International Journal of Advanced Trends in Engineering and Management
(ISSN 2583-7052)

Vol. IV, No. 12, pp. 1-16, December

Hybrid Seagull Optimization Algorithm and Hawk-RNN for accurate prediction of Road Traffic Accident Severity

D. Karthikeyan[1*], P S Anu Rakhi[2], Sowmiya R[3]

[1]Assistant Professor, Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai 603203, India. karthikd@srmist.edu.in
[2]Assistant Professor, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai. anurakhips@veltech.edu.in
[3]Assistant Professor, Department of Electronics and Communication Engineering, Rohini College of Engineering and Technology, Tamil Nadu, India, sow25miya@gmail.com

[1*]Corresponding Author E-mail: karthikd@srmist.edu.in

ABSTRACT: Predicting traffic accident severity is crucial for preventing accidents and ensuring the safety of vulnerable road users. This study presents a robust and intelligent framework for predicting Road Traffic Accident Severity (RTAS) using a deep learning model enhanced with bio-inspired optimization techniques. The method initiates with the collection of a comprehensive road traffic accident dataset containing detailed information on various accident-related factors. Data Pre-processing involves handling missing values and performing Exploratory Data Analysis (EDA) to ensure data quality and reveal hidden patterns. To refine input features, the Seagull Optimization Algorithm (SOA) is used for optimal feature selection, effectively reducing dimensionality while preserving critical information. The issue of class imbalance common in accident severity datasets is addressed using SMOTE (Synthetic Minority Over-sampling Technique) to synthetically balance the classes. The proposed framework is a novel deep architecture called Hierarchical Attention Wave Kernel Recurrent Neural Network (Hawk-RNN), which integrates temporal dynamics, hierarchical attention, and wave kernel transformations to model complex dependencies in the data. The model is evaluated using standard performance metrics such as accuracy, precision and recall, value of 92%, respectively. Experimental results demonstrate that the proposed approach significantly outperforms traditional models in accurately predicting accident severity, thereby offering a valuable tool for proactive road safety and traffic management.

Keywords: Road Traffic Accident Severity, Handling Missing Values, Exploratory Data Analysis, Seagull Optimization Algorithm, Synthetic Minority Over-sampling Technique, Hierarchical Attention Wave Kernel Recurrent Neural Network.

1. Introduction

According to the World Health Organization (WHO) global status report on road safety, which assessed data from 175 countries, approximately 1.35 million people lose their lives annually due to traffic accidents. Traffic-related fatalities are identified as the eighth leading cause of death across all age groups [1]. Consequently, a more recent WHO article, published in 2024, highlights that over half of these fatalities involve vulnerable road users, including pedestrians, cyclists, and motorcyclists [2-3]. Furthermore, it is reported that 93% of road traffic deaths occur in low- and middle-income countries, despite these regions only owning about 60% of the world's vehicles [4]. Whereas, with the increasing number of vehicles on the road and the constant entry of new drivers into traffic each day, the frequency and severity of traffic accidents continue to rise, highlighting the urgent need for effective and proactive solutions [5]. So, the economic cost of traffic accidents exceeds 500 billion dollars globally, representing a significant burden on national economies [6]. These losses account for approximately 2% of the Gross National Product (GNP) in developed countries, 1.5% in middle-income countries, and 1% in low-income countries. Whereas, these percentages are estimated to rise further by the 2030s, indicating an urgent need for effective road safety interventions [7-8]. To address these challenges, various pre-processing, optimization techniques and deep neural network are employed to prediction of RTAS.

To feature the road traffic accident and predict RTAS, an optimization technique is essential. Several existing methods are used to optimize the RTAS. Bayesian Optimization (BO) is utilized as a predictive model for RTAS to maximize accuracy or minimize loss functions effectively. It's improve predictive accuracy by finding optimal balances between over fitting and under fitting. However, BO becomes inefficient with very high-dimensional parameter spaces, and computationally expensive for large datasets [9]. Additionally, for RTAS prediction Geohash algorithm is utilized. In order to express latitude and longitude coordinates in a compressed format, it transforms geographic location data into a string for traffic flow. Geohash algorithm to understand and analyze geospatial patterns in accident occurrences and severities. Yet, it is critical for predicting accident severity and lost at coarser geohash granularity [10]. Moreover, Genetic Algorithm (GA) identifies road traffic, optimizing for accuracy, offering transparent insights into accident patterns. GA manage nonlinearity and interactions well without requiring convexity, creating to a good fit for modeling multifaceted factors in traffic accidents . Nevertheless, its fitness evaluation lacks proper validation, and it's over fit training data, reducing generalizability [11]. To overcome the above limitation proposed SOA technique is utilized as effectively reducing dimensionality for further Processing.

Several existing techniques have been employed to model the RTAS prediction system. Random Forest (RF) classifier is utilized as a predictive model in the accident zone. RF handles large numbers of data and complex interactions with minimal preprocessing for safety analysis. Yet, it struggle to capture sequence or spatial-temporal dependencies unless the input data explicitly encode time or location information [12]. Consequently, injury severities of motor vehicle crashes are predicted using Multi-Layer Perceptron (MLP). This model predicts the accident zone and traffic flow with higher accuracy. Nevertheless, MLP is struggle to automatically learn complex interactions without physical data [13]. Furthermore, a Conditional Random Field multi-branch Spatial-Temporal Graph Convolution Network (CRFAST-GCN) to capture the day, time and week for road accident. CRFAST-GCN successfully handles the complex spatial-temporal dynamics effectively and achieves higher performance. However, it increases model complexity, upturns risk of over fitting on limited data [14]. Moreover to enhance the road safety Convolutional Neural Network with Bidirectional Long Short-Term Memory network (CNN-BiLSTM) is utilized. For traffic accident severity prediction CNN-BiLSTM produces with higher accuracy. CNN effectively extracts spatial patterns such as road geometry, traffic density forecasts whereas the BiLSTM seizures temporal dependences such as sequential traffic dynamics or crash precursors. Yet, to exploit spatial-temporal representation, model need large labeled datasets that include spatial road images and temporal sequences, which is unavailable in all accident databases [15]. To overcome the limitations of existing modeling, a proposed Hawk-RNN are implemented for the modeling process to get high accuracy and efficiency of RTAS prediction.

2. Related Work

Yang *et al* [16] (2022) have proposed a Multi-task Deep Neural Network (DNN) for predicting different levels of injury, death, and property loss severity. Multi tasks DNN have the precise analysis of traffic accident severity. DNNs reduce the need for large labelled datasets and help prevent over fitting. Its captures shared patterns across tasks, often outperforming models trained separately for accident zone. Yet, this model is more complex, requiring careful modification and validation.

Prencipe *et al* [17] (2025) have presented an Artificial Neural Network as well as a Pattern Recognition (ANN-PR) model for predicting the traffic severity exposure of each zone in a specific urban area. ANN-PR is used as a speed reduction system in the traffic flow and in the accident zone. This model learns complex, nonlinear mappings between accident features and severity levels. However, struggle to trace specific input variables lead to predictions, which undermines trust in safety-critical applications.

Jaradat *et al* [18] (2024) have developed a Bi-LSTM classifier to enhance crash severity classification. Bi-LSTM processes text in both forward and backward directions, capturing comprehensive context and dependencies in crash descriptions especially effective for unstructured text describing accident events. However, this classifier typically demands more memory, longer training time, and possibly more computational resources especially with large datasets.

