Early Detection Warning Systems and Intensity Estimation for Tropical Cyclones

Aditya Harikrish - 2020111009
Srijan Chakraborty - 2020115001
Anusha Nath Roy - 2020101124
Santhoshini Thota - 2021101097
Chinmay Pateria - 2021114013

Abstract

Due to a global rise in sea surface temperatures, there is a trend of more rapid intensification of cyclones and higher intensities overall. This elicits the need for a robust early warning and detection system. In this report, we present two complementary models - TwistCNN and TwistRF. TwistCNN utilizes convolutional neural networks to analyze satellite imagery and estimate current cyclone wind speeds. By leveraging deep learning on infrared imagery, it can identify patterns indicative of wind speeds. TwistRF employs random forest algorithms to forecast the peak future intensity of cyclones based on their current parameters like latitude, longitude and wind radii.

Integrating these models offers a comprehensive cyclone forecasting solution, enabling real-time wind estimation from satellite data and predicting future peak intensity. This combination provides insights into current and future cyclone behavior to facilitate disaster preparedness strategies.

The report outlines the methodologies and datasets of TwistCNN and TwistRF. It discusses future work including model integration, incorporating additional data sources, expanding geographic coverage, real-time updating, and integration with disaster management systems. Through refinement and collaboration, these models can enhance cyclone resilience and mitigate impacts.

Table of Contents

Introduction

Cyclones pose a significant threat to coastal regions worldwide, characterised by their destructive winds and torrential rains, often resulting in catastrophic consequences for communities and infrastructure. The ability to predict cyclones accurately and in a timely manner is paramount for minimising their impact. Early detection warning systems coupled with advanced intensity estimation techniques play a pivotal role in providing crucial alerts and preparing vulnerable populations for impending cyclonic events.

In recent decades, the frequency and intensity of cyclones, particularly Category 4 and above hurricanes in the Atlantic Ocean, have exhibited alarming trends. The 20-year average number of such severe hurricanes has doubled since 2000, highlighting the pressing need for robust early warning systems and improved methods for estimating cyclone intensity.

By analysing historical data and using machine learning, we can gain insights into the evolving landscape of cyclone forecasting and mitigation strategies. Ultimately, the goal is to empower decision-makers, emergency responders, and communities to better anticipate, prepare for, and respond to cyclonic events, thereby mitigating the devastating impacts on lives and infrastructure.

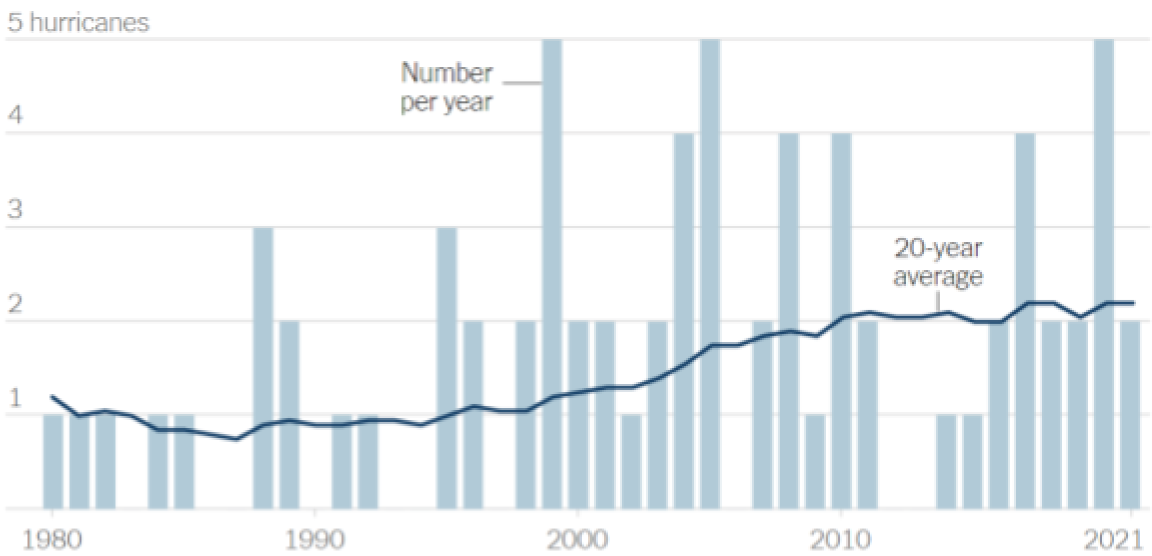


Figure 1: The 20-year average number of annual Category 4+ hurricanes in the Atlantic Ocean has doubled since 2000.

Objectives & Scope

Scope

TwistRF, the model that predicts the peak intensity of a tropical storm based on its current and past parameters, has been trained on data sourced from the National Hurricane Center (NHC), which is the USA's agency responsible for tracking and predicting tropical weather systems in regions of the North Atlantic and Northeast Pacific Oceans. The data is from 1851 to 2014.

