ESCAP/WMO
Typhoon Committee

Available online at www.sciencedirect.com

ScienceDirect

Tropical Cyclone Research and Review 14 (2025) 1–12

Tropical Cyclone Research and Review

www.keaipublishing.com/tcrr

Ensemble deep learning models for tropical cyclone intensity prediction using heterogeneous datasets

Dikshant Gupta, Menaka Pushpa Arthur*

School of Computer Science and Engineering, Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India

Available online 15 February 2025

Abstract

The prediction of the Tropical Cyclone (TC) intensity helps the government to take proper precautions and disseminate appropriate warnings to civilians. Intensity prediction for TC is a very challenging task due to its dynamically changing internal and external impact factors. We proposed a system to predict TC intensity using CNN-based ensemble deep-learning models that are trained by both satellite images and numerical data of the TC. This paper presents a thorough examination of several deep-learning models such as CNN, Recurrent Neural Networks (RNN) and transfer learning models (AlexNet and VGG) to determine their effectiveness in forecasting TC intensity. Our focus is on four widely recognized models: AlexNet, VGG16, RNN and, a customized CNN-based ensemble model all of which were trained exclusively on image data, as well as an ensemble model that utilized both image and numerical datasets for training. Our analysis evaluates the performance of each model in terms of the loss incurred. The results provide a comparative assessment of the deep learning models selected and offer insights into their respective prediction loss in the form of Mean Square Error (MSE) as 194 in 100 epochs and execution time 1229 s to forecasting TC intensity. We also emphasize the potential benefits of incorporating both image and numerical data into an ensemble model, which can lead to improved prediction accuracy. This research provides valuable knowledge to the field of meteorology and disaster management, paving the way for more resilient and precise TC intensity forecasting models.

Keywords: Ensemble model; Transfer learning model; Deep learning model; CNN; TC intensity prediction; AlexNet; VGG16; RNN

1. Introduction

Tropical cyclones (TC) are powerful weather events that go by different names depending on the region, such as hurricanes in the Atlantic and northeastern Pacific, typhoons in the northwestern Pacific, and TCs in the Indian Ocean (Kapoor et al., 2023). They are considered to be one of the most destructive natural disasters on our planet. With their ferocious winds, torrential rains, and the potential for storm surges, TCs pose severe threats to coastal communities, ecosystems, and infrastructure. The increasing frequency and intensity of these cyclonic systems in a changing climate make accurate TC intensity prediction a matter of paramount importance. TC intensity prediction is a complex task, as it requires a deep understanding of the underlying atmospheric and oceanic dynamics. Accurate predictions are essential not only for public safety but also for efficient evacuation planning, resource allocation, and disaster response. Traditional numerical weather prediction models, driven by sophisticated mathematical equations and physical parameterizations, have significantly improved our ability to forecast the tracks of TCs. However, predicting the intensity of these storms remains a formidable challenge, often resulting in underestimations or overestimations of their destructive potential. Over the past few years, machine learning has transformed several domains, most notably computer vision, natural language processing, and healthcare. In particular, the advent of deep learning has

* Corresponding author.
E-mail address: menakapushpa.a@vit.ac.in (M.P. Arthur).
Peer review under the responsibility of Shanghai Typhoon Institute of China Meteorology Administration.

https://doi.org/10.1016/j.tcrr.2025.02.001

revolutionized the way these fields are approached, opening up new avenues for research and development. This transformative technology has also found applications in meteorology and climatology. Deep learning models, such as Convolutional Neural Networks (CNNs) (Matsuoka et al., 2018; Wang et al., 2020; Zhang et al., 2021; Wang & Li, 2023; Xu et al., 2022; Krizhevsky et al., 2012; Simonyan & Zisserman, 2014) and RNNs (Kapoor et al., 2023; Biswas et al., 2021), have shown remarkable promise in extracting complex patterns and features from large and complex datasets.

Among all the reviewed literature, CNN was most commonly used for predicting the intensity of TCs. One such example Mehra et al. (Mehra et al., 2020), which uses a CNN model to estimate wind speed from infrared satellite imagery is called 'Deepti'. It is trained on a dataset of over 100,000 TC images and wind speed measurements. It is used to generate real-time intensity estimates for TCs around the world. Chen et al (Chen et al., 2021) proposed an alternative approach to estimate TC (TC) intensity using satellite remote sensing, and this approach involves a novel hybrid model based on CNN. The model consists of three fine-grained models, a coarse-grained model, a backpropagation (BP) model, and a classification model. The authors evaluate the proposed model on a dataset of satellite images of TCs with known intensities. The model has the advantage of being able to learn the intricate connections among the TC cloud structure, and the TC intensity from the satellite images in an end-to-end manner. Alahmadi et al (Alahmadi & Alzahrani, 2022) proposed a CNN architecture and trained on a dataset of HURDAT2 which provides wind speed data and infrared satellite imagery. This work fine-tuned a pre-trained VGG19 model to predict the extent of damage caused by hurricanes. The VGG19 model achieves an accuracy of 98 % in classifying the damage level of buildings as flooded/damaged or undamaged.

Arulmozhi et al. (Arulmozhi & Sivakumar, 2023) designed a model to classify satellite cloud images with 94 % accuracy using a deep convolutional neural network (DCNN). This model is trained by 1 million satellite cloud images of the Meteorological & Oceanographic Satellite Data Archival Centre, India. Rai et al. (Rai et al., 2020) used the Brovey transform to fuse the panchromatic band with the three RGB bands of the Landsat 8 OLI image, resulting in a fused image with a spatial resolution of 15 m. Principal component analysis (PCA) is then applied to the fused image to reduce its dimensionality. Finally, a CNN is used to classify the PCA-reduced image into different land cover classes. The proposed method was evaluated on four different datasets of Landsat 8 OLI images, containing three, four, five, and seven land cover classes, respectively.

A new method for predicting the path and strength of cyclones has been suggested by Kapoor et al. (Kapoor et al., 2023), which involves utilizing variational recurrent neural networks (VRNNs). VRNNs are well-suited for cyclone trajectory and intensity prediction because they can learn the complex temporal dynamics of cyclones while also accounting for the uncertainty associated with these predictions. Biswas et al.(Biswas et al., 2021) designed a model to predict the strength of TCs in the North Indian Ocean by using a stacked Bidirectional LSTM architecture. This model estimates the maximum surface sustained wind speed (MSWS) of the cyclones. The model that has been proposed can accurately predict MSWS, up to 72 h ahead of time.

Chen et al. (Chen & Lin, 2023) argued that deep learning provides an opportunity to improve TC RI (TC Rapid Intensification) prediction by concurrently dealing with environmental and TC-related parameters specified by humans and details derived from satellite imagery. An ensemble of 20 deep-learning models is suggested. Their purpose is to anticipate the TC intensity distributions at +24 h. The ensemble approach is evaluated on a dataset of Western Pacific TCs from 2005 to 2019. Lee et al. (Lee et al., 2021) suggested a distinct proximity model based on random walk to determine likeness among images of TCs. The model takes into account both the spatial and temporal relationships between the images. For the spatial relationships, the model assumes that cyclones with similar intensities have similar cloud patterns. For the temporal relationships, the assumption made by the model is that the rate of change in the strength of a TC is slow over a period of time. A sizable dataset of TC images is used to train the model and their corresponding intensity values. Wang et al (Wang & Huang, 2019) argued that deep learning can be used to address this challenge by automatically extracting important features from TC images. The proposed model uses two types of attention mechanisms: self-attention and sequence attention. The model can learn long-range dependencies in the TC images through self-attention, while sequence attention enables it to learn the temporal evolution of the TC. The model is trained on a benchmark dataset of TC images and RI labels.

