**Predictive Crash Analytics for Traffic Safety using Deep Learning**

**Karthik Sivakoti**

[**karthiksivakoti@utexas.edu**](mailto:karthiksivakoti@utexas.edu)

**The University of Texas at Austin, Masters in AI, Department of CS**

**Abstract**

Traditional automated crash analysis systems heavily rely on static statistical models and historical data, requiring significant manual interpretation and lacking real-time predictive capabilities. This research presents an innovative approach to traffic safety analysis through the integration of deep learning and real-time data processing for crash risk assessment and prediction. My system has achieved a Mean Average Precision (mAP) of 0.893 across diverse conditions, demonstrating 92.4% accuracy in risk prediction and 89.7% precision in hotspot identification. The architecture incorporates temporal pattern analysis, spatial risk mapping, and real-time environmental condition monitoring through an advanced dashboard interface. Through extensive training using 500,000 crash records for the year 2023 (courtesy of Pennsylvania Department of Transportation’s publicly available dataset), my solution shows marked improvements in prediction accuracy while significantly reducing computational overhead compared to traditional statistical methods. This research contributes to intelligent transportation systems by introducing a scalable, precision-centric solution that enhances operational efficiency and public safety in modern traffic management.

1. **Introduction**

The development of predictive crash analysis systems represents a critical advancement in modern transportation infrastructure management. Traditional methods rely heavily on retrospective statistical analysis, which often fails to capture the dynamic nature of crash risks and the complex interactions between various contributing factors (Wang et al., 2023). Recent developments in deep learning and real-time data processing have created opportunities for revolutionary improvements in this field, particularly in developing predictive rather than reactive approaches to traffic safety (Rahman & Singh, 2023; Baek et al., 2022).

1. **Related Work**

The evolution of crash analysis systems has undergone several significant phases, each marking important technological advancements.

**2.1 Early Approaches in Crash Analysis**

Early research in crash analysis primarily focused on statistical modeling using limited variables. Thompson et al. (2023) demonstrated that traditional statistical approaches achieved moderate success in identifying crash patterns, with accuracy rates of 75-80% under optimal conditions. However, these systems struggled significantly with real-time prediction and complex pattern recognition. The work of Chen & Li (2022) further highlighted how these early systems required extensive manual intervention, particularly during adverse weather conditions or high-traffic scenarios.

**2.2 Machine Learning Integration**

The integration of machine learning marked a significant advancement in crash analysis capabilities. Studies by Kim et al. (2023) showed that initial machine learning implementations improved prediction accuracy to 82-85%, though still maintaining significant hardware dependencies. Zhou & Chen (2022) further developed these approaches by implementing ensemble learning techniques, achieving accuracy rates of 87% in controlled environments. However, these systems continued to face challenges with real-time processing and environmental adaptability.

**2.3 Deep Learning Advancements**

Recent years have seen significant advancement in the application of deep learning to crash analysis. Transformative work by Yang & Zhang (2022) introduced attention mechanisms in crash prediction models, achieving accuracy rates of 89% through advanced feature extraction techniques. This was further enhanced by Wang et al. (2023)'s implementation of transformer architectures, which demonstrated superior performance in handling temporal dependencies in crash patterns.

Particularly notable is the work of Liu et al. (2023), who developed a multi-modal approach combining computer vision and sensor data. Their system achieved 90% accuracy in crash prediction but required substantial computational resources and complex hardware configurations. Our work builds upon these foundations while addressing the limitations of computational overhead and hardware dependencies.

**2.4 Current Challenges**

Current research in crash analysis faces several critical challenges, as identified by recent studies (Park et al., 2023; Sun et al., 2023), which includes the integration of real-time data streams with predictive models, particularly in handling the volume and variety of incoming data. Environmental adaptability continues to be a crucial factor, with many systems showing degraded performance under adverse weather conditions. These challenges present opportunities for innovation in both model architecture and implementation strategies.

1. **Methodology**

Based on our analysis of the codebase, our system implements a sophisticated pipeline for crash risk prediction and analysis. The architecture consists of several key components working in concert to provide real-time risk assessment and visualization.

**3.1 Data Integration Layer**

The data integration layer, implemented in data\_loader.py, handles the ingestion and preprocessing of crash data from the Pennsylvania Department of Transportation. The code reveals a robust ETL process:

|  |
| --- |
| def process\_chunk(chunk, db: Session): records = [] for idx, row in chunk.iterrows(): lat, lon, is\_estimated = get\_coordinates(row) crash = Crash( crash\_datetime=datetime(...), location=from\_shape(Point(lon, lat), srid=4326), severity=determine\_severity(row), weather=str(safe\_int(row['WEATHER1'])), road\_condition=str(safe\_int(row['ROAD\_CONDITION'])) ) records.append(crash) |