DEEP LEARNING IN CYCLONE FORECASTING

MS (Research) Thesis

By

Manish Kumar Mawatwal

DEPARTMENT OF ASTRONOMY, ASTROPHYSICS AND SPACE ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

JUNE, 2023

DEEP LEARNING IN CYCLONE FORECASTING

A THESIS

Submitted in fulfilment of the
requirements for the award of the degree
of

Master of Science (Research)

by
MANISH KUMAR MAWATWAL

DEPARTMENT OF ASTRONOMY, ASTROPHYSICS AND SPACE ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY INDORE
JUNE, 2023

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **DEEP LEARNING IN CYCLONE FORECASTING** in the fulfilment of the requirements for the award of the degree of **MASTER OF SCIENCE (RESEARCH)** and submitted in the **DEPARTMENT OF ASTRONOMY, ASTROPHYSICS AND SPACE ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from August 2021 to June 2023 under the supervision of Dr Saurabh Das, Associate Professor.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

28-06-24
Signature of the student with date
MANISH KUMAR MAWATWAL

This is to certify that the above statement made by the candidate is correct to the best of my/our knowledge.

28/06/24

Signature of the Supervisor of
MS (Research) thesis (with date)
Dr Saurabh Das

MANISH KUMAR MAWATWAL has successfully given his MS (Research) Oral Examination held on 28th June, 2024.

Signature of Chairperson (OEB)

28/06/24

Signature of Thesis Supervisor

28/06/24

Signature of Convener, DPGC

28.6.24

Signature of Head of Discipline

(Acting HOD)

1

ACKNOWLEDGEMENTS

To my family and friends

Abstract

Prediction of severe weather events such as Tropical Cyclones (TCs) has always been challenging for climate researchers. A disastrous cyclone striking in the coastal region causes serious hazards to human life and economic losses. Intensity predictions are difficult because of the complicated physical mechanisms of TC dynamics and the way they interact with upper-ocean and atmospheric circulation. This research gives importance to estimating TC intensity to identify different categories of cyclones.

We attempted to predict TC intensity using Convolutional Neural Networks (CNNs) by proposing a simple and robust architecture for TC intensity estimation. The results yielded better performance than the state-of-the-art techniques with reduced computation time. In addition, we presented a visualisation portal in a production system that displays Deep Learning (DL) output and contextual information for end users. CNN model is trained and tested with classified cyclone data for cyclone identification. The model comprises a binary classifier, a multiclass classifier, a YOLOv3 based cyclone detector and a regression module. The model is tuned for the North Indian Ocean (NIO) region with binary classification accuracy of 98.4% ($\pm$0.003), multiclass classification accuracy of 63.83% ($\pm$1.3), and Root Mean Square Error (RMSE) of 16.2 ($\pm$0.9) knots.

Another major problem is Rapid Intensification (RI) classification of TCs, which is difficult to forecast accurately in a timely manner. The research also proposes a Machine Learning (ML) based classification framework for RI predictions that utilizes Support Vector Machines (SVM) in conjunction with the Synthetic Minority Over-Sampling Technique (SMOTE) to handle class imbalance. The proposed framework has the potential to be a useful tool for forecasters to issue timely warnings and prepare for the impacts of RI events.

Keywords: Convolutional Neural Network, Tropical Cyclone, Object Detection, Machine Learning, Deep Learning, North Indian Ocean, Intensity Estimation, Rapid Intensification

LIST OF PUBLICATIONS

In Refereed Journals

1. **Manish Mawatwal**, Saurabh Das, "An End-to-End Deep Learning Framework for Cyclone Intensity Estimation in North Indian Ocean Region using Satellite Imagery", Journal of the Indian Society of Remote Sensing (accepted, 2024).

2. Tushar Sharma, **Manish Mawatwal**, Saurabh Das, "An Explainable Machine Learning Approach for Predicting Rapid Intensification in Tropical Cyclones", (under review).

In Refereed Conferences

1. **Manish Mawatwal**, Saurabh Das, "North Indian Ocean Tropical Cyclone Intensity Prediction using Machine Learning Classifiers and Analysis", Regional Conference on Radio Science, Indore, 2022 (accepted).

2. **Manish Mawatwal**, Saurabh Das, "Deep Learning Regression and Classification of Tropical Cyclones based on HURSAT data," URSI AT-RASC, Spain, 2024 (accepted).

3. **Manish Mawatwal**, Saurabh Das, "North Indian Ocean Tropical Cyclone Detection Using YOLOv5," WCONF, Raipur, 2024 (accepted).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

TC	Tropical Cyclone
NIO	North Indian Ocean
ML	Machine Learning
GAN	Generative Adversarial Network
MSW	Maximum Sustained Wind
RMSE	Root Mean Square Error
YOLO	You Only Look Once
NWP	Numerical Weather Prediction
CNN	Convolutional Neural Network
DC-GAN	Deep Convolutional Generative Adversarial Network
DL	Deep Learning
RSMC	Regional Specialised Meteorological Centres
CIMSS	Cooperative Institute for Meteorological Satellite Studies
NIO	North Indian Ocean
ADT	Advanced Dvorak Technique
SVM	Support Vector Machine
SMOTE	Synthetic Minority Oversampling Technique
RI	Rapid Intensification

Chapter 1

Introduction

TC is the generic term for a low-pressure system over tropical or subtropical waters, with organisation (i.e., thunderstorm activity) and winds at low levels circulating either anti-clockwise (Northern Hemisphere) or clockwise (Southern Hemisphere). The whole storm system may be five to six miles high and 300 to 400 miles wide, although sometimes it can be even bigger.

TC, also known as hurricanes, typhoons, or cyclones depending on the region, are powerful and destructive weather phenomena that have captured the attention of scientists and the public alike. As our understanding of climate change deepens, evidence suggests that these TCs are influenced by the changing climate [1].

Recognizing this connection is crucial in addressing the challenges posed by these storms and mitigating their impact on human lives and the environment.

Climate change has the potential to affect TC in several ways. One of the primary concerns is the increase in sea surface temperatures due to global warming. Warmer ocean waters provide the necessary fuel for TCs to intensify rapidly, leading to stronger and more destructive storms. Additionally, climate change can influence atmospheric circulation patterns, which can impact the formation, track, and intensity of these storms.

Accurate prediction of TC is essential for managing their potential risks and reducing their impact. Advancements in meteorological science and technology have significantly improved our ability to forecast these storms, enabling timely evacuation, resource allocation, and emergency response. By accurately predicting the track and intensity of a TC, authorities can issue appropriate warnings and implement evacuation measures, potentially saving countless lives.

Furthermore, improved prediction capabilities empower policymakers to make informed decisions regarding climate change mitigation and adaptation strategies. By understanding the potential risks associated with TC in a changing climate, policymakers can develop resilient infrastructure,

implement effective land-use planning, and design appropriate disaster response and recovery mechanisms. Accurate predictions also assist in optimizing resource allocation, aiding in the preparation and distribution of relief supplies, and reducing the economic and social impact of these storms. To enhance TC prediction, ongoing research and investment in meteorological science are essential. This includes improving Numerical Weather Prediction (NWP) models, enhancing observational networks, and developing more advanced satellite technologies. By combining these efforts with a comprehensive understanding of the relationship between TC and climate, we can better anticipate, prepare for, and mitigate the impacts of these powerful storms.

In summary, the interplay between TC and climate change underscores the importance of accurate prediction in addressing the associated challenges. By leveraging improved forecasting capabilities, we can save lives, protect infrastructure, and develop effective strategies to mitigate and adapt to the impacts of TCs. A comprehensive approach that combines scientific research, technological advancements, and informed policymaking is vital in building resilience and safeguarding communities against the forces of nature in a changing climate.

1.1 Motivation

The NIO region, including the Arabian Sea (~24%) and Bay of Bengal (~76%) accounts for approximately 6% of the global TCs. With a staggering number of 1942 disasters recorded over the past 50 years, these natural phenomena have caused over 1408 billion dollars in economic losses. Tragically, these cyclones have resulted in the loss of 800,000 lives, equivalent to an alarming average of 43 deaths every day. Among the countries most affected by tropical cyclones, Bangladesh stands out as the hardest hit. The impact on this South Asian nation has been severe. The daily damages amount to a staggering 78 million dollars (https://public.wmo.int/en/our-mandate/focus-areas/natural-hazards-and-disaster-risk-reduction/tropical-cyclones).

The coastal region is the most affected by storm surge, which leads to colossal damage [2]. If analysts can accurately estimate TC intensity, it can prove to be a deciding factor in the amount of loss. Creating faster and more reliable estimates of storms may therefore lead to better decision-making. Improving TC prediction remains an important task for meteorologists due to the severe impact they can have on human life, properties, and economies. The immense toll these cyclones take on human lives and livelihoods as shown in Fig. 1.1 serves as a reminder of the urgent need to prioritize disaster preparedness, mitigation, and resilience measures to protect vulnerable communities from such catastrophic events.

Accurate predictions of TC forecasting will provide actionable information on TC hazards to mitigate damages and loss of life. There is growing interest in applying AI and ML techniques to improve the accuracy of operational meteorological tasks, including estimating hurricane wind speed.

1.2 Research Gap

TC is classified based on Maximum Sustained Wind (MSW) speed, which is measured within the Radius of Maximum Wind from the cyclone centre. MSW measurements were taken from land stations, ships, and aircraft. The majority of the lifecycle of a TC takes place on the open ocean, and measurements can be taken only after landfall. This leads to sparse and

Figure 1.1: Ruins of Tacloban in the Philippines after the passage of Typhoon Haiyan in November 2013. (Source: Encyclopedia of the Environment)

regional data. In large spans of ocean or remote areas of land surface, in situ observations are difficult; hence, data from satellite observations are the primary source of TC information. A solution could be the use of high temporal frequency satellite images, which have global coverage and provide continuous necessary information for intensity estimation.

There are many difficulties in TC forecasts, including understanding of the physical mechanisms and complex interactions with ocean and surrounding atmosphere environments. ML can provide new ways to improve accuracy and efficiency of TC prediction [3].

Cyclone's evolution depends on the altitude and time, which leads to difficulty in modelling. At present, national forecasts are done by consensus methods combining different dynamical models. Statistical forecasting models perform poorly regarding dynamical models even though the database of past hurricanes is constantly growing. A fixed region of water body has some limitations. First, the tracked storm must stay in the region, but the storms often cross oceans, forcing the selection of an extensive region, even if memory issues constrain it. Second, learning local phenomena on a large scale and non-centred images can be difficult. Upper ocean feedback has important effects on TCs, but few operational numerical forecast models take it into consideration, which reduces the performance of the models. However, human intervention in analysing the similarity in cloud patterns means the technique is still not perfect. Researchers are relying on computers to detect similarity in images. State-of-the-art satellite image-based intensity estimation needs to lower noise between neighbouring prediction values and reduce the Root Mean Square Error (RMSE).

Air-sea interaction and other processes in the TC core are highly non-linear, suggesting that non-linear methods can improve statistically based intensity forecast models [4].

Because of the complexity and randomness of weather factors, meteorologists and forecasters have found it difficult to predict weather events using computational models. The availability and volume of data, as well as the computational time required, depend on the complexity of the variables. Observations from weather stations, satellites, and weather balloons are used to source the data, which is then used by NWP models.

However, these models have limitations, particularly when forecasting over longer periods, as the interpolation of large volumes of data takes more computational time. As the number of processes increases, the modelling time for forecasting also increases. The problem lies in the existing techniques used in estimating cyclone intensities and the need to categorise them to comprehend and mitigate the economic losses and human lives.

