



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: <https://doi.org/10.22214/ijraset.2025.71886>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Predicting the Severity of Road Accidents Using Deep and Transfer Learning Approaches

Ms. P. Sudha¹, A. P. Vanitha²

¹AP/IT, Department of Information Technology, Vivekanandha College of Engineering for Women, Tiruchengode, Tamil Nadu, India

²PG Scholar, Department of Information Technology, Vivekanandha College of Engineering for Women, Tiruchengode, Tamil Nadu, India

Abstract: Road traffic accidents are a major concern worldwide, often resulting in serious injuries and fatalities. Rapid and accurate assessment of accident severity can significantly aid emergency response and resource allocation. This project presents a comprehensive, automated system for predicting the severity of road traffic accidents using both tabular (structured) and image (unstructured) data. For tabular data, models including Multilayer Perceptron (MLP), XGBoost and TabNet were implemented to classify severity levels. For image-based classification, deep convolutional neural networks such as MobileNetV2 and ResNet50 were used, leveraging transfer learning to enhance performance. An EfficientNetB7 model was employed to auto-label unlabelled accident images, improving dataset consistency and reducing manual labelling effort. The system was evaluated using accuracy, precision, recall and F1-score, with performance visualized through confusion matrices and bar charts. The results demonstrate high classification accuracy across all models. This intelligent, multi-input framework serves as a practical tool for real-time accident severity assessment, with potential applications in emergency response systems and smart traffic management.

Index Terms: Road Traffic Accident Severity, Deep Learning, EfficientNetB7, Multilayer Perceptron (MLP), ResNet, MobileNet, Auto-labeling, Transformer Models, Real-time Accident Prediction, Road Safety, Data Imputation, Behavioural Analytics.

I. INTRODUCTION

Road traffic accidents (RTAs) pose a serious threat to public safety, resulting in substantial loss of life and property every year. The ability to accurately predict the severity of an accident at the time of occurrence can significantly enhance emergency response, hospital preparedness, and overall traffic management. Traditional methods of severity assessment often rely on manual reporting or limited data inputs, leading to delays and inconsistencies. With advancements in machine learning and deep learning, it is now possible to build intelligent systems that can analyze diverse data sources such as structured accident records and real-time accident scene images to automatically classify the severity level of a crash. This project explores a multimodal approach to accident severity prediction by combining tabular data (e.g., time, location, weather conditions) and image data (e.g., vehicle damage) for more accurate and reliable assessment. The proposed system utilizes powerful models such as MLP, XGBoost and TabNet for tabular data analysis, while MobileNetV2 and ResNet50 are used for classifying accident images.

Additionally, EfficientNetB7 is employed for auto-labeling images when severity annotations are missing, reducing manual effort and improving data quality. The system is evaluated using key classification metrics such as accuracy, precision, recall and F1 score. Visual comparisons and confusion matrices further validate the model's performance. By integrating machine learning with real-time data processing and visual interpretation, this work provides a scalable, automated framework for road accident severity prediction, with potential applications in intelligent transport systems, emergency management and smart city initiatives.

II. LITERATURE REVIEW

[1] Omar Ibrahim Aboulola focused on addressing the ongoing issue of road traffic accidents by developing and evaluating multiple transfer learning-based predictive models. Techniques such as MLP, CNN, LSTM, ResNet, EfficientNetB4, InceptionV3, Xception and MobileNet were assessed, with MobileNet demonstrating the highest accuracy of 98.17%. The study also utilized Shapley values for interpretability, offering valuable insights into the features influencing accident severity and enhancing transparency in model predictions for informed safety interventions. [2] Nuntaporn Klinjun investigated Thailand's high rate of road fatalities, revealing that over 66% of deaths occur at the crash site. Analyzing forensic data from 25 critical incidents involving 407 individuals and 47 vehicles, the study applied the Haddon matrix to identify key risk contributors such as excessive speed, driver fatigue, non-use of seat belts, overloaded or modified vehicles and poor road design.