This research paper focuses on the application of deep learning techniques to TC intensity prediction. The research delves into the usage of three renowned deep learning models – CNN, AlexNet and VGG16 - which are equipped with exceptional image recognition and feature extraction capabilities. We have used these models to analyze TC imagery and evaluate their effectiveness in extracting vital meteorological information from visual data. Our work also incorporated an RNN model, which has the capability to learn the temporal evolution of cloud patterns, wind structures, and other relevant features in satellite images (Jaya & Srinivasan, 2019). This allows the model to capture the dynamic nature of TCs, which ultimately leads to a more nuanced understanding of their intensity changes over time.

a. Major Objectives:

- To evaluate the performance of deep learning models in the context of TC intensity prediction. This involves assessing their capacity to learn and exploit relevant features from the diverse data sources available.
- Our work aims to quantify the loss of these models and investigate their potential to outperform traditional numerical models.
- Beyond the standalone evaluation of pre-trained AlexNet and VGG16, our work explores the synergistic potential of

the ensemble model that combines both image and numerical data.

In this paper, the following structure is followed: Section 2 elaborates on the methodology employed, including data collection, pre-processing, feature selection, and the architecture of the deep learning models under investigation. Section 3 presents the results of our experiments, supported by quantitative metrics and visualizations, shedding light on the performance of AlexNet, VGG16, and the mixed model, and in Section 4, conclusions were drawn.

2. Data and Methodology

The TC Intensity prediction system consists of the following sub-components which are given in Figure 1.

2.1. Dataset Description

The dataset used in our work was more than 3 GB in size, containing various numerical values like latitude, longitude, Eye Location, and corresponding images. The whole dataset was in h5 format. The H5 file format, short for Hierarchical Data Format version 5 (Koranne & Koranne, 2011), is a versatile and widely used data storage format in the scientific and data analysis communities. The dataset was curated from a diverse range of sources, including GridSat which features geostationary satellite images taken over an extended period of time, and CMORPH which provides data transmitted via spatial propagation from low-orbit satellites. The information contained in the dataset is of utmost importance, as it includes key details such as the maximum sustained wind speed (measured in knots), the size of the TC, the minimum sea-level pressure, and the precise location of the cyclone's eye. The images were obtained from multiple locations, including the Atlantic Ocean, Indian Ocean, West Pacific, East Pacific, and Southern Hemisphere. The dataset consisted of 4580 images, each present in 4 different channels such as infrared channel, water vapor channel, visible channel and passive microwave channel (Menzel et al., 2016; Georgiev et al., 2016; Kovordányi & Roy, 2009) in Figure 2. The images in the dataset with resolution of 201*201 size, with a radius of 7° in both latitude and longitude. Furthermore, the TC's center was positioned in the middle of the vector. The resolution of the images was 7/100° lat/lon.

2.2. Data Preprocessing/Augmentation

Data preprocessing involves converting unprocessed data into an organized format that is appropriate for analysis and training. The significance of preprocessing lies in its ability to standardize the input data, thereby facilitating the effective learning of patterns and relationships within the data. The accuracy and efficiency of the resulting model heavily relies on the quality of the preprocessing stage.

The process of data augmentation in Figure 3 involves creating more training samples through the application of diverse transformations to the existing dataset. When it comes to predicting TC intensity through image data, incorporating rotation and enhancement as data preprocessing and augmentation techniques can prove to be incredibly advantageous. TCs can take on a variety of shapes and orientations in satellite imagery, and their evolution can be captured from different rotational perspectives. Randomly rotating the training images

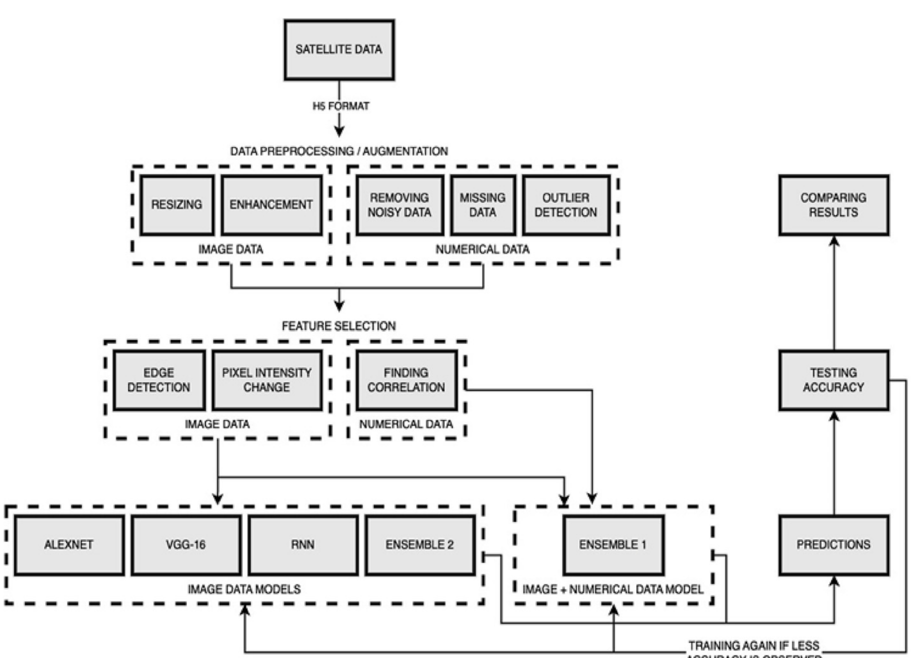

Fig. 1. Proposed TC intensity prediction system.