DL and ML techniques are successful in Computer Vision, Natural Language Processing, pattern recognition, object detection, etc., but their application to TC intensity prediction is in its nascent stage. A TC dataset using automated algorithms from satellite images can reduce these issues. Hence, ML/DL techniques could extract features from satellite images just like researchers determine cloud patterns, which are related to TC intensity ranges. ML/DL can use these features further as predictors to estimate TC intensity.

1.3 Objectives

The broader aim is to estimate the TC intensity in the NIO region, and forecast the intensity using satellite data. Exploratory data analysis and visualisation of cyclones in NIO basin are also done to understand the cyclonic activity in the concerned region.

The major contributions of the proposed study are:

- Development of an improved DL model for estimating the intensity and classification of the storm using the infrared satellite imagery dataset, primarily over the NIO region.
- Development of an explainable ML approach to detect RI events in the lifecycle of TC.

1.4 Thesis Contribution

The widespread use of image recognition and object detection techniques with higher accuracy has prompted us to explore the use of these methods on satellite images for estimating TC intensity and detecting RI events. We developed an efficient CNN model that reduces estimation time and increases prediction accuracy. This research presents a CNN architecture designed specifically for TC intensity estimation over the NIO region and a

ML technique for classification of RI events over different regions. A pipeline has been created that classifies the cyclone image into five different classes, giving predicted intensity of the cyclone and a bounding box of the cyclone's location on the image as well. Interpretation of the models was also attempted so that the important features can be identified.

1.7 Organization of the Thesis

Chapter 1 presents Motivation, Research Gap, Objectives and Thesis contribution.

Chapter 2 presents Literature Review.

Chapter 3 presents Dataset description for the present research.

Chapter 4 discusses various ML and DL algorithms and frameworks used in the research.

Chapter 5 presents the Methodology and Results of the proposed DL framework for cyclone detection, classification, and intensity estimation.

Chapter 6 presents the Methodology and results of ML models for RI events classification over different basins.

Chapter 7 discusses Limitations, Conclusion and Future Scope.

Chapter 2

Literature Review

TCs are essential features of the Earth's atmosphere, as they transfer heat and energy between the equator and the cooler regions nearer the poles. It is of utmost importance to forecast TC in advance and take measures to prevent devastating effects. Determining an accurate TC intensity value will help to provide early warning.

Any TC that develops within the NIO region between 100°E and 45°E is monitored by the Indian Meteorological Department (IMD) New Delhi. The cyclone is classified based on the intensity; Table 2.1 shows the Saffir-Simpson Hurricane Wind Scale which acts as a standard for defining wind speed ranges. An average of 90 tropical cyclones (TCs) form annually over the tropical ocean waters [5]. TC causes significant socio-economic damage to human life and assets, and the damage is directly correlated with the intensity of the cyclone [6].

Table 2.1: TROPICAL CYCLONE INTENSITY SCALE

Cyclone Category	Sustained Wind Speed Range (knots)
Depression	17-27
Deep Depression	28-33
Cyclonic Storm	34-47
Severe Cyclonic Storm	48-63
Very Severe Cyclonic Storm	64-89
Extremely Severe Cyclonic Storm	90-119
Super Cyclonic Storm	≥120

2.1 Cyclone Classification, Regression and Detection

TC intensity forecasts can be divided into three different categories: (1) Dynamical models, also known as Physics based models, (2) Statistical models, and (3) Statistical-Dynamical models, a combination of dynamical model with statistics. TC analysts and forecasters have long been using the

Dvorak technique as a standard tool for estimating TC intensity from satellite imagery. Another method uses enhanced infrared to quantitatively estimate the intensity of a tropical system. Based on the cloud patterns in satellite imagery, it shows a sign of cyclogenesis before the storm reaches tropical storm intensity. Based on these patterns, a series of rules and careful analysis, a forecast is made.

Intensity estimation of TC from satellite images assumes that similar intensities have similar patterns, there are hidden correlations between TC images. The method is to define cloud patterns, construct image features and analyse their similarity to other images [7].

The deviation angle variation technique uses infrared images and the variance is used to estimate the intensity values. The constraint being the centre of TC images should be visibly marked which makes it difficult to tune parameters across multiple regions [8].

Dvorak technique estimated the intensity via human interpretation of the cyclone shape when direct measurements were not available. Techniques vary depending on the length and curvature of the storm bands and the intensity is estimated by relating between the features. Some satellites provide images of temperature, clouds, water vapour, and precipitation which can be indirectly used for estimating TC intensity [9].

While the spatial patterns in infrared satellite imagery strongly relate to TC intensity, researchers found that parts of the Dvorak technique are subjective, leading to two different intensity estimates, decreasing the efficiency of the process [10].

Advanced Dvorak Technique uses passive microwave data from aircraft to estimate the intensity, but the performance was worse for weaker storms [11].

Statistical models are also, found to perform poorly in handling complex and nonlinear relationships between TC related predictors [12].

As in-situ measurements over large spans of ocean or remote areas of land surface are difficult, data from satellite observations are also used as the alternate primary source of TC information. Satellites measure temperature, clouds, water vapor, and precipitation through active or passive mode and can be indirectly used for estimating TC intensity [13].

Intensity predictions are particularly very challenging due to the complicated physical mechanisms in TC dynamics and the way they interact with upper ocean and atmospheric circulation [14].

In recent times, ML and DL models were also attempted to predict the cyclone intensity. [15] has applied neural networks for this purpose to the cyclones over Western North Pacific and [16] on the North Indian Ocean cyclones. An ensemble base CNN classifier on outgoing longwave radiation for classifying TC is done in [17]. Four state-of-the-art U-net models were developed in [18] for the detection of regions of interest for cyclones. There are attempts to estimate TC intensity from satellite data by using CNN in recent times [19].

[20] introduces a dynamic competitive neural network classifier for predicting the maximum potential intensity of TC using a 10-year dataset of Western North Pacific cyclones and monthly Mean Sea Surface Temperature. The model incorporates attribute selection and a Binary Trigger to optimize network training. The experiments conducted demonstrate the promising performance of the proposed model, highlighting its potential for accurate and efficient maximum potential intensity prediction.

Presently, data is sourced from varied types of observations and act as a boundary condition for the NWP models. A drawback of the NWP model is that as the prediction period increases, the computation time increases to handle the interpolation of large amounts of data, and the errors continue to accumulate until new observations are made. Further, upper ocean feedback has important effects on TCs, but at present, only a few operational numerical forecast models take it into consideration [21].

[22] compared intensity estimates of TC in the NIO by the Joint Typhoon Warning Center (JTWC) and the Regional Specialized Meteorological Centre New Delhi (RSMCND), highlighting differences in intensity estimation based on cyclone intensity, intensity trend, and translation speed. Results show that JTWC estimates are generally higher than RSMCND estimates, with Mean Absolute Differences (MAD) and Root Mean Square Differences (RMSD) of 9.7 knots and 13.3 knots respectively. Location estimation differences and standard deviations are also quantified. The study

emphasizes the need for a robust objective method to ensure uniformity in intensity estimates across agencies, aiding operational forecasters in better monitoring and post-storm analysis.

[23] proposed a CNN approach for estimating TC intensity without the need for explicit feature extraction. Transfer learning experiments using a VGG19 model pre-trained on ImageNet were conducted on grayscale infrared images of TCs from geostationary satellites in the Western North Pacific region. The retrained model achieved comparable performance to existing feature-based approaches, with a RMSE of 13.23 knots. Additionally, the model learned generic TC features that were previously identified as important indicators of TC intensity.

[24] presents a novel ML based method for estimating the intensity or maximum sustained wind speed of TC using infrared satellite imagery. The approach utilizes support vector regression and statistical features that capture the uniformity of temperature bands within a hurricane. The proposed method achieves a low prediction error (approximately 10 knots) and performs comparably to SATCON consensus, demonstrating its potential for accurate intensity prediction. The scheme is also analysed in relation to errors in hurricane centre annotation and aircraft reconnaissance data.

The research conducted by [25], focuses on utilizing deep CNN for the task of estimating typhoon intensity. The study was conducted in collaboration with the Shanghai Typhoon Institute of the China Meteorological Administration. To conduct their analysis, the researchers utilized data from the Northwest Pacific region, specifically sourced from the website of the Earth Observation Research Center operated by the Japan Aerospace Exploration Agency. The proposed deep CNN achieved significant accuracy in its predictions, with a top-1 classification accuracy of 81.4% and a top-2 classification accuracy of 93.3%. These results demonstrate the effectiveness of the algorithm in accurately estimating the intensity of typhoons.

[26] proposes a two-branch CNN model (TCIENet) to estimate TC intensity using infrared and water vapor images in the northwest Pacific basin. The model achieves the best performance with an overall RMSE of 5.13 m/s and mean absolute error (MAE) of 4.03 m/s when using input images of size 60x60 pixels. The correlation between rainfall intensity and estimation errors

suggests that the TCIENet model performs best for tropical storms and super typhoons but shows degraded performance for moderate intensities and tropical depressions.

In the study by [27], a modified CNN was employed for the task of cyclone detection and classification. The researchers utilized KALPANA-I satellite images obtained from the Meteorological and Oceanographic Satellite Data Archival Centre (MOSDAC) operated by the IMD. The proposed algorithm achieved a remarkable accuracy of 97% in distinguishing between storm and non-storm images. Moreover, it achieved a 45% accuracy in multiclass classification, encompassing the identification of cyclone eye. The cyclone detection aspect of the algorithm exhibited a confidence level of 84%. Therefore, the research addressed the challenges of classifying storm and non-storm images, detecting cyclone eyes, and achieving multiclass classification in the context of satellite imagery, demonstrating promising results.

[28] introduces a deep-learning-based diagnostic model for accurately estimating TC intensity using infrared satellite imagery. The model achieves a RMSE of 13.24 knots, providing an objective estimate of intensity. Additionally, the article presents a visualization portal that displays the deep learning output and contextual information, making it user-friendly and innovative in the field of TC analysis.

[29] used ML algorithms (XGBoost and Decision Tree) to estimate TC grade and Maximum Sustained Wind (MSW). With an accuracy of 88% for cyclone grade and a RMSE of 2.3 for MSW, the models demonstrate promising results, improving to an average accuracy of 98.84% for higher grade categories.

[30] focused on the classification of TC intensity using a cloud intensity classification technique based on feature extraction and pattern recognition. The research utilizes images of ten cyclones from 2013 to 2018 over the Bay of Bengal and Arabian Sea basins. By employing ML classifiers, including Random Forest, the proposed method achieves an accuracy of 86.66% for intensity classification. Additionally, the RMSE for estimating MSW speed with the Random Forest classifier is found to be 9.84 knots, demonstrating the feasibility of the approach for intensity classification from infrared

images of TC. Most of these early works were carried over the Atlantic basin, and there is a serious dearth of satisfactory models over the North Indian Ocean (NIO) region. The major reason for limited performance of the DL models in this region is due to limited amounts of past data and complex tropical/equatorial weather phenomena.

[31] study explored the use of simple and computationally inexpensive ML models to enhance MSW estimates of TCs using Advanced Dvorak Technique (ADT) retrieved features from satellite imagery. The models are trained and validated using ADT TC analysis parameters and best track datasets from 2005 to 2016 globally. The proposed model, named "artificial intelligence (AI) enhanced ADT (AiDT)", demonstrates significant improvement over ADT, achieving a global RMSE of 7.7 to 8.2 knots for the 2017-2018 datasets and showing a 30% to 23% reduction in error compared to ADT.