TwistCNN, the model that estimates current wind speeds using satellite imagery, has been trained on images sourced from INSAT-3D, an Indian weather satellite. According to ISRO, "INSAT-3D is designed for enhanced meteorological observations, monitoring of land and ocean surfaces, generating vertical profile of the atmosphere in terms of temperature and humidity for weather forecasting and disaster warning." The data is from 2012 to 2021 and consists of satellite images over the Indian Ocean, along with known wind speeds.

A summary of the scope is as follows:

Model	Temporal Scope	Spatial Scope	Data Source
TwistRF	1851 - 2014	Pacific and Atlantic Oceans	NHC
TwistCNN	2012 – 2021	Indian Ocean	INSAT-3D

Objectives

The primary objective of this report is to develop and evaluate an advanced predictive analytics model that integrates data from the INSAT-3D satellite imagery database and the NHC Hurricane Database. This model aims to enhance cyclone prediction capabilities by leveraging satellite metadata and historical cyclone patterns, ultimately contributing to more effective disaster preparedness and mitigation strategies.

This report aims to bridge insights from satellite-based observations and historical cyclone records to advance cyclone prediction capabilities. By synthesising findings from these complementary datasets, this study contributes to the development of more robust early warning systems and risk mitigation strategies tailored to cyclone-prone regions. The outcomes and recommendations derived from this analysis are intended to inform policymakers, meteorologists, and disaster management authorities, facilitating proactive measures to enhance cyclone preparedness and response efforts.

TwistCNN

- Exploration of satellite imagery from the INSAT-3D database spanning from 2012 to 2021.
- Examination of satellite-derived parameters and features used for cyclone prediction, such as sea surface temperatures, cloud patterns, and atmospheric conditions.
- Identification of correlations between satellite observations and cyclone behaviour to enhance early warning systems and disaster preparedness.

TwistRF

- Utilisation of historical cyclone data from the National Hurricane Center database covering the period from 1851 to 2014.
- Investigation into cyclone characteristics, including frequency, intensity, and tracks, within the Pacific and Atlantic basins.

- Development and validation of predictive models using cyclone metadata to forecast cyclone occurrence and behaviour, including peak intensity.

By achieving these objectives, this report aims to advance the field of cyclone prediction by demonstrating the value of integrating satellite data with historical cyclone records. The outcomes and insights derived from this study will serve as a valuable resource for meteorologists, disaster response agencies, policymakers, and researchers working towards enhancing cyclone resilience and safety worldwide.