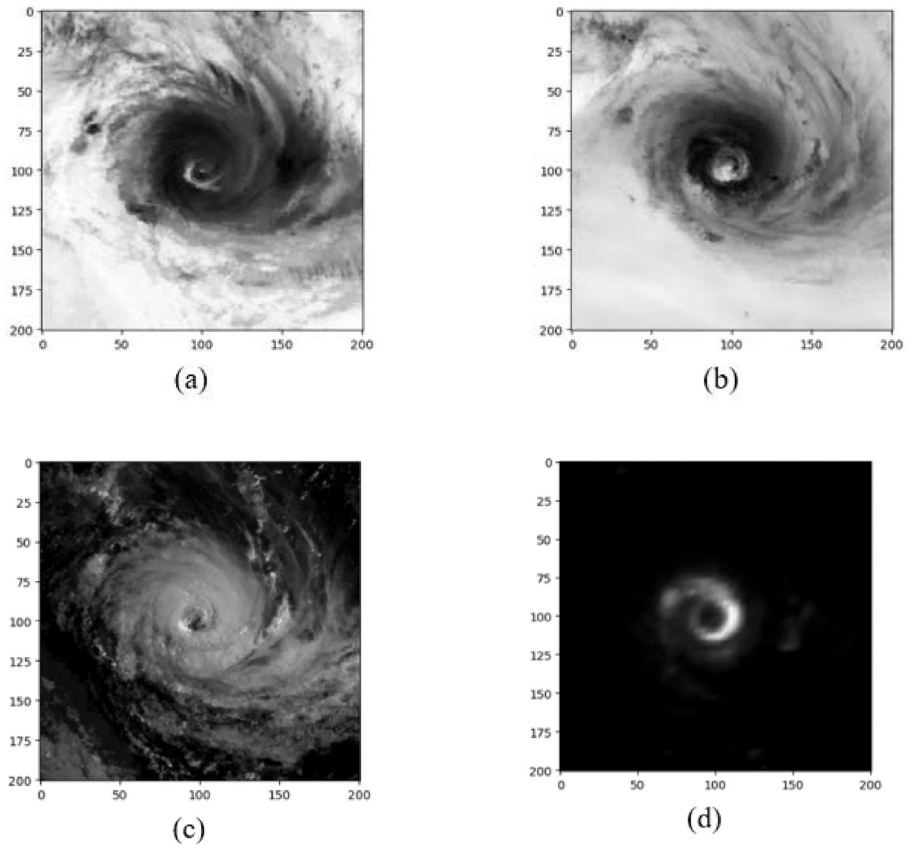

Fig. 2. TC image samples with different categories:(a) Infrared (b) Water Vapor (c) Visible (d) Passive Micro Waves.

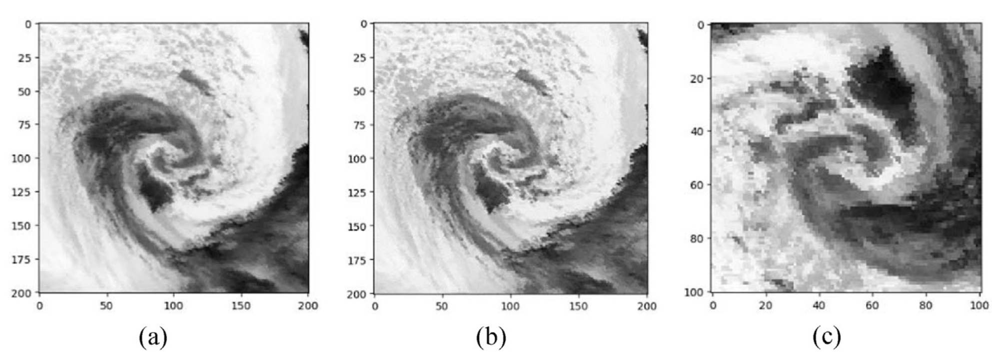

Fig. 3. Image preprocessing/augmentation samples: (a) Original Image (b) Enhanced Image (c) Rotated Image.

can make the model more adaptable to variations in TC orientation, improving its ability to recognize TCs with different spatial orientations. This, in turn, leads to better generalization when faced with unfamiliar cyclonic patterns. Additionally, features such as cloud patterns and wind structures may vary in intensity and visibility, but the model can learn to recognize and adjust to these variations through enhancement.

To ensure the robustness and effectiveness of the predictive model, outlier detection was conducted, removed noisy data, and addressed missing data for numerical features. In TC intensity prediction, outliers in numerical features may represent erroneous or anomalous readings, potentially leading to inaccurate predictions. By identifying and removing these outliers through diverse techniques, it was ensured that the model was trained on a representative and reliable dataset. This improves the model's ability to discern genuine patterns and relationships within the numerical data, resulting in more accurate TC intensity predictions. Noisy data, characterized by random errors or inconsistencies, can introduce unwanted variability and hinder the learning process of the model. Removing noisy numerical data helps create a cleaner and more stable training dataset, resulting in a model that is less susceptible to being misled by irrelevant or erroneous information. Addressing missing data is crucial for maintaining the integrity of the model's learning process and enhancing its capability to make reliable TC intensity predictions, especially when certain numerical features may carry vital information.

2.3. Feature Selection

The process of feature selection is an essential step in optimizing model performance, decreasing complexity, and improving interpretability. It entails selecting a relevant subset of features from the original set of variables shown in Figure 4 and Figure 5. In the specific context of TC intensity prediction, feature selection is critical in identifying the most informative input features from a diverse pool of variables. This system utilized advanced techniques, namely the Prewitt operator for edge detection and the modified Sobel operator for pixel intensity change (Zhou et al., 2019; Agarwal, 2015). These innovative methods offer numerous benefits, such as highlighting TC structures, capturing spatial patterns, and reducing dimensionality, resulting in a refined set of distinctive features. By emphasizing relevant information and improving model generalization, these processed features play a crucial role in generating precise predictions. Furthermore, the interpretability of these features aligns with meteorological understanding, making them a valuable tool for decision-making in disaster management scenarios.

Correlation analysis was used to select features in numerical data for predicting TC intensity. This assessment examines the correlation and direction of links between numerical characteristics and the outcome factor, determining the most significant characteristics and minimizing repetition by identifying those that are highly correlated. This process streamlines the model, potentially improving its generalization performance, while mitigating collinearity issues and enhancing model interpretability. By selecting features based on correlation, the model can focus on the most informative ones, leading to quicker training times and improved training efficiency. This approach also enhances the model's robustness and ensures its ability to make accurate predictions across diverse conditions by carefully selecting a subset of features and managing data dimensionality.

It can be observed from Fig. 6(a) that there is a weaker negative correlation between size and pressure, this means that larger TCs tend to have lower pressure. The correlation coefficient between intensity and size is 0.663, i.e. there is a moderately strong positive relationship between these two variables. The correlation coefficient between intensity and pressure is −0.951, i.e. there is a very strong negative relationship between these two variables. In other words, as intensity increases, pressure tends to decrease significantly. Fig. 6(b) gives similar inferences to the heatmap, the only difference is the plotting of the scatter plot.

2.4. Model Training and Testing

So as shown in Fig.1, in our work training and testing was conducted for 4 different types of models (Alexnet, VGG16, Ensemble, RNN). AlexNet is a powerful CNN architecture that boasts eight layers, five of which are convolutional and three are fully connected (Zhang et al., 2021). The Rectified Linear Units (ReLU) activation function is used throughout. AlexNet's advantages include parallelization, efficient GPU utilization, and dropout for regularization to prevent overfitting. When it comes to predicting TC intensity using image datasets, AlexNet shines by extracting intricate features from visual data. The convolutional layers excel in recognizing patterns such as cloud formations and wind patterns, which are key to predicting TC intensity (O'Shea & Nash, 2015). Additionally, its transfer learning capabilities make it possible to leverage pre-trained models on large datasets, leading to better performance even with limited labeled TC data. With CNNs' inherent spatial dependency capturing and real-time processing capabilities, AlexNet is an invaluable asset in TC intensity prediction tasks, where accurate and timely assessments are critical.