The research advocated for stricter traffic laws and more comprehensive road safety policies. [3] Badr Ben Elallid reviewed recent developments in autonomous vehicle systems powered by AI and intelligent transportation technologies. Emphasizing deep learning and reinforcement learning, the study proposed a structured taxonomy of AI methods for critical AV tasks like perception, planning and decision-making. It highlighted the advantages of AI over traditional approaches and discussed current limitations and future research opportunities to improve autonomous mobility systems. [4] Linchao Li addressed the complexities of detecting road incidents using limited real-time data and highly imbalanced samples. The study proposed a hybrid architecture combining Generative Adversarial Networks (GANs) for balancing the dataset and a Temporal-Spatially Stacked Autoencoder (TSSAE) to model traffic flow dynamics. Tested on real highway data from California's I-80, the approach showed enhanced detection performance, with recommendations to integrate weather data and improve explainability in future versions. [5] Karishma Pawar developed a deep learning-based traffic monitoring system using a one-class classification approach to detect road accidents from surveillance videos. The framework, built with spatio-temporal and LSTM-based autoencoders, was trained on normal traffic behavior to identify deviations (accidents) as anomalies. It achieved strong results on standard datasets and outlined future plans for hardware deployment on PYNQ boards and adaptive learning to adjust to evolving traffic patterns. [6] Elif Cicek evaluated multiple machine learning classifiers Decision Trees, MLP, SVC, Case-Based Reasoning and Naive Bayes on accident injury severity data from NHTSA. The MLP model outperformed the others, especially in recognizing key risk factors like alcohol use, lack of seat belt usage and speeding. The research noted class imbalance issues and leveraged both Chi-square testing and Shapley values for feature relevance, calling for more balanced datasets and interpretable models. [7] Jing Gan employed the Deep Forest algorithm, an ensemble of decision trees, to classify accident severity levels using the UK's road safety dataset. This model demonstrated better stability, reduced hyperparameter sensitivity, and high accuracy across datasets of varying sizes. The approach was proposed as a promising alternative for related applications such as travel time prediction and traffic pattern analysis, with further improvements expected in feature optimization. [8] Ibrahim Aldhari analyzed traffic incidents in Saudi Arabia's Qassim Province over a two-year span, aiming to forecast injury severity. Among the models tested Random Forest, Logistic Regression and XGBoost the XGBoost classifier performed best, achieving 71% accuracy for multi-class and 94% for binary classification. SHAP values revealed influential factors like road types, lighting, accident time, and crash outcomes. The study suggested hybrid neural networks for future enhancements. [9] Maher Ibrahim Sameen designed an optimized RNN model for predicting accident severity on Malaysia's North-South Expressway, using six years of crash records. Combining LSTM with dense layers, the model achieved a validation accuracy of 71.77%, outperforming both MLP and Bayesian Logistic Regression. Techniques such as dropout and SGD were used to improve generalization. A sensitivity analysis showed time of day and road direction as critical predictors, with future plans to incorporate spatial-temporal context. [10] Arshad Jamal compared the predictive power of XGBoost with traditional models like Decision Trees, Random Forest and Logistic Regression using a dataset of over 13,000 rural highway crashes in Saudi Arabia. XGBoost led the performance rankings with 93% accuracy, ahead of Decision Trees (88%), Random Forest (84%) and Logistic Regression (63%). The key predictors included vehicle type, road surface, crash type and weather. The study emphasized addressing data imbalance and including demographic factors for deeper crash analysis.

III. METHODOLOGIES

A. Data Collection Module

The Data Collection module gathers two types of inputs: (1) a tabular dataset in CSV format containing structured attributes such as accident time, location, weather conditions, vehicle type and road surface; and (2) an image dataset composed of real-world accident scene images stored in train, test and validation subfolders. These datasets are verified, loaded into memory, and prepared for processing.

B. Data Preprocessing Module

In the Data Preprocessing module, the tabular data undergoes cleaning by filling missing numerical values with median values and categorical values with a constant such as "missing". Categorical features are encoded using One-Hot Encoding, while labels (e.g., severity levels like low, medium, high) are encoded using Label Encoding. Features are then standardized using the formula $x' = (x - \mu) / \sigma$, ensuring consistent input ranges across all models. For images, preprocessing includes resizing (typically to 96×96 or 128×128), normalizing pixel values by dividing by 255 and applying data augmentation techniques like flipping, rotation and zooming to enhance model generalization. If the image dataset lacks severity labels, the system employs the EfficientNetB7 model to auto-label each image based on learned patterns.

