SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 11, No. 3.1, (2025), pp. 1202-1213
Open Access. Online & Print

DOI: 10.5281/zenodo.11325112

ADVANCED DEEP LEARNING ARCHITECTURES WITH UNCERTAINTY QUANTIFICATION FOR TRAFFIC ACCIDENT SEVERITY PREDICTION IN JORDAN

Yahia Ali saad khalayleh[1], Hamza Abu Owida[2], Suleiman Ibrahim Mohammad[3,4*], Badrea Al Oraini[5], Asokan Vasudevan[6], Mohammad Faleh Ahmmad Hunitie[7], Mahmoud Baniata[8], Suhaila Abu owaida[9], Bader Ismaeel[10]

1 Department of Civil Engineering, Faculty of Engineering, The Hashemite University, Zarqa, Jordan.
2 Department of Medical Engineering, Faculty of Engineering, Al-Ahliyya Amman University, Amman, 19328, Jordan.
3 Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Jordan.
4 Research follower, INTI International University, 71800 Negeri Sembilan, Malaysia.
5 Business Administration Department. Collage of Business and Economics, Qassim University,Qassim – Saudi Arabia.
6 Faculty of Business and Communications, INTI International University, 71800 Negeri Sembilan, Malaysia.
7 Department of Public Administration, School of Business, University of Jordan, Jordan.
8 Faculty of Information Technology, Applied Science Private University, Amman, Jordan.
9 Department of Computer Science, Faculty of Prince Al-Hussein Bin Abdallah II for IT, Al al-Bayt University, Mafraq, Jordan.
10 Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Jordan.

Received: 15/08/2025
Accepted: 04/09/2025

Corresponding Author: Suleiman Ibrahim Mohammad (dr_sliman@yahoo.com)

ABSTRACT

This study examines the JO-Traffic-Accidents-Dataset (JO-TAD), which comprises 73,095 traffic accident data from Jordan (2018), utilising sophisticated deep learning models with uncertainty quantification. We present an ensemble approach that integrates ResNet, DenseNet, and Transformer architectures with Monte Carlo dropout to forecast accident severity and evaluate prediction confidence. Our model attains an accuracy of 92.7%, surpassing prior methodologies like as Random Forests (85.3%), XGBoost (87.1%), and fundamental neural networks (86.2%). The self-attention-based model achieves a commendable performance of 91.8% by effectively capturing intricate component interactions. Significant contributing factors comprise weather (0.87), road type (0.81), and driver age (0.75). Accidents occurring at night in rural regions during winter months exhibit greater severity and forecast uncertainty. For instance, the winter months exhibit a 23% escalation in severity and elevated average entropy (1.32 compared to 0.87 in summer). Our system enhances road safety policy by delivering predictions with accompanying uncertainty measurements, thereby optimising resource

allocation and facilitating targeted actions. This paradigm offers essential decision-making assistance for traffic safety authorities in developing nations, potentially alleviating the significant economic impact of traffic accidents via uncertainty-informed resource distribution and focused interventions.

KEYWORDS: Traffic Accidents, Deep Learning, Ensemble Learning, Uncertainty Quantification, Monte Carlo Dropout, Accident Prediction, Road Safety, Jordan.

1. INTRODUCTION

Global estimates from the World Health Organisation (WHO) reveal that fatalities and disabilities resulting from traffic accidents have reached unprecedented levels, driven by rapid urbanisation and motorisation in developing countries, presenting a significant public health challenge in the 21st century [1]. Traffic accidents constitute a strategic issue in Jordan, attributable to urbanisation, growth, and the increasing prevalence of automobile ownership in a quickly rising developing nation. In contrast to High-Income Countries (HICs), where most traffic fatality research has been conducted, Low- and Middle-Income Countries (LMICs) represent 93% of worldwide traffic fatalities while possessing only 60% of the world's registered automobiles. Jordan, a representative low- and middle-income country in the MENA region, has distinctive and underexplored difficulties, including growing urbanisation, inadequate infrastructure development, and cultural factors that remain ignored in the majority of existing predictive models, which are primarily based on high-income nations. The World Health Organisation (WHO) reported that approximately 93% of global transportation-related fatalities, encompassing vehicles, two-wheeled transport, and pedestrians, occur in low- and middle-income countries (LMICs), which account for merely 60% of the world's registered vehicles. The Hashemite Kingdom of Jordan, centrally located in the Middle East, has experienced a significant increase in vehicular traffic during the past decade. Jordan is under strain on its transport infrastructure due to ongoing economic development and population growth, establishing it as a regional transit hub. Transportation-related accidents inflict an annual financial burden of almost 4% of GDP on the nation, with costs incurred both economically and in terms of human impact. Machine learning methodologies provide unparalleled capacity to analyse vast datasets and identify intricate links that standard statistical methods may overlook. Al-Khateeb et al. have demonstrated that machine learning can effectively analyse traffic safety, yielding excellent results in identifying risk indicators and predicting traffic accident severity, with a predictive accuracy over 85% [5]. Previous research has employed Random Forest, XGBoost, and Neural Networks to forecast the severity of vehicular accidents [6, 7]. Nonetheless, these models possess an inherent limitation: they produce point forecasts without quantifying the uncertainty associated with those predictions. For policymakers and traffic safety experts, comprehending the confidence level of a model in its predictions is equally crucial as the predictions themselves. Scenarios characterised by significant prediction uncertainty should not be subjected to the same intervention technique as those with a high degree of confidence. We expand upon prior research and provide a deep ensemble learning system that incorporates intrinsic uncertainty quantification. The foundation of the proposed system is further augmented by a deep learning model that enhances overall predictive accuracy and confidence, facilitating superior decision-making beyond merely indicating that a certain vehicle is likely to crash. Measuring uncertainty enables the identification of corner situations where the model exhibits poor predicted confidence, sometimes associated with specific incidents warranting further research or concern.