The Visual Geometry Group created the VGG16, which is widely respected for its simplicity and effectiveness. The VGG16 has a complex architecture, with 16 wt layers, comprising 13 convolutional layers and 3 fully connected layers (Xu et al., 2022). The key to its success lies in the consistent use of small 3x3 convolutional filters throughout the network, allowing for a deeper representation while keeping the parameter count manageable. This uniform design makes the architecture easy to understand and interpret. VGG16's structure, with its deep convolutional layers, makes it a powerful feature extractor, capable of capturing intricate patterns in images. In the context of TC intensity prediction, VGG16's ability to discern fine-grained details and spatial relationships in satellite imagery is advantageous. Its deep convolutional layers can automatically learn hierarchical features related to TC indicators, and the model can benefit from transfer learning by pre-training on large visual datasets. VGG16's versatility and strong feature extraction capabilities

$$\begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix}$$

(a)

$$\begin{vmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{vmatrix}$$

(b)

Fig. 4. Operators used for feature selection: (a) Prewitt (Zhou et al., 2019) (b) Modified Sobel (Agarwal, 2015).

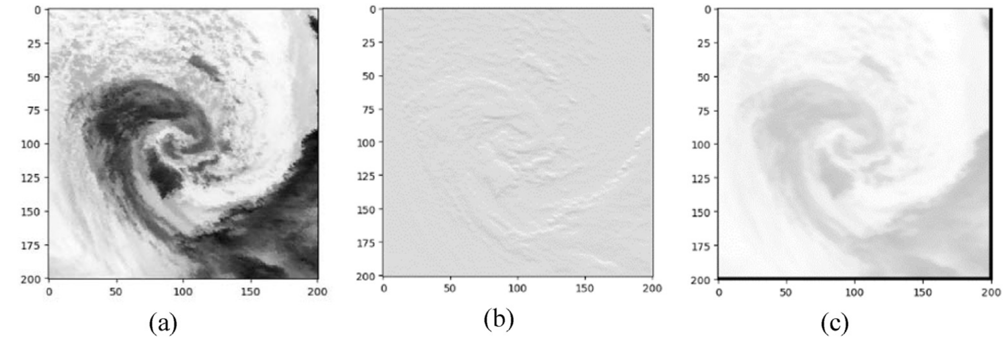

Fig. 5. Image samples after feature selection: (a) Original (b) Edge Detection (c) Pixel Intensity Change.

make it a valuable asset for TC intensity prediction tasks, contributing to accurate and reliable predictions based on image data.

Ensemble models represent a highly effective approach in machine learning that involves combining multiple individual models to achieve a more robust and accurate prediction. The models' structure usually consists of different base models, like decision trees, neural networks, or support vector machines. Given the complexity and variability of factors influencing TC intensity, combining diverse models helps capture a broader range of patterns and features present in satellite imagery. Ensembles are particularly effective in reducing overfitting, enhancing generalization, and providing more reliable predictions, especially when individual models may struggle with certain aspects of the data, such as intricate cloud formations or subtle wind patterns. Additionally, ensemble methods, such as bagging and boosting, contribute to increased stability and resilience in the face of noisy or incomplete datasets, ultimately improving the accuracy and robustness of TC intensity predictions. Our work employs two distinct ensemble models to effectively process our data. The initial model utilizes a CNN-based framework to analyze image data and a straightforward dense neural network to handle numerical data values. Our second ensemble model incorporates 3, CNN-based architectures, one for original images, another for edge-detected images, and a third for images in which pixel intensity changes were observed.

RNNs are a highly specialized type of neural network that excels at processing sequential data. Their unique architecture features recurrent connections, which allow them to retain information over time. Unlike traditional feedforward neural networks, RNNs are particularly well-suited for tasks that involve sequences, natural language processing and time series analysis are examples of such technologies. The key advantage of RNNs is their ability to capture temporal dependencies. This means that they can take into account previous inputs when processing current ones, making them ideal for predicting TC intensity using image datasets. Understanding the evolution of meteorological patterns over time is essential for this task, and RNNs excel at modeling dynamic processes in satellite imagery. By effectively tracking how TCs develop and change, RNNs are a valuable tool for predicting and mitigating the impact of these dangerous weather events.

During training and testing, mean squared error (MSE) served as the chosen metric for evaluation. MSE is a commonly used measure in machine learning to assess the average squared difference between predicted and actual values. Throughout the training process, some models utilized the K-fold method, a

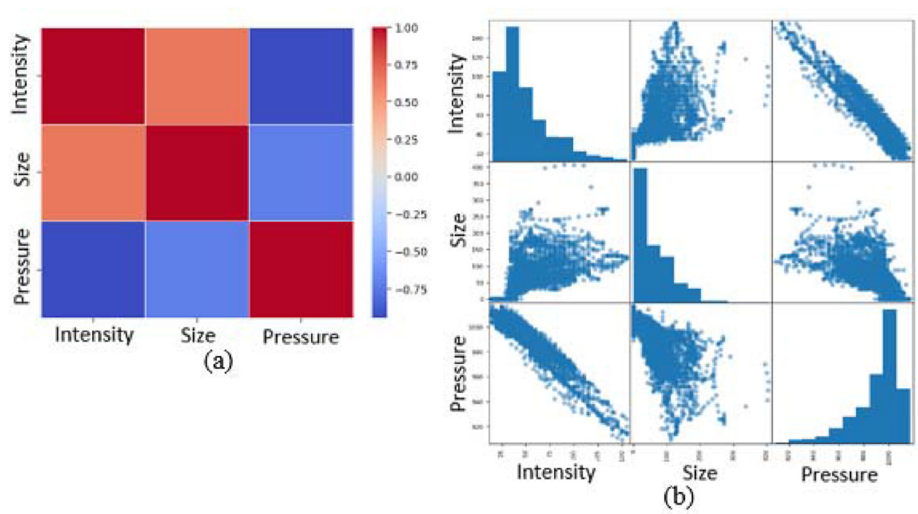

Fig. 6. Visualization of Features Correlation in the TC numerical dataset (a) Correlation Heatmap (b) Scatter Plot Matrix.

cross-validation technique that involves dividing the dataset into K folds of equal size. The model undergoes K rounds of training and evaluation, where each round involves using a distinct fold as the testing set and the remaining folds for training (Wong & Yeh, 2019). During each iteration, a different fold is chosen as the testing set, and the remaining folds are used for training. This produces more accurate estimates of the model's performance by averaging the results from

Table 1
Training results and other details for Alexnet

Input Image	Training Loss (MSE)	Testing Loss (MSE)	Execution Time (in seconds)	No. of Epochs
Original	1.57	277	2227	600
Rotated	5.3	495	2356	600
Enhanced	3.4	302	2248	600
Edge detected	2.7	260	748	600
Pixel intensity changes	4.6	254	756	600