[32] introduced a DL based Multilayer Perceptron (MLP) model for predicting TC intensity. The model surpasses other statistical-dynamical models and achieves comparable results to the HWFI model in operational and real-time forecasting scenarios. Additionally, a lightweight MLP coupled with a synthetic TC track model generates realistic TC intensity distribution, making it a promising approach for operational and climate study purposes.

[33] examined the efficiency of ML classifiers in predicting the intensity of TCs in the NIO using best track data. The classifiers, including Naive Bayes, Logistic regression, MLP, Sequential minimal optimization, C4.5 decision tree, Random Trees, and Random forests, were compared based on correctly classified instances. The results demonstrate high classification accuracy of 97─99% with ML classifiers, highlighting their potential for improving TC intensity prediction in a computationally efficient manner.

2.3 Rapid Intensification

RI was defined by [34] as events of rapid increase in maximum sustained surface wind speed by at least 30 knots (approximately 35 miles per hour) over a 24-hour period.

Deeper understanding and improvement in RI detection has been the top priority of the global agencies including National Hurricane Center (NHC) [35].

There are several factors that can contribute to RI in a TC. These include warm ocean waters, high levels of moisture in the atmosphere, low vertical wind shear (i.e., the difference in wind speed and direction at different heights in the atmosphere), and a well-defined low-level circulation. RI can be particularly dangerous, as it can result in a storm reaching major hurricane or typhoon status with little warning. This can leave communities vulnerable and unprepared, and can result in devastating impacts such as high winds, storm surge, and heavy rainfall. As a result, predicting and monitoring RI is a key area of focus for TC forecasting and emergency management efforts.

RI is difficult to forecast, and seasons with higher number of RI cases have larger annual average forecast errors. Researchers have explored the role of ocean characteristics, inner-core processes, and environmental interactions in TC intensity change [36]. However, most studies have focused on only one of the above areas, leaving many questions unanswered about the precise physical mechanisms responsible for RI.

Researchers have been exploring various techniques to improve RI prediction, including statistical models, ML and DL methods [37].

[38] study contributes to the existing literature by utilizing an unsupervised learning approach to quantify the distinguishability between rapid intensification (RI) and non-RI environments for Atlantic Ocean tropical cyclones (TCs) from 2004 to 2016. By combining principal component analysis with k-means cluster analysis on TC-centered Global Forecast System analysis (GFSA) grids, consisting of various environmental variables, the study demonstrates improved separability compared to operational RI forecast predictors for multiple RI definitions. The findings highlight the importance of mid- and upper-level relative humidity in identifying short-term RI onset and weaker absolute vorticity in long-term, higher-magnitude RI. Additionally, the study identifies other valuable predictors, such as optimal thermodynamic RI ingredients along the TC's mean trajectory, suggesting their potential use in RI prediction. Traditional statistical analysis techniques have been used to find the associations

between rapidly intensifying TCs and their environmental properties, but they are limited in their ability to capture the complex interactions between multiple non-linear factors for example, air-sea interaction.

In recent years, ML-based approaches have shown some promising results. Some works utilized SVM for the classification of RI cases [39].

[40] introduced an automatic knowledge discovery framework to identify spatiotemporal precursors to RI from TC environmental fields. The framework involves composite analysis of RI and non-RI fields, clustering to detect homogeneous regions around cyclone centers, statistical analysis to determine candidate precursors, and comparison with existing predictors. Applying the framework to factors such as zonal wind, relative humidity, and vertical shear, the automatically discovered precursors demonstrate comparable or superior capability for estimating the probability of RI compared to manually labeled predictors in existing schemes like the Statistical Hurricane Intensity Prediction Scheme (SHIPS). The study used Long Short-Term Memory (LSTM) networks to predict RI of TCs.

[41] utilized MLP models to forecast the intensity of tropical cyclones in the Atlantic and eastern Pacific regions. They used environmental predictors, which were derived from numerical model forecast data or from the statistical-dynamical SHIPS database.

More recently, [42] found that a CNN approach is more skilful at predicting RI than the current operational RI models. The study introduced "I-RI," a probabilistic model based on a CNN, to predict rapid intensification (RI) in North Atlantic and eastern North Pacific TCs. The model's RI probability output is compared to two operational RI guidance methods—empirical and deterministic—for different lead times. The results indicate that in North Atlantic TCs, I-RI exhibits higher skill in predicting RI over 12- and 24-hour lead times compared to the operational RI guidance. In eastern North Pacific TCs, I-RI outperforms the empirical guidance but falls short of the deterministic guidance at all thresholds. However, for TCs north of 15°N, where deterministic skill is lower, I-RI demonstrates superior performance over more than half of the RI thresholds compared to the deterministic operational guidance. Additionally, the study emphasizes the value of utilizing two-dimensional structures within TC satellite imagery and their

evolution, as captured by the convolutional neural network, as effective indicators of RI over 12-24 hours, surpassing existing scalar assessments based on satellite brightness temperature.

[43] introduced a novel method for analyzing the potential for RI change in TC using satellite observations of precipitation obtained from microwave radiometers. The approach involves condensing environmental and vortex information using a low wavenumber representation of a Rain Index (RaIn) derived from passive microwave observations. A DL, multilayer neural network (NN) is trained using rain and wind changes over the next 24 hours. The resulting NN demonstrates excellent performance in identifying RI, defined as a hurricane wind speed increase of over 30 knots within a 24-hour period. Analysis of the NN structure provides valuable insights into the physics of TCs and offers potential improvements for model forecasting. Furthermore, the research highlights the influence of environmental conditions up to 1050 km from the TC center, impacting RI through processes such as absolute angular momentum inflow, wind shear stabilization, and the steering of upper tropospheric outflow jets. These findings can be leveraged to develop a real-time RI discriminant for operational applications.

[44] employed a sophisticated AI system, their study significantly advances RI prediction performance compared to previous efforts. With approximately 21-50% improvement in Probability of Detection (POD) while reducing the False Alarm Rate (FAR), it surpasses the capabilities of existing studies. Moreover, the study uncovers essential SHIPS variables that have been overlooked in previous research by utilizing variable importance scores. These results establish a solid baseline for future endeavours in RI prediction, providing a framework for the identification of new predictors using more intricate AI techniques. The study emphasizes the necessity of comprehensively exploring and harnessing the complete SHIPS database, shedding light on its untapped potential for further advancements in TC RI prediction studies.

[45] introduced a CNN, which is applied to ECMWF ERA-Interim reanalysis data to identify new features relevant to RI. The study expands the existing AI system to incorporate large-scale environmental conditions, refining

features such as specific humidity, vorticity, horizontal wind, and ozone that aid RI prediction. By combining these new features with SHIPS database inputs, the RI prediction performance is significantly improved, enhancing Kappa, POD, and FAR metrics by 43%, 23%, and 30% respectively compared to a modern classification model using only SHIPS inputs. This literature review highlights the potential of deep learning techniques in extracting informative features for improved RI prediction in TCs.

[46] study reveals a linear correlation between TC's 24-hour subsequent intensity change (DV24) and deviations in satellite observations of inner-core precipitation, ice water content, and outflow temperature from threshold values specific to neutral TCs with consistent intensity. These threshold values exhibit a linear relationship with TC intensity. By incorporating machine learning techniques and combining inner-core precipitation with predictors used by the National Hurricane Center (NHC) for probabilistic RI forecast guidance, the proposed model surpasses the NHC operational RI consensus. During 2009-2014 in the Atlantic basin, the model achieves a Peirce Skill Score improvement of 37%, 12%, and 138% for DV24 $\geq$ 25, 30, and 35 knots, respectively. The probability of detection also increases by 40%, 60%, and 200% compared to the operational RI consensus, while maintaining only a minimal increase in the false alarm ratio of 4%, 7%, and 6%. This literature review underscores the efficacy of incorporating machine learning methods and inner-core precipitation observations to enhance RI prediction accuracy in TCs.

[47] examined deterministic and probabilistic intensity models used by the National Hurricane Center (NHC) to assess RI forecast progress. Historically, deterministic models showed limited RI utility, but since 2015, dynamical models have demonstrated better performance for the Atlantic, while statistical models excel for the eastern North Pacific. Probabilistic RI guidance has shown modest skill since its introduction in 2001, with the DTOPS tool currently being the most effective among NHC's probabilistic models. The Hurricane Forecast Improvement Program has introduced a new RI metric to measure programmatic progress, revealing a ~20-25% improvement in RI forecasts since the baseline period of 2015-2017. This

literature review highlights the ongoing efforts to enhance RI forecasting capabilities for TCs.

[48] study presents an automatic and objective method utilizing deep residual network (ResNet) and long short-term memory (LSTM) models to identify the RI trend in typhoons based on satellite images from 2005 to 2018 in the Northwest Pacific and South China Sea. The method incorporates typhoon lifecycle indication and demonstrates effective forecasting and identification of RI trends. Application of the method to operational typhoon satellite cloud images in 2019 shows its capability to capture sudden changes in typhoon intensity, achieving a threat score of 0.24 in independent sample estimation.

[49] study investigates the influence of tropical cyclone (TC)–trough interactions on TC intensity changes, focusing on rapid intensification (RI). Through the use of ML techniques and clustering algorithms, upper-tropospheric troughs involved in TC–trough interactions were classified into three distinct clusters. Composite analyses of RI and non-RI TCs were conducted, examining the upper-tropospheric potential vorticity structure, TC convective structure, and TC environment. The findings indicate that RI episodes are associated with shorter zonal wavelengths and greater displacements of upper-tropospheric troughs compared to non-RI episodes. RI occurrences were more frequent when an upper-tropospheric cutoff low was located approximately 500-1000 km southwest of the TC. RI was also linked to environments with reduced ventilation of the TC warm core by low-entropy air. While potential trough-induced forcing for convection did not show a strong relationship with RI, RI episodes exhibited heightened convective activity within the TC inner core as observed by satellite imagery.