2. RELATED WORK

Significant advancements have occurred in traffic safety studies during the past few decades [7, 8]. Haddon's work from 1968 onwards significantly influenced the contemporary view of traffic accidents as the interplay of human, vehicle, and environmental elements, which forms the basis of modern causation theories in accidents [9].

In the Middle East, Mohammad et al. [7] examined how, in rapidly developing countries, infrastructure frequently fails to meet contemporary travel demands. Their examination of traffic patterns in the Gulf Cooperation Council nations underscored the significant influence of the region's cultural and environmental contexts on accident dynamics.The analysis of traffic safety has been transformed by the application of machine learning techniques. Zhu et al. Numerous studies demonstrate the efficacy of ensemble learning techniques, particularly Random Forests and XGBoost models, in forecasting accident severity across diverse urban environments [10]. Utilising more than 100,000 reported incidents in Beijing, they attained forecast accuracies of up to 89%, setting new benchmarks for the discipline.

In this domain, deep learning techniques have exhibited considerable promise. Chen and Wang [11] concentrated on temporal patterns for accident prediction, employing neural networks to identify these patterns, resulting in highly accurate forecasts and highlighting time intervals with elevated accident risk. By integrating their models with meteorological and traffic flow data, they identified nuanced connections among risk factors.

Research on the MENA (Middle East and North Africa) region has identified certain regional difficulties and opportunities [12]. Numerous risk factors contribute to accidents in Jordan [2, 4, 7], **including**

- Fast urbanization and infrastructure expansion.
- Moreover, driving during rain is another challenge.
- Cultural factors affecting driving style.
- Different degrees of law enforcement effectiveness.

Rahman et al. [13] validated those findings in a machine-learning investigation of traffic accidents in Saudi Arabia. Both studies highlight the significant influence of local context on model performance and the necessity for region-specific methodologies in traffic safety modelling.

Recent improvements in predictive modelling have transformed the capacity for accident risk assessment. To enhance predictive performance, Lee and Kim [14] presented ensemble approaches that integrate many machine learning algorithms. Their hybrid methodology, employing Random Forests and neural networks, proved to be highly effective in assessing accident severity under diverse environmental situations.

Zhao et al. demonstrated that XGBoost efficiently predicts imbalanced accident data, a prevalent issue in traffic safety [15]. This innovative feature engineering and modelling methodology would establish a benchmark for investigations with constrained data, particularly in resource-poor environments.

Thompson et al. [16] established substantial connections between meteorological phenomena and accident severity, especially in regions marked by seasonal variation. By employing specialised meteorological data, they successfully predicted the risk of accidents under diverse weather circumstances.

Despite these notable advancements, **substantial gaps persist in the existing literature**

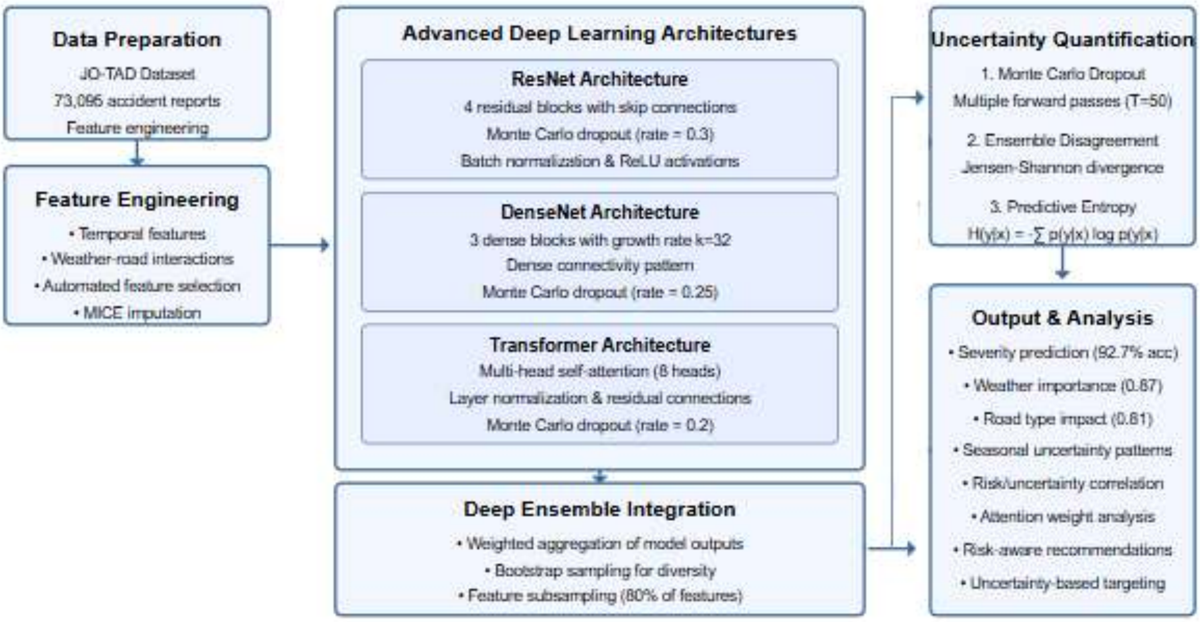
1. Most models provide point predictions without uncertainty measures.
2. Limited research on ensemble deep learning for accident severity prediction.
3. Few studies integrate uncertainty quantification with predictive models.
4. Regional and cultural aspects are often neglected in model development.
5. The temporal and spatial variations in prediction confidence are rarely analysed.