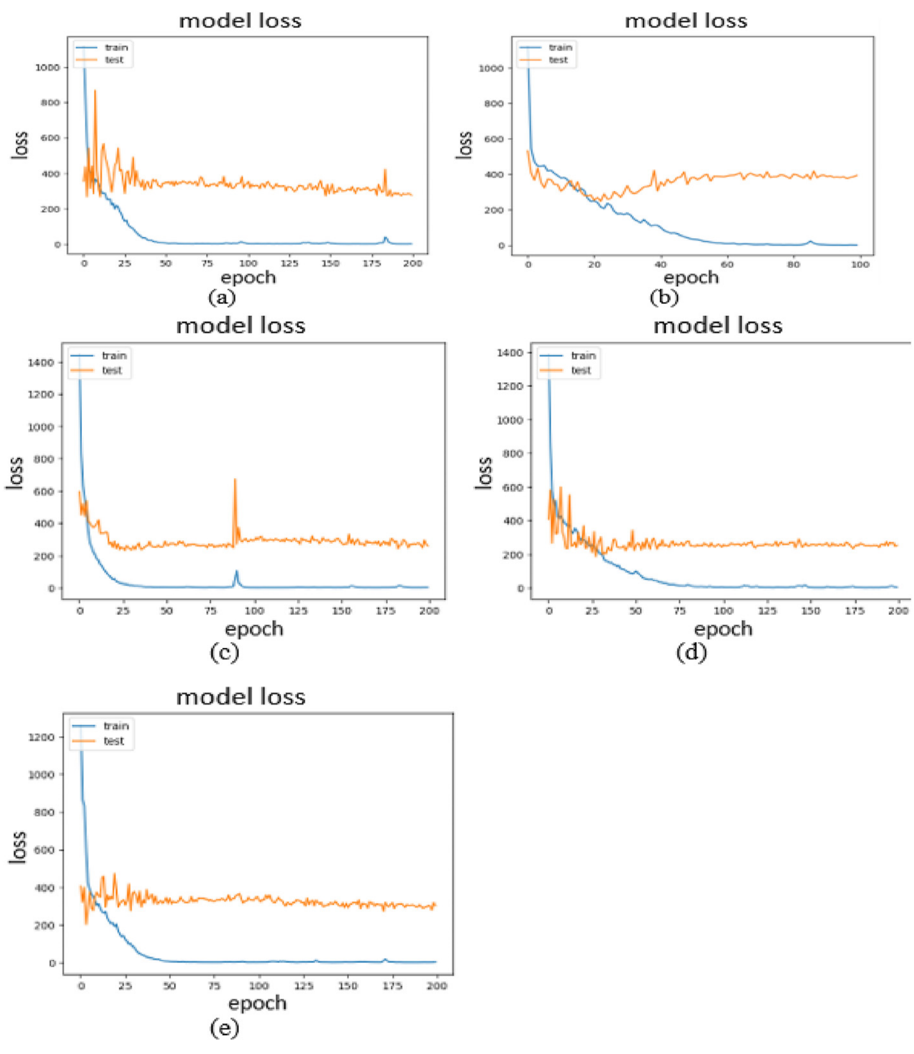

Fig. 7. Training and Testing results for Alexnet with different input images: (a) Original (b) Rotated (c) Edge Detected (d) Pixel Intensity Changes (e) Enhanced.

Table 2
Training results and other details for VGG16

Input Image	Training Loss (MSE)	Testing Loss (MSE)	Execution Time (in seconds)	No. of Epochs
Original	224	373	17,550	90
Rotated	158	329	16,250	90
Enhanced	218	385	14,188	90
Edge detected	58	348	14,666	90
Pixel intensity changes	337	544	13,719	90

each iteration, ensuring that it can handle diverse data. This is particularly important when working with satellite images of TCs, which can differ greatly due to weather patterns, lighting conditions, and geographical features. K-fold training also detects potential overfitting or underfitting issues by assessing the model's consistency across multiple folds (Wong & Yeh, 2019). Moreover, it maximizes the usage of available data for both training and testing, a crucial factor when working with limited datasets.

3. Results & Discussions

In this section, the performance of the proposed system is analyzed using the dataset with 4580 satellite images and numerical data. The TC intensity prediction system is majorly designed with two options in the input data: (i) Only satellite images or (ii) Fusion of satellite images and numerical data of the TC. Under satellite input images, the first section analyzes the performance of the TC intensity prediction system with a single deep learning model such as RNN or AlexNet or VGG16 over 3 CNN-based transfer learning models. The Second section analyzes the fusion of input categories such as TC satellite images and numerical data to train the ensemble of Deep Learning and Machine Learning models correspondingly. This section also compares the proposed system with existing systems designed with deep learning models such as ConvGRU (Zhang et al., 2022) and ConvLSTM (Tong et al., 2022) over the various input data. The image preprocessing methods are also observed to know their impact on the accuracy of TC intensity prediction.