Adewopo *et al* [19] (2024) have proposed an Inflated 3D ConvNet (I3D) followed by ConvLSTM2D layers for accident detection in smart city traffic. I3D-CONVLSTM2D model, generated to precisely detect accidents by efficiently capturing the spatiotemporal characteristics of video data. I3D ConvNet increases traditional 2D CNNs into 3D to model both space and time, whereas ConvLSTM2D layers effectively capture motion dynamics in accident zone. Nevertheless, computational, memory demand, data imbalance sensitivity and complex.

Wei and Xu [20] (2024) have presented a Spatio-Temporal Conv-Long Short-Term Memory Autoencoder (STCLA) to enhance road safety with an accident prediction and prevention system. Convolutional layer capture spatial layout with Conv-LSTM units that learn temporal dynamics effectively recognizing moving patterns and anomalies in traffic. However, STCLA requires intensive computation and memory due to encoding, convolution, and recurrent operations across video sequences.

The contribution of the proposed work as follows:

- Data-pre-processing stage incorporating handling missing values and EDA, for improving data quality.
- SOA is employed to efficiently and effectively reduce dimensionality for improved data balancing.
- SMOTE is used to efficiently balance the data by addressing minority class imbalance before modeling
- Proposed Hawk-RNN architecture is designed as a model that transforms complex data. It enhances predictive capabilities and holds significant potential for deployment in intelligent transportation systems.

3. Proposed Work

The proposed block diagram in figure 1 introduces a novel deep learning framework optimized for RTAS prediction through the integration of bio-inspired feature selection and class-balancing techniques. Data pre-processing technique such as handling missing values finds the missing values from the Road Traffic

Accident dataset and EDA used for capture hidden patterns. To enhance feature relevance and reduce dimensionality, the SOA performs smart feature selection, preserving critical predictors. Class imbalance, a common challenge severity labels are unevenly distributed, is addressed through SMOTE, which synthetically balances minority classes.

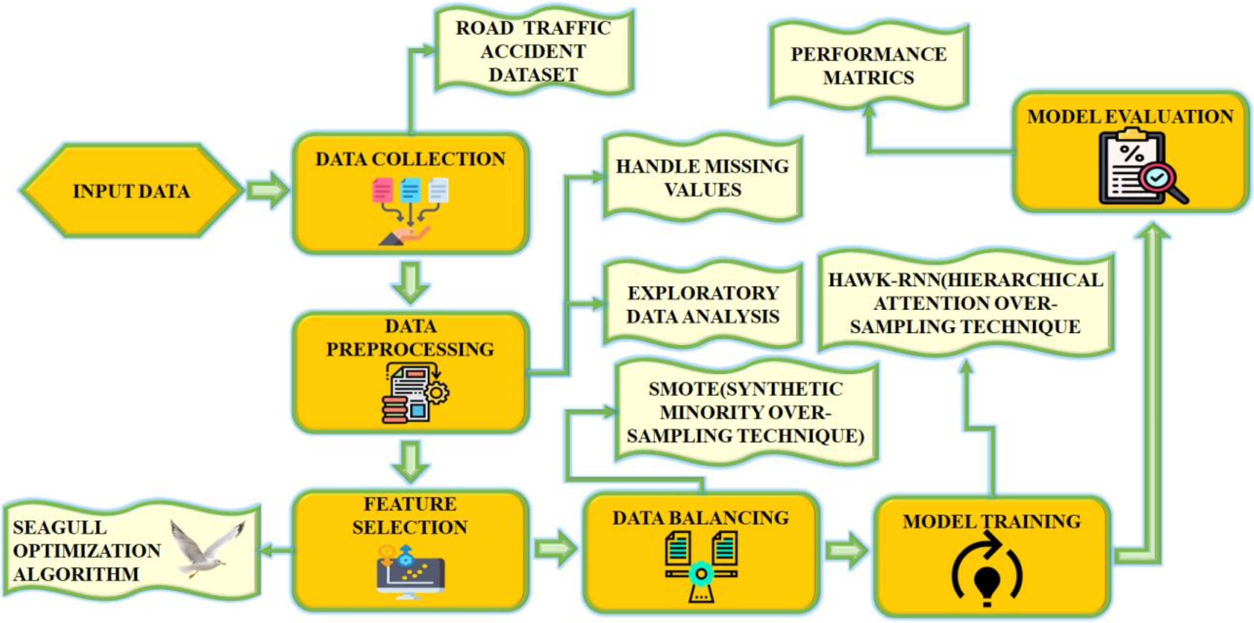

Figure 1: *Block diagram of proposed work*

The proposed predictive model, Hawk-RNN, merges temporal learning (through RNN), hierarchical attention mechanisms, and wavelet-inspired kernel transformations to model rich, multi-scale dependencies in the data. The Hawk-RNN determines advanced predictive capability combined with efficient input preparation and optimization to deliver a comprehensive and intelligent accident severity prediction system.

3.1 Data Pre-processing for RTAS Prediction

For RTAS prediction two pre-processing methods are utilized; handling missing values and EDA.

3.1.1 Handling Missing Values

In the RTAS Prediction task, handling missing values is a crucial preprocessing step to ensure data quality and model reliability. Missing values arise due to various reasons such as human error during data entry, sensor failures, and incomplete reports. To address this, the road traffic accident dataset is first examined to identify the extent and pattern of missing data. For numerical features like vehicle speed, number of vehicles involved, or time of the accident, imputation techniques such as interpolation are applied based on the distribution and nature of the data. For categorical variables like weather condition, road type, or light condition, the mode (most frequent value) is commonly used to fill in missing entries. In cases where a feature contains a significant proportion of missing values, it is removed to prevent distortion in model training. Proper handling of missing values ensures that the predictive model is trained on a complete and consistent dataset, reducing biases and improving the overall performance of the severity prediction system.

3.1.2 Exploratory Data Analysis

EDA in RTAS prediction system plays a vital role in understanding the structure, trends, and patterns within the dataset. EDA begins with summarizing the dataset to check for data types, distributions, and inconsistencies. The target variable, typically accident severity levels (such as minor, serious, or fatal), is analyzed for class distribution to assess imbalance, which is common in such datasets. Univariate analysis is conducted to examine the distribution of individual features like vehicle speed, number of casualties, weather conditions, time of accident, and road surface conditions using histograms, bar plots, and box plots. Bivariate analysis explores relationships between the severity of accidents and other variables such as plotting accident severity against weather, time of day, or road type to identify potential patterns or contributing factors. Correlation matrices are used to detect multicollinearity among numerical features. EDA also helps uncover anomalies, outliers, and inconsistencies in the data, it is addressed before model training. Overall, EDA provides valuable insights that guide feature selection, data balancing strategies, and the design of predictive models for more accurate and interpretable accident severity optimization. The pre-processed data is fed to the optimization technique for predicting the critical predictors.

3.2 Feature Selection by Seagull Optimization Algorithm

In the RTAS Prediction framework, feature selection plays a crucial role in improving model accuracy, reducing over fitting and minimizing computational complexity. To achieve this, the SOA is employed as a bio-inspired metaheuristic technique that efficiently selects the most relevant features from the road traffic accident dataset. Inspired by the dynamic movement and hunting behavior of seagulls, SOA balances exploration and exploitation to search the feature space for optimal subsets. Seagulls come in a wide variety of shapes, sizes, and masses such as omnivores, seagulls consume fish, amphibians, reptiles, insects, and earthworms. The majority of seagulls have white feathers, and incredibly intelligent birds. Both fresh and salt water are suitable for seagull consumption in figure 2.