Chapter 3

Dataset

Fig. 3.1 shows the cyclonic regions of the world; wherein the cyclones are termed hurricanes in America or typhoons in Japan, our interest of study is North Indian Oceanic region, preferably from 40E to 100E and 10S to 30N which is shown in Fig. 3.2. The cyclones in the Northern Hemisphere move
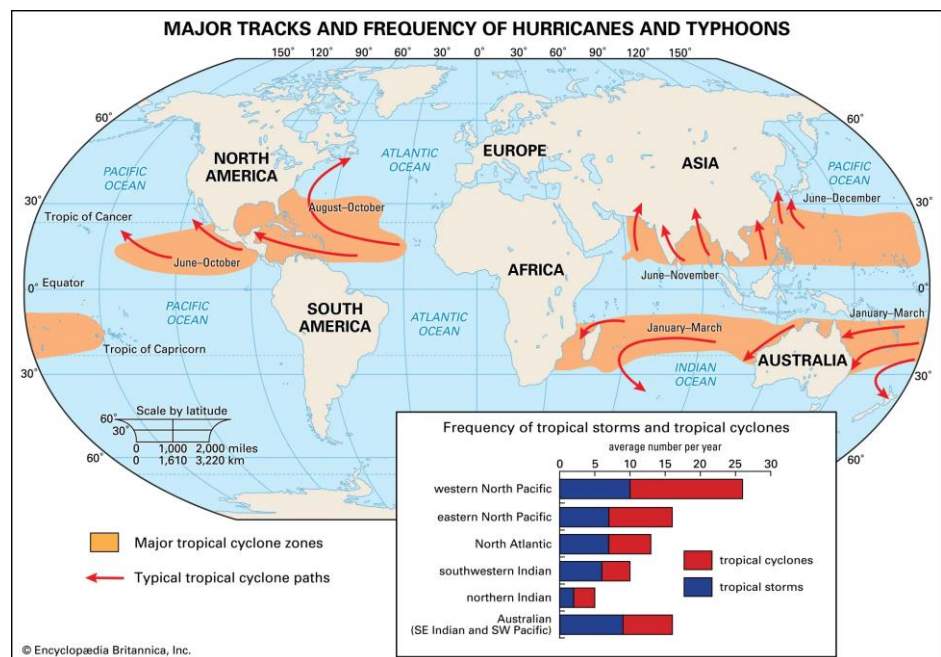

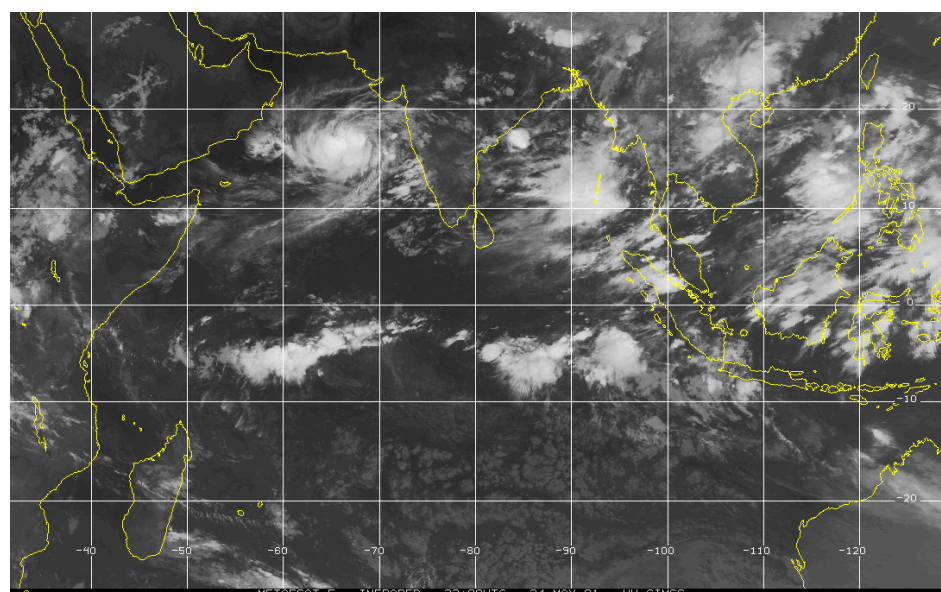Figure 3.1: Different cyclonic regions of the world (Source: Britannica)

Figure 3.2: NIO region of study (Source: CIMSS Tropical Cyclone Archive)

upwards towards the poles and cyclones in the Sothern Hemisphere move downwards towards the poles.

3.1 Cyclone Intensity Data

Best Track Data as shown in Fig. 3.3 represents the "best" location (storm track) estimates for TCs throughout the lifetime of the storm, but also the intensity (central pressure and wind speed) estimates at each point along the track. Because these storms are complex and highly dynamic and observations are limited and subject to uncertainty, scientists need to conduct post-event analysis of all available data to ensure that each storm is appropriately represented in the best track data set. Best track data has a wide range of applications across disciplines. It contains information like basin of origin, name of the cyclone, date (DD/MM/YYYY), time (UTC), latitude, longitude, tropical cyclone number, estimated central pressure, maximum sustained surface wind, pressure. Best track data is available on the RSMC website

(https://rsmcnewdelhi.imd.gov.in/report.php?internal_menu=MzM=).

This is provisional Best Track. The Final will Be uploaded after review of the panel.											
Serial Number of system during year	Basin of origin	Name	Date(DD/MM/YYYY)	Time (UTC)	Latitude (lat)	longitude (Long)	CI No [or "T. No"]	Estimated Central Pressure (hPa) [or "E.C.P"]	Maximum Sustained Surface Wind (kt)	Pressure Drop (hPa)[or "delta P"]	Grade (text)
1	BOB			0000	5.3	84.0	1.5	1004	25	3	D
	BOB			0300	5.6	83.7	1.5	1004	25	3	D
	BOB		03-03-2022	0600	6.1	83.4	1.5	1004	25	3	D
	BOB			1200	7.1	83.3	1.5	1004	25	3	D
	BOB			1800	7.8	83.1	1.5	1004	25	3	D
	BOB			0000	8.5	83.0	1.5	1004	25	3	D
	BOB			0300	8.9	82.9	1.5	1004	25	3	D
	BOB		04-03-2022	0600	9.2	82.8	1.5	1003	25	4	D
	BOB			1200	10.0	82.7	2.0	1002	30	5	DD
	BOB			1800	10.3	82.6	2.0	1002	30	5	DD
	BOB			0000	11.0	82.6	2.0	1002	30	5	DD
	BOB			0300	11.4	82.6	2.0	1002	30	5	DD
	BOB		05-03-2022	0600	11.6	82.6	2.0	1002	30	5	DD
	BOB			1200	11.7	82.5	2.0	1002	30	5	DD
	BOB			1800	11.5	82.4	1.5	1003	25	4	D
	BOB			0000	11.2	82.2	1.5	1004	25	3	D
	BOB		06-03-2022	0300	Weakened into a Well-Marked Low Pressure Area over Southwest Bay of Bengal						

Figure 3.3: Snapshot of Best Track Data taken from RSMC New Delhi

3.2 Cyclone Image Data

CIMSS TC image database as shown in Fig. 3.4 contains images of all the locations of the world. Long wave infrared images of NIO are taken from CIMSS Tropical Data Archive captured by Meteosat-5/7/8 for the period 2000-2022. The images which match the RSMC New Delhi best track data of the NIO region are filtered. Each image contains the name of the satellite (e.g., Meteosat-8), the type of the image (Infrared), time, date. The temporal

Figure 3.4: Screenshot of webpage from CIMSS Tropical Cyclone Archive

resolution is 3 hours or 6 hours. The images have a 10km spatial resolution. This removes the un-complying and noisy characteristics and improves the efficiency of the model. Table 3.1 shows the list of cyclones with their names and year of occurrence from the year 2000 to 2022 which have been analysed.

3.3 SHIPS Model Data

The Statistical Hurricane Intensity Prediction Scheme (SHIPS) was developed in response to the limited objective guidance for intensity prediction at the NHC. John Kaplan and Mark DeMaria from the Hurricane Research Division initiated the project after recognizing the need for improved intensity forecasting during Hurricane Joan in 1988. Building upon previous statistical intensity forecasting efforts, the SHIPS model integrates predictors from climatology, persistence, the atmosphere, and ocean to estimate changes in TC MSW. The first real-time runs of SHIPS began in 1990, initially providing forecasts up to 48 hours. Over the years, the model has undergone advancements, transitioning from a "statistical-synoptic" to a "statistical-dynamical" approach by incorporating predictors from atmospheric forecast models in addition to analyses. Despite showing skill compared to climatology and persistence forecasts, SHIPS has faced challenges in accurately predicting rapidly intensifying cyclones. To address this, the rapid intensity index (RII) was developed to estimate the probability of RI within the next 24 hours. SHIPS provides intensity forecasts for the Atlantic, eastern and central North Pacific, while a similar model called the Statistical Typhoon Intensity Prediction Scheme (STIPS) was developed for other regions. SHIPS and STIPS employ linear regression techniques, with the Logistic Growth Equation Model (LGEM) offering an alternative by utilizing a nonlinear differential equation to forecast intensity changes, accounting for time variations in predictors and overcoming some limitations of the linear assumptions in SHIPS. SHIPS developmental data is publicly available at the RAMMB website (http://rammb.cira.colostate.edu/research/tropical_cyclones/ships/index.asp).

3.4 General Characteristics of TC over NIO basin

Fig 3.5 shows the frequency of cyclones which shows two periods in a year when cyclone occurs, Mar-Apr and the other in Aug-Oct. The Best Track data was used to analyse this trend. Fig. 3.6(a) shows the Probability Density of cyclones passing through certain latitudes in the NIO region, it peaks around 14° N; (b) shows the probability density of cyclones passing through certain longitude in the NIO region, there are two peaks, one in Arabian Sea

and other in Bay of Bengal; (c) shows the Probability Density of cyclones with respect to Maximum Wind Speed. It shows that the cyclones are generally of low intensities with the peak between 20 to 40 knots. There are two cyclonic seasons in NIO, one around Mar-Apr and the other around Aug-Nov.
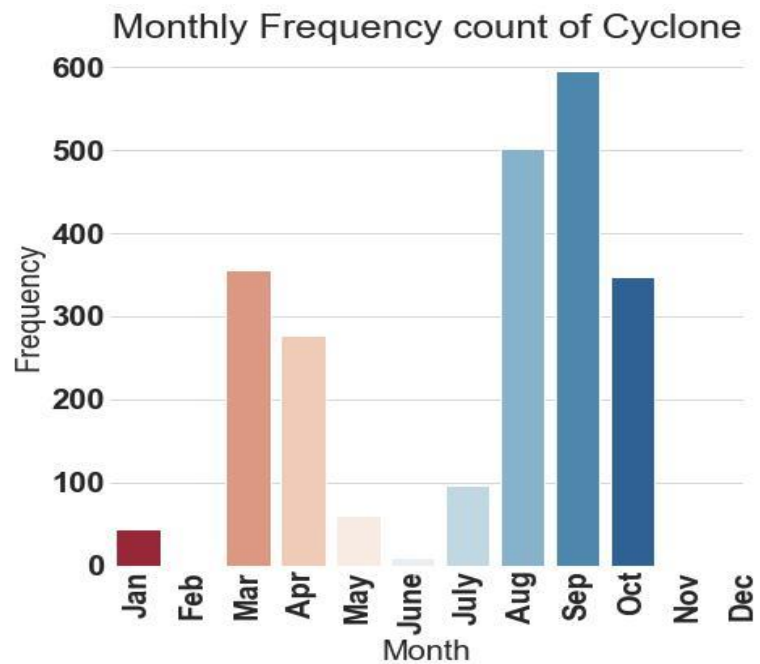
Figure 3.5: Monthly frequency count showing two different cyclonic seasons in NIO.
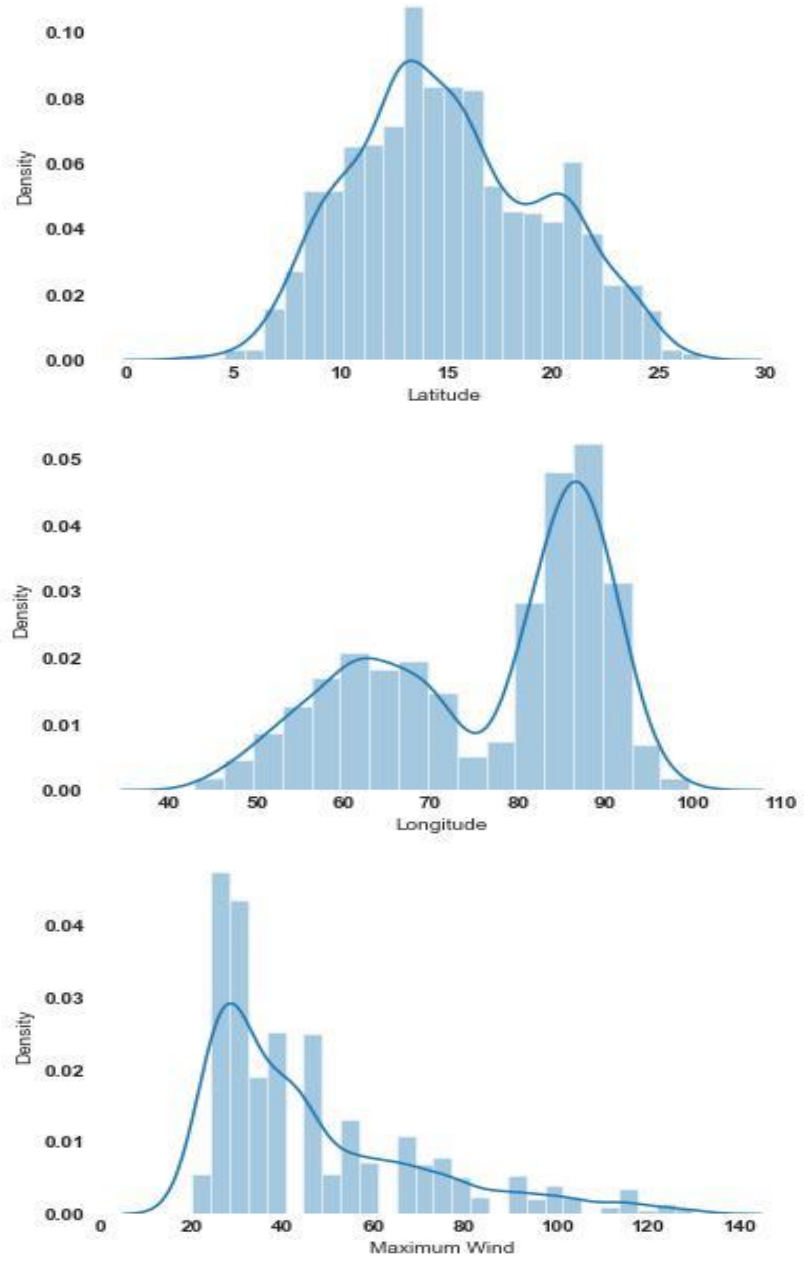
Figure 3.6: (a) Probability Density vs Latitude, (b)Probability Density vs Longitude, (c)Probability Density vs Maximum Wind Speed