C. Data Splitting and Feature Engineering module

The Data Splitting and Feature Engineering module partitions the dataset into training, validation and test sets usually at a 70:15:15 ratio using stratified sampling to maintain class balance. Feature engineering then extracts additional features such as Accident_Hour from time-related fields, or interactions between weather and road type. These engineered features are merged with the original data using horizontal stacking or sparse matrices depending on dimensionality.

D. Training and Testing Module

The Training and Testing module plays a central role in developing and validating models that predict the severity of road traffic accidents. In this module, both tabular and image-based datasets are used to train multiple machine learning and deep learning models. For tabular data, models such as Multilayer Perceptron (MLP), XGBoost and TabNet are employed. MLP is a feedforward neural network capable of learning complex patterns in structured data. XGBoost is a powerful ensemble technique based on decision trees that improves prediction accuracy through boosting. TabNet, on the other hand, is a deep learning model specifically designed for tabular data, utilizing attention mechanisms to select relevant features dynamically, enhancing both performance and interpretability. For image data, transfer learning techniques are applied using pre-trained convolutional neural networks such as MobileNetV2 and ResNet50. These models are fine-tuned on the accident image dataset to classify severity levels based on visual features. During training, the models learn to map input features or image patterns to the target classes (e.g., low, medium, high severity), while during testing, their performance is evaluated on a separate, unseen dataset to assess generalization. The training process also includes mechanisms such as early stopping and learning rate adjustments to prevent overfitting and optimize learning efficiency. This module ensures that the final models are both accurate and robust, ready for deployment in real-time or large-scale accident severity prediction systems.

E. Evaluation and Visualization Module

Finally, the Evaluation and Visualization module measures model performance through standard classification metrics.

Metric	Formula
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
F1 Score	$\frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$

Results from all models are visualized using bar charts and confusion matrices, allowing side-by-side performance comparisons for both tabular and image-based classification tasks. These modules collectively form a powerful, end-to-end predictive system capable of real-time accident severity assessment, offering significant value to emergency response systems, smart city infrastructure, and road safety policy planners.

IV. DATA MODELS

A. Multilayer Perceptron (MLP)

The Multilayer Perceptron (MLP) is a type of feedforward neural network that consists of an input layer, one or more hidden layers and an output layer. It is suitable for handling structured data and solving classification or regression problems. The input features are passed through fully connected layers. Each neuron computes a weighted sum of its inputs, adds a bias and applies an activation function (e.g., ReLU, sigmoid). The final layer outputs probabilities (via softmax) for multi-class classification. The network is trained using backpropagation to minimize a loss function (e.g., categorical cross-entropy). Each layer in an MLP performs the operation: $h^{(l)} = f(W^{(l)} \cdot h^{(l-1)} + b^{(l)})$ where $h^{(l)}$ is the output of layer l , $W^{(l)}$ and $b^{(l)}$ are weights and biases and f is the activation function.

B. XGBoost

XGBoost is a scalable, optimized gradient boosting framework that uses decision tree ensembles to solve supervised learning problems with high accuracy and speed. Models are built sequentially, each learning from the errors of the previous one. Trees are added in a forward stage-wise fashion to minimize the loss. Regularization is applied to control model complexity and prevent overfitting. Objective Function:

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t)}) + \sum_{k=1}^t \Omega(f_k)$$

$$\text{where } \Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_j w_j^2,$$

T is the number of leaves and w_j are leaf weights. This regularized loss improves generalization.

C. TabNet

TabNet is a deep learning model designed specifically for tabular data. It employs a sequential attention mechanism to choose which features to focus on during decision-making. Applies sparse feature selection using attention masks at each decision step. Transforms the selected features via a shared feature transformer network. Combines outputs across multiple decision steps to form the final prediction. Provides interpretability by identifying which features influenced predictions most. TabNet computes feature masks M_i using soft attention: $M_i = \text{Sparsemax}(P_i - 1 \cdot H_i)$ where $P_i - 1$ is the prior feature importance and H_i is the output from the attentive transformer. The final prediction aggregates transformations from all steps.