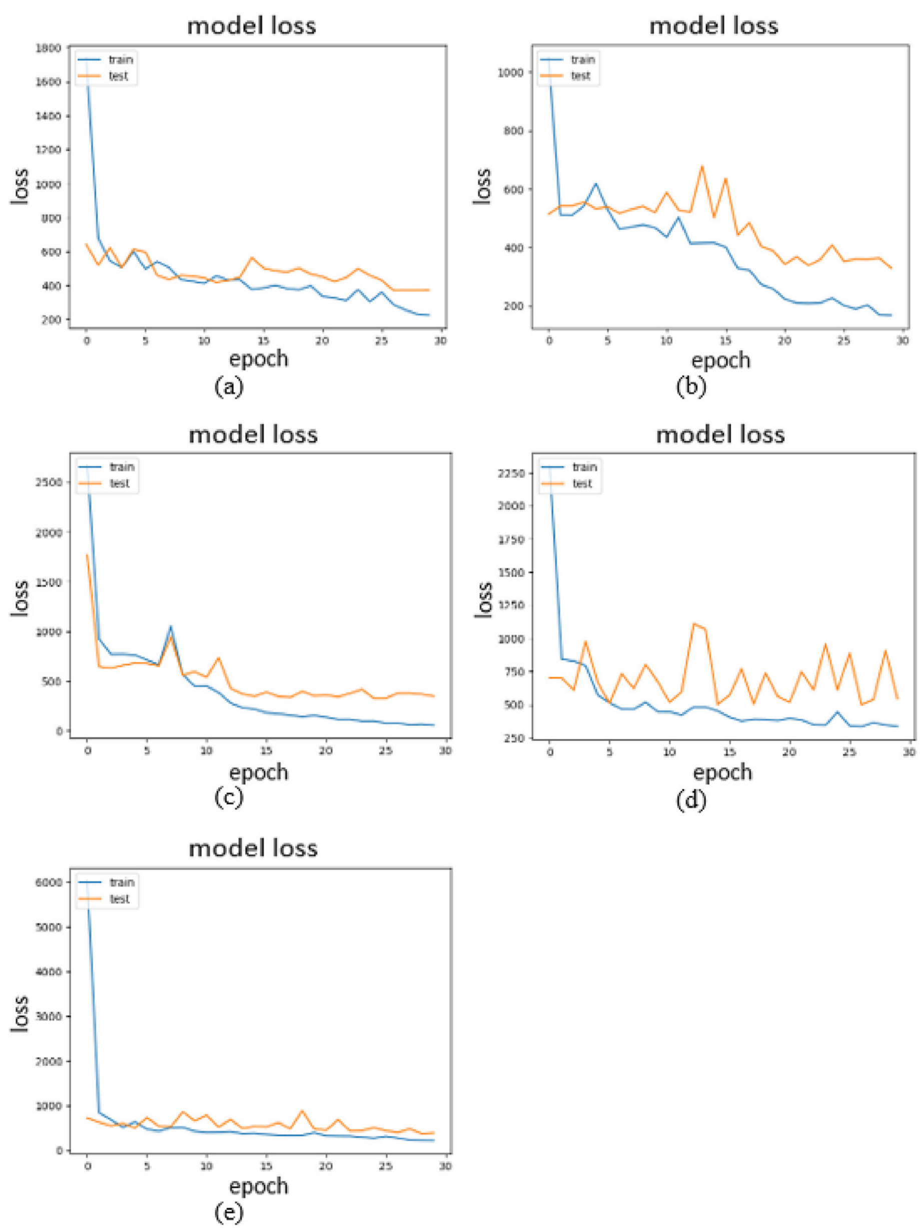

Fig. 8. Training and Testing results for VGG16 with different input images: (a) Original (b) Rotated (c) Edge Detected (d) Pixel Intensity Changes (e) Enhanced.

3.1. Analysis of Deep Learning model with satellite images

3.1.1. Alexnet

Table 1 reveals that the model utilizing images with pixel intensity changes out-performed the others, not only in terms of loss but also in training time. Additionally, the graph for this model displayed minimal fluctuations (as evident in Fig. 7). This could be because pre-processing steps, such as rotation, edge detection, or enhancement, may introduce artificial features that do not necessarily aid in the prediction task. Also, as the pixel intensity model took less execution

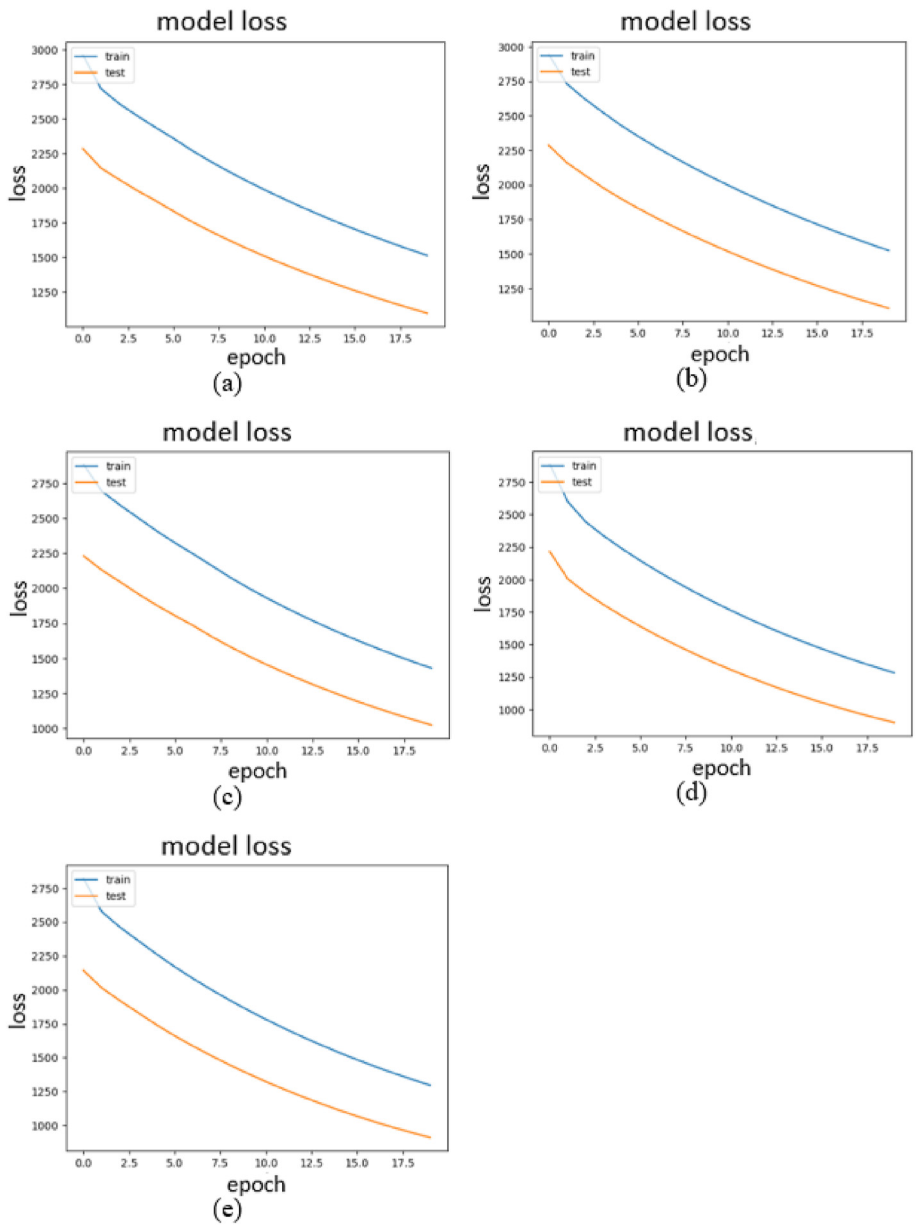

Fig. 9. Training and Testing results for RNN with different input images: (a) Original (b) Rotated (c) Edge Detected (d) Pixel Intensity Changes (e) Enhanced.

Table 3
Training results and other details for RNN

Input Image	Training Loss (MSE)	Testing Loss (MSE)	Execution Time (in seconds)	No. of Epochs
Original	1513	1097	8284	20
Rotated	1525	1107	8570	20
Enhanced	1297	911	8259	20
Edge detected	1429	1024	8145	20
Pixel intensity changes	1282	900	8399	20

Table 4
Training results and other details for Ensemble 2 (Ensemble of CNN Models on Satellite Images)

Input Images	Training Loss (MSE)	Testing Loss (MSE)	Execution Time (in seconds)	No. of Epochs
Original Edge detected Pixel intensity changes	78	295	5470	100

time, it could be trained for even more epochs and may achieve better results, while being less computationally intensive.

3.1.2. VGG16

In Table 2, it is evident that the model fed with rotated images showcased the best performance in terms of loss. However, it consumed more time to train compared to the other models. On the other hand, the model with the least execution time, which was fed with pixel intensity changes, resulted in the worst loss value. This is evident from the graph of this model in Fig.8, which shows an insignificant decrease in loss value after epoch 5. In summary, VGG provides satisfactory loss values but demands high computation power and time.

3.1.3. RNN

Observing Fig 9 and Table 3, it is evident that the RNN exhibited a consistent decrease in loss value per iteration across all cases. However, it is noteworthy that the model demonstrated superior performance with enhanced and pixel-intensity change images despite this similarity. Additionally, the testing values consistently fell below the training values, indicating that the model may have been underfitting. As a result, it can be inferred that more training iterations may lead to improved outcomes.