Methodology

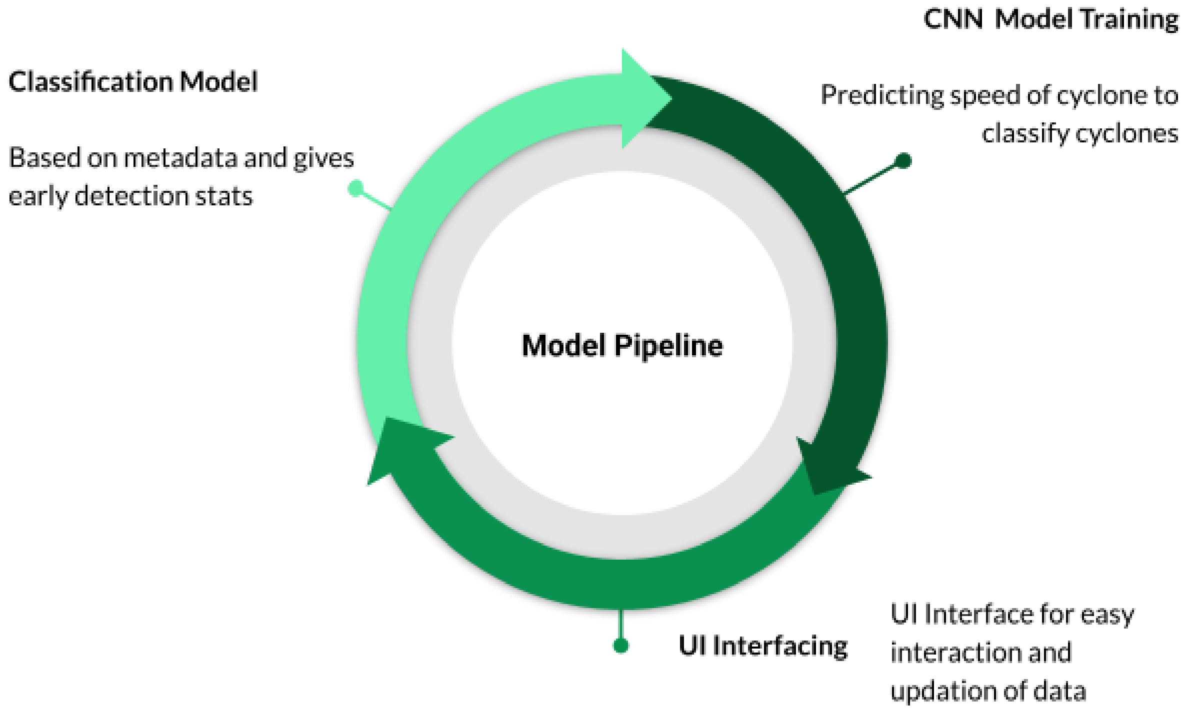


Figure 2: Model Pipeline

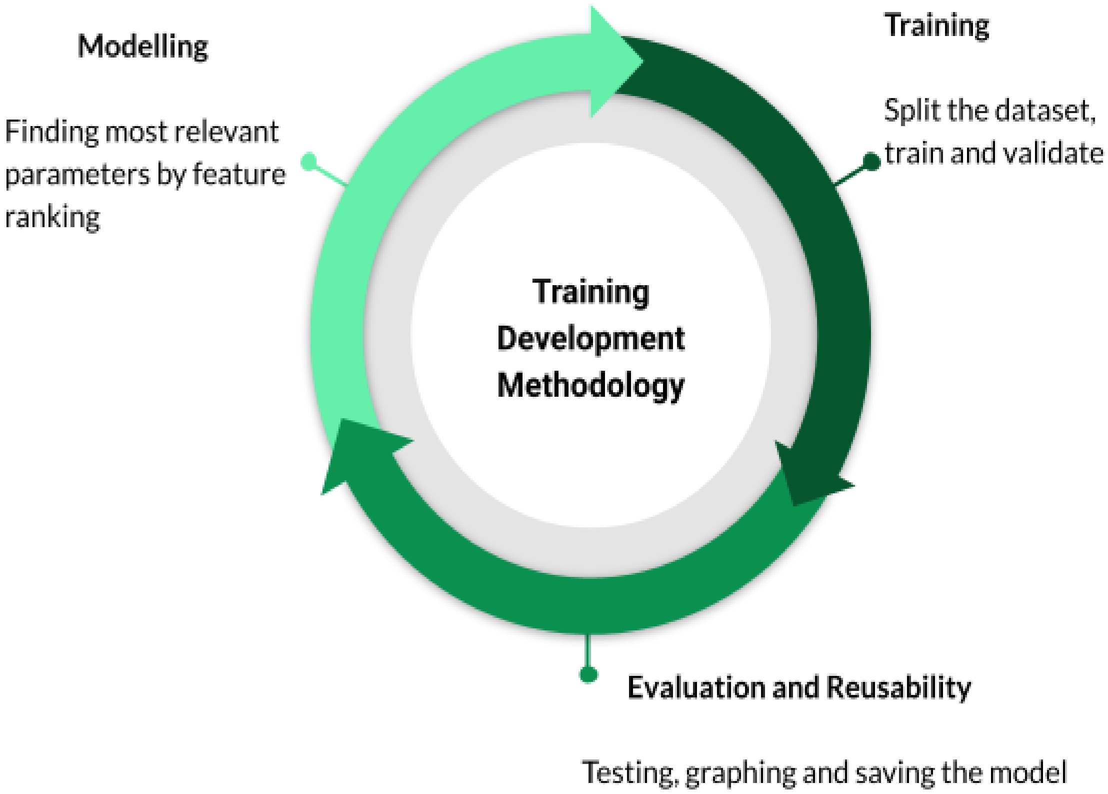


Figure 3: Development methodology of the model

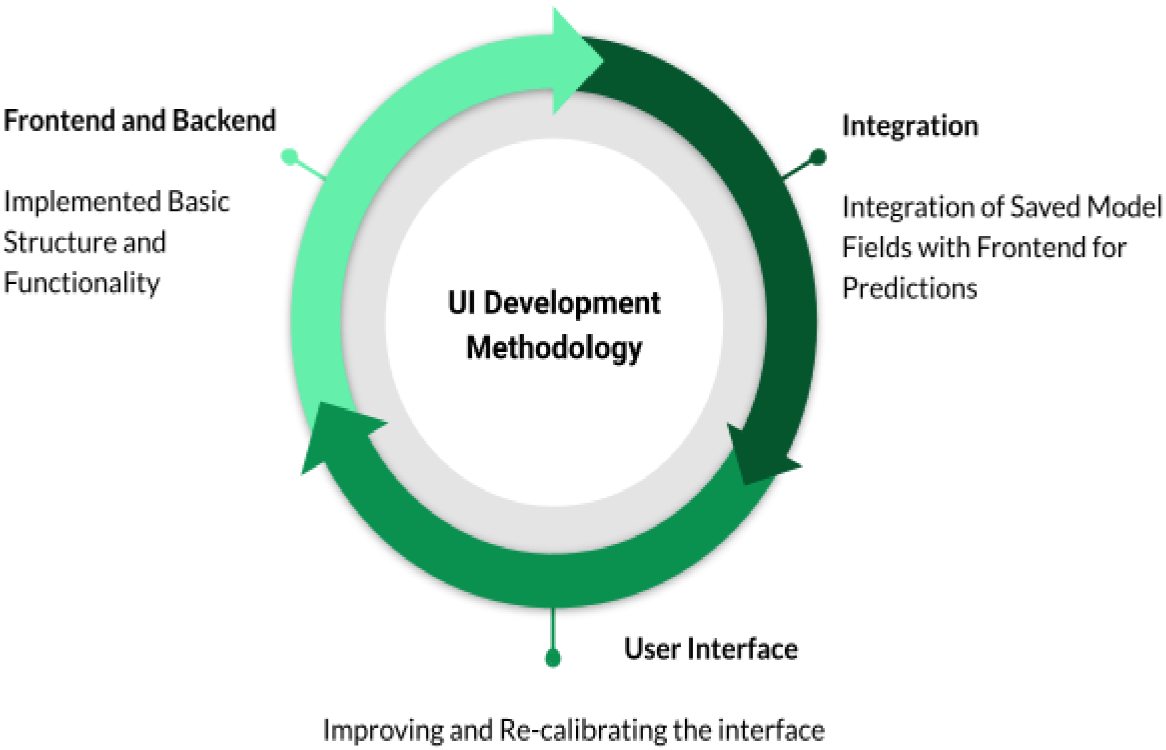


Figure 4: Development methodology of the UI

TwistCNN

Background

A convolutional neural network (CNN) is a type of artificial neural network specifically designed to process and analyze visual data, such as images and videos. They are widely used in computer vision tasks like image classification, object detection, and image segmentation. A overview of the components and operations of CNNs are as follows.