Table 3.1: LIST OF CYCLONE NAME WITH THEIR YEAR OF OCCURRENCE

Year	Cyclone Number
2000	04B
2001	01A
2003	01B, 03B
2004	01A, 05A (Agni), 01B
2005	06B (Fanoos), 05B (Baaz)
2006	02B (Mala)
2007	02A (Gonu), 06B (Sidr)
2008	01B (Nargis), 04B (Rashmi), 05B (Khai-Muk), 06B (Nisha)
2009	01B (Bijli), 02B (Aila), 05B (Ward)
2010	01B (Laila), 02A (Bandu), 03A (Phet), 04B (Giri), 05B (Jal)
2011	03A (Keila), 06B (Thane)
2012	02B (Nilam), 01A (Murjan)
2013	01B (Mahasen), 02B (Phailin), 04B (Helen), 05B (Lehar), 06B (Madi)
2014	01B, 02A, 03B (Hudhud), 04A (Nilofar), 05B
2015	01A (Ashobaa), 02B (Komen), 03A, 04A (Chapala), 05A (Megh)
2016	01B (Roanu), 02A, 03B (Kyant), 04B (Nada), 05B (Vardah)
2017	01B (Maarutha), 02B (Mora), 03B (Ockhi), 04B
2018	01A (Sagar), 02A (Mekunu), 04B (Daye), 05A (Luban), 06B (Titli), 07B (Gaja), 08B (Phethai)
2019	01B (Fani), 02A (Vayu), 03A (Hikaa), 04A (Kyaar), 05A (Maha), 06A (Pawan), 07A
2020	01B (Amphan), 02A (Nisarga), 03A (Gati), 04B (Nivar), 05B (Burevi)
2021	01A (Tauktae), 02B (Yaas), 03B (Gulab), 03A (Shaheen)
2022	01B, 02B (Asani)

Chapter 4

Overview of selected ML and DL Techniques

This section provides concise definitions and explanations of key terms, concepts, and acronyms used throughout the study. It serves as a reference guide, ensuring a shared understanding and effective communication for readers. The techniques discussed in this chapter are not the exhaustive set of all the techniques used, but it gives a fair idea about some of them.

4.1 Deep Learning Techniques

4.1.1 Convolutional Neural Network

A CNN is a DL algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a CNN is much lower as compared to other classification algorithms. CNN has the ability to learn filters/characteristics. The architecture of a CNN is analogous to that of the connectivity pattern of neurons in the Human Brain and was inspired by the organisation of the Visual Cortex. Different convolution operations are used to learn the weights of convolution filters, which take the input and generate the feature maps and apply ReLU for non-linearity. The non-linearity in the network is captured by the ReLU activation function. The pooling layer reduces spatial size. To minimise the amount of parameters and generate more abstract features, max pooling techniques are applied.

CNN work on convolution and pooling operations. The convolution operation is a mathematical function that is useful in identifying the feature map. A feature map is a set of key portions of an image. Pooling operation is used for dimensionality reduction of the image. The filter is applied to the input image and identifies the image features such as shape, eye. When the filter is applied to the first 3×3 part of an input image, the features are equivalent or higher to the threshold value. The filter slides through the

image horizontally and vertically and keeps on updating the features of the image. Each convolution layer finalises the feature vector based on activation functions like ReLU, Sigmoid, Tanh and Leaky ReLU. While training the network, the optimizer adjusts each neuron's weights to reduce each epoch's loss. The last layer comprises a fully connected layer that consists of a final feature vector. Adam optimization technique is for computing the adaptive learning rate for each parameter. It is also observed that this technique works well in tuning the learning rate.

The layers are ordered as follows: input layer, convolution layer with activation function, pooling layer, and lastly fully connected layers. To obtain a tiny amount of translational invariance at each level, max pooling or average pooling is applied to the output of convolution processes. Regularisation techniques are employed to reduce data overfitting, while dropouts are used to randomly change the network design. In Keras, Adam is most typically employed as an optimizer. Dropout lengthens the training convergence time, whereas batch normalisation speeds up convergence and improves model training efficiency.

4.1.2 Multilayer Perceptron

MLP is a fundamental type of feedforward Artificial Neural Network (ANN) widely used in ML. With its multiple layers of interconnected nodes or neurons, MLP is capable of learning complex patterns and making predictions. Each neuron in an MLP performs a weighted sum of its inputs, followed by the application of an activation function, which introduces non-linearity to the model. MLPs employ a forward propagation algorithm, where data flows through the network from the input layer, through the hidden layers, and finally to the output layer. The hidden layers allow the MLP to learn and extract higher-level representations of the input data, enabling it to solve problems such as classification, regression, and pattern recognition. MLPs have been successfully applied in various domains, including image and speech recognition, natural language processing, and financial forecasting. They serve as a foundational model in neural network architectures and continue to be a versatile tool in the field of ML.

4.2 Object Detection Techniques

4.2.1 YOLOv3

YOLOv3 (You Only Look Once, Version 3) is a state-of-the-art, real-time object detection algorithm that identifies specific objects in videos, live feeds, or images. YOLO uses features learned by Darknet to detect an object. Darknet has 53 convolutional layers with consecutive 3x3 and 1x1 convolution layers followed by a skip connection to help the activations propagate through deeper layers without gradient diminishing. The methodology for detecting and locating an object is to split the image into various segments and supply each segment to model. To eliminate time and cost for computing, the You Only Look Once (YOLO) algorithm is preferred. It was proposed to deal with the problems in object recognition models at that time. YOLO takes an image as input and resizes it to 448x448 by keeping the aspect ratio the same and performing padding. The image is then passed on the CNN network. This model has 24 convolution layers, 4 max-pooling layers followed by 2 fully connected layers. For the reduction of the number of layers, we use 1x1 convolution that is followed by 3x3 convolution. This architecture uses Leaky ReLU as its activation function in the whole architecture except the layer where it uses linear activation function. Fig. 4.1 shows the architecture diagram of YOLOv3. Non-Maximum Suppression (NMS) is a technique used in numerous computer vision tasks. It is a class of algorithms to select one entity (e.g., bounding boxes) out of many overlapping entities. Most object detection algorithms use NMS to whittle down many detected bounding boxes to only a few. Thousands of windows (anchors) of various sizes and shapes are generated. These windows supposedly contain only one object, and a classifier is used to obtain a probability/score for each class. Once the detector outputs a large number of bounding boxes, it is necessary to filter out the best ones. Darknet is mainly for object detection, and have different architecture, features than other DL frameworks. It is faster than many other Neural network architectures. YOLO object detection utilises Darknet in background, and the Darknet repository is cloned. Cloning basically means you want to get a local

Layer	Filters size	Repeat	Output size
Image			416×416
Conv	32 3 × 3/1	1	416×416
Conv	64 3 × 3/2	1	208×208
Conv	32 1 × 1/1	⌈Conv⌉	208×208
Conv	64 3 × 3/1	⌊Conv⌋ × 1	208×208
Residual		⌊Residual⌋	208×208
Conv	128 3 × 3/2	1	104×104
Conv	64 1 × 1/1	⌈Conv⌉	104×104
Conv	128 3 × 3/1	⌊Conv⌋ × 2	104×104
Residual		⌊Residual⌋	104×104
Conv	256 3 × 3/2	1	52×52
Conv	128 1 × 1/1	⌈Conv⌉	52×52
Conv	256 3 × 3/1	⌊Conv⌋ × 8	52×52
Residual		⌊Residual⌋	52×52
Conv	512 3 × 3/2	1	26×26
Conv	256 1 × 1/1	⌈Conv⌉	26×26
Conv	512 3 × 3/1	⌊Conv⌋ × 8	26×26
Residual		⌊Residual⌋	26×26
Conv	1024 3 × 3/2	1	13×13
Conv	512 1 × 1/1	⌈Conv⌉	13×13
Conv	1024 3 × 3/1	⌊Conv⌋ × 4	13×13
Residual		⌊Residual⌋	13×13

Conv

Con2d Layer → BN Layer → LeakyReLU Layer
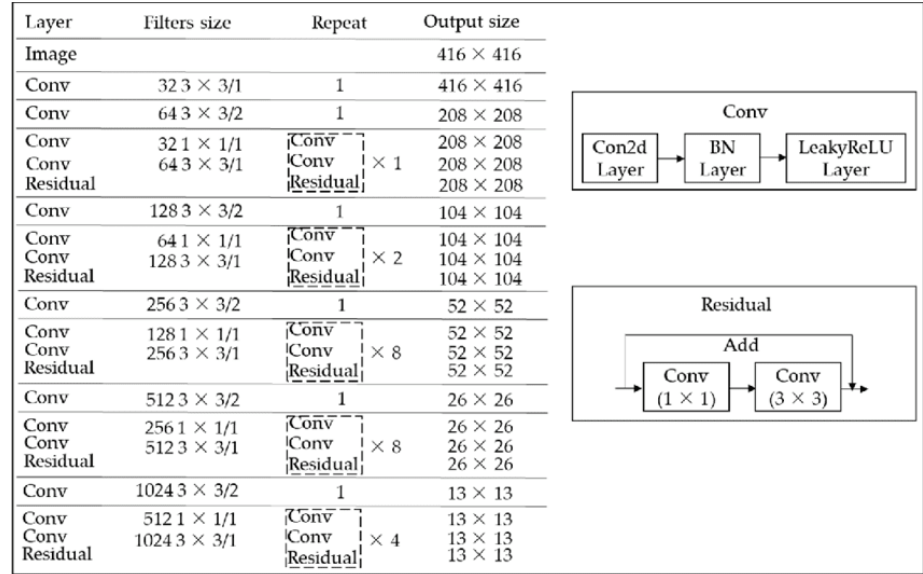
Residual

Add

Conv (1 × 1) → Conv (3 × 3)

Figure 4.1: Architecture diagram of YOLOv3

copy of the code present in the repository. After cloning you can then do whatever changes you like in the code and then you can pull the changes back to the repository. LabelImg is a graphical image XML files annotation tool. Annotations are saved as in YOLOv3 format. The file format used in this research is YOLO. The generated .txt file has the class name and the four coordinates.