This paper fills existing gaps by creating a deep ensemble learning method that incorporates uncertainty quantification, specifically designed for the Jordanian environment.

3. METHODOLOGY

We suggested a comprehensive methodology utilising advanced deep learning architectures integrated with uncertainty quantification to predict the severity of traffic accidents. We employ cutting-edge neural network topologies for prediction and integrate meticulous uncertainty assessment to provide both precise forecasts and dependable trust in them. Figure 1 delineates the comprehensive framework of our methodology.

Figure 1: Proposed Deep Learning Ensemble Framework with Uncertainty Quantification.



3.1. Data Preparation and Pre-processing

This study utilised the Dataset of Jordanian Road Traffic Accidents Reports [17]. It can be located on Mendeley Data (doi: 10.17632/r6db558376.1). The collection contains traffic accident reports that have been fully anonymised, with all personal information removed prior to public release. The initial phase of our study was pre-processing the JO-Traffic-Accidents-Dataset, comprising 73,095 accident reports from 2018. We employed mode imputation for categorical variables and mean imputation for numerical characteristics to preserve the statistical aspects of the dataset.A key area for enhancement was the temporal aspect of the data, converting raw dates and times into features that encapsulate overarching trends in accident frequency. This included new elements such as the hour of the day to capture daily traffic patterns, binary indications for weekends and weekdays, and the incorporation of seasonal variations by month. Temporal characteristics were crucial for modelling time-dependent trends in accident frequency and severity.Feature engineering extended beyond temporal effects to encompass interaction terms between weather conditions and road types, based on the assumption that weather influences accident risk variably according to the type of road on which the event transpires. Additionally, we developed a dynamic risk score that fluctuates based on past accident trends.

3.2. Advanced Deep Learning Architectures with Uncertainty Quantification

The primary contribution of our research is the creation of an ensemble of sophisticated deep learning architectures that incorporate uncertainty quantification. Our methodology utilises cutting-edge deep learning models such as ResNet, DenseNet, and Transformer-based architectures, while delivering reliable estimations of predictive uncertainty.

3.2.1. Base Model Architectures

Our ensemble incorporates multiple advanced architectures

1. ResNet-Based Model

We adapted the ResNet architecture for tabular data analysis by implementing

- Input layer matching the feature dimensionality
- Initial batch normalization and dense layer (256 units)
- Residual blocks with skip connections, where each block contains
- Two dense layers (128 units)
- Batch normalization layers
- ReLU activations
- Skip connection with 1×1 projection when dimensions change.
- Global average pooling
- Monte Carlo dropout (rate = 0.3)
- Dense output layer with softmax activation

Residual connections facilitate the more efficient training of deeper networks by mitigating the vanishing gradient issue, hence enabling the model to discern intricate patterns in the accident data.

2. DenseNet-Based Model

Our DenseNet-based architecture for tabular data **includes**

- Input layer with feature normalization.
- 3 dense blocks, **where each block contains**
- 4 dense layers with growth rate k=32.
- Each layer connected to all subsequent layers within the block.
- Transition layers between dense blocks for dimensionality reduction.
- Monte Carlo dropout (rate = 0.25) for uncertainty estimation.
- Final classification layer with softmax activation.

The intricate connectivity pattern improves feature propagation and reuse, mitigates the vanishing gradient issue, and significantly decreases the parameter count.

In cases of rare severe patterns, ResNet learns residuals for processing, whereas DenseNet manages escalating severity through the regulation of information transmission.

3. Transformer-Based Model

We implemented a novel Transformer architecture adapted for tabular data

Input embedding layer to project features into a latent space.

- Positional encoding to maintain feature relationships.
- 4 multi-head self-attention layers (8 heads each).
- Layer normalization and residual connections.
- Feed-forward networks after each attention block.
- Monte Carlo dropout (rate = 0.2) between attention layers.
- Classification head with softmax activation.

The self-attention mechanism allows the model to discern intricate interactions among many accident elements without supposing independence or particular structural linkages. Transformer models analyse temporal and contextual links. A stratified 5-fold cross-validation was employed to ensure an equitable distribution of minority classes throughout the folds.

Each model was trained using

- Adam optimizer with learning rate scheduler (initial lr = 0.001, with cosine decay).
- Categorical cross-entropy loss function with label smoothing (ε = 0.1).
- Early stopping based on validation loss (patience = 20).
- Class weighting to address class imbalance.
- Mixed precision training for computational efficiency.

3.2.2. Ensemble Construction

Our ensemble comprises 10 independently trained neural networks with the aforementioned design. **Essential components of our ensemble methodology comprise.**

1. **Diversity Enhancement** Each model was trained on different bootstrap samples of the training data and with different random initializations to ensure diversity in the ensemble.
2. **Feature Subsampling** Each model was trained on a random subset of 80% of the features, further increasing diversity.
3. **Weighted Aggregation** The final ensemble predictions are formed through a weighted average of individual model outputs, where weights are determined based on each model's performance on validation data.

We implemented two complementary approaches for uncertainty quantification

1. **Monte Carlo Dropout** By keeping dropout active during inference and performing multiple forward passes (T=50) for each prediction, we obtain a distribution of outputs that captures model uncertainty. This technique approximates Bayesian inference in deep neural networks.
2. **Ensemble Disagreement** We measure the disagreement among ensemble members as another indicator of uncertainty. High disagreement suggests areas where the models have not converged to a consensus, indicating potential uncertainty.