1. Convolutional Layers: Convolutional layers are the core building blocks of CNNs. They consist of filters (also called kernels) that slide over the input image to perform convolution operations. See the figure below for a convolution operation.

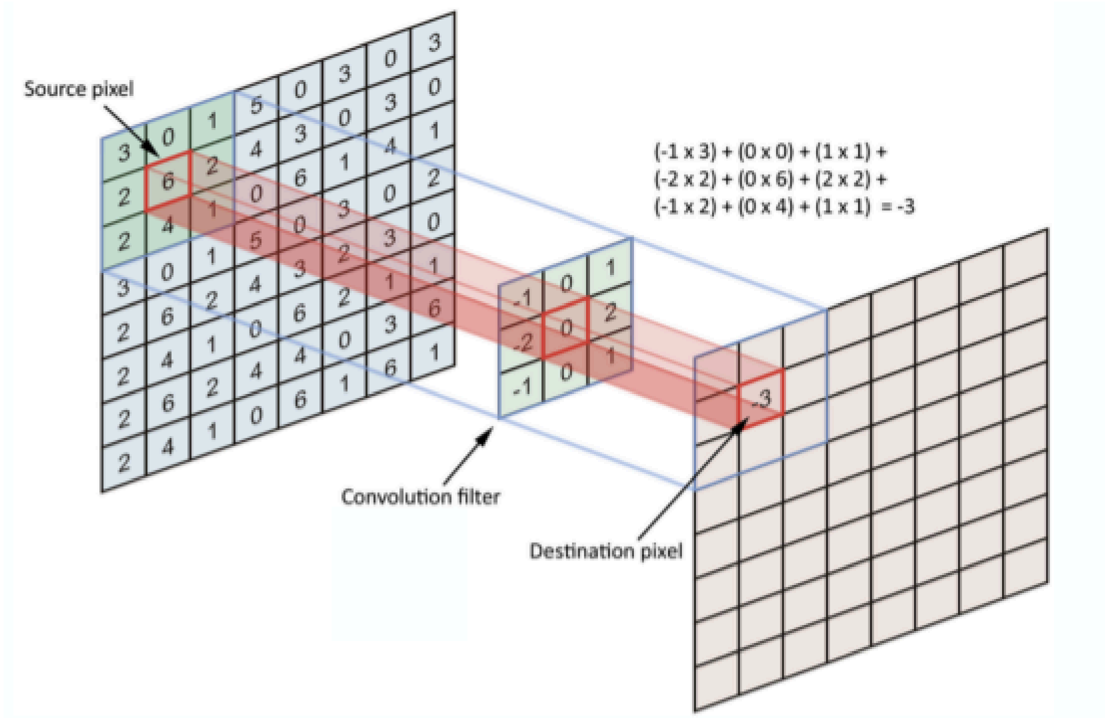


Figure 5: A convolution operation. *Credits: https://towardsdatascience.com*

   Each filter extracts certain features from the input image by performing element-wise multiplication between the filter weights and the corresponding pixels in the receptive field of the image, then summing up the results. By using multiple filters, CNNs can learn to detect various features at different spatial locations in the input image. The output of a convolutional layer is a feature map that represents the presence of specific features within the input image.