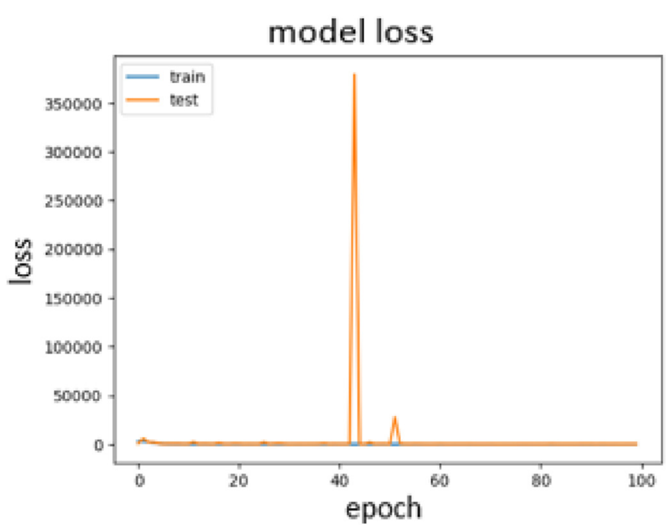

Fig. 10. Training and Testing results for Ensemble 2 (Ensemble of CNN Models on Stellite Images).

3.2. Analysis of ensemble models

3.2.1. Analysis of ensemble deep learning models with satellite images

Upon analyzing the data presented in Table 4 and Figure 10, it became clear that th ensemble model is developed using alternate image types, such as rotated, edge-detected, and pixel-intensity adjusted images. Despite these efforts, there was no discernible improvement in either the loss values or the execution time, as evidenced by the data in Table 4. Furthermore, similar spikes were observed during the training of this model, indicating that further improvement was unlikely.

3.2.2. Analysis of ensemble DL and ML models on fusion dataset

By examining Table 5, it can be seen that the model utilizing enhanced images yielded the best results in terms of both loss and execution time. Comparing the ensemble's outcomes with those of other architectures, it is clear that it achieved superior loss values with significantly less computational effort. Additionally, as seen in Fig 11, loss values for non-enhanced images experienced sporadic spikes during testing. This could indicate that the model struggled to establish connections between the images and numerical data, resulting in increased loss values when encountering slightly different data values.

3.3. Comparison of proposed and existing systems

Upon examining Table 6, it is evident that the ensemble model that utilized enhanced images outperformed the rest and did so in a remarkably swift manner, indicating its computational

Table 5
Training results and other details for Ensemble 1 (Image + Numerical data)

Input Image	Training Loss (MSE)	Testing Loss (MSE)	Execution Time (in seconds)	No. of Epochs
Original	56	255	1500	100
Rotated	63	424	1323	100
Enhanced	56	194	1229	100
Edge detected	56	207	1331	100
Pixel intensity changes	69	365	1209	100

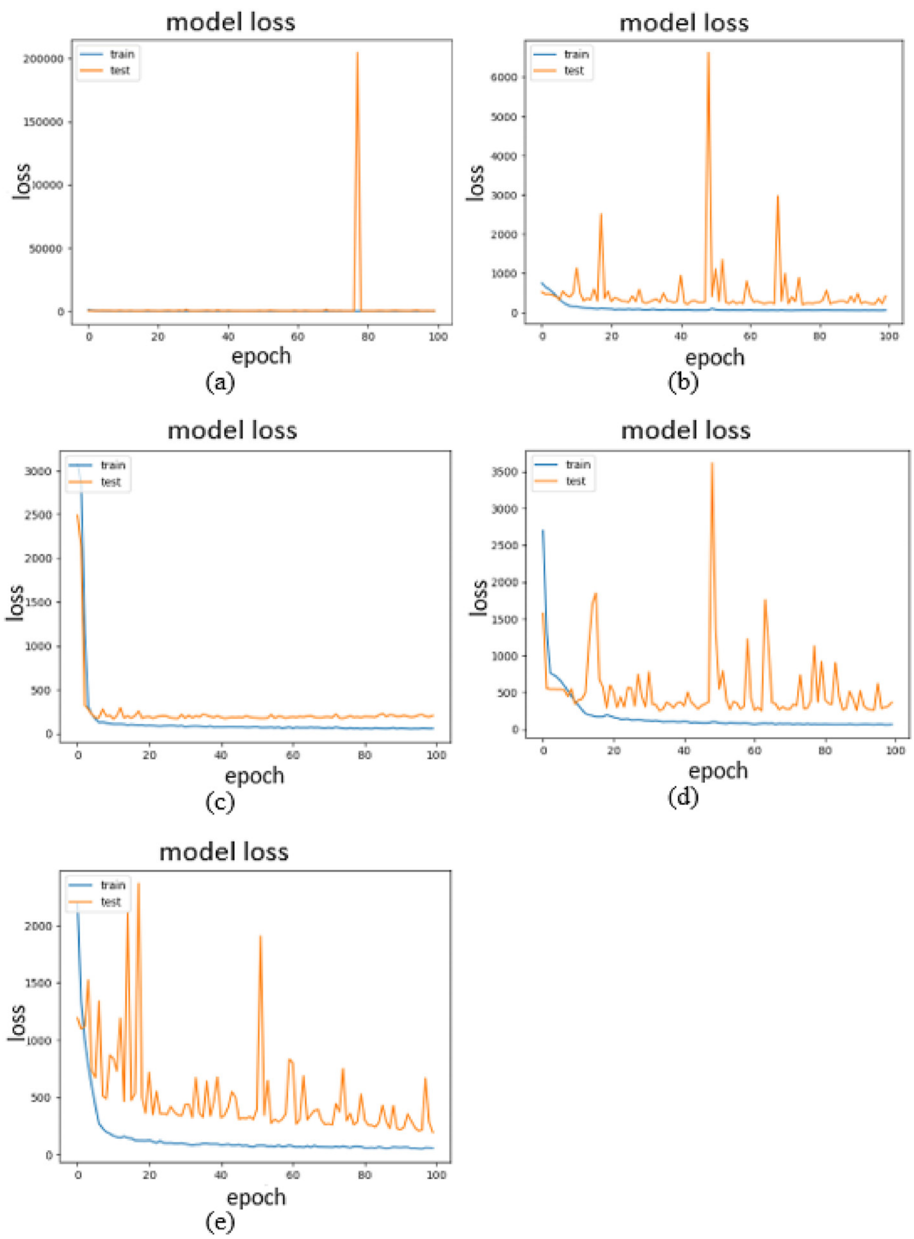

Fig. 11. Training and Testing results for Ensemble 1 (Image + Numerical Data) with different input images: (a) Original (b) Rotated (c) Edge Detected (d) Pixel Intensity Changes (e) Enhanced.