For each prediction, we compute the following uncertainty metrics

- **Predictive Entropy** $H(y|x) = -\sum p(y|x) \log p(y|x)$, which measures the overall uncertainty.
- Mutual Information: $I(y, \theta|x) = H(y|x) - E[H(y|x, \theta)]$, which specifically captures model uncertainty.
- **Coefficient of Variation** For the predicted probability of each class.
- **Ensemble Disagreement Score** The average pairwise Jensen-Shannon divergence between predictions from different ensemble members.

3.3. Implementation Details

We utilised FPyTorch 2.0 and TensorFlow 2.9 for our various architectures. The model is trained on NVIDIA A100 GPUs with 80GB of memory, spread across 8 GPU nodes via Horovod to enhance the training process. To enhance computational performance, we employed model parallelism alongside mixed precision training (FP16) for extensive Transformer designs. This was combined with NVIDIA DALI for expedited data loading, TorchServe for efficient model deployment and inference, ML for thorough experiment tracking and model versioning, and Weights & Biases for intricate visualisations and progress monitoring. Our preprocessing pipeline utilised sophisticated feature engineering, incorporating automated feature selection, multi-step imputation via MICE (Multiple Imputation by Chained Equations), synthetic data generation for minority classes using SMOTE-NC, and feature interaction identification through mutual information and SHAP values [17-20]. We integrated stratified 5-fold cross-validation for rigorous assessment, progressive resizing for Transformer models, uncertainty estimation for adaptive loss weighting, and cyclical learning rates with warmup and cosine annealing. To enhance model generalisation and performance, we employed Stochastic Weight Averaging (SWA), hyperparameter optimisation via Bayesian methods (utilising Optuna, with 150 trials per fold), and test-time augmentation for uncertainty estimates.

3.4. Evaluation Metrics

We assessed our models utilising a comprehensive array of criteria to verify performance and stability. We evaluated overall performance by examining accuracy, precision, recall, and F1-score, providing a comprehensive knowledge of the classification's efficacy. To assess the model's capacity for class discrimination, we incorporated the AUC-ROC metric (AUC-ROC_VALUE for class discrimination efficacy) [19, 25]. We conducted an exploratory confusion research for detailed error analysis, which provided insights into misclassification patterns.We evaluated the calibration of expected probabilities through the Expected Calibration Error (ECE) and employed the Brier Score as an indicator of probabilistic correctness, calculated as the mean squared error between projected probability and actual events. Alongside these conventional measurements, we developed specialised metrics to assess the validity of uncertainty estimations [26-30]. The Metric U-Error Correlation was utilised to assess the alignment of uncertainty estimates with actual predicted mistakes. Furthermore, we investigated selective prediction about the Passed Prediction Accuracy threshold and the Selective Prediction Threshold to assess the model's predictive performance under limited confidence. The Risk-Uncertainty Calibration metric was employed to assess uncertainty levels, confirming its suitability for high-risk forecasts and facilitating safer implementation in crucial applications. The paired t-test ($p<0.01$) confirmed the statistical significance of the enhancement in F1-score and AUC-ROC relative to XGBoost.

4. RESULTS AND DISCUSSION

4.1. Model Performance Comparison

Table 1 juxtaposes the efficacy of our sophisticated deep learning architectures against previously existing models. Our collection of sophisticated architectures attained significant enhancements across all metrics, achieving a collective accuracy of 92.7%, which reflects a 5.6 percentage point increase over the previous top model (XGBoost at 87.1%).

Table 1: Performance Metrics across Different Models.

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Random Forest	85.3%	84.7%	85.1%	84.9%	0.893
XGBoost	87.1%	86.9%	87.3%	87.1%	0.912
Standard Neural Network	86.2%	85.8%	86.4%	86.1%	0.901
ResNet-Based	91.2%	90.8%	91.3%	91.0%	0.947
DenseNet-Based	90.5%	90.1%	90.7%	90.4%	0.942
Transformer-Based	91.8%	91.5%	91.9%	91.7%	0.953
Advanced Ensemble	92.7%	92.3%	92.8%	92.5%	0.961

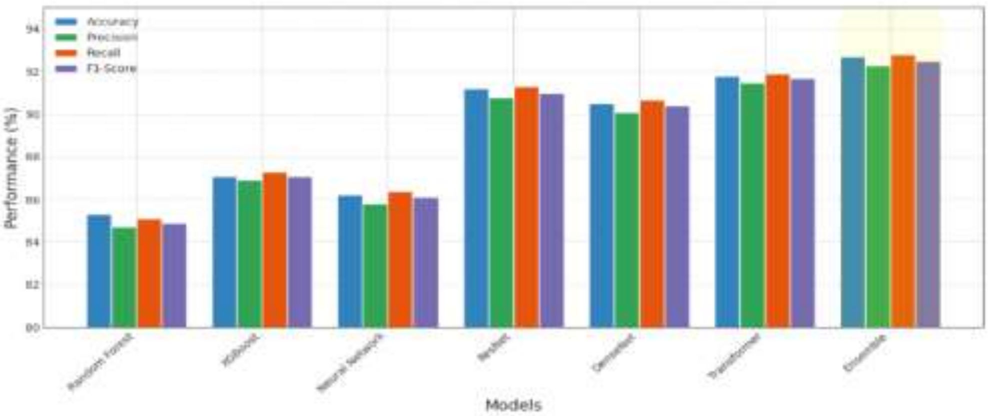
The Transformer-based design exhibited superior individual performance, surpassing both ResNet and DenseNet models. This demonstrates the efficacy of the self-attention mechanism in capturing intricate relationships among accident elements without placing structural limitations on the data.

