International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

AI-Based Accident Severity Ranking and Emergency Response Automation System

Prithvi Nanjundan S[1], Dr. Pulkit Dwivedi[2]

[1]Department of Data Science & Artificial Intelligence, PES University, Bengaluru, Karnataka, India
Email: *prithvinanjundan[at]gmail.com*

[2]Department of Data Science & Artificial Intelligence, PES University, Bengaluru, Karnataka, India
Email: *pulkitdwivedi90[at]gmail.com*

Abstract: ***Road accidents require rapid assessment and coordinated communication with insurance and emergency services. This paper presents an AI-Based Accident Severity Ranking and Emergency Response Automation System that integrates image preprocessing, contextual accident-scene understanding, license-plate information extraction, deep-learning-based severity classification, severity-driven decision logic, and automated email notification. The system classifies an accident into Rank 1 (minor), Rank 2 (moderate or injury-related), or Rank 3 (severe) and maps the predicted rank to progressively broader response actions. The implementation uses a dataset of approximately 6,498 accident-related images with a 70:15:15 train-validation-test split. Phase 1 used MobileNetV2 as the baseline severity model, while Phase 2 introduced MobileNetV3 and a Gemini-3.6-Flash vision-language model for contextual scene description. The Phase 2 MobileNetV3 model achieved 93.35% accuracy, 94.26% precision, 93.35% recall, and a 93.80% F1-score. The integrated application combines the classifier output, accident-scene description, visible vehicle information, decision logic, a graphical user interface, and severity-specific email generation to provide an end-to-end accident-response workflow.***

Keywords: accident severity classification, MobileNetV3, vision-language model, emergency response automation, license plate recognition, intelligent transportation systems.

1. Introduction

Road accidents often require several activities to occur in sequence: the scene must be interpreted, accident severity must be determined, the involved vehicle should be identified when possible, and the appropriate insurance or emergency service must be informed. When these tasks are handled through separate manual processes, the resulting delay can be particularly important in cases involving injuries, fire, smoke, or multiple vehicles. The objective of this work is therefore not limited to accident detection; it is to connect accident understanding directly to severity-dependent action.

Recent accident-severity research has applied machine learning, deep learning, transformer models, and explainable artificial intelligence to structured crash records and sensor data [1]–[4], [9], [10], [12]. Fusion-based approaches have also shown the value of combining complementary learning models for severity prediction [13]. However, a prediction label by itself does not convey all accident-scene context required for operational response. The proposed framework extends severity classification with vision-language-based scene interpretation, vehicle registration extraction, a severity-to-service decision engine, and automated email communication.

The system uses three operational severity levels. Rank 1 represents minor visible damage and primarily initiates insurance communication. Rank 2 represents more significant or injury-related accidents and extends the response to ambulance or hospital services. Rank 3 represents major accidents involving conditions such as extensive damage, multiple vehicles, fire, smoke, or other severe indicators and can trigger coordinated communication with insurance, medical, police/traffic, and fire services. This mapping makes the classification output directly actionable within the application.

The main contribution of the work is an integrated pipeline that combines quantitative severity prediction with contextual accident-scene description and response automation. Phase 2 further replaces the MobileNetV2 baseline with MobileNetV3 and adds Gemini-3.6-Flash for detailed accident-scene description. The resulting application provides image input, analysis, result visualization, service selection, and email-notification status in a single graphical workflow.

2. Related Work

Road-accident severity prediction has traditionally relied on structured attributes such as road type, driver characteristics, environmental conditions, vehicle attributes, and crash location. Recent work has expanded this direction through transformer learning and explainable AI [1], machine-learning analysis of driver, environmental, and road conditions [2], improved severity-prediction models [3], and comparative machine-learning evaluation on regional crash data [4]. Review studies similarly report broad adoption of machine learning for traffic-accident severity prediction [7].

Class imbalance is an important issue because severe crashes are often less frequent than minor events. Amiri et al. specifically examine accident-severity prediction under class imbalance [10], while explainability-based feature identification has also been studied to improve interpretation of severity models [9]. In safety-oriented applications, aggregate accuracy alone is therefore insufficient; class-wise precision, recall, F1-score, and the confusion matrix are needed to identify critical underestimation errors.