Table 6
The best performers selected from each of the four models

Model	Input Image	Training Loss (MSE)	Testing Loss (MSE)	Execution Time (in seconds)	No. of Epochs (Best)
ConvGRU (Zhang et al., 2022)	Enhanced	120	480	540	30
ConvLSTM (Tong et al., 2022)	Enhanced	470	780	600	30
Alexnet	Original	2.7	260	748	60
VGG	Rotated	158	329	16,250	90
RNN	Pixel intensity changes	1282	900	8399	20
Ensemble model 1 (image + numerical data)	**Enhanced**	**56**	**194**	**1229**	**100**

efficiency. In comparison, the VGG and RNN models required significantly more time to train but yielded inferior results compared to the aforementioned architectures. This section also compares the proposed system with existing systems designed with deep learning models such as ConvGRU (Zhang et al., 2022) and ConvLSTM (Tong et al., 2022) over the various input data. The image preprocessing methods are also observed to know their impact on the accuracy of TC intensity prediction.

4. Conclusion

Our research has revealed that the ensemble model, which merges image and numerical data, holds great potential in enhancing the precision of TC intensity forecasts, surpassing the pre-trained models (Alexnet, VGG16). This is attributed to the model's ability to assimilate both visual and numerical information, resulting in a more comprehensive understanding of the intricate factors that affect TC intensity. Additionally, the model is efficient in terms of execution time, making it computationally effective. Furthermore, while RNN requires significant computational power, our findings suggest that longer training iterations on a larger dataset could lead to improved loss values.

Moreover, a variety of image types and processing techniques were utilized to train and evaluate models. By integrating information from various sources, these models can leverage the unique strengths of each data type to generate more comprehensive and contextually informed predictions, resulting in greater forecast accuracy. Our findings indicate that edge detection and pixel intensity change images produced superior results with reduced execution times, likely due to their ability to identify critical features of the original images and simplify the model training process.

In conclusion, our study has achieved notable advancements in predicting TC intensity through the application of deep learning techniques. We highly recommend the integration of numerical data in conjunction with image data to enhance the precision and scope of future forecasts. Our research serves as a stepping stone for re-fining TC prediction methodologies and highlights the promising role that deep learning models can play in aiding our comprehension and management of natural calamities.

References

Agarwal, R., 2015. Modified bit-planes sobel operator: a new approach to edge detection. Int. J. Comp. Appl. 117 (7).

Alahmadi, A.H., Alzahrani, A., 2022. A novel deep learning based model for tropical intensity estimation and post-disaster management of hurricanes. Appl. Sci. 11 (9), 4129.

Arulmozhi, K., Sivakumar, B., 2023. Deep convolutional network based machine intelligence model for satellite cloud image classification. Int. J. Eng. Appl. Sci. Tech. 8 (3), 254–258.

Biswas, K., Kumar, S., Pandey, A.K., 2021. Intensity prediction of TCs using long short-term memory network. arXiv preprint arXiv: 2107.03187.

Chen, S.H., Lin, I.I., 2023. A deep learning ensemble approach for predicting TC rapid intensification. Atmos. Sci. Lett. 24, e1151.

Chen, C., Zhang, H., Zhao, Y., 2021. A CNN-based hybrid model for TC intensity estimation in meteorological industry. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, pp. 730–733. https://doi.org/10.1109/IGARSS47739.2021.9554172.

Georgiev, C., Santurette, P., Maynard, K., 2016. Weather Analysis and Forecasting: Applying Satellite Water Vapor Imagery and Potential Vorticity Analysis. Academic Press.

Jaya, V.K., Srinivasan, S., 2019. Satellite cloud image classification for cyclone prediction using Dichotomous Logistic Regression Based Fuzzy Hypergraph model. Future Generation Comp. Syst. 93, 624–639.

Kapoor, A., Negi, A., Marshall, L., Chandra, R., 2023. Cyclone trajectory and intensity prediction with uncertainty quantification using variational recurrent neural networks. Environ. Model. Softw. 165, 105654.

Koranne, S., Koranne, S., 2011. Hierarchical data format 5: HDF5. Handbook of Open Source Tools, pp. 191–200.

Kovordányi, R., Roy, C., 2009. Cyclone track forecasting based on satellite images using artificial neural networks. ISPRS J. Photogrammetry Remote Sensing 64 (6), 513–521.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25.

Lee, Y.-J., Hall, D., Liu, Q., Liao, W.-W., Huang, M.-C., 2021. Interpretable TC intensity estimation using Dvorak-inspired machine learning techniques. Eng. Appl. Artif. Intelligence 104, 104233.

Matsuoka, D., Nakano, M., Sugiyama, D., Uchida, S., 2018. Deep learning approach for detecting TCs and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model. Prog. Earth Planet. Sci. 5 (1), 80. https://doi.org/10.1186/s40645-018-0245-y.

Mehra, A., Chawla, N.V., Savin, S., 2020. Deepti: deep-learning-based TC intensity estimation system. In: 2020 IEEE Aerospace Conference. IEEE, pp. 1–7. https://doi.org/10.1109/AERO.2020.9173951.

Menzel, W.P., Tobin, D.C., Revercomb, H.E., 2016. Infrared remote sensing with meteorological satellites. In: Advances in Atomic, Molecular, and Optical Physics, 65. Academic Press, pp. 193–264.

O'Shea, K., Nash, R., 2015. An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.

Rai, A.K., Mandal, U., Kumar, A., 2020. Landsat 8 OLI satellite image classification using convolutional neural network. Proced. Comp. Sci. 167, 987–993.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Tong, B., Wang, X., Fu, J.Y., Chan, P.W., He, Y.C., 2022. Short-term prediction of the intensity and track of TC via ConvLSTM model. J. Wind Eng. Ind. Aerodynamics 226. https://doi.org/10.1016/j.jweia.2022.105026.

Wang, X., Huang, H., 2019. Attention-based deep TC rapid intensification prediction. In: Proceedings of the 28th International Conference on Computer Communication and Networks (ICCCN'19), pp. 1–9.

Wang, C., Li, X., 2023. Deep learning in extracting TC intensity and wind radius information from satellite infrared images—a review. Atmos. Oceanic Sci. Lett. 16 (2), 100373.

Wang, X., Wang, W., Yan, B., 2020. TC intensity change prediction based on surrounding environmental conditions with deep learning. Water 12 (10), 2685. https://doi.org/10.3390/w12102685.

Wong, T.T., Yeh, P.Y., 2019. Reliable accuracy estimates from k-fold cross validation. IEEE Trans. Knowledge Data Eng. 32 (8), 1586–1594.

Xu, X.Y., Shao, M., Chen, P.L., Wang, Q.G., 2022. TC intensity prediction using deep convolutional neural network. Atmosphere 13 (5), 783.

Zhang, C., Wang, X., Lu, X., 2021. TC intensity classification and estimation using infrared satellite images with deep learning. IEEE J. Selected Top. Appl. Earth Observations Remote Sensing 14 (11), 9320–9334.

Zhang, Zhe, Yang, Xuying, Shi, Lingfei, Wang, Bingbing, Du, Zhenhong, Zhang, Feng, Liu, Renyi, 2022. A neural network framework for fine-grained TC intensity prediction. Knowledge-Based Syst. 241. https://doi.org/10.1016/j.knosys.2022.108195.

Zhou, R.G., Yu, H., Cheng, Y., Li, F.X., 2019. Quantum image edge extraction based on improved Prewitt operator. Quan. Inf. Process. 18, 1–24.