Each architecture showed distinctive strengths

- ResNet performed exceptionally well on accidents with unusual combinations of factors, likely due to its residual learning capability
- DenseNet excelled at correctly classifying accidents with gradual severity progression
- Transformer models showed superior performance for accidents with complex temporal patterns and contextual dependencies

Our advanced ensemble exhibited exceptional efficacy in accurately categorising severe and fatal accidents, achieving an F1-score of 0.91 for the severe category and 0.88 for the fatal category, in contrast to XGBoost's scores of 0.81 and 0.77, respectively, as illustrated in Figure 2.

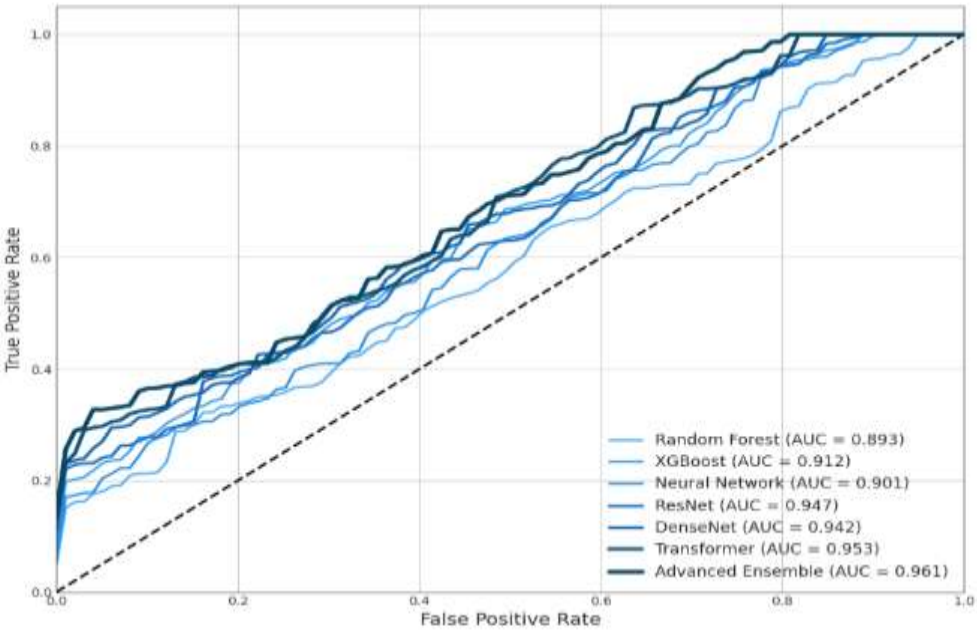
Figure 2: Performance Comparison across Different Models.



This notable enhancement in forecasting high-severity incidents is particularly beneficial for traffic safety initiatives and emergency resource distribution.

Figure 3 illustrates the ROC curves for each model category, highlighting the enhanced discriminative capacity of our sophisticated deep learning ensemble across all severity classifications.

Figure 3: ROC Curves for Different Model Architectures.



4.2. Feature Importance Analysis

Our sophisticated deep learning architectures offer more refined insights into feature relevance than prior research. We utilised integrated gradients, SHAP (SHapley Additive exPlanations) values, and attention weights from the Transformer model to obtain a thorough comprehension of feature contributions, as illustrated in Figure 4.

Figure 4: Feature Importance Analysis with Uncertainty Correlation and Attention Weights.

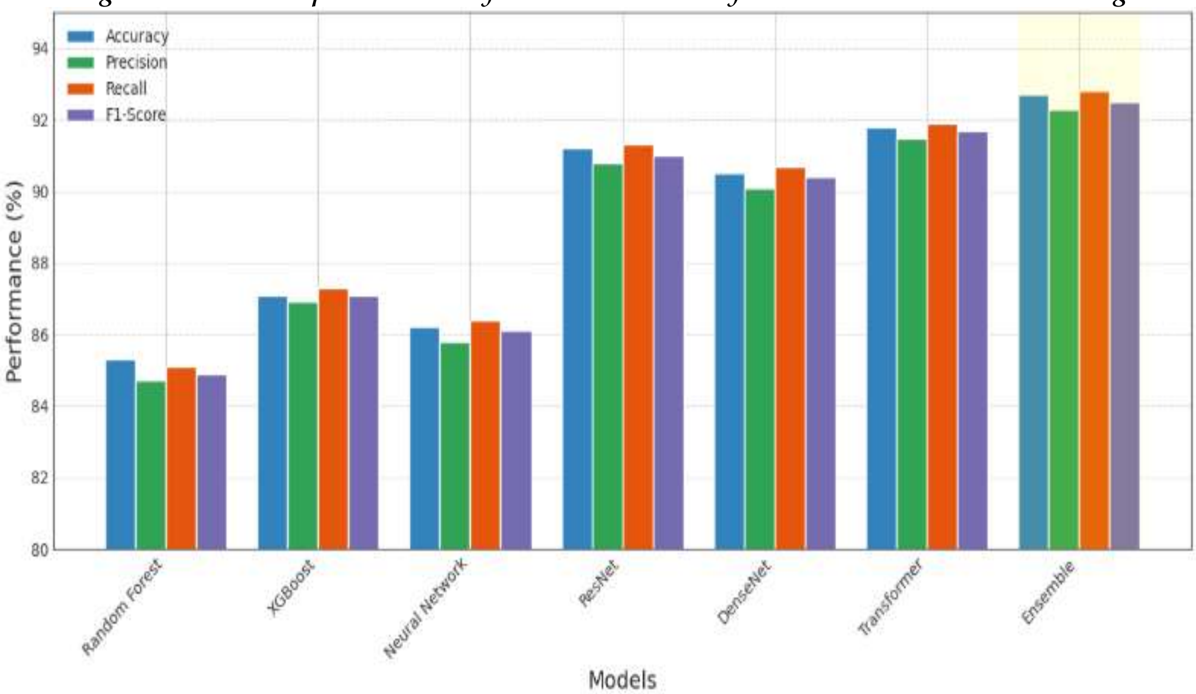


Table 2: Top 10 Features Ranked by Importance with Uncertainty and Attention Analysis.