4.2.2 YOLOv5

YOLOv5 is an advanced object detection algorithm that has gained significant popularity in the field of computer vision. Developed as an evolution of the YOLO (You Only Look Once) series, YOLOv5 represents a more streamlined and efficient approach to real-time object detection. It combines the advantages of high accuracy and remarkable speed, making it a preferred choice for various applications such as autonomous driving, surveillance systems, and robotics. YOLOv5 introduces several architectural improvements, including a lightweight backbone network, a feature fusion module, and a more robust training pipeline. These enhancements enable YOLOv5 to achieve impressive detection performance on a wide range of object classes while maintaining real-time inference speeds. With its ease of use, superior accuracy, and remarkable efficiency, YOLOv5 continues to push the boundaries of object detection capabilities, empowering researchers

and developers to tackle complex computer vision tasks with greater precision and speed.

4.2.3 Mask R-CNN

Mask R-CNN (Mask Region-based Convolutional Neural Network) is a state-of-the-art deep learning architecture designed for instance segmentation tasks. Mask R-CNN builds upon the Faster R-CNN framework by extending it to also generate pixel-level masks for each object instance detected. By integrating a branch for semantic segmentation with the existing object detection capabilities, Mask R-CNN achieves precise localization and accurate pixel-level segmentation of objects in an image. This architecture has been widely adopted for a variety of computer vision applications, including object recognition, image segmentation, and instance-level image understanding. Its ability to simultaneously detect and segment objects makes Mask R-CNN a powerful tool in computer vision research and has significantly advanced the field of instance segmentation.

4.3 Machine Learning Techniques

4.3.1 Support Vector Machine

Support Vector Machine (SVM) is a powerful supervised ML algorithm used for classification and regression tasks. SVMs are particularly effective in dealing with complex, high-dimensional datasets. The key principle behind SVM is to find an optimal hyperplane that best separates the data points belonging to different classes. This hyperplane is chosen to maximize the margin between the classes, allowing for better generalization and improved performance on unseen data. SVMs can handle both linearly separable and non-linearly separable datasets by using different types of kernels, such as linear, polynomial, or radial basis function (RBF) kernels. Additionally, SVMs are robust against overfitting and can handle noisy data effectively. Due to their versatility and solid theoretical foundation, SVMs have found applications in various domains, including image recognition, text categorization, and bioinformatics.

4.3.2 Long Short-Term Memory

LSTM is a type of Recurrent Neural Network (RNN) architecture that excels at capturing and modelling sequential data. LSTM addresses the vanishing gradient problem of traditional RNNs by incorporating memory cells and gating mechanisms. The key feature of LSTM is its ability to selectively remember and forget information over long sequences, making it effective for tasks such as natural language processing, speech recognition, and time series analysis. With its recurrent connections and internal memory, LSTM can capture both short-term dependencies and long-term dependencies in data, enabling it to effectively model temporal dynamics. The LSTM architecture has become a fundamental building block in deep learning, empowering researchers and practitioners to tackle complex sequential tasks with improved accuracy and robustness.

4.4 Class Balancing Techniques

4.4.1 SMOTE

Imbalance data is a case where the classification dataset class has a skewed proportion. An imbalance class creates a bias where the ML model tends to predict the majority class.

SMOTE, which stands for Synthetic Minority Over-sampling Technique, is a widely used data augmentation method in the field of ML and imbalanced classification tasks. SMOTE addresses the issue of class imbalance by generating synthetic minority class samples to balance the dataset. It works by randomly selecting a minority class sample and finding its nearest neighbors. Synthetic samples are then created by interpolating between the selected sample and its neighbors. This technique helps to overcome the bias towards the majority class and enhances the performance of classifiers by providing more representative training data. SMOTE has been proven to be effective in various domains and continues to be a valuable tool in handling imbalanced datasets.

SMOTE is an oversampling technique where the synthetic samples are generated for the minority class. This algorithm helps to overcome the

overfitting problem posed by random oversampling. It focuses on the feature space to generate new instances with the help of interpolation between the positive instances that lie together. SMOTE is used to synthesize data where the features are continuous and a classification problem. SMOTE works by utilizing a k-nearest neighbour algorithm to create synthetic data. SMOTE first starts by choosing random data from the minority class, then k-nearest neighbours from the data are set. Synthetic data would then be made between the random data and the randomly selected k-nearest neighbour.

4.4.2 Data Augmentation

Data augmentation is a common technique used in DL to artificially increase the size of a training dataset by applying various transformations to the existing data samples. By introducing modified versions of the original data, data augmentation helps to enhance the generalization capability of deep learning models. The transformations can include geometric changes such as rotation, scaling, and translation, as well as alterations in colour, contrast, or noise levels. This technique allows the model to learn from a wider range of variations and variations that are likely to occur in real-world scenarios. By augmenting the data, the model becomes more robust, reducing the risk of overfitting and improving its ability to accurately classify or predict on unseen data. Data augmentation is particularly valuable when the available dataset is limited, as it effectively expands the dataset and enables more diverse and representative training samples.

4.5 Explainable ML Techniques

4.5.1 SHAP

SHAP (SHapley Additive exPlanations) is a Python library that provides a unified framework for interpreting the output of ML models. It is based on the concept of Shapley values from cooperative game theory, which allows us to assign a value to each feature in a model's prediction. SHAPely Python, as the name suggests, refers to the Python implementation of SHAP. This powerful library enables users to understand the impact of individual features

on model predictions, providing insights into the model's decision-making process. By using SHAP, analysts and data scientists can gain a deeper understanding of complex ML models and communicate their findings effectively. With its comprehensive set of tools and intuitive API, SHAPely Python has become a popular choice for explainable AI and interpretability tasks in various domains. SHAP values consider the interaction between features and provide a unified measure of feature importance that is consistent across different feature sets. The mean absolute SHAP value for a feature represents the average magnitude of the impact of that feature on the model's output.

4.5.2 LIME

LIME (Local Interpretable Model-agnostic Explanations) is a technique used to interpret the predictions of ML models and provide explanations for their decisions in a local context. It aims to address the black-box nature of many complex ML models by generating locally faithful and human-interpretable explanations. LIME works by perturbing the input data and observing how the model's predictions change. It then constructs a simpler, interpretable "local" model around the perturbed instances to approximate the behaviour of the original model. This local model can be easily understood by humans and used to explain the reasons behind the model's predictions. LIME is a powerful tool for increasing transparency and trust in ML systems, enabling users to understand and validate the decisions made by these models.

4.6 Hyperparameter Tuning

4.6.1 Grid Search CV

Grid Search CV (Cross-Validation) is an algorithm used for hyperparameter tuning in ML models. Hyperparameters are model parameters that are set before the learning process begins, and they control the behaviour and performance of the model. Grid Search CV automates the process of systematically searching through a predefined set of hyperparameters to find the optimal combination that yields the best model performance.

The algorithm works by exhaustively evaluating all possible combinations of hyperparameter values within a given grid or search space. It performs a cross-validation evaluation for each combination of hyperparameters to estimate the model's performance on unseen data. Cross-validation helps to mitigate the risk of overfitting and provides a more robust assessment of the model's generalization ability.

Grid Search CV is typically implemented using nested loops. The outer loop iterates over each hyperparameter, while the inner loop performs cross-validation to evaluate the model. The algorithm compares the performance metrics (such as accuracy, precision, or mean squared error) across different hyperparameter combinations and selects the combination that achieves the best performance.

4.7 Evaluation Metrics

To evaluate models, the metrics used are Mean Absolute Error (MAE), Mean Squared Error (MSE), and RMSE. The equations for the same are

$$MAE = \frac{1}{N}\sum_{i=1}^{N}(\overline{x_p} - x_i) \qquad\qquad 4.1$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\overline{x_p} - x_i)^2 \qquad\qquad 4.2$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\overline{x_p} - x_i)^2} \qquad\qquad 4.3$$

$\overline{x_p}$ is the predicted intensity value, x_i is the actual intensity value, N denotes the number of samples.

Accuracy measures the proportion of correctly classified samples out of the total number of samples. It provides an overall assessment of the model's performance but may not be suitable for imbalanced datasets.

Precision is the measure of a positive predicted value and is defined as

$$Precision = \frac{TP}{TP+FP} \qquad\qquad 4.4$$

Recall or True Positive Rate is defined as

$$Recall = \frac{TP}{TP+FN} \qquad\qquad 4.5$$

True Positive (TP) is the outcome where the model correctly predicts the positive class. False Positive (FP) is the outcome where the model incorrectly predicts the positive class.

F1 score is the harmonic mean of precision and recall and is defined as

$$F1\ score = \frac{2*(Precision*Recall)}{Precision+Recall} \qquad 4.6$$

Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) measure the trade-off between true positive rate and false positive rate at different classification thresholds. They provide insights into the model's performance across different decision boundaries.

Probability of Detection (POD), also known as the true positive rate or hit rate, measures the effectiveness of a detection system in correctly identifying the presence of a target or signal when it is actually present. It quantifies the proportion of positive instances that are correctly detected. Mathematically, the probability of detection (PD) can be expressed as:

$$POD = \frac{TP}{TP+FN} \qquad 4.7$$

where TP represents the number of true positives (correct detections) and FN represents the number of false negatives (missed detections). PD ranges from 0 to 1, where a value of 1 indicates perfect detection performance.

False Alarm Rate (FAR), also known as the False Positive Rate, measures the rate at which the detection system incorrectly identifies the presence of a target or signal when it is not actually present. It quantifies the proportion of negative instances that are incorrectly classified as positive. Mathematically, the false alarm rate (FAR) can be expressed as:

$$FAR = \frac{FP}{FP+TN} \qquad 4.8$$

where FP represents the number of false positives (incorrect alarms) and TN represents the number of true negatives (correct rejections). FAR ranges from 0 to 1, where a value of 0 indicates perfect performance (no false alarms).

Mean Average Precision (mAP) is commonly used in object detection tasks, including the YOLO algorithm. It quantifies the accuracy of the model by considering both precision and recall. The mAP is calculated by averaging

the average precision (AP) scores across multiple object categories. The formula for calculating AP involves computing the precision-recall curve, where precision is plotted against different levels of recall. The area under this curve represents the AP score. By averaging the AP scores across all categories, the mAP provides a comprehensive measure of the overall detection performance of the YOLOv5 model.