Deep and hybrid learning approaches have also been explored. Shafique et al. study machine-learning prediction of sensor-related accident severity for autonomous-vehicle safety [12], and Manzoor et al. present a decision-level fusion framework combining machine- and deep-learning models [13]. Hospital recommendation has additionally been linked with severity analysis in deep-learning-based work [6]. These studies motivate accurate severity prediction but do not by themselves provide the complete image-to-context-to-response workflow targeted here.

The present system therefore focuses on the research gap between accident analysis and operational response. It combines visual preprocessing, scene description, vehicle identification, severity classification, severity-based agency selection, and notification generation. This multimodal and action-oriented design is intended to provide more useful accident information than a prediction-only system.

3. System Design and Methodology

a) *Overall Architecture*

The implemented architecture is modular. Accident images or frames are accepted through the user interface and validated before preprocessing. The processing pipeline includes image enhancement and normalization, YOLO-based object-level accident analysis, feature extraction, contextual description generation using a cloud-based vision-language model, severity classification, information fusion, decision making, and email notification. Downstream modules support report generation, storage, monitoring, and stakeholder-facing output.

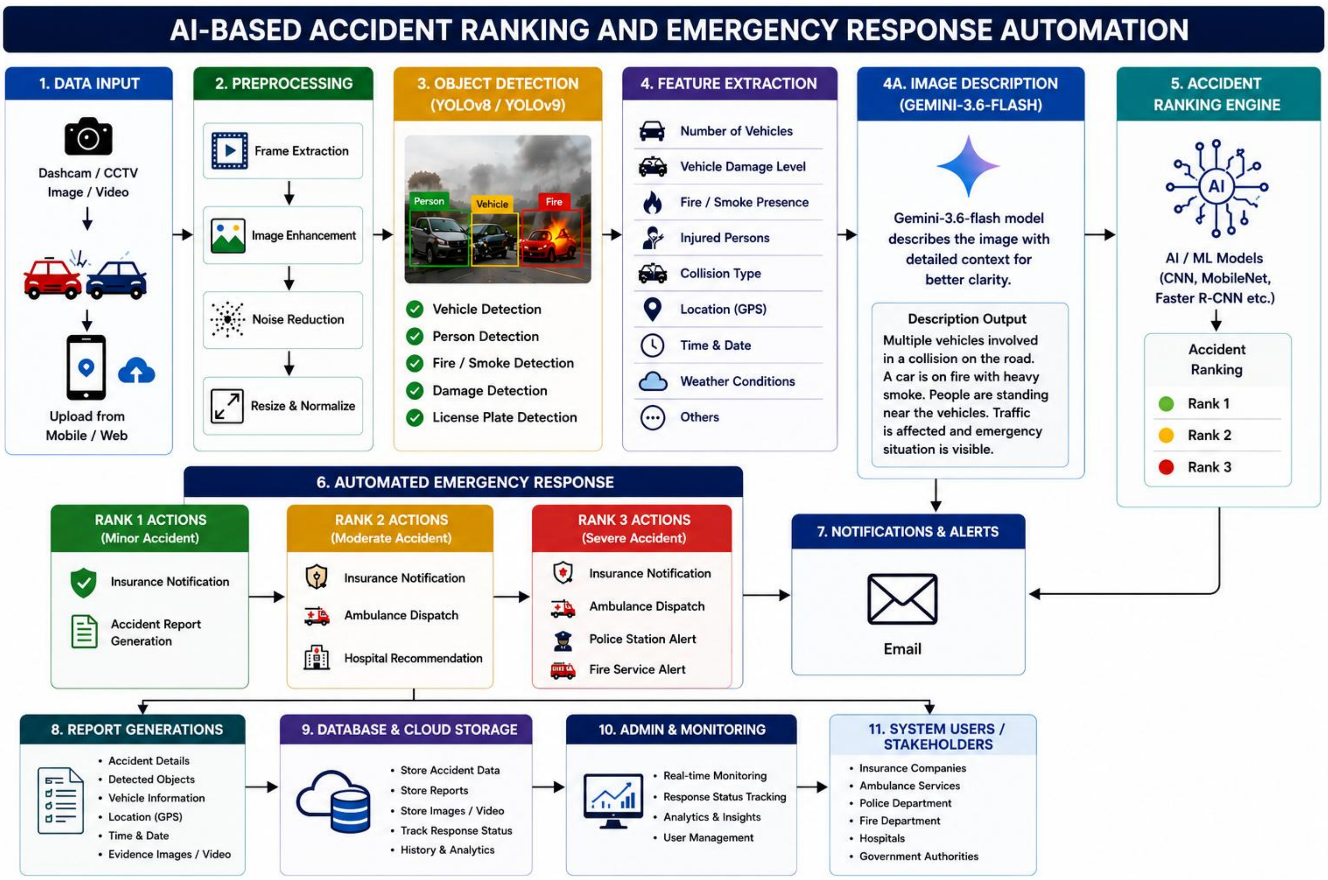

Figure 1: Overall architecture of the AI-based accident ranking and emergency response automation system

b) *Dataset and Severity Definition*

The project uses approximately 6,498 accident-related images containing accident scenes, vehicle damage, fire or smoke conditions, and injury-related situations. The images are divided into training, validation, and test subsets in a 70%, 15%, and 15% ratio, respectively. The test subset is kept independent for final model evaluation. Severity labels are assigned using observable accident characteristics so that each class has a direct operational meaning.

Table I: Severity Ranks and Primary Response Mapping

Rank	Interpretation	Primary response
Rank 1	Minor accident	Insurance
Rank 2	Moderate / injury-related	Insurance + Ambulance/Hospital
Rank 3	Major / severe accident	Insurance + Ambulance/Hospital + Police/Traffic + Fire

c) *Image Preprocessing*

Accident images can vary in resolution, illumination, contrast, compression, camera angle, and noise. Preprocessing therefore standardizes the input before model inference. The implemented workflow includes image validation, resizing to the model input size, denoising, contrast or image enhancement where appropriate, and normalization. Data augmentation and class balancing can be applied only to the training subset to reduce class imbalance without leaking augmented or near-duplicate samples into validation or test data.

