greece

Introduction

Road traffic crashes remain a major **global safety issue** (WHO).
Risky behaviours such as **speeding** and **harsh braking** increase crash risk.

Machine learning can classify risky behaviour, but:
- Naturalistic driving data are highly imbalanced
- Normal driving dominates
- High-risk events are rare

→ Leads to biased predictions and poor detection of critical cases

- cGANs generate synthetic minority-class samples
- Improve detection of rare high-risk events

Objectives

- Evaluate **cGAN augmentation** for class imbalance
- Assess performance vs. overfitting
- Examine impact on interpretability (SHAP)

➡️ Support **safe use of synthetic data**

Data Overview

Dataset: Naturalistic driving data (i-DREAMS)
Countries: Belgium (43 drivers) & United Kingdom (26 drivers)
Duration: 4-month monitoring period
Total: 265,000+ minutes of driving

Data Acquisition:
Vehicles equipped with OBD-II systems & Mobile sensors
Recorded variables: Speed, Headway, Braking

Data Structure:
Driving segmented into 30-second intervals
Classified into: Normal, Dangerous, Avoidable Accident

Methodology

Data Preparation:
i. Removed missing values
ii. Applied **Min–Max normalization**
iii. Problem: **Binary classification** (Normal vs. Avoidable Accident)

Data Augmentation (cGAN):
i. Generated synthetic high-risk samples
ii. Generator: Noise + risk label → synthetic data
iii. Discriminator: Evaluates authenticity + class consistency

Model Development:
- 80/20 train–validation split
- Stratified 5-fold cross-validation

Models Evaluated
- XGBoost–Random Forest (hybrid ensemble)
- RNN + AdaBoost (temporal patterns)
- GAN-based classifier

Evaluation & Interpretability:
- Metrics: Accuracy, Precision, Recall, F1-score
- SHAP analysis:
  i. Before & after augmentation
  ii. Assess impact on decision patterns

Results

Before Augmentation (Original Data):
- XGBoost–Random Forest
  - Accuracy: >90% (Belgium & UK)
- RNN–AdaBoost
  - Moderate performance
- GAN-based classifier
  - Lower accuracy
  - Limited minority-class representation

After cGAN Augmentation:
- Accuracy increased across all models
- Recall improved significantly
  - Better detection of high-risk events
- GAN-based model
  Improved to ~90% accuracy

Key Concern:
- Hybrid models reached near-perfect accuracy
- → Indicates potential overfitting

Table 2: *Performance Comparison of Classification Models*

Dataset	Model	Accuracy	Precision	Recall	F1-score
Belgium (original dataset)	XGBOOST & RF	93%	93%	93%	93%
	RNN & AdaBoost	83%	82%	83%	82%
	GANS	76%	64%	76%	66%
UK (original dataset)	XGBOOST & RF	92%	92%	92%	91%
	RNN & AdaBoost	85%	84%	85%	84%
	GANS	79%	80%	79%	74%
Belgium (Augmented dataset)	XGBOOST & RF	100%	100%	100%	100%
	RNN & AdaBoost	100%	100%	100%	100%
	cGANS	90%	90%	90%	90%
UK (Augmented dataset)	XGBOOST & RF	100%	100%	100%	100%
	RNN & AdaBoost	100%	100%	100%	100%
	cGANS	91%	90%	91%	91%

Overfitting and Generalization

Key Observation:
- Models reached ~100% accuracy after augmentation
- Strong indication of overfitting to synthetic patterns

Performance Behavior:
- Recall improved significantly
- Precision improved only slightly
- → Increased sensitivity, but limited improvement in true positive accuracy

Model Stability:
- Performance varied across cross-validation folds
- Synthetic data showed lower variability
- → Reduced complexity → risk of memorization

Implications:
- Models may learn artificial structures, not real behavior
- Creates a trade-off:
- 📊 Higher metrics
- vs. ⚠️ Lower real-world reliability

- Critical issue for safety-sensitive applications

Feature Importance & Model Interpretability

Key Predictors (Pre-Augmentation):
- Speed-related variables → strongest influence
  - Mean speed
  - High-level speeding indicators
- Harsh braking → critical for detecting Avoidable Accidents

Impact of Data Augmentation:
- Feature importance shifted after GAN augmentation
- Models relied more on synthetic feature patterns

Implications:
- Augmentation may alter learned risk representations
- Improved balance ≠ unchanged model logic
- Raises concerns about:
  - Interpretability
  - Robustness

Discussion

- **Reaching near-100% accuracy** after augmentation signals overfitting to synthetic patterns, not real improvement. Recall rose sharply, but precision barely improved, models flag more risk events without better separating true from false positives. Results also varied across cross-validation folds, hinting at memorization rather than generalization.

- **Speed-related variables (mean speed, high-level speeding)** were the top pre-augmentation predictors; harsh braking was key for flagging Avoidable

- Better class balance doesn't mean unchanged model logic, models leaned more on synthetic patterns post-augmentation, raising interpretability and robustness concerns. In safety-critical systems, these matter as much as accuracy.

Conclusions

Key Findings:
- Conditional GANs reduce class imbalance in risky driving data
- Improve:
  - 📈 Recall
  - 📊 Overall accuracy

Limitations:
- ⚠️ Risk of overfitting with synthetic data
- ⚠️ Potential distortion of feature importance

Interpretation:
- Near-perfect accuracy may indicate memorization, not generalization

Implications:
- In safety-critical systems:
  - Robustness and Interpretability
  - are as important as Performance

Contact Information: Eleni Maria Theodoraki, Research Associate NTUA
Email: e_theodoraki@mail.ntua.gr