2. Activation Function: After the convolution operation, an activation function is applied element-wise to the resulting feature map. The activation function introduces non-linearity into the network, allowing it to learn complex patterns and relationships in the input data. TwistCNN uses ReLU (Rectified Linear Unit) as its activation function.

3. Pooling Layers: Pooling layers are used to reduce the spatial dimensions (width and height) of the feature maps while preserving important information. The most common type of pooling operation is max pooling, where the maximum value within each pooling

window is retained, effectively downsampling the feature maps. Pooling helps to make the network more robust to variations in the input data and reduces the computational complexity of subsequent layers.

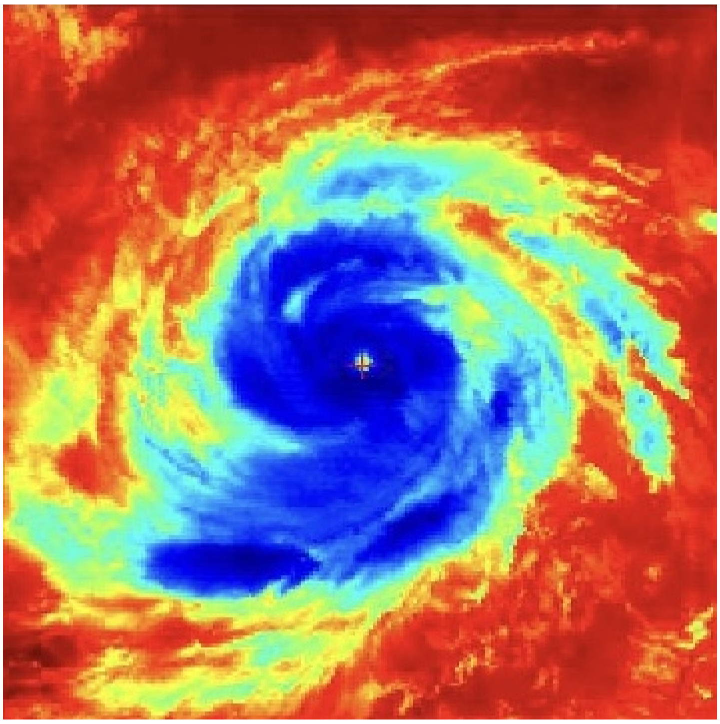
4. Fully Connected Layers: After several convolutional and pooling layers, the feature maps are flattened into a one-dimensional vector and passed to one or more fully connected layers. Fully connected layers perform transformations on the flattened feature vector using learnable weights and biases. These layers enable the network to learn high-level representations and make predictions based on the extracted features.

5. Output Layer: The output layer of a CNN typically consists of one or more neurons representing the classes or categories that the network is trained to recognize. The activation function used in the output layer depends on the nature of the task. For multi-class classification, softmax activation is commonly used, while for binary classification, sigmoid activation may be used.

6. Training: CNNs are trained using supervised learning, where they are presented with input images along with corresponding labels (ground truth). During training, the network adjusts its weights and biases using optimization algorithms such as gradient descent and backpropagation to minimize the difference between its predictions and the true labels. The process of training involves iteratively feeding batches of training samples through the network, computing the loss (error) between the predicted outputs and the ground truth labels, and updating the network parameters to minimize this loss.

Dataset

The INSAT-3D training dataset consists of the satellite images (in infrared) and the respective known wind speeds. The training-validation-testing split is conducted in an 80-10-10 ratio for TwistCNN. We chose the best performing model after repeating the above split at random 5 times.



Infrared imaging is also called thermal imaging. We chose this because in general, there exists a relationship between temperature and wind speed in cyclones. Stronger winds are often associated with regions of intense convective activity, which tend to have higher temperatures due to the release of latent heat during condensation. By measuring the temperature distribution using infrared imaging, one can infer the regions of higher wind speeds within the cyclone. The eye and eyewall regions of a cyclone have distinct thermal signatures. The eye, being relatively calm and clear, may exhibit warmer temperatures compared to the surrounding eyewall, where intense convective activity occurs. The thermal structure of these regions provides valuable information about the cyclone's intensity and wind speed distribution.

Model

Using PyTorch, we created a set of 10 repeating, consecutive sets of the following three layers:

1. a convolution layer,
2. a batch normalisation layer, and
3. a ReLU (non-linear activation) layer.

So, there are 30 layers in total. Each convolution layer uses a kernel size of 3, a stride of 1 and a padding of 1. To train the model, the inputs given were a vectorised representation of the images using numpy. The outputs of these input vectors are the intensities of the winds in knots.

TwistRF

Background

A decision tree is an directed acyclic graph (i.e., a tree) that is used in classification algorithms. The tree splits the dataset based on a series of questions. Each node represents a question. The children of a node represent the subsets of the data represented by the node itself.