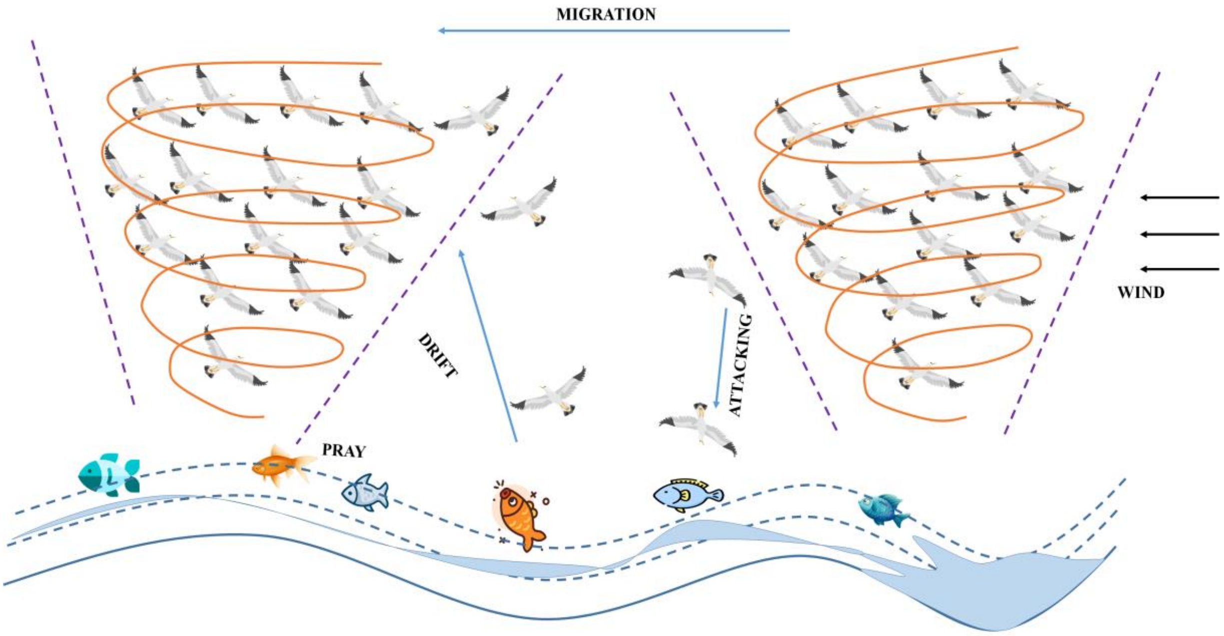

Figure 2. *Basic SOA*

Initially, each seagull (solution) represents a candidate feature subset encoded in a binary vector, where '1' indicates selected features. The algorithm updates each seagull’s position based on spiraling movements and leadership behavior, allowing it to explore new combinations and converge towards the global best solution that enhances model performance. Figure 3 shows the flow chart for SOA.

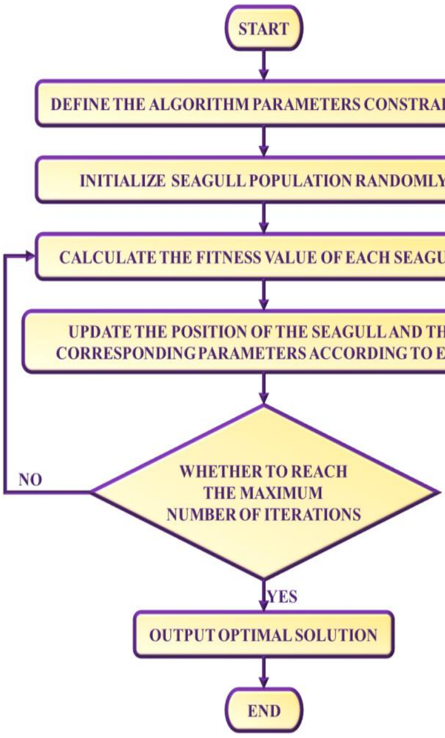

Figure 3. *Flow chart for SOA*

Table 1 illustrates the pseudo code for SO algorithm to find the optimization problems in the RTAS prediction system.

Table 1: *Pseudo Code for SOA*

```
Algorithm :SO
Input: $P_s$ is the seagull population
Output: $P_{bs}$ is the best position for the search agent with the best fitness value using equ. (3)
Random initialization of $(t)$, $B$, and Maximum_Iteration
Set the value of the $f_c = 2, u = 1,$ and $v = 1, t = 0.$
While ($t$ <Maximum_Iteration)
{Compute Fitness value($P(t)$)
Create the random number($rd$) $= (0,1)$
Create the random number $(\theta) = (0,2\pi)$
Compute $r = u * e^{\theta v}$
Estimate the distance $(D(t))$
Estimate the new position $p_{best}$
Update the position of the optimal seagull and fitness value
$t = t + 1;$
} End while
Output: Best position $p_{best}$ of the seagull as well as fitness value
```

During the migration, the program simulated the flow of a flock of gulls from one location to another. In order to avoid collisions between nearby search agents, the new search agent position is calculated using an additional variable, A.

$$C_s = A \times P_s \tag{1}$$

Where, C_s symbolizes the location of the search agent that avoids collisions with other search agents, P_s represents the search agent's current position, x denotes the current iteration, and A denotes the search agent's movement behavior inside a specified search space.

$$A = f_c - (x \times (f_c / Max_{iteration})) \tag{2}$$

Where, the frequency of using variable A is controlled by introducing f_c , which is linearly reduced from f_c to 0. The search agents drive in the direction of the best neighbour after avoiding neighbour collisions.

$$M_s = B \times (P_{bs}(x) - P_s(x)) \tag{3}$$

Where M_s stands for search agent, P_s locations in relation to search agent P_{bs}. B randomized behaviour is in charge of appropriately striking a balance between exploration and exploitation. B is computed as follows:

$$B = 2 \times A^2 \times rd \tag{4}$$

Where rd indicated random integer that falls between 0 and 1. In comparison to the top search agent, the search agent adjusts its ranking by:

$$D_s = |C_s + M_s| \tag{5}$$

Where D_s is the distance between the search agent and the search agent that best fits the query. To balance the optimized data for RTAS prediction system, utilized a data balancing technique called SMOTE.

3.3 Data Balancing by Synthetic Minority Over-sampling Technique

In the RTAS Prediction system, class imbalance is a common challenge, where severe or fatal accident cases are significantly fewer compared to minor ones. This imbalance biases the predictive model toward the majority class, reducing its ability to correctly identify rare but critical cases. To address this issue, the SMOTE is applied as a data balancing method. Figure 4 shows the SMOTE processing, SMOTE generates synthetic samples for the minority class, which enhances diversity and improves model learning.