Chapter 5

DL Framework for Cyclone Intensity Estimation

This chapter is based on the manuscript titled "An End-to-End Deep Learning Framework for Cyclone Intensity Estimation in North Indian Ocean Region using Satellite Imagery" submitted to Journal of the Indian Society of Remote Sensing. The chapter focuses on the dataset used for multiclass classification, regression analysis and detection which includes infrared images taken from the Meteosat-8 satellite spanning from 2000 to 2022. The images are processed through cropping, grayscale conversion, and noise reduction techniques to improve cyclone classification. The chapter describes the methodology involving CNN, data augmentation, and model evaluation using binary classification, multiclass classification, and intensity prediction. It highlights the effectiveness of the proposed model in simultaneously detecting cyclonic regions, classifying cyclones into multiple classes, and predicting their intensity. The performance of the model is compared to existing studies, showcasing its superior classification capabilities. The chapter concludes with a case study of a specific cyclone, demonstrating the accuracy of the model's intensity prediction.

5.1 Dataset

For multiclass classification and regression analysis, the images were taken from 2000 to 2022. Fig. 5.1 (a) shows the image with no cyclone and (b) shows image with a cyclone. Both are infrared images taken from Meteosat-8 satellite in January, 2019 and November, 2017, respectively. Image spans from 10 E to 110 E and 30 S to 25 N. Fig. 5.1(b) shows a cyclonic region in the Arabian Sea at the south tip of India. For binary cyclone classification, image data of January 2019, 2020 and 2021 are used for class category 'no', keeping in mind that no cyclone occurs during that time. For class category 'yes', the higher intensity images (intensity $\geq$ 40 knots) from multiclass classification dataset are used. Number of images used in class 'without

cyclones' is 708 and 'with cyclones' is 942. The non-cyclonic images do not have cloud patches.

5.2 Data Processing

The images of cyclones in NIO are first cropped to an average dimension of 310 x 310 manually (average aspect ratio is maintained at 1) as shown in Fig. 5.2. The cyclone eye is not in the centre, which makes the model more robust, as the region contains non-cyclonic parts as well. Cropping speeds up the data augmentation and model training process by preserving all of the rain bands and cloud cover of cyclones.

The reduction of image size is done only to reduce the processing time. The images are then converted into grayscale. The images are pre-processed such
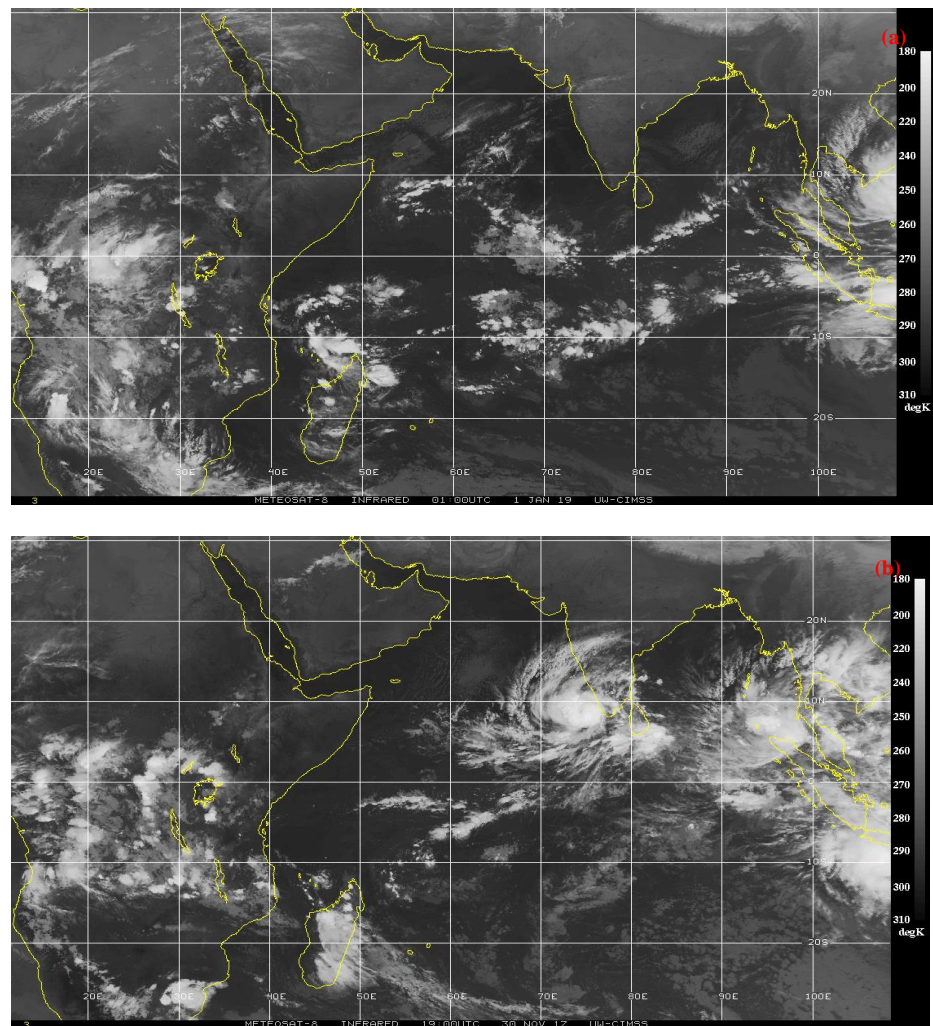

Figure 5.1: (a)No Cyclone and (b) Cyclone

that all the pixels below 128 threshold value were converted to 0 to reduce noise from the images and improve cyclone classification.

Table 5.1 shows the data split of images for different models. Actual 1884 images are divided into 5 different classes with their corresponding wind speeds as shown in Table 5.2 for multiclass classification. Very high cyclone intensity ($\geq$120) images are not used in regression problem as the number of samples are very low (15 samples only).

5.3 Methodology

All the CNN models are fed with 310x310 image size. Data Augmentation by rotation range of 10 degrees has been used to increase the number of samples for training as shown in Fig. 5.3. Data Augmentation was done only on the training dataset. Fig. 5.4 (a) shows the raw sample image before pre-processing, (b) Depression and (c) Very Severe Cyclone, and (d) shows an image with no cyclone. Since we have very limited dataset, the images have been split into training, validation and test dataset in the ratio 0.7:0.15:0.15 maintaining the ratios of different cyclones. In our proposed model Adam optimizer has yielded best results. Aspect ratio is kept approximately 1 so that it doesn't stretch or resize the cyclonic images.

	FileName	Size	Width	Height	Aspect Ratio
0	20140611.18-45.jpg	(332, 333)	332	333	1.00
1	20130514.06-40.jpg	(329, 334)	329	334	0.99
2	20130515.09-45.jpg	(334, 332)	334	332	1.01
3	20141026.18-45.jpg	(332, 332)	332	332	1.00
4	20160519.18-40.jpg	(336, 333)	336	333	1.01

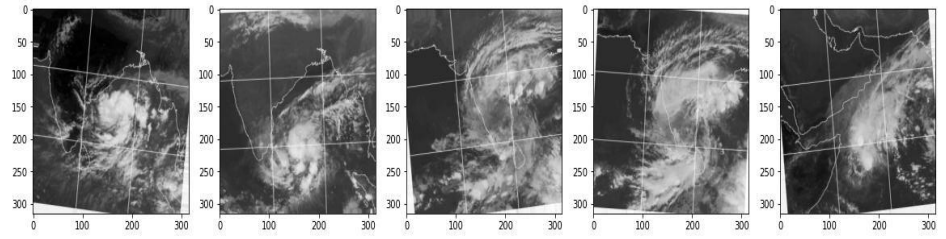Figure 5.2: Sample image files showing aspect ratio

Figure 5.3: Augmented Data Sample Multiclass Classification

However, it should be noted that as the cyclone images in few hours doesn't change much, there is always a possibility of leakage of the data between training and testing samples. To avoid this, the evaluation is also carried separately with a testing dataset of 2020-22 while keeping 2000-19 as training dataset. This, however, doesn't guarantee the proper representation of all the cyclone classes.

Here, four different models are evaluated as shown in Fig. 5.5. As already mentioned, it involves four different steps: first, the input image is given to a binary classifier, second, the cyclone is detected and cropped using pre-trained weights on YOLOv3 architecture, third the image is given to a multiclass-classifier and fourth, an intensity predictor estimates the intensity of the image.
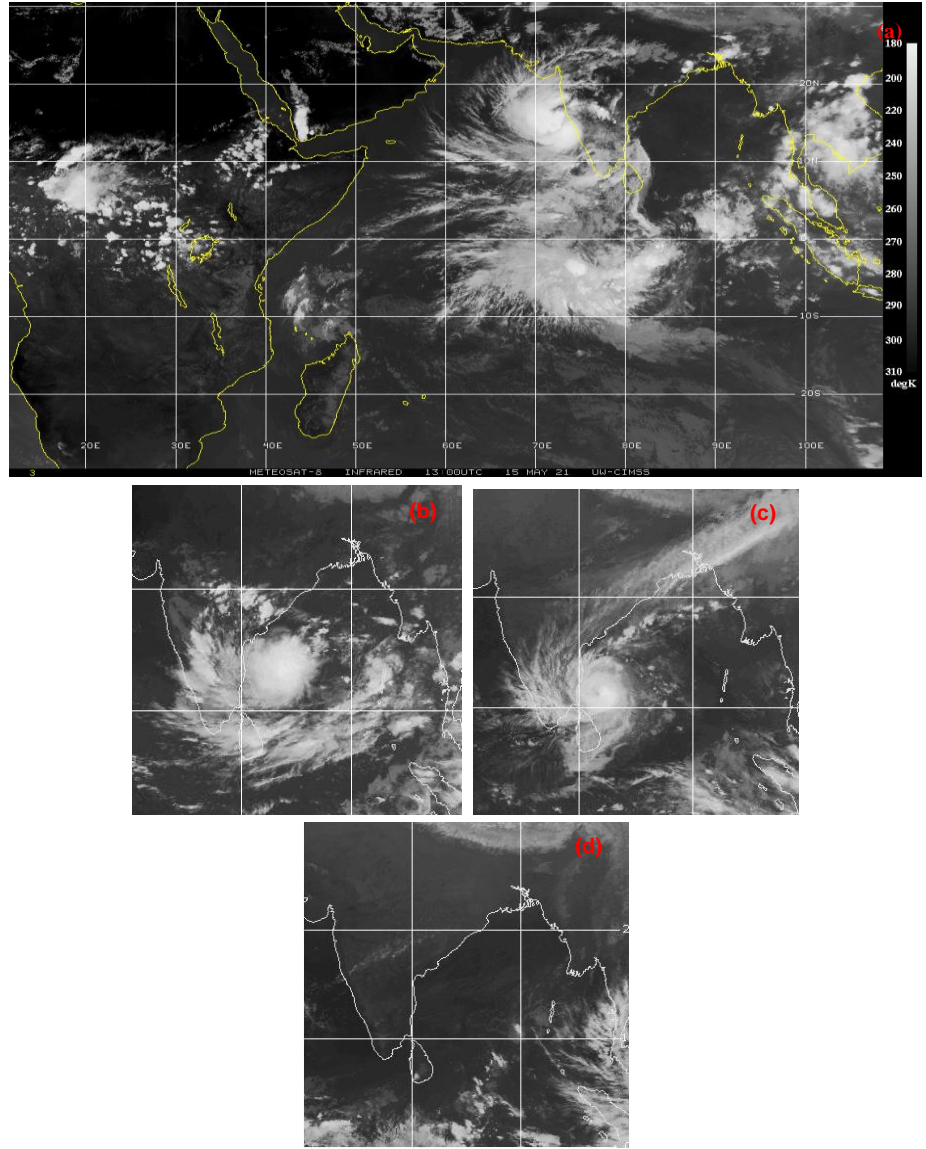
Figure 5.4: (a) Raw image from CIMSS dataset (b) Depression and (c) Very Severe Cyclone, and (d) shows an image with no cyclone