Care is required because aggressive enhancement can amplify noise or alter relevant visual evidence. The selected preprocessing configuration should therefore be based on validation and test behavior rather than the number of operations applied. The project also identifies low illumination, blur, occlusion, and domain variation as important sources of error.

d) Vision-Language Accident Scene Description

A vision-language model (VLM) is used to generate human-readable contextual information that complements the numerical severity prediction. During development, Qwen2.5-VL, LLaMA-based vision models, BLIP, and a cloud-based Gemini model were considered. The integrated Phase 2 workflow uses Gemini-3.6-Flash because it provided a suitable combination of accident-context description quality and response speed during project experimentation.

The VLM prompt is structured to focus on accident-relevant evidence rather than generic scene captioning. Requested fields include the number of visible vehicles, vehicle damage, persons, possible injury-related visual evidence, fire, smoke, collision characteristics, and an overall scene summary. The generated text is then converted into structured attributes such as vehicle count, damage status, person visibility, fire, smoke, and collision condition. Explicit target labels such as "Rank 3" should be excluded from training descriptions when the text is used as model input, preventing the classifier from learning the label directly rather than the underlying accident evidence.

e) Severity Classification

Phase 1 used MobileNetV2 as the baseline severity classifier. Phase 2 replaces it with MobileNetV3 to improve classification performance while preserving the lightweight characteristics desirable for future deployment. For each input, the final classification layer produces probabilities for the three accident ranks. The predicted class is the class with the maximum softmax probability

$$p_i = exp(z_i) / \Sigma_j \, exp(z_j), \quad i \in \{1, 2, 3\}.$$

The classification output is evaluated using accuracy, precision, recall, F1-score, and confusion-matrix analysis. Rank 2 and Rank 3 recall are especially important because underestimating a severe accident can suppress required emergency escalation. The model-selection process therefore considers both predictive performance and suitability for integration with the VLM and notification pipeline.

f) License Plate Extraction and Information Fusion

When a registration plate is sufficiently visible, the system localizes the plate region, crops it, performs suitable preprocessing, recognizes the characters, and normalizes the result into a usable registration string. If the plate is not visible or is too blurred, occluded, damaged, or strongly angled, the system records the information as unavailable rather than generating an unsupported registration number. Plate information serves a different purpose from severity classification: it provides vehicle identity for reports and service-specific notifications.