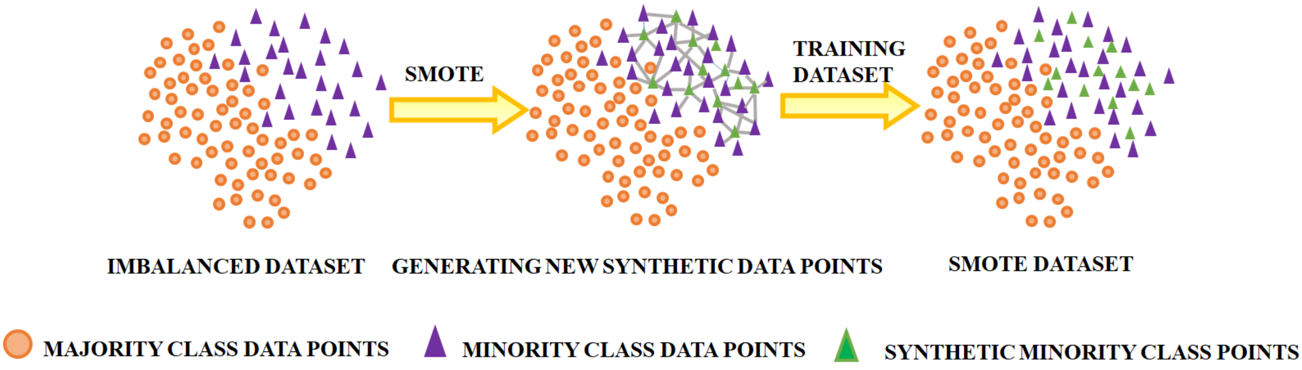

Figure 4: *SMOTE Processing*

SMOTE works by selecting a minority class sample and identifying its nearest neighbours. New synthetic data points are created by interpolating between the sample and its neighbours. The equation for generating a synthetic instance x_{new} is:

$$x_{new} = x_i + \lambda . (x_{nn} - x_i) \tag{6}$$

Where, x_i denotes original minority class instance, x_{nn} is one of its k -nearest neighbours, $\lambda \in [0,1]$ is a random number, x_{new} is the newly generated synthetic sample. By applying

SMOTE, the road accident dataset becomes more balanced, which allows models such as Hawk-RNN to learn patterns from all severity levels more effectively. SMOTE are eventually leads to improved classification accuracy, especially for the underrepresented and more critical accident severity categories, thereby creating the predictive system more robust and equitable. The balanced data is fed to the proposed Hawk-RNN technique to predict the road accident in the traffic.

3.4 Hierarchical Attention Wave Kernel Recurrent Neural Network

Hawk-RNN is an advanced deep learning model tailored for analysing complex sequential or time-series data, used in RTAS Prediction systems. This model effectively integrates three powerful components hierarchical attention, wave kernel transformation, and Recurrent Neural Networks (RNNs) to capture both short-term and long-term dependencies in traffic-related data.

3.4.1 Hierarchical Attention

Hierarchical Attention is a multi-level attention mechanism that consents a model to focus on the most informative parts of input data at different levels of generalisation. In the RTAS Prediction system, hierarchical attention is particularly valuable because it helps the model identify and emphasize critical features (road type, time of day, weather, or driver behaviour) across both individual and grouped data levels (e.g., trips, regions, or time windows).

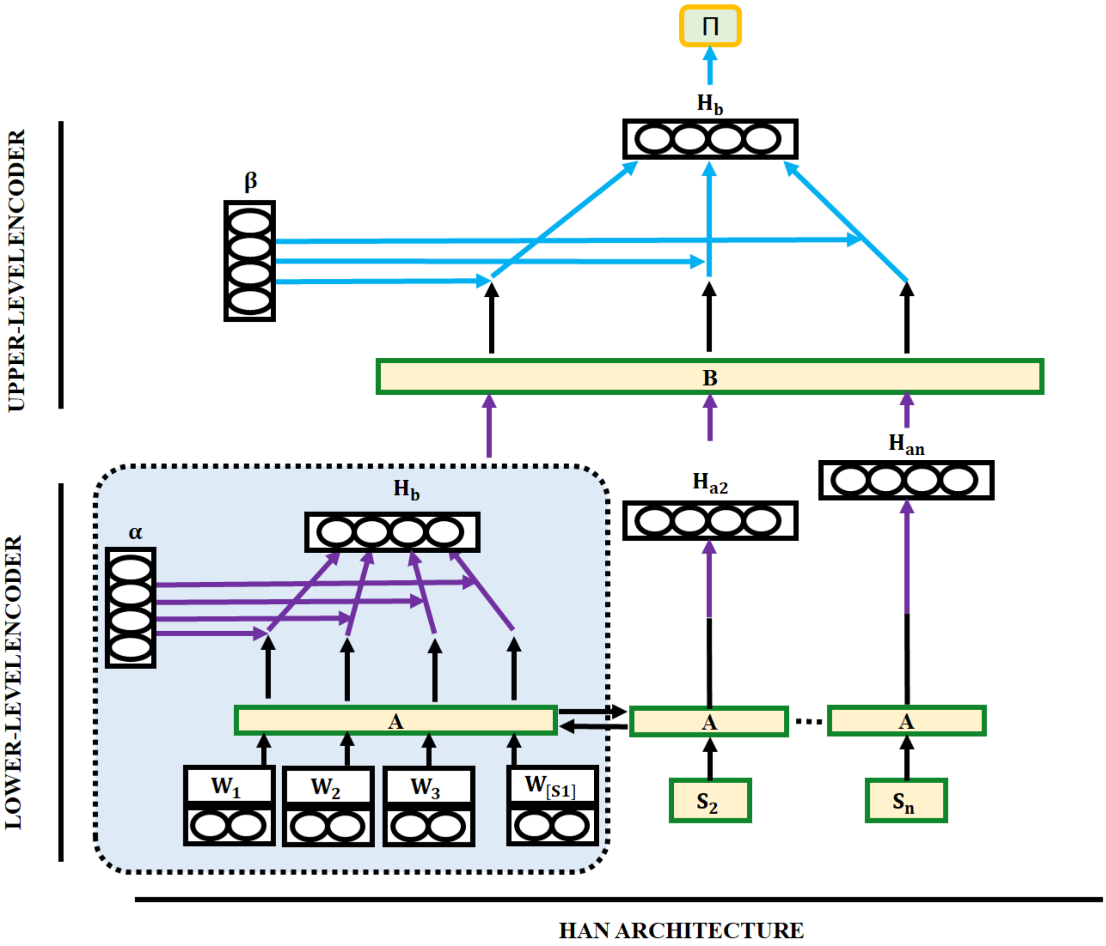

Figure 5. *Hierarchical Attention architecture*

In accident severity prediction, the input data structured hierarchically for instance, individual features like speed and weather, and thus it's grouped into higher-level categories like environmental conditions or traffic background. The hierarchical attention mechanism in figure 5 assigns attention weights first at the lower level (within groups) and then at a higher level (between groups), ensuring the model captures both fine-grained and coarse-grained patterns that influence accident severity. For identifying traffic group let x_{ij} represent the j^{th} feature in the i^{th} group. The feature level attention within each group is given by

$$u_{ij} = \tanh(W_1 x_{ij} + b_1) \tag{7}$$

$$a_{ij} = \frac{\exp(u_{ij}^T v_1)}{\sum_j \alpha_{ij} x_{ij}} \tag{8}$$

$$g_i = \sum_j \alpha_{ij} x_{ij} \tag{9}$$

The group level attention across group for accident severity prediction is given by

$$u_i = \tanh(W_2 g_i + b_2) \tag{10}$$

$$\beta_i = \frac{\exp(u_i^T v_2)}{\sum_j \alpha_{ij} x_{ij}} \tag{11}$$

$$v = \sum_i \beta_i g_i \tag{12}$$

Where, β_i is the attention weight of the i^{th} group and v is the final context vector capturing all relevant information. Hierarchical attention benefits prioritize influential factors like smooth roads in bad weather or speeding in low visibility. By learning to assign importance to features both within individual categories and across different categories, the model more effectively captures complex relationships in the data. This results in more accurate and interpretable predictions of accident severity.