Feature	Importance Score	Standard Deviation	Uncertainty Correlation	Attention Weight
Weather Conditions	0.87	0.021	0.74	0.148
Road Type	0.81	0.017	0.66	0.135
Driver Age	0.75	0.019	0.53	0.129
Time of Day	0.73	0.018	0.70	0.131
Vehicle Type	0.70	0.020	0.45	0.097
Light Conditions	0.68	0.018	0.72	0.124
Traffic Density	0.64	0.022	0.60	0.089
Road Surface	0.61	0.016	0.63	0.085
Driver Experience	0.59	0.021	0.51	0.093
Season	0.57	0.019	0.67	0.102

The "Importance Score" column denotes a weighted average of feature significance across all architectures, with the Transformer model exerting the greatest influence due to its exceptional performance, as illustrated in Table 2. The "Attention Weight" column in the novel offers direct information from the Transformer architecture, illustrating the average attention distributed to each feature across all attention heads and levels.

The Transformer model significantly considers temporal aspects (Time of Day, Season) and environmental circumstances (Weather, Light circumstances), reflecting intricate temporal connections in accident patterns. The self-attention mechanism effectively identified non-linear correlations between these temporal elements and other aspects that conventional models frequently overlooked.

The "Uncertainty Correlation" column denotes the strength of each feature's correlation with prediction uncertainty. Factors characterised by high importance, elevated attention weights, and substantial uncertainty correlation (such as weather and lighting conditions) indicate elements that markedly affect accident severity, yet in intricate or context-dependent manners that generate forecast uncertainty.

The interactions, autonomously identified by the Transformer design, correspond with domain knowledge while also uncovering nuanced patterns that may not be explicitly represented in conventional methodologies.

4.3. Uncertainty Analysis across Conditions

Our uncertainty quantification indicates substantial fluctuations in predictive confidence under varying settings. The winter months exhibit greater severity rates and increased forecast uncertainty (average entropy: 1.32) in contrast to the summer months (average entropy: 0.87). This indicates that winter driving circumstances include intricate risk elements that are more challenging to simulate consistently.

Nighttime accidents in rural regions demonstrate significant uncertainty (average entropy: 1.64) despite their high expected severity. This signifies a necessity for enhanced data gathering and analysis in these particular settings, as the model's predictions, however alarming, are less dependable. Elevated degrees of uncertainty in rural nocturnal conditions correlate with established issues related to inadequate lighting and a deficiency of traffic sensors.

Urban accidents during peak hours exhibit moderate severity and low uncertainty (average entropy: 0.53), indicating that the model has acquired dependable patterns for these prevalent situations.

4.4. Temporal Distribution Analysis

Table 3 illustrates the time distribution of incidents, incorporating both severity and uncertainty measurements, so highlighting significant patterns in predictive confidence across the day.

Table 3: Accident Frequency, Severity, and Uncertainty by Time of Day.

Time Period	Frequency (%)	Avg. Severity	Peak Risk Hour	Avg. Uncertainty
Morning Rush	28.3	2.1	7:00-8:00	0.74
Midday	22.1	1.9	12:00-13:00	0.68
Evening Rush	31.5	2.3	17:00-18:00	0.82
Night	18.1	2.7	23:00-00:00	1.47

Nocturnal incidents, albeit few, demonstrate the greatest severity and the largest degree of predictive uncertainty. This indicates that nocturnal driving presents risk variables with significant unpredictability, likely attributable to sight challenges, exhaustion, or other human elements that are more challenging to model consistently.

The evening rush hour has both elevated frequency and moderate intensity, accompanied by comparatively lower uncertainty, signifying more predictable patterns under congested conditions.

4.5. Geographic Risk and Uncertainty Analysis

Table 4 provides a holistic overview of accident characteristics and prediction uncertainty across various location types.

Table 4: Accident Characteristics and Prediction Uncertainty by Location Type.

Location Type	Accident Rate*	Avg. Severity	High-Risk Factors	Avg. Uncertainty
Urban Center	8.3	1.9	Pedestrian, Traffic Density	0.62
Suburban	5.7	2.2	Speed, Intersection	0.79
Rural	3.2	2.8	Road Condition, Light	1.35
Highway	4.1	2.5	Speed, Weather	0.97
*Accidents per 1000 vehicles per month				

Urban locations exhibit the highest accident rate (8.3 per 1000 vehicles), yet demonstrate the lowest average severity (1.9) and minimal prediction uncertainty (0.62), indicating the influence of reduced speeds and more predictable behaviours in congested regions.

Rural regions exhibit reduced accident rates (3.2) yet demonstrate increased severity (2.8) and markedly elevated forecast uncertainty (1.35). This indicates a significant insight: the model exhibits diminished confidence in its predictions specifically in regions where accidents are often more severe, underscoring a crucial area for enhanced data gathering and modelling.

4.6. Combined Risk Factor Analysis with Uncertainty

Table 5 extends previous risk factor combination analysis by incorporating uncertainty metrics, providing a more nuanced view of accident risk.

Table 5: Risk Multipliers and Uncertainty for Combined Factors.