Before decision making, the application fuses the available severity class and confidence with the VLM description, structured accident indicators, license-plate information, and fire/smoke or person-related evidence. The resulting structured record is forwarded to the decision engine.

g) Severity-Based Decision and Notification Logic

The decision engine converts the severity prediction into service- specific action. Rank 1 selects insurance communication. Rank 2 selects insurance and medical response. Rank 3 selects insurance, medical, and police/traffic response, while fire-service notification can be conditioned on fire, smoke, or another configured major hazard. The use of contextual evidence prevents all severe cases from receiving an identical generic alert when a service-specific condition is absent.

Separate email templates are maintained for insurance, ambulance/hospital, police/traffic, and fire services. Each template includes only information relevant to the receiving agency, such as registration number, accident rank, damage summary, visible persons, fire/smoke status, collision description, and notification status. The system is intended as decision support and emergency notification; it does not infer medical diagnoses from images.

4. Implementation

The application is implemented primarily in Python 3.10 or above. OpenCV is used for image preprocessing; PyTorch or TensorFlow and Torchvision support model execution; NumPy and Pandas support numerical and structured data handling; Scikit-learn is used for accuracy, precision, recall, F1-score, and confusion-matrix evaluation; an OCR or plate-recognition module provides registration extraction; Gmail/SMTP is used for email delivery; and a GUI framework integrates image selection, pipeline execution, results, logs, and notification status.

The GUI makes the system an integrated application rather than a collection of independent scripts. The user selects an accident image, verifies the preview, starts analysis, and then receives the generated scene description, available registration information, severity prediction, confidence, selected response services, and email status in one interface. Fig. 2 shows the completed accident-analysis dashboard used in the project.

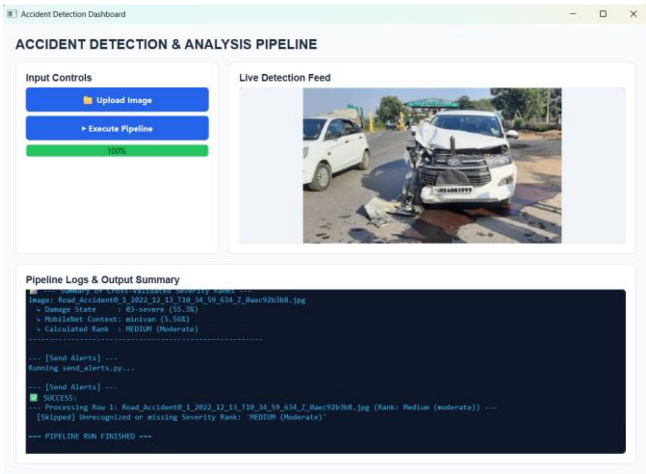

Figure 2: Completed accident-analysis and emergency-response pipeline in the integrated GUI.

5. Results and Discussion

a) *Severity Classification Results*

The classification framework was evaluated in two stages. In Phase 1, MobileNetV2 provided the strongest baseline performance among the evaluated models. Phase 2 introduced MobileNetV3 and retained AlexNet, InceptionV3, and Xception as comparison architectures. The reported Phase 2 MobileNetV3 result is 93.35% accuracy, 94.26% precision, 93.35% recall, and 93.80% F1-score.

Table II: Classification Performance Reported in the Project

Phase	Model	Accuracy	Precision	Recall	F1-score
1	MobileNetV2	70.73%	92.12%	70.73%	76.68%
1	AlexNet	70.15%	92.11%	70.15%	76.24%
1	InceptionV3	69.55%	92.07%	69.55%	75.76%
1	Xception	69.33%	92.05%	69.33%	75.57%
2	MobileNetV3	93.35%	94.26%	93.35%	93.80%

Relative to the Phase 1 MobileNetV2 baseline, Phase 2 improves accuracy and recall by 22.62 percentage points, F1-score by 17.12 percentage points, and precision by 2.14 percentage points. The relative accuracy increase from 70.73% to 93.35% is approximately 31.98%. These results indicate a substantial improvement in the measured severity-classification stage.