3.4.2 Wave Kernel

The Wave Kernel is capturing both spatial and temporal characteristics of traffic data. Real-world traffic data such as speed, weather conditions, and time of day often exhibit non-linear patterns and localized fluctuations. The wave kernel transforms these input features into a representation that preserves in cooperation local detail and global trends creating it especially useful designed for analysing periodic or oscillatory behaviour in accident patterns.

3.4.3 Wave Kernel Transformation

The wave kernel is inspired by the principles of wavelet transforms, where features are decomposed into multiple frequency bands. It enables the model to localize important variations over time and space, improving predictive sensitivity to critical conditions (e.g., sudden speed drops or weather changes). Let $x(t)$ be a time-series input signal (such as vehicle speed or traffic density over time). The wave kernel function $K_{wave}(x, x')$ between two input vectors $x \text{ and } x'$ is defined as:

$$K_{wave}(x, x') = \exp\left(-\frac{(\| x - x' \|)^2}{2\sigma^2}\right) . \cos(\omega \| x - x' \|) \tag{13}$$

Where, $\| x - x' \|$ is the Euclidean distance between two input vectors, σ is a scale parameter controlling the width of the Gaussian envelope, ω is the frequency parameter that captures periodicity in the data. The exponential term provides locality sensitivity, allowing the model to focus on nearby data points in feature space (e.g., accidents under similar weather conditions). The cosine term captures cyclic or oscillatory trends (like daily traffic flow variations or weather cycles).

3.4.4 Recurrent Neural Network (RNN)

In the RTAS Prediction system, a RNN is employed to effectively model sequential dependencies and temporal patterns within the road traffic accident dataset. To retain memory of previous inputs, making them ideal for analysing time-series data such as hourly accident logs, traffic flow, weather conditions, and road characteristics. Figure 6 shows the Hierarchal Attention RNN architecture it is designed for predicting RTAS. The process activates at the Input Layer, where features (such as time, weather, road type, and vehicle details) are received and grouped into meaningful categories. These inputs are then passed to the Embedded Layer, where each feature is converted into a dense vector representation, allowing the model to capture semantic relationships between them. The output of this layer is structured hierarchically in the Hierarchical Structure Layer, enabling the model to learn feature importance both within and across categories.

The structured inputs are fed into the Attention Module, which resides within a larger network composed of a Sequence Neural Layer and a Block Neural Layer. The Sequence Neural Layer uses recurrent units (like RNNs or LSTMs) to model temporal dependencies across the input sequence, whereas the Block Neural Layer captures higher-level patterns across grouped features.

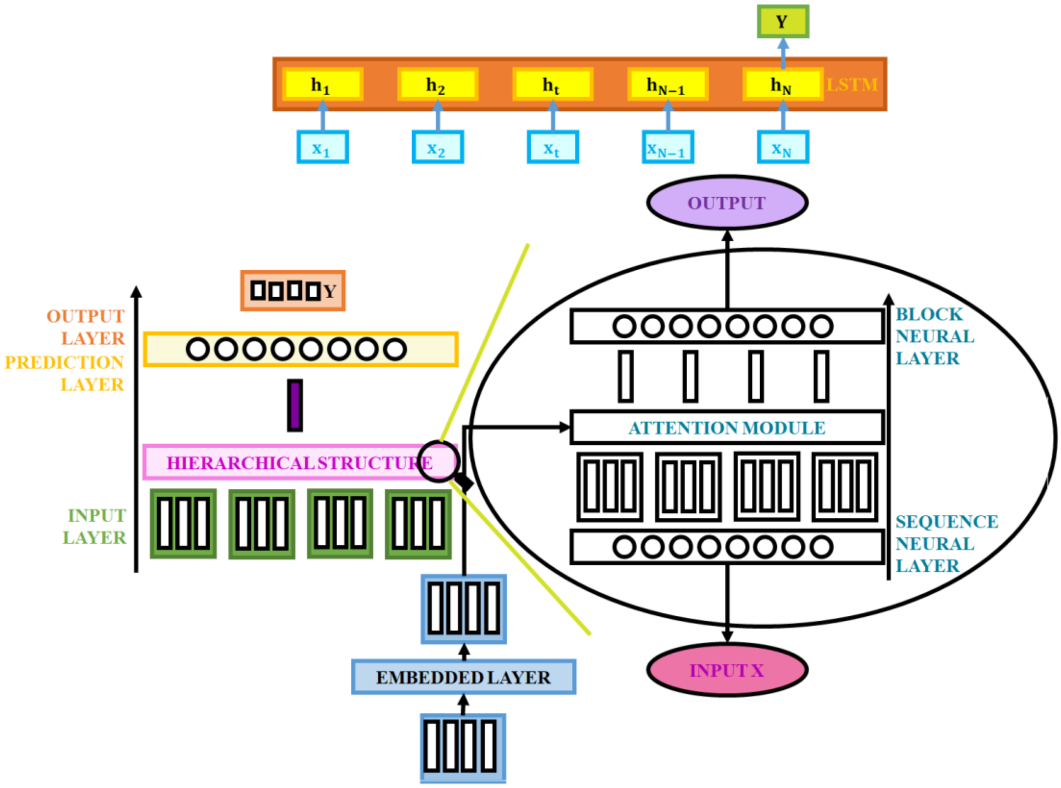

Figure 6: *Hierarchical Attention RNN*

An RNN processes input sequences one time step at a time, maintaining a hidden state that captures information about the sequence. Given an input sequence $\{x_1, x_2, \ldots x_t)$ the hidden state h_t at time t is calculated as:

$$h_t = \sigma(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (14)$$

Where, x_t is the input at time t, h_{t-1} is the hidden state from the previous time step, W_{xh} *and* W_{hh} are weight conditions, b_h is the bias term, σ indicates non-linear activation function. RNNs learn from sequential traffic features (e.g., vehicle speed, congestion levels, time-of-day patterns) to predict accident severity classes (e.g., minor, moderate, severe). The temporal modelling helps capture how past events influence the likelihood and severity of future accidents enabling more accurate and context-aware predictions.

The Hawk -RNN is a sophisticated deep learning framework for predicting RTAS by capturing both hierarchical feature relationships and temporal dependencies within the data. The model leverages a RNN structure, which excels at processing sequential data such as time-stamped accident records. To enhance its learning capacity, a wave kernel function is integrated into the RNN to model complex, non-linear interactions between features, allowing the network to better understand intricate patterns in variables like time of day, weather conditions, vehicle types, and road environments. Additionally, a hierarchical attention mechanism is employed to assign different levels of importance to various features and time steps first focusing on individual features within each time segment, and then on the relevance of each section to the overall prediction. This two-level attention approach ensures that the model captures critical local details and maintains a global understanding of the accident situation. As a result, the Hawk-RNN offers enhanced accuracy and interpretability, making it a powerful tool for identifying the severity of traffic accidents and supporting proactive traffic management strategies.