D. MobileNetV2

MobileNetV2 is a lightweight convolutional neural network designed for mobile and real-time applications. It balances accuracy and computational efficiency using depthwise separable convolutions and linear bottlenecks. Input image is passed through a stack of inverted residual blocks. Each block includes a depthwise convolution followed by a pointwise convolution. Linear bottlenecks preserve low-dimensional representations. Global average pooling is applied before classification. Key operation: $\text{Output} = \text{ReLU6}(\text{BN}(\text{ConvDW}(x))) \rightarrow \text{BN}(\text{ConvPW}(x))$ Where ConvDW is depthwise convolution and ConvPW is pointwise convolution. The architecture reduces parameters while maintaining high accuracy.

E. ResNet50 (Residual Network)

ResNet50 is a deep convolutional neural network known for its residual connections, which allow it to train very deep architectures by solving the vanishing gradient problem. Each block learns a residual function $F(x)$ rather than directly learning the output. The residual is added to the block's input via skip connections. Deep layers are stacked without degrading performance, enabling better feature extraction. Residual block computes: $y = F(x, \{W_i\}) + x$ where $F(x)$ is the residual function and x is the identity input. This helps gradients flow efficiently during training.

F. EfficientNetB7

EfficientNetB7 is part of the EfficientNet model family, which scales CNNs uniformly in depth, width and resolution using a compound coefficient for optimal accuracy and efficiency. Uses mobile inverted bottleneck blocks and squeeze-and-excitation modules. The compound scaling formula ensures balanced increases in model dimensions. Designed to achieve state-of-the-art performance on image classification with fewer parameters. Scaling formula: $\text{depth} = \alpha^\phi$, $\text{width} = \beta^\phi$, $\text{resolution} = \gamma^\phi$ subject to $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$, where ϕ is the compound coefficient. EfficientNetB7 uses the highest scaling in the family, making it accurate but computationally intensive.

V. COMPARATIVE ANALYSIS OF TABULAR AND IMAGE-BASED MODELS

The system incorporates and compares multiple models trained on tabular and image datasets for road traffic accident severity prediction. For tabular data, the models MLP, XGBoost, and TabNet show exceptionally high performance. As seen in the bar chart, both MLP and XGBoost achieved perfect scores across all metrics (accuracy, precision, recall and F1-score), with values reaching 1.000. TabNet closely followed, with marginally lower recall and F1-score values of 0.972 and 0.985 respectively. These results demonstrate that traditional and deep learning models for structured data can yield near-perfect classification under well-prepared data conditions and proper feature engineering.

For image classification, the performance of MobileNetV2 and ResNet50 was comparatively lower. Both models performed similarly in terms of accuracy and precision, reaching approximately 0.78–0.79. However, precision for both models was significantly lower (around 0.60), suggesting challenges in differentiating severity classes visually. Despite this, recall and F1-scores remained reasonably balanced, highlighting the model's general ability to identify critical features in accident scenes. MobileNetV2 performed slightly better in generalization, while ResNet50 maintained consistency across classes due to its deeper architecture.

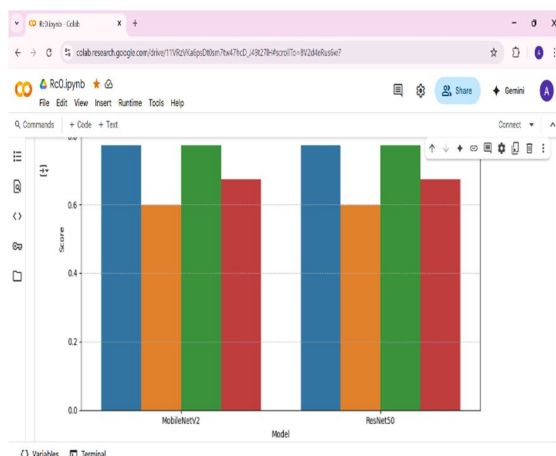


Fig.1 Comparative barchart of image models

In summary, tabular models demonstrated superior performance in this multimodal framework, with XGBoost and MLP achieving nearly perfect predictions. On the other hand, image models were less consistent, likely due to variability in visual data, imbalance, or fewer training samples. Therefore, combining predictions from both data sources can enhance decision-making reliability in real-world accident assessment systems.