The confusion matrix remains important because the operational cost of errors is asymmetric. A Rank 3 → Rank 1 error is more critical than a Rank 1 → Rank 2 error because the former can suppress emergency services for a major event. The report therefore emphasizes class-wise metrics and Rank 2/Rank 3 recall. However, final numerical confusion-matrix counts were not supplied in the project document and are not inferred here.

b) *Vision-Language and Structured Scene Understanding*

The VLM extends the application beyond a class-and-confidence output. A useful description can state that two vehicles are involved in a frontal collision, identify extensive front-end deformation, and explicitly indicate whether fire or smoke is visible. These observations can be transformed into structured fields that are usable by the decision engine. The project evaluates VLM candidates using accident-scene understanding, vehicle information, damage description, person/injury-related description, fire/smoke recognition, relevance, response time, and structured-extraction success.

Gemini-3.6-Flash is selected in Phase 2 for the integrated workflow. The project document provides qualitative selection evidence but does not provide final measured average VLM response time or a field-level correctness score. Consequently, the present paper reports the integration and selection criteria without inventing quantitative VLM metrics. A rigorous future evaluation should report structured-field accuracy across a held-out set together with end-to-end latency.

c) *Decision and Notification Validation*

Table III: Severity-To-Service Decision Mapping

Rank	Insurance	Amb./ Hospital	Police/ Traffic	Fire
Rank 1	Yes	No	No	No
Rank 2	Yes	Yes	No / Conditional	No / Conditional
Rank 3	Yes	Yes	Yes	Conditional / Yes when fire or major hazard exists

End-to-end scenario validation follows the same operational logic. A minor accident should generate only insurance communication; a moderate or injury-related accident should generate insurance and ambulance/hospital communication; and a major accident with fire or smoke should generate the coordinated Rank 3 response. This separates decision-logic validation from classifier accuracy because the service mapping can also be tested with manually supplied ranks.

The final application demonstrates integration of input handling, accident analysis, severity ranking, contextual output, and email generation within a single GUI. The report includes generated email evidence for service-specific alerts. Quantitative delivery-time measurements and plate-recognition accuracy are not provided as final measured results, so they remain items for future experimental reporting.

d) *Failure Cases and Practical Limitations*

The system is affected by the visibility limits of image-based analysis. Poor lighting can reduce classifier, VLM, and plate performance; motion blur can remove useful features; an occluded or damaged plate can make vehicle identity unavailable; complex multi-vehicle scenes can create ambiguous associations; reflections, dust, or exhaust can be confused with fire or smoke; and accidents near the Rank 1/Rank 2 or Rank 2/Rank 3 boundaries can produce low-confidence predictions. These failure modes should be reported together with successful cases rather than hidden by aggregate accuracy.

Cloud-based VLM inference also introduces Internet dependency and external API latency. In addition, the system can only use information observable in the supplied image; a person may be injured inside a vehicle without being visible. For these reasons, the framework should support rather than replace emergency professionals and should expose the underlying contextual observations alongside the predicted rank.

6. Conclusion and Future Work

This work presents an end-to-end accident severity ranking and emergency response automation framework integrating image preprocessing, contextual scene understanding, license-plate extraction, deep-learning severity classification, severity-based decision making, email notification, and GUI-based result presentation. The Phase 2 MobileNetV3 implementation reports 93.35% accuracy, 94.26% precision, 93.35% recall, and 93.80% F1-score, improving substantially over the Phase 1 MobileNetV2 baseline. The integration of Gemini-3.6-Flash adds human-readable accident context such as vehicle damage, vehicle count, persons, fire, smoke, and collision characteristics, which complements the numerical severity prediction.

Future work will extend the current image-based system to live dashcam or CCTV video, investigate an offline VLM for improved availability, privacy, and reduced cloud dependency, incorporate vehicle-specific and NCAP-related safety information into severity assessment, and add SMS and WhatsApp alongside email for multi-channel communication. Further evaluation should also report class-wise confusion-matrix values, plate-recognition accuracy, structured VLM field correctness, notification-delivery success rate, and end-to-end latency on a fixed test protocol.