CNN is a network architecture for deep learning which learns directly from data, eliminating the need for manual feature extraction. Early stopping, Dropouts, Kernel Regularizers and Learning rate scheduler was used to prevent overfitting the model. Table 5.3 shows the tuned hyperparameters used for different classifiers for the proposed model. Table 5.5 shows the proposed CNN model used for binary, multiclass and regression analysis.
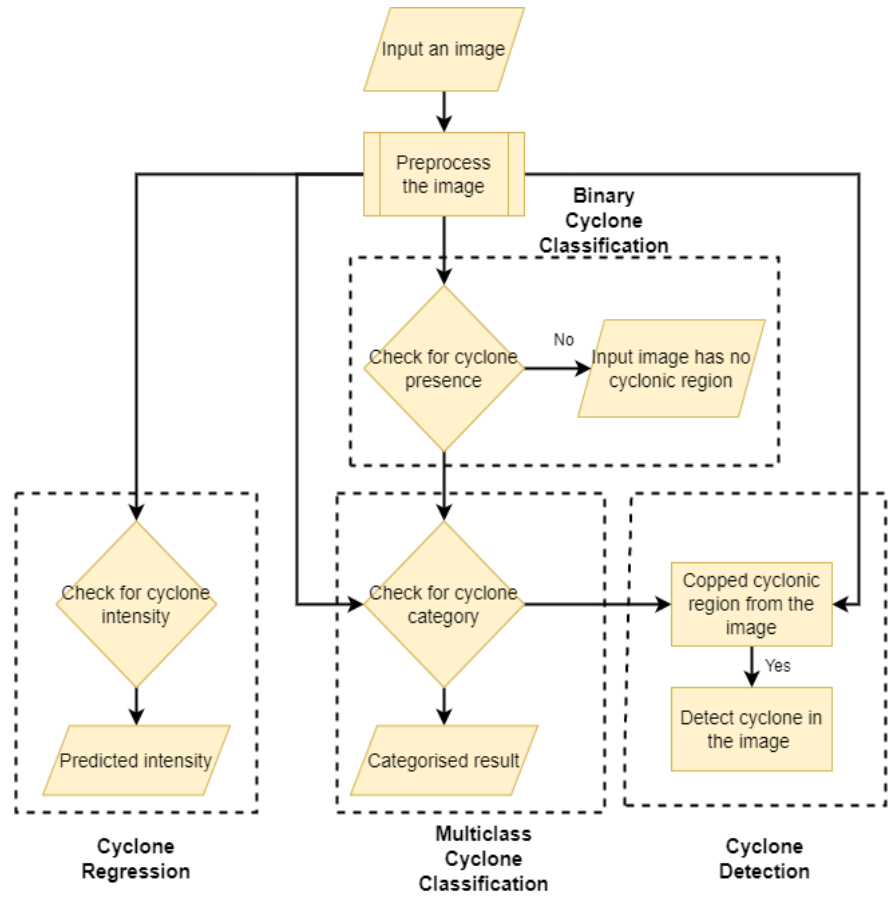
Figure 5.5: Flowchart of the implemented pipeline

Table 5.1: DATA SPLIT

	Binary Classification	Multiclass Classification	Regression
Training	1154	1316	1308
Validation	247	280	348
Testing	249	288	281
Total	1650	1884	1869

Table 5.2: CATEGORY OF CYCLONES (MULTICLASS
CLASSIFICATION)

Category	Wind speed	Actual Images
Depression (D)	20-25 knots	352
Deep Depression (DD)	30-35 knots	534
Cyclonic Storm (CS)	40-50 knots	396
Severe Cyclonic Storm (SCS)	55-65 knots	257
Very Severe Cyclonic Storm (VSCS)	70+ knots	305

Table 5.3 HYPERPARAMETER TUNING FOR DIFFERENT ML
MODELS

	Binary Classifier	Multiclass classifier	Regression
Epochs	30	50	40
Optimiser	Adam	Adam	Adam
Loss function	Binary Cross-entropy	Categorical Cross-entropy	Mean Squared Error (MSE)
Initial Learning rate	0.001	0.0001	0.001
Training /Validation/Testing batch size	2/1/1	2/1/1	2/1/1

Training for cyclone detection by YOLOv3 is done by annotating images
with the cyclonic regions with a bounding box and the text files containing
the class and coordinates of the bounding box in YOLO format. A custom
weights file is generated using transfer learning with Darknet weights to train
on our custom dataset. Our dataset had only one target, i.e., cyclones. Non
maximum suppression is set to 0.3 to filter the best bounding box out of
many overlapping boxes. Minimum probability to eliminate weak predictions
is 0.3. The predicted bounding box size varies depending on the size of the
cyclone. Table 5.7 shows the comparative study of the existing cyclone
prediction models in the NIO region with the proposed model.

Table 5.4: PROPOSED CNN MODEL

Layer	Filters/Dropout value/Units/Output Shape	Kernel size/ Pool size	Activation
Convolution	32 (308, 308, 32)	3,3	ReLU
Max Pooling	(154, 154, 32)	2,2	
Dropout	0.4		
Convolution	64 (152, 152, 64)	3,3	ReLU
Max Pooling	(76, 76, 64)	2,2	
Dropout	0.4		
Convolution	128 (74, 74, 128)	3,3	ReLU
Max Pooling	(37, 37, 128)	2,2	
Dropout	0.4		
Convolution	128 (35, 35, 128)	3,3	ReLU
Max Pooling	(17, 17, 128)	2,2	
Dropout	0.4		
Flatten			
Dropout	0.4		
Dense	512		ReLU
Dropout	0.4		
Dense	1(BC)[a] (R)[a], 5 (MC)[a]		Sigmoid (BC)[a], SoftMax (MC)[a], Linear(R)[a]

[a]Terms - BC: Binary Classification; MC: Multiclass Classification; R: Regression

5.4 Results

Compared to existing studies, our proposed pipeline simultaneously detects the cyclonic region, classifies and predicts the intensity. The models were evaluated separately since most of the mis-classifications are found to be within single class difference and hence, the intensity value may be a better representation of the capabilities of the DL model. The proposed model is able to classify cyclones in five different classes as compared to most of the existing studies which classify only in one or two classes. The false positive rate is 0.007% for binary classification. Table 5.6 shows the multiclass testing dataset metrics for individual classes. It can be seen that the model performs best for Very Severe Cyclonic Storm and worst for Cyclonic storm. The well-defined structure in VSCS may be the reason for the high

performance. Fig 5.6 shows the feature maps of the first layer of CNN. It shows the learning pattern. The purpose of feature maps is to identify meaningful features in the input data, such as edges, corners, or texture patterns. These features are then used to classify or identify the input image.

A case study of a specific cyclone 'ASANI' from the year 2022 (Timestamp: 2022-05-07, 12:00 hours) has been done. The pipeline predicted the intensity of the cyclone is 33 knots (Actual: 30 knots). The cyclonic region is also detected correctly. Fig 5.7 shows the multiple detections from YOLOv5 architecture. Fig 5.8 shows the detected cyclonic region from YOLOv3. Predefined CNN architectures provide general architectural recommendations for deep learning practitioners to handle wide variety of problems and develop newer and data specific architectures when the training dataset is limited. Hence, these approaches were also tested with the same dataset.

It is to be noted that Advanced Dvorak Technique (ADT) has shown much better RMSE than the present model. It is, however, also to be noted that the performance of all the previous studies using Deep Learning techniques as well the pre-defined architectures evaluated in this work performed poor than the ADT. This is one of the major drawbacks of the Deep learning models as the performance directly depends on the number of training data available. As NIO region has very small dataset at present, the performance may improve once more data is made available. Another reason for high RMSE is that the best track intensity estimation from IMD is only available with 5 knots resolution.

Feature maps provide insight into the internal representations for input in the CNN layers of the model. Another important observation from Table 5.8 is that the evaluation of all the DL models based on data-wise split shows higher accuracy than the year-wise split. As this data-wise split results were only available from previous published works in this domain, we keep these statistics as well for comparison. However, one should be careful in interpreting this result since data similar to training images may be present in testing samples as the cyclone images doesn't always change fast. Table 5.6 shows the final results of the proposed model compared to other predefined models.
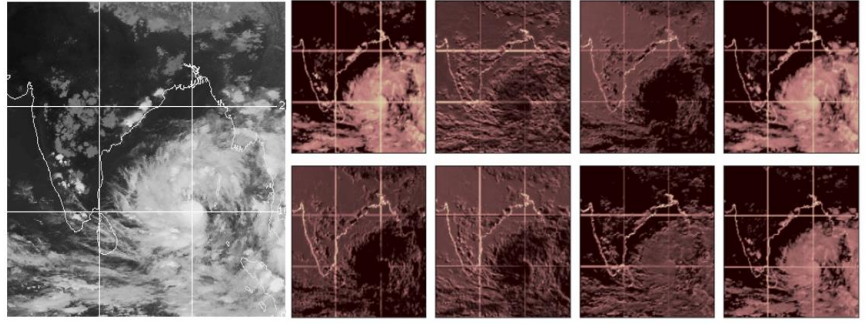
Figure 5.6: (a) shows original image and (b) shows the feature maps of the first layer of CNN.
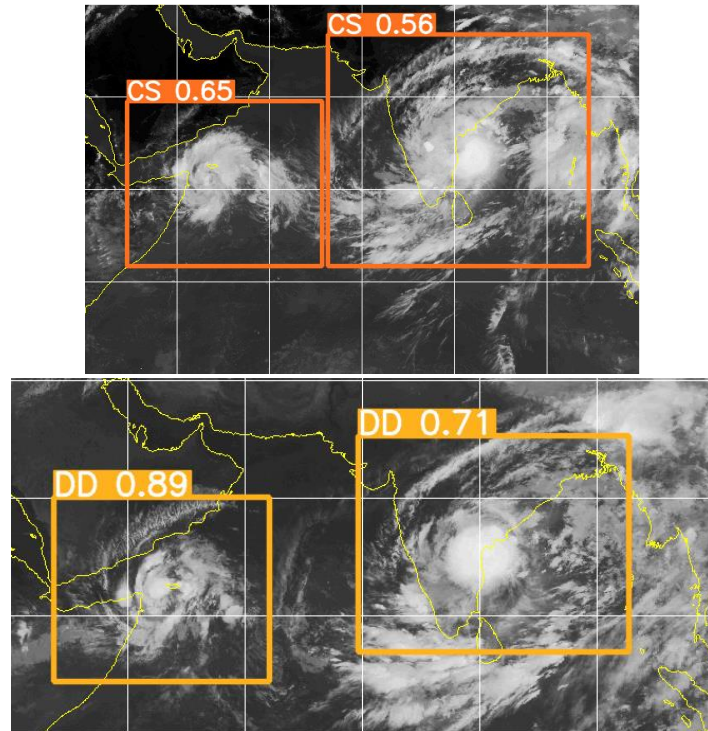

Figure 5.7: Multiple cyclone detections using YOLOv5 model. Two images detected cyclones, with the top image highlighting "CS" and the bottom image highlighting "DD".

Table 5.5: EVALUATION METRICS FOR MULTICLASS CLASSIFICATION MODEL, EVALUATED ON DIFFERENT CATEGORIES OF CYCLONIC ACTIVITIES