A decision tree is a simple model and is highly interpretable. Decision trees mimic human decision-making processes and are easy to understand and interpret. They provide a clear explanation of how decisions are made based on the values of input features. Decision trees can handle both numerical and categorical data without requiring feature scaling or one-hot encoding. This makes them versatile and applicable to a wide range of datasets, including our use case that has numerical data. Additionally, decision trees are robust to outliers and missing values in the data. They can handle missing values by simply ignoring them during the splitting process.

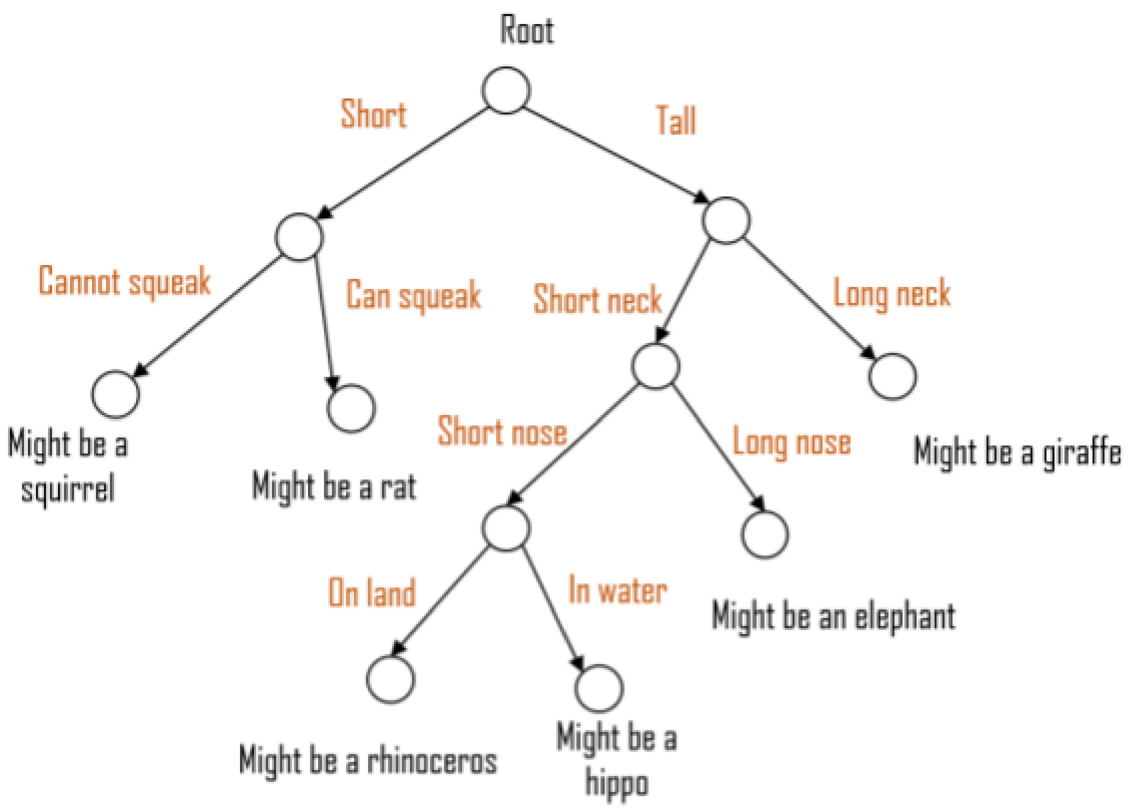


Figure 6: A simple example of a decision tree

A random forest is an ensemble learning method based on decision trees. It builds multiple decision trees during training and combines their predictions to make more accurate and robust predictions. Each decision tree in a random forest is trained on a random subset of the training data (bootstrap sample) and a random subset of features. This randomness helps to decorrelate the trees and reduce overfitting. At each node of the decision tree, the algorithm selects the best split based on a criterion, such as Gini impurity or entropy. Entropy measures the uncertainty or disorder in the data. The goal is to minimize entropy at each split, leading to more homogeneous subsets of data. We chose a random forest over a decision tree for its following advantages:

1. Random forests reduce overfitting compared to individual decision trees by combining predictions from multiple trees. By averaging the predictions of many trees, random forests achieve better generalization performance on unseen data.
2. Random forests reduce the variance of individual decision trees by randomly sampling both data points and features for each tree. This randomness helps to decorrelate the trees and reduce the risk of overfitting.
3. Random forests provide a measure of feature importance based on how much each feature contributes to reducing impurity or entropy across all trees. This information can be valuable for feature selection and understanding the underlying patterns in the data.
4. Random forests are robust to noise and outliers in the data. Since they aggregate predictions from multiple trees, they are less sensitive to individual noisy data points or outliers.
5. Training and prediction with random forests can be easily parallelized across multiple processors or machines. Each tree in the forest can be trained independently, allowing for efficient use of computational resources.