4. Result and Discussion

In result and discussion section the RTAS are predicted using the proposed Hawk-RNN. For identifying accident in road the road traffic accident dataset is taken from the kaggle.com implemented using Python software.

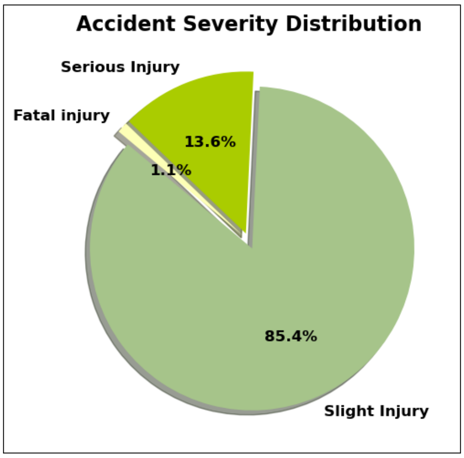

Figure 7: *Accident severity distribution*

Figure 7 shows the accident severity distribution for RTAS prediction system. The pie chart illustrates the distribution of accident severity within the road traffic accident dataset, revealing a common but critical pattern in traffic safety analysis slight injuries dominate with 85.4%, indicating the vast majority of crashes result in minor harm. Serious Injuries account for 13.6%, representing a smaller but significant proportion. And fatal Injuries are the rarest at 1.1%, reflecting a heavily skewed and highly imbalanced dataset.

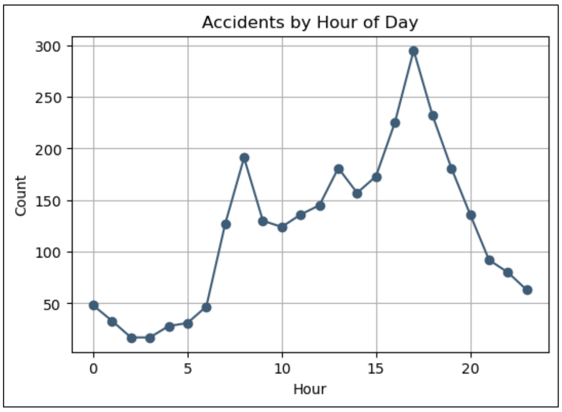

Figure 8: *Accident by hour of day*

Figure 8 shows the accident by hour of day for RTAS prediction system. The line plot of accidents by hour of day shows clear and meaningful temporal patterns early morning (midnight to ~6 AM) accident counts are very low, ranging between 10–50 events, reflecting minimal traffic activity. Morning rush (around 7–9 AM) there's a sharp increase to approximately 130–190 accidents, coinciding with peak commuter traffic. Daytime (10 AM–2 PM) accident numbers moderate, fluctuating between 120–160, as midday traffic steadies. During the afternoon rush hours (3–6 PM), there is a noticeable flow in accidents, rising between 220 and 295 incidents, creating it the busiest period on the roads. In contrast, during the evening hours (7–11 PM), the number of accidents declines significantly to a range of 60 to 140, reflecting reduced traffic volume.

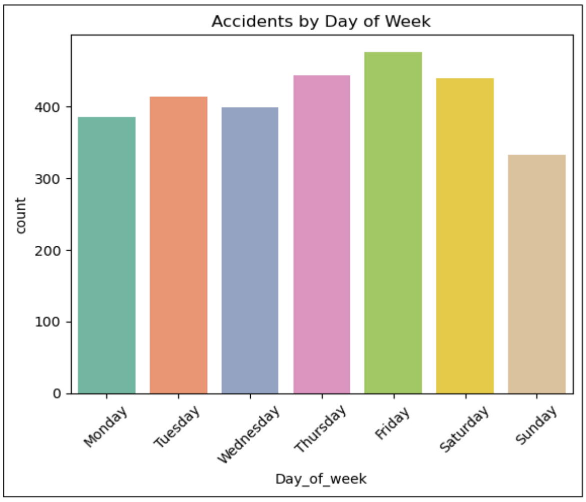

Figure 9: *Accident by day of week*

Figure 9 shows the accident by day of week for RTAS. The bar chart displays the distribution of accidents by day of the week, revealing clear trends: Friday records the highest number of accidents, making it the peak day for road incidents. Thursday and Wednesday follow closely, with slightly fewer accidents than Friday. Monday, Tuesday, and Saturday also show elevated crash counts, but lower than the mid-to-late-week peak. Sunday has the fewest accidents, considerably lower than other weekdays.

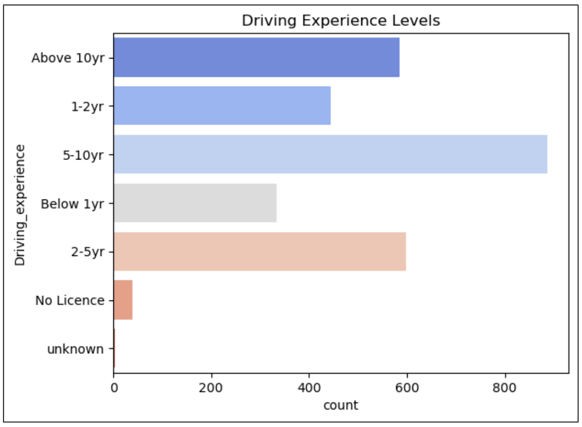

Figure 10: *Driving experience level*

Figure 10 shows the driving experience level for RTAS. The 5–10 years' experience group shows the highest count nearly 900 incidents. This is followed by the 2–5 years group (~600 incidents), then above 10 years (~550), 1–2 years (~500), and below 1 year (~320). A small number of unlicensed drivers (~50) and unknown experience (<10) round out the data.

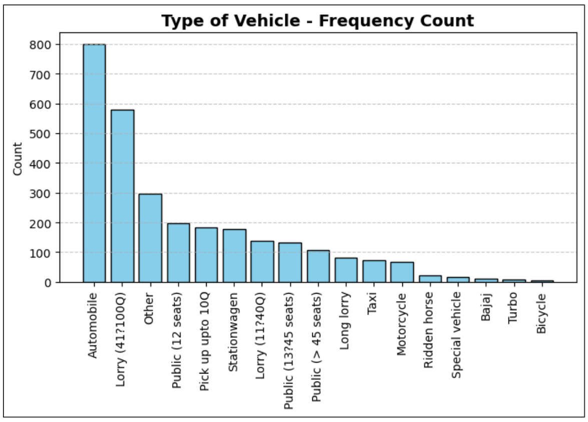

Figure 11: *Types of vehicles in frequency count*

Figure 11 shows the types of vehicles in frequency count. The frequency of different vehicle types involved in recorded traffic accidents. Automobiles (800 cases) and Lorries (~580) dominate the incident counts, reflecting their high road usage. Other vehicles (~300), public vehicles with 12 seats (~200), pickup up to 1 (~180), station wagons (~170), Lorries (~140), public vehicles with more than 45 seats (~130), long lorries (~110), and taxis (~80) follow in decreasing frequency. Motorcycles and ridden horses each appear around 70 cases, whereas special vehicles, Bajaj, Turbo, and bicycles are relatively rare (each under 25 cases)

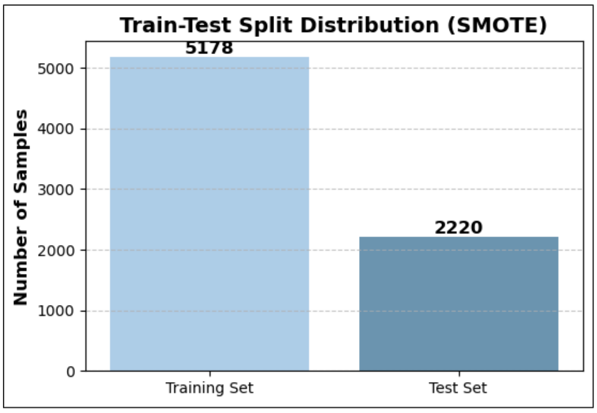

Figure 12: *Train-test split distribution (SMOTE)*

Figure 12 shows the train-test split distribution (SMOTE). The X-axis indicates the test and train set and Y-axis indicates the number of samples. The train–test split after applying SMOTE, with 5,178 samples used for training and 2,220 samples held out for testing.