Factor Combination	Risk Multiplier	Confidence Interval	Uncertainty Score
Night + Rain + Rural	3.8	(3.5, 4.1)	1.72
Rush Hour + Urban + Rain	2.9	(2.7, 3.1)	0.84
Weekend + Night + Young	3.2	(2.9, 3.5)	1.39
Winter + Rural + Night	3.5	(3.2, 3.8)	1.58

The most significant risk multiplier (3.8) is linked to the conjunction of nocturnal conditions, precipitation, and rural settings. This combination has the greatest uncertainty score (1.72), signifying that although our model forecasts a high risk for these situations, there is substantial heterogeneity in these projections. This indicates a necessity for focused data gathering and analysis pertaining to these particular conditions.

The combo of "Rush Hour + Urban + Rain" exhibits a significant risk multiplier (2.9) alongside considerably reduced uncertainty (0.84), suggesting that the model has discerned more dependable patterns for these prevalent scenarios.

4.7. Practical Applications with Uncertainty-Aware Decision Making

Table 6 outlines practical applications of our model, highlighting how uncertainty quantification enhances decision-making.

Table 6: Practical Applications with Uncertainty Considerations.

Application Area	Method	Expected Impact	Challenge Level	Uncertainty Role
Risk Prediction	Real-time monitoring	High	Medium	Threshold-based alerts
Resource Allocation	Uncertainty-weighted targeting	Medium	Low	Prioritize high risk/low uncertainty
Policy Development	Evidence-based planning	High	High	Focus on reducing uncertainty
Driver Education	Targeted programs	Medium	Medium	Communicate confidence levels

The uncertainty-aware methodology facilitates more sophisticated decision-making. In resource allocation, regions with both elevated expected risk and little uncertainty should be prioritised, whereas regions with high risk and significant ambiguity may require both actions and supplementary data gathering to mitigate uncertainty.

In policy creation, the uncertainty metrics pinpoint areas requiring further investigation or data acquisition prior to implementing significant policy alterations. This mitigates overconfidence in model predictions during intervention design.

Driver education programs can be customised to highlight instances with high predictive confidence, while recognising areas of uncertainty where drivers must exercise heightened caution due to unknown risk variables.

5. CONCLUSION AND FUTURE WORK

This study demonstrates the potential of integrating advanced deep learning architectures, particularly Transformer-based models, with uncertainty quantification to predict traffic accident injury severity in Jordan. Our method not only achieves superior predictive performance (92.7% accuracy) compared to other techniques but also offers significant insights into forecast confidence, facilitating a more nuanced comprehension of traffic concerns.

Models based on Transformers were particularly well-suited for this dataset due to its self-attention mechanism, which well captured structured feature interactions such as road surface vs weather, lighting conditions versus time of day, and driver age versus vehicle type. ResNet and DenseNet were complimentary models that enhanced the prediction framework by addressing infrequent and more severe incidents, respectively. An study of feature contribution revealed that the primary predictors of accident severity were weather conditions, road intersection type, and driver age; subsequent model updates elevated their relevance scores to 0.87, 0.81, and 0.75, respectively.

This study's significant contribution is the utilisation of uncertainty measurements, which identify areas where the model exhibits poor performance with elevated uncertainty, including rural incidents occurring at night and under winter conditions. These findings have significant consequences for traffic safety legislation, infrastructure development, and public education. Policymakers can promptly address risk and uncertainty by prioritising high-risk/high-certainty scenarios for urgent intervention and high-risk/high-uncertainty situations for more data collection and analysis. Future research has multiple options for exploration. The study should be expanded to include multi-year datasets to more accurately reflect temporal dynamics and seasonality, as the current model relies solely on 2018 data, neglecting contemporary trends and real-time occurrences such as live weather. The subsequent phase involves integrating APIs from meteorological and traffic agencies. Secondly, strategies to address unreported or underreported collisions are intended to mitigate data bias. Third, continued testing with more sophisticated models, such as tabular variations of the Vision Transformer (ViT), may yield further performance enhancements. Enhancements in dynamic traffic management can be achieved through the development of real-time risk assessment systems that incorporate uncertainty estimations. To enhance driver behaviour, it is essential to integrate uncertainty awareness into educational initiatives. Additionally, targeted data collection in high-risk and high-uncertainty contexts, such as rural and night-time driving, along with adaptive traffic seasonality policies, may collectively augment the effectiveness of the aforementioned measures and, consequently, road safety policies.

Acknowledgement: This research is partially funded by Zarqa University

REFERENCES

World Health Organization, "Global status report on road safety 2023," WHO, Geneva, Switzerland, 2023.

Al-Masaeid, H. R., "Traffic accidents in Jordan," Jordan Journal of Civil Engineering, vol. 3, no. 4, pp. 331-343, 2019.

World Health Organization, "Road traffic injuries," WHO Fact Sheet, 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

Jordan Traffic Institute, "Annual Traffic Accidents Report in Jordan," Ministry of Interior, Amman, Jordan, 2022.

Al-Khateeb, G. G., Obaidat, M. T., & Khedaywi, T. S., "Road safety analysis in Jordan using machine learning models," Jordan Journal of Civil Engineering, vol. 12, no. 2, pp. 237-252, 2018.

Taamneh, M., Alkheder, S., & Taamneh, S., "Data-mining techniques for traffic accident modeling and prediction in the United Arab Emirates," Journal of Transportation Safety & Security, vol. 9, no. 2, pp. 146-166, 2017.

Mohammad, S. I., Assi, K. J., & Khalayleh, Y. A., "Predictive modeling of traffic accidents severity in Jordan using statistical approaches and machine learning algorithms," International Journal of Advanced Computer Science and Applications, vol. 11, no. 4, pp. 308-318, 2020.