Dataset

The dataset consists of several measurements of tropical cyclones. It has 22 features: *ID*, *name*, *date of the observation*, *time of the observation*, *event*, *status of the cyclone*, *latitude*, *longitude*, *maximum wind speed*, *minimum pressure*, *low wind NE*, *low wind SE*, *low wind SW*, *low wind NW*, *moderate wind NE*, *moderate wind SE*, *moderate wind SW*, *moderate wind NW*, *high wind NE*, *high wind SE*, *high wind SW*, and *high wind NW*.

Low wind NE refers to the minimum distance of an area northeast of the eye with a wind speed 34 knots. Similarly, moderate winds and high winds refer to wind speeds of 50 and 64 knots. (1 knot = 1.852 km/hr)

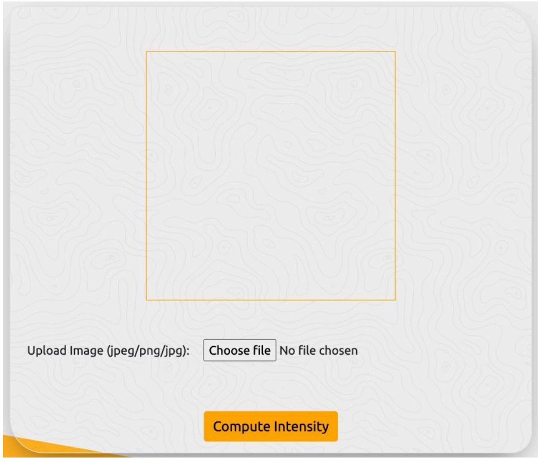
Model

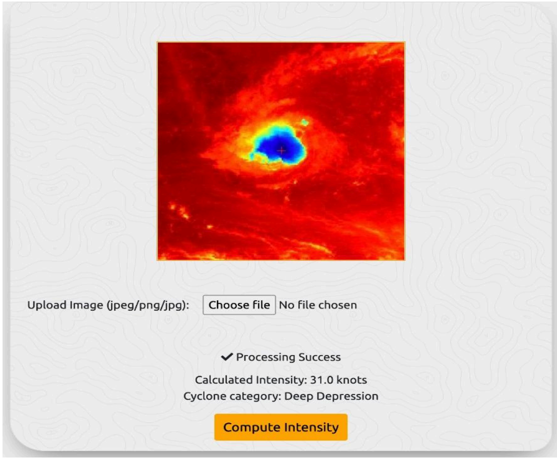
The random forest identified 6 features that contribute the greatest in estimating the intensity of the cyclone, based on the information gain through the splitting of the datasets based on said features. They are: latitude, longitude, low wind SW radius, moderate wind NE radius, moderate wind SE radius, and high wind northwest radius. Choosing more features in the decision tree did not result in any significant increase in accuracy, and would contribute to overfitting the model to the training dataset. Thus, we limited ourselves to these 6 features.

User Interface

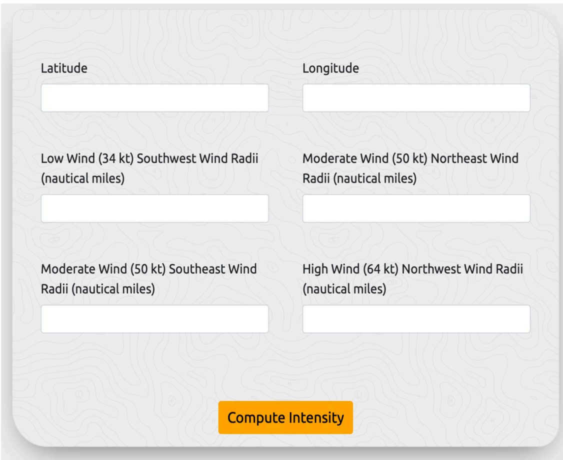
The model can be accessed using an easy-to-use UI which does not require any technical knowledge to run the model. The user needs to input the satellite imagery or metadata as needed.

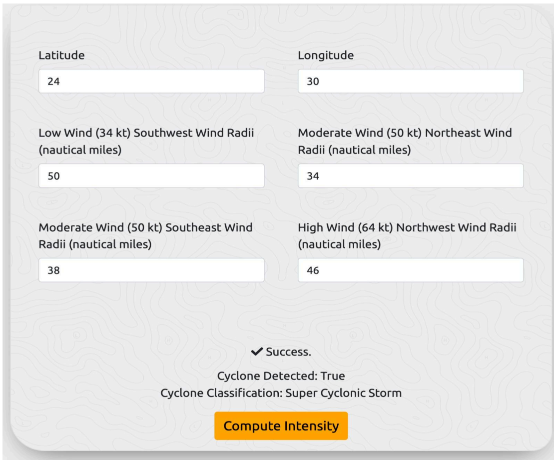
To use TwistCNN, the user simply has to upload a satellite image to the webpage and press a button.





To use TwistRF, the user needs to enter the following 6 parameters and click the button below to view the estimated peak classification of the cyclone.