References

[1] O. I. Aboulola, E. A. Alabdulqader, A. A. Alarfaj, S. Alsubai, and T.-H. Kim, "An automated approach for predicting road traffic accident severity using transformer learning and explainable AI technique," IEEE Access, vol. 12, pp. 61062–61072, 2024, doi: 10.1109/ACCESS.2024.3380895.

[2] A. K. Mengistu, A. E. Gedefaw, N. D. Baykemagn, A. D. Walle, T. Z. Yehuala, M. A. Alemayehu, M. A. Messelu, et al., "Predicting car accident severity in Northwest Ethiopia: A machine learning approach leveraging driver, environmental, and road conditions," Scientific Reports, vol. 15, Art. no. 21913, 2025, doi: 10.1038/s41598-025-08005-2.

[3] J. Tang, Y. Huang, D. Liu, L. Xiong, and R. Bu, "Research on traffic accident severity level prediction model based on improved machine learning," Systems, vol. 13, no. 1, Art. no. 31, 2025, doi: 10.3390/systems13010031.

[4] I. Mohamad, S. Jomnonkwao, and V. Ratanavaraha, "Machine learning predictive performance in road accident severity: A case study from Thailand," Results in Engineering, vol. 26, Art. no. 104833, 2025, doi: 10.1016/j.rineng.2025.104833.

[5] M. Banson, M. K. Ametefe, K. Darko, J. C. B. Dakubo, M. Iddrisu, and T. Dakurah, "Comparative analysis of traumatic brain injury severity in motorcycle and car accident victims treated at Korle-Bu Teaching Hospital," Journal of the West African College of Surgeons, vol. 15, pp. 75–82, 2025, doi: 10.4103/jwas.jwas_188_23.

[6] A. R. Gowthami, G. Abhilash, M. L. Tripura, K. V. Krishna, and A. A. K. Reddy, "Road accident severity & hospital recommendations using deep learning techniques," Fringe Multi-Engineering Proceedings, vol. 1, no. 4, pp. 54–64, 2025, doi: 10.69996/fmep.2025030.

[7] N. Hamdan and T. Sipos, "Advancements in machine learning for traffic accident severity prediction: A comprehensive review," Periodica Polytechnica Transportation Engineering, vol. 53, no. 3, pp. 347–355, 2025, doi: 10.3311/PPtr.40369.

[8] W. Cao, X. Wang, Y. Feng, J. Zhou, and Z. Yang, "Improving maritime accident severity prediction accuracy: A holistic machine learning framework with data balancing and explainability techniques," Reliability Engineering & System Safety, vol. 266, Part A, Art. no. 111648, 2026, doi: 10.1016/j.ress.2025.111648.

[9] J. Alotaibi, "Enhancing traffic accident severity prediction: Feature identification using explainable AI," Vehicles, vol. 7, no. 2, Art. no. 38, 2025, doi: 10.3390/vehicles7020038.

[10] M. A. Amiri, S. Afshari, and A. Soltani, "Machine learning approaches to traffic accident severity prediction: Addressing class imbalance," Machine Learning with Applications, vol. 22, Art. no. 100792, 2025, doi: 10.1016/j.mlwa.2025.100792.

[11] Y. Jiang, X. Qu, W. Zhang, W. Guo, J. Xu, W. Yu, and Y. Chen, "Analyzing crash severity: Human injury severity prediction method based on transformer model," Vehicles, vol. 7, no. 1, Art. no. 5, 2025, doi: 10.3390/vehicles7010005.

[12] R. Shafique, F. Rustam, S. Murtala, A. D. Jurcut, and G. S. Choi, "Advancing autonomous vehicle safety: Machine learning to predict sensor-related accident severity," IEEE Access, vol. 12, pp. 25933–25948, 2024, doi: 10.1109/ACCESS.2024.3366990.

[13] M. Manzoor, M. Umer, S. Sadiq, A. Ishaq, S. Ullah, H. A. Madni, and C. Bisogni, "RFCNN: Traffic accident severity prediction based on decision level fusion of machine and deep learning model," IEEE Access, vol. 9, pp. 128359–128371, 2021, doi: 10.1109/ACCESS.2021.3112546.