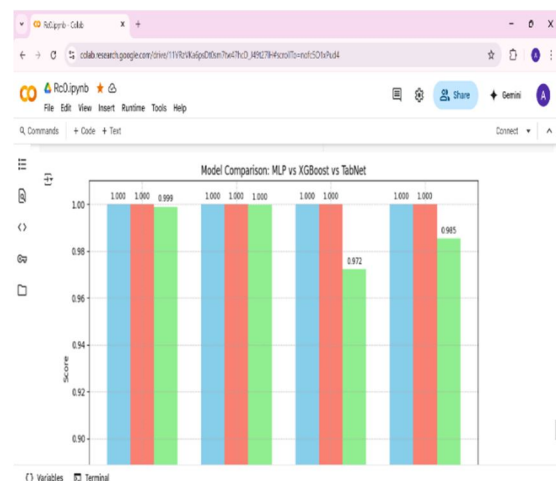


Fig.2 Comparative barchart of tabular models

VI. RESULT AND DISCUSSION

The experimental results of this project clearly demonstrate the effectiveness of combining machine learning and deep learning models to predict the severity of road traffic accidents using both tabular and image data. For structured data, three models Multilayer Perceptron (MLP), XGBoost, and TabNet were trained and evaluated using standard classification metrics. Among them, XGBoost and MLP achieved perfect scores across all key indicators (accuracy, precision, recall and F1-score), each reaching a value of 1.000. These results highlight their robustness when applied to well-preprocessed and engineered tabular features, and showcase their potential for real-world deployment in traffic monitoring systems. TabNet also delivered strong results, with slightly lower recall and F1 scores (0.972 and 0.985 respectively), while offering the added benefit of interpretability due to its attention-based feature selection.

On the other hand, for unstructured image data, MobileNetV2 and ResNet50 were utilized using transfer learning. Although their overall accuracies remained relatively strong (0.78–0.79), their precision values dropped significantly to around 0.60, indicating some difficulty in confidently distinguishing among accident severity levels solely based on visual input. Nevertheless, both models maintained balanced recall and F1-scores, suggesting their capability to detect severe crashes, albeit with occasional misclassification. EfficientNetB7 was employed as an auxiliary model to auto-label accident images where severity classes were not predefined, improving dataset usability and facilitating a smoother training pipeline. Visual comparison using bar charts confirmed the superior performance of tabular models over image-based models in terms of consistency and reliability. These findings emphasize the complementary nature of structured and visual data in accident severity prediction. While tabular models offer precision and stability, image-based models can provide valuable contextual cues, especially in cases where structured data is incomplete or unavailable. Overall, the system demonstrates a powerful, multimodal framework that combines deep learning and ensemble methods to deliver high-performing, scalable predictions, which can be effectively used to support emergency response decision-making and intelligent transport systems.

VII. CONCLUSION

In conclusion, this project successfully developed an end-to-end, automated system for predicting the severity of road traffic accidents using both tabular and image data. By integrating multiple machine learning and deep learning models including MLP, XGBoost, TabNet, MobileNetV2 and ResNet50 the system demonstrated strong performance, especially in structured data classification, with some models achieving near-perfect accuracy. The use of EfficientNetB7 for auto-labeling unlabeled accident images further streamlined the data preparation process and enhanced model training. Evaluation through accuracy, precision, recall and F1-score confirmed the effectiveness of the models, and comparative analysis highlighted the strengths of both data modalities. Looking forward, this framework holds immense potential for real-time deployment in smart traffic systems and emergency response platforms. Future enhancements may include integrating more advanced vision models such as Vision Transformers (ViTs), expanding datasets for better generalization, incorporating temporal or geospatial data for context-aware predictions, and deploying the system via web or mobile interfaces for live monitoring. Furthermore, combining this predictive system with IoT and edge computing devices could enable low-latency accident analysis and immediate alert generation, contributing significantly to safer and smarter transportation infrastructure.