Results & Discussion

TwistCNN achieved an accuracy of 82.6%. The negligible mean squared error (MSE) of 3.7% indicates that the model's predictions are not only accurate but also have minimal deviation from the true wind speed values. This low error rate highlights the model's robustness and reliability in wind speed estimation. It is worth noting that the model's performance was evaluated on a held-out test set, which was not used during training. This ensures that the reported accuracy and error metrics are representative of the model's generalization capabilities on unseen data.

The success of TwistCNN can be attributed to the powerful feature extraction capabilities of convolutional neural networks (CNNs). By leveraging the hierarchical structure of CNNs, the model can automatically learn to identify relevant patterns and features in the satellite imagery, such as temperature gradients, cloud formations, and other meteorological indicators, that are indicative of cyclone wind speeds.

Furthermore, the use of infrared imagery as input data has proven to be advantageous. The thermal signatures captured by infrared sensors provide valuable information about the cyclone's structure, intensity, and convective activity, which are closely related to wind speed distribution.

TwistRF attained a remarkable accuracy of 92%. The MSE was negligible at 3%. This indicates high precision in cyclone intensity prediction. By combining multiple decision trees, the model can effectively capture complex patterns and mitigate the inherent limitations of individual decision trees, such as overfitting and high variance. In case of misclassification, the model always overestimated the intensity of the cyclone. This is in line with our expectation that we are never caught off guard with a cyclone more intense that predicted.

Overall, the high accuracy and low error rates achieved by both TwistCNN and TwistRF demonstrate their effectiveness in addressing the challenges of cyclone wind speed estimation and peak intensity prediction, respectively. These models can serve as valuable tools for early warning systems and disaster preparedness efforts, enabling more informed decision-making and potentially saving lives and minimizing damage caused by cyclonic events.

Conclusion

TwistCNN's ability to accurately estimate real-time wind speeds from satellite imagery is a crucial component in the early detection and monitoring of cyclonic events. By harnessing the feature extraction capabilities of deep learning, the model can effectively identify patterns and thermal signatures that are indicative of wind speed distribution, providing valuable information for initial evacuation warnings and emergency response planning.

Complementing TwistCNN, TwistRF's remarkable accuracy in predicting peak cyclone intensity based on current parameters offers a powerful tool for long-term planning and resource allocation. By capturing the underlying relationships between cyclone metadata and eventual peak intensity, the model can guide decision-makers in anticipating the potential severity of an event and implementing appropriate mitigation strategies.

The integration of these two models creates a comprehensive cyclone forecasting system, bridging the gap between real-time monitoring and future intensity predictions. This synergy enables a holistic understanding of cyclone behavior, facilitating more effective coordination among stakeholders and enhancing the overall resilience of communities in the face of these natural disasters. The usage of the app requires no technical knowledge.

Future work

Integration of TwistCNN and TwistRF

While TwistCNN and TwistRF were developed independently, integrating them into a unified system could provide a comprehensive solution for cyclone forecasting and disaster management. By combining the real-time wind speed estimates from satellite imagery with the predicted peak intensity from historical data, decision-makers and emergency responders could gain a more holistic understanding of the cyclone's current and future behavior.

The integrated system could leverage the strengths of both models, enabling proactive disaster preparedness and response strategies. For example, the wind speed estimates from TwistCNN could trigger initial evacuation warnings, while the peak intensity predictions from TwistRF could guide resource allocation and long-term planning.

Incorporation of Additional Data Sources

While the current models utilize satellite imagery and historical cyclone data, incorporating additional data sources could further improve their accuracy and robustness. For instance, integrating meteorological data, such as temperature, pressure, and humidity measurements, could provide valuable context for the models to make more informed predictions.

Furthermore, incorporating data from numerical weather prediction models and ensemble forecasts could enhance the models' ability to forecast cyclone trajectories and intensity changes over time.

Integration with Disaster Management Systems

Ultimately, the goal of these models is to support effective disaster preparedness and response efforts. To achieve this, the integrated TwistCNN and TwistRF system should be seamlessly integrated with existing disaster management systems and decision support tools.

This integration would enable real-time sharing of cyclone forecasts, wind speed estimates, and intensity predictions with emergency responders, local authorities, and affected communities. It would also facilitate the coordination of evacuation plans, resource allocation, and relief efforts based on the model's outputs.

References

1. NOAA Hurricane Database on Kaggle
2. INSAT 3D Infrared Raw Cyclone Images 2013-2021 on Kaggle
3. NHC Hurricane Database (HURDAT): The New York Times article

4. Machine Learning in Tropical Cyclone Forecast Modeling: A Review [MDPI]: https://www.mdpi.com/2073-4433/11/7/676)

5. Can we use AI to improve the prediction of rapid intensification in tropical cyclones?: https://www.frontiersin.org/articles/10.3389/fmars.2023.1296274

6. INSAT-3D