4.1 Performance Metrics

Performance metrics like accuracy, sensitivity and specificity which indicate the likelihood RTAS prediction the probabilities calculated by the Hawk-RNN. The performance metrics in this study explained in table 2.

Table 2: *Performance Metrics*

Performance Metrics	**Values**
Accuracy	92%
Precision	92%
Recall	92%

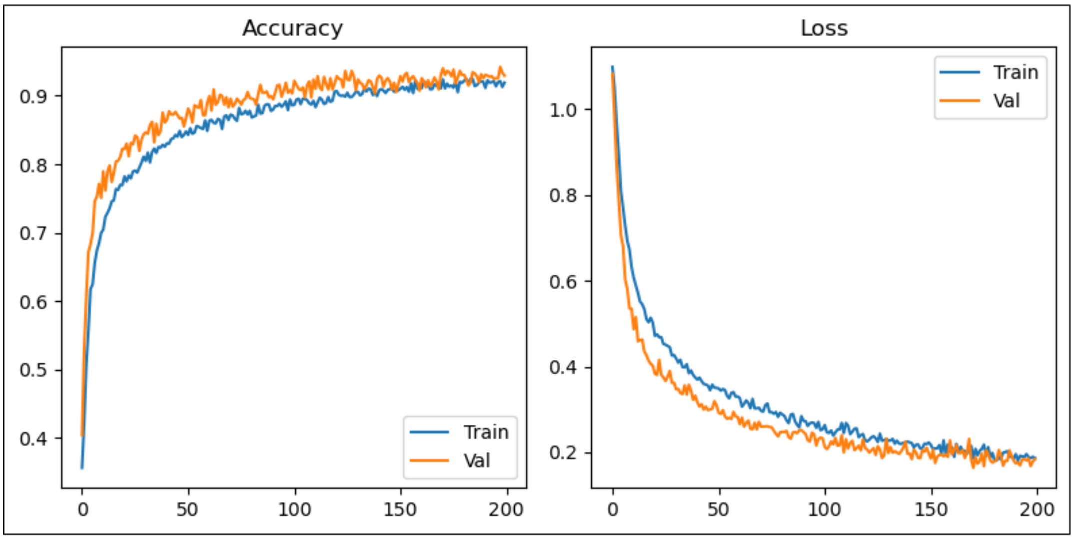

Figure 13: *Accuracy and loss for RTAS*

Figure 13 shows the accuracy and loss for RTAS. X-axis indicates the epochs and Y-axis indicates the accuracy. The accuracy value is trained using the Hawk-RNN to classify the model to get the higher accuracy of 92%. Here in loss X-axis shows the epochs and Y-axis shows the loss. The training loss is going on decreasing and the validation loss is also decreasing it is calculated up to 200 epochs.

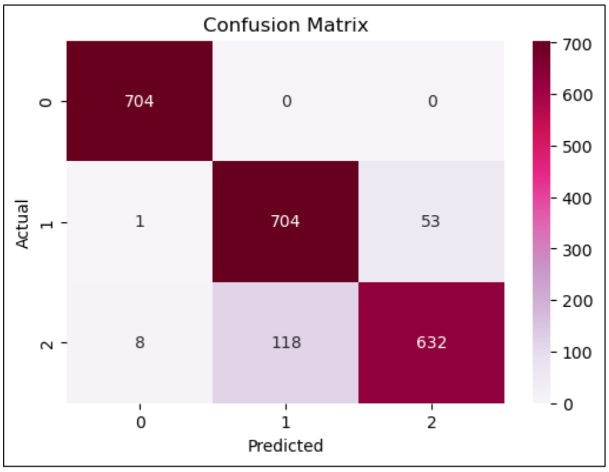

Figure 14: *Confusion matrix*

Figure 14 shows the confusion matrix for RTAS prediction system. Proposed method evaluates the model's performance for three classes. The class o and class 1 have the count of 704 and class 2 have the count of 632 respectively.

Figure 15 shows the multi-class ROC curve for proposed Hawk-RNN continues to have the highest AUC value of 99 for class 0 and 97 for class1 and class 2 demonstrating the overview and flexibility of RTAS approach.

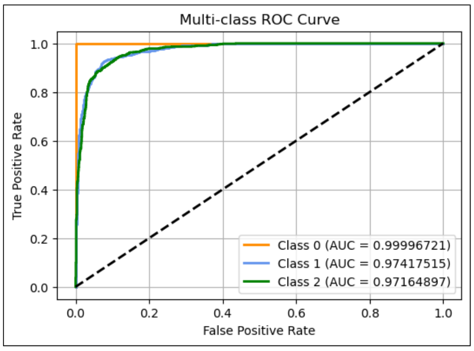

Figure 15: *Multi-class ROC curve*

4.2 Accuracy comparison for the Proposed Method

The comparison of the suggested method with existing method is presented in table 3. The RF [14] and ANN-PR [17] achieves an accuracy of 88.5%, 90% and the proposed Hawk-RNN achieves a higher accuracy of 92% compared to the existing method. The proposed Hawk-RNN have better performance for RTAS prediction system.

Table 3: *Comparison for the Proposed Method*

Method	Accuracy
RF	88.5%
ANN-PR	90%
Hawk-RNN	92%

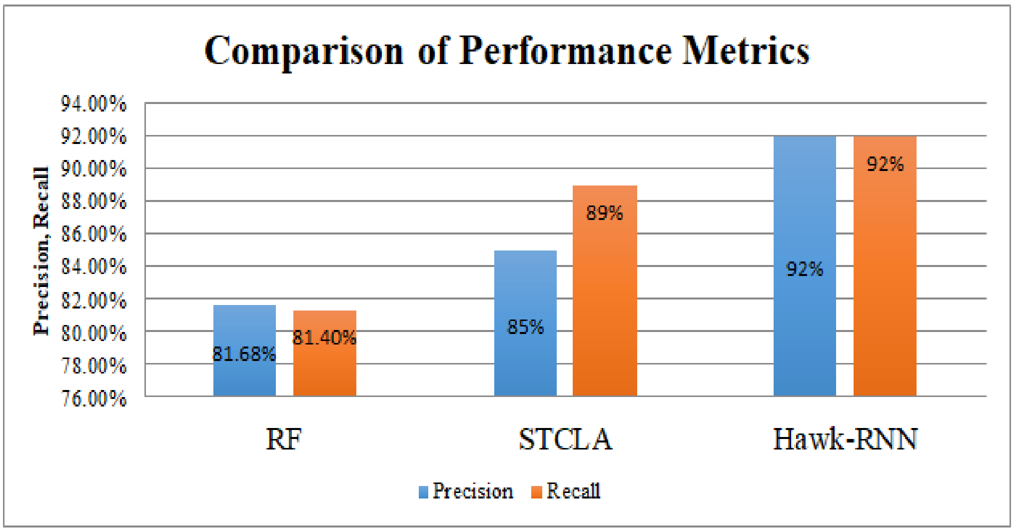

Figure 16: *Comparison of performance metrics*

Figure 16 shows the comparison of performance metrics for precision and recall value of proposed work and the existing work. The RF have precision value of 81.6% and recall value of 81.4% [14], STCLA have the precision value of 85% and recall of 89% [20] and proposed Hawk-RNN have the precision value of 92% and recall of 92% for RTAS prediction system. The proposed precision and recall value is high compared to the existing method.