REFERENCES

- [1] Omar Ibrahim Aboulola, Ebtisam Abdullah Alabdulqader, Aisha Ahmed Alarfaj, Shtwai Alsubai and Tai-Hoon Kim, (Member ieee) An Automated Approach for Predicting Road Traffic Accident Severity Using Transformer Learning and Explainable AI Technique 6 May 2024.
- [2] N. Klinjun, M. Kelly, C. Praditsathaporn, and R. Petsirasan, "Identification of factors affecting road traffic injuries incidence and severity in Southern Thailand based on accident investigation reports," *Sustainability*, vol. 13, no. 22, p. 12467, 2021.
- [3] B. B. Elallid, N. Benamar, A. S. Hafid, T. Rachidi, and N. Mrani, "A comprehensive survey on the application of deep and reinforcement learning approaches in autonomous driving," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7366–7390, Oct. 2022.
- [4] L. Li, Y. Lin, B. Du, F. Yang, and B. Ran, "Real-time traffic incident detection based on a hybrid deep learning model," *Transp. A, Transp. Sci.*, vol. 18, no. 1, pp. 78–98, 2022.
- [5] K. Pawar and V. Attar, "Deep learning-based detection and localization of road accidents from traffic surveillance videos," *ICT Exp.*, vol. 8, no. 3, pp. 379–387, Sep. 2022.
- [6] E. Cicek, M. Akin, F. Uysal, and R. Topcu Aytas, "Comparison of traffic accident injury severity prediction models with explainable machine learning," *Transp. Lett.*, vol. 15, no. 9, pp. 1043–1054, Oct. 2023.
- [7] J. Gan, L. Li, D. Zhang, Z. Yi, and Q. Xiang, "An alternative method for traffic accident severity prediction: Using deep forests algorithm," *J. Adv. Transp.*, vol. 2020, pp. 1–13, Dec. 2020.
- [8] I. Aldhari, M. Almoshaogeh, A. Jamal, F. Alharbi, M. Alinizzi, and H. Haider, "Severity reduction of highway crashes in Saudi Arabia using machine learning techniques," *Appl. Sci.*, vol. 13, no. 1, p. 233, Dec. 2022.
- [9] M. Sameen and B. Pradhan, "Severity prediction of traffic accidents with recurrent neural networks," *Appl. Sci.*, vol. 7, no. 6, p. 476, Jun. 2017.
- [10] A. Jamal, M. Zahid, M. T. Rahman, H. M. Al-Ahmadi, M. Almoshaogeh, D. Farooq, and M. Ahmad, "Injury severity prediction of traffic crashes with ensemble machine learning techniques: A comparative study," *Int. J. Injury Control Saf. Promotion*, vol. 28, no. 4, pp. 408–427, Oct. 2021.
- [11] R. E. AlMamlook, K. M. Kwayu, M. R. Alkasisbeh, and A. A. Frefer, "Comparison of machine learning algorithms for predicting traffic accident severity," in *Proc. IEEE Jordan Int. Joint Conf. Electr. Eng. Inf. Technol. (JEEIT)*, Apr. 2019, pp. 272–276.
- [12] F. Malin, I. Norros, and S. Innamaa, "Accident risk of road and weather conditions on different road types," *Accident Anal. Prevention*, vol. 122, pp. 181–188, Jan. 2019.
- [13] M. K. Gebru, "Road traffic accident: Human security perspective," *Int. J. Peace Develop. Stud.*, vol. 8, no. 2, pp. 15–24, 2017.
- [14] V. Nuri Sumantri, A. I. Rifai, and F. Ferial, "Impact of inter-urban street lighting on users perception of road safety behavior: A case of jalan majalengka-rajagaluh," *Citizen, J. Ilmiah Multidisiplin Indonesia*, vol. 2, no. 5, pp. 703–711, Dec. 2022.



- [15] C. Gutierrez-Osorio and C. Pedraza, "Modern data sources and techniques for analysis and forecast of road accidents: A review," J. Traffic Transp. Eng., vol. 7, no. 4, pp. 432–446, Aug. 2020.
- [16] R. Roscher, B. Bohn, M. F. Duarte, and J. Garcke, "Explainable machine learning for scientific insights and discoveries," IEEE Access, vol. 8, pp. 42200–42216, 2020.
- [17] R. C. Staudemeyer and E. Rothstein Morris, "Understanding LSTM—A tutorial into long short-term memory recurrent neural networks," 2019, arXiv:1909.09586.
- [18] F. He, T. Liu, and D. Tao, "Why ResNet works? Residuals generalize," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 12, pp. 5349–5362, Dec. 2020.
- [19] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 1800–1807.
- [20] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Proc. Adv. Neural Inf. Process. Syst., vol. 30, 2017, pp. 1–10.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)