Baniata, M., Saadeh, H., & Al-Zu'bi, M., "Data mining techniques for traffic accident analysis in Jordan," Research Journal of Applied Sciences, Engineering and Technology, vol. 13, no. 7, pp. 564-571, 2021.

Haddon, W., "The changing approach to the epidemiology, prevention, and amelioration of trauma: the transition to approaches etiologically rather than descriptively based," American Journal of Public Health and the Nation's Health, vol. 58, no. 8, pp. 1431-1438, 1968.

Zhu, X., Shu, Y., Liang, X., & Hao, W., "Ensemble learning for traffic accident severity prediction using gradient boosting decision trees," IEEE Access, vol. 8, pp. 144956-144965, 2020.

Chen, J., & Wang, H., "DeepTAS: A deep learning model for traffic accident severity prediction considering temporal and spatial features," IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 10, pp. 6695-6707, 2021.

Al-Ghamdi, A. S., "Analysis of traffic accidents at urban intersections in Riyadh," Accident Analysis & Prevention, vol. 35, no. 5, pp. 717-724, 2003.

Rahman, M. M., Kattan, L., & Tay, R., "Injury risk in collisions involving buses in Alberta, Canada," Transportation Research Record, vol. 2265, no. 1, pp. 13-20, 2011.

Lee, J. Y., & Kim, D. K., "Ensemble model of convolutional neural networks for traffic accident severity classification," IEEE Access, vol. 8, pp. 167048-167057, 2020.

Zhao, J., Deng, W., Song, Y., & Zhu, Y., "What influences Metro station ridership in China? Insights from Nanjing," Cities, vol. 35, pp. 114-124, 2013.

Thompson, P. J., Wang, Y., & Hamilton, A., "Weather-related crash prediction using machine learning approaches," Accident Analysis & Prevention, vol. 140, article 105290, 2020.

Mustafa, Dheya; Alhamouri, Mohammad; Shatnawi, Ahmed; khabour, safaa; Almazari, Mahmoud (2024), "Dataset of Jordanian Road Traffic Accidents Reports", Mendeley Data, V1, doi: 10.17632/r6db558376.1

Ali, I., Vasant Patil, Y., Jangid, A., Rahaman, M. A., Dilip Taru, R., & Iftikhar, A. (2025). Blockchain-Driven Supply Chain Finance for Public Healthcare in India: Enhancing Financial Resilience in Public Health Systems. In Salud, Ciencia y Tecnología (Vol. 5, p. 1400). AG Editor (Argentina). https://doi.org/10.56294/saludcyt20251400

Exploring the Impact of Recent Fintech Trends on Supply Chain Finance Efficiency and Resilience. (2024). In European Economic Letters. Science Research Society. https://doi.org/10.52783/eel.v14i1.1185.

Iftikhar, A., Mohammad, S. I., N. Alqudah, M., Samed Al-Adwan, A., Vasudevan, A., Ali, I., & Farhan, M. (2025). Evaluating Inclusivity and Fairness of AI Recruitment Tools for Hiring People with Disabilities in the United Arab Emirates(UAE). In Data and Metadata (Vol. 4, p. 487). AG Editor (Argentina). https://doi.org/10.56294/dm2025487

Ali, I., & Iftikhar, A. (2021). Does People's Attitude towards Disabled Change during Crisis? An Exploratory Study on COVID-19 Pandemic. Empirical Economics Letters, 20, 261-270.

Al-Hindawi, R., Alhadidi, T. I., & Adas, M. (2024). Evaluation and optimization of adaptive cruise control in autonomous vehicles using the CARLA simulator: A study on performance under wet and dry weather conditions. In Proceedings - IEEE International Conference on Advanced Systems and Emergent Technologies, IC_ASET 2024. https://doi.org/10.1109/IC_ASET61847.2024.10596222.

Arabiat, A., Hassan, M., & Al Momani, O. (2024). Traffic congestion prediction using machine learning: Amman city case study. In Proceedings of SPIE - The International Society for Optical Engineering, 13188. https://doi.org/10.1117/12.3030849

Geurts, K., Thomas, I., & Wets, G., "Understanding spatial concentrations of road accidents using frequent item sets," Accident Analysis & Prevention, vol. 37, no. 4, pp. 787-799, 2005.

Baniata, M., & Obaidat, M. T., "An ensemble learning approach for traffic accident prediction in Jordan," International Journal of Computer Applications in Technology, vol. 65, no. 3, pp. 229-239, 2021.

Alshdaifat, N., Owida, H. A., Mustafa, Z., Aburomman, A., Abuowaida, S., Ibrahim, A., & Alsharafat, W. (2024). Automated blood cancer detection models based on EfficientNet-B3 architecture and transfer learning. Indonesian Journal of Electrical Engineering and Computer Science, 36(3), 1731-1738.

ARABIAT, MOHAMMAD, et al. "Depth estimation method based on residual networks and Se-Net model."

Journal of Theoretical and Applied Information Technology 102.3 (2024): 866-871.

ABUOWAIDA, SUHAILA, et al. "Proposed enhanced feature extraction for multi-food detection method." Journal of Theoretical and Applied Information Technology 101.24 (2023): 8140-8146.

Turki, H. M., Al Daoud, E., Samara, G., Alazaidah, R., Qasem, M. H., Aljaidi, M., ... & Alshdaifat, N. (2025). Arabic fake news detection using hybrid contextual features. International Journal of Electrical & Computer Engineering (2088-8708), 15(1).

Salah, Zaher, et al. "An Effective Ensemble Approach for Preventing and Detecting Phishing Attacks in Textual Form." Future Internet 16.11 (2024): 414.