5. Conclusion

In this study, the proposed Hawk-RNN is utilized to predict the RTAS to achieve high prediction accuracy. The data pre-processing find the missing values using handling missing values and the hidden patterns are predicted using EDA. The use of the SOA for feature selection and SMOTE for class balancing ensures high-quality inputs for the model. The novel Hawk-RNN architecture, enriched with hierarchical attention and wave kernel mechanisms, enables the system to capture temporal, contextual, and periodic patterns in the data effectively. The overall performance metrics the improved accuracy, precision and recall of 92%, confirm the superiority of the approach in predicting accident severity levels. This system enhances predictive capabilities and holds significant potential for deployment in intelligent transportation systems, enabling authorities to respond quickly and implement preventive safety measures.

References

1. Kannaiya Raja, Karthikeyan Kaliyaperumal, L. Velmurugan, and Subash Thanappan., Year: 2023, “Forecasting road traffic accident using deep artificial neural network approach in case of Oromia Special Zone”, Soft Computing, Vol: 27, No: 21, pp. 16179 – 16199.

2. Saeid Pourroostaei Ardakani, Xiangning Liang, Kal Tenna Mengistu, Richard Sugianto So, Xuhui Wei, Baojie He, and Ali Cheshmehzangi, Year: 2023, “Road car accident prediction using a machine-learning-enabled data analysis”, Sustainability, Vol: 15, No: 7, pp. 5939.

3. Omar Ibrahim Aboulola, Ebtisam Abdullah Alabdulqader, Aisha Ahmed AlArfaj, Shtwai Alsubai, and Tai-Hoon Kim, Year: 2024, “An automated approach for predicting road traffic accident severity using transformer learning and explainable AI technique”, IEEE Access.

4. Yang Yang, Kun He, Yun-peng Wang, Zhen-zhou Yuan, Yong-hao Yin, and Man-ze Guo, Year: 2022, "Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods", Physica A: Statistical Mechanics and Its Applications, Vol: 595. pp. 127083

5. Chuangang Cheng, Shuyan Chen, Yongfeng Ma, Fengxiang Qiao, and Zhuopeng Xie, Year: 2025, "Crash severity prediction and interpretation for road determinants based on a hybrid method", Journal of Transportation Safety & Security, Vol: 17, No: 1, pp. 30 – 56.

6. Elif Cicek, Murat Akin, Furkan Uysal, and ReyhanMerve Topcu Aytas, Year: 2023, "Comparison of traffic accident injury severity prediction models with explainable machine learning", Transportation letters, Vol: 15, No: 9, pp. 1043 – 1054.

7. Ibrahim Aldhari, Meshal Almoshaogeh, Arshad Jamal, Fawaz Alharbi, Majed Alinizzi, and Husnain Haider, Year: 2023, "Severity prediction of highway crashes in Saudi Arabia using machine learning techniques", Applied Sciences, Vol: 13, No: 1, pp. 233.

8. Ronghui Zhang, Yang Liu, Zihan Wang, Junzhou Chen, Qiang Zeng, Lai Zheng, Hui Zhang, and Yulong Pei, Year: 2025, "Innovative prediction and causal analysis of accident vehicle towing probability using advanced gradient boosting techniques on extensive road traffic scene data", Accident Analysis & Prevention, Vol: 211, pp. 107909.

9. Hussam Hijazi, Karim Sattar, Hassan M. Al-Ahmadi, Sami El-Ferik, Year: 2024, "Comparative Study for Optimized Deep Learning-Based Road Accidents Severity Prediction Models", Arabian Journal for Science and Engineering, Vol: 49, No: 4, pp. 5853 – 5873.

10. Mesut Ulu, Erdal Kilic, Yusuf Sait Türkan, Year: 2024, "Prediction of Traffic Incident Locations with a Geohash-Based Model Using Machine Learning Algorithms", Applied Sciences, Vol: 14, No: 2, pp. 725.

11. Luis Pérez-Sala, Manuel Curado, Leandro Tortosa, Jose F. Vicent, Year: 2023, "Deep learning model of convolutional neural networks powered by a genetic algorithm for prevention of traffic accidents severity", Chaos, Solitons & Fractals, Vol: 169, pp. 113245.

12. Shakil Ahmed, Md Akbar Hossain, Sayan Kumar Ray, Md Mafijul Islam Bhuiyan, Saifur Rahman Sabuj, Year: 2023, "A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance", Transportation research interdisciplinary perspectives, Vol: 19, pp. 100814.

13. Karim Sattar, Feras Chikh Oughali, Khaled Assi, Nedal Ratrout, Arshad Jamal, Syed Masiur Rahman, Year: 2023, "Transparent deep machine learning framework for predicting traffic crash severity", Neural Computing and Applications, Vol: 35, No: 2, pp. 1535 – 1547.

14. Chunyan Diao, Dafang Zhang, Wei Liang, Kuan-Ching Li, Yujie Hong, Jean-Luc Gaudiot, Year: 2022, "A novel spatial-temporal

multi-scale alignment graph neural network security model for vehicles prediction", IEEE Transactions on Intelligent Transportation Systems, Vol: 24, No: 1, pp. 904 – 914.

15. Fares Alhaek, Weichao Liang, Taha M. Rajeh, Muhammad Hafeez Javed, Tianrui Li, Year: 2024, "Learning spatial patterns and temporal dependencies for traffic accident severity prediction: A deep learning approach", Knowledge-Based Systems, Vol: 286, pp. 111406.

16. Zekun Yang, Wenping Zhang, Juan Feng, Year: 2022, "Predicting multiple types of traffic accident severity with explanations: A multi-task deep learning framework", Safety science, Vol: 146, pp. 105522.

17. Luigi Pio Prencipe, Simona De Bartolomeo, Leonardo Caggiani, Michele Ottomanelli, Year: 2025, "An injury severity-based methodology for assessing priority areas for shared micromobility accident risk mitigation", Travel Behaviour and Society, Vol: 39, pp. 100962.

18. Shadi Jaradat, Richi Nayak, Alexander Paz, Mohammed Elhenawy, Year: 2024, "Ensemble learning with pre-trained transformers for crash severity classification: A Deep NLP Approach", Algorithms, Vol: 17, No: 7, pp. 284.

19. Victor A Adewopo, Nelly Elsayed, Year: 2024, "Smart city transportation: Deep learning ensemble approach for traffic accident detection", IEEE Access.

20. Xu Wei, Year: 2024, "Enhancing road safety in internet of vehicles using deep learning approach for real-time accident prediction and prevention", International Journal of Intelligent Networks, Vol: 5, pp. 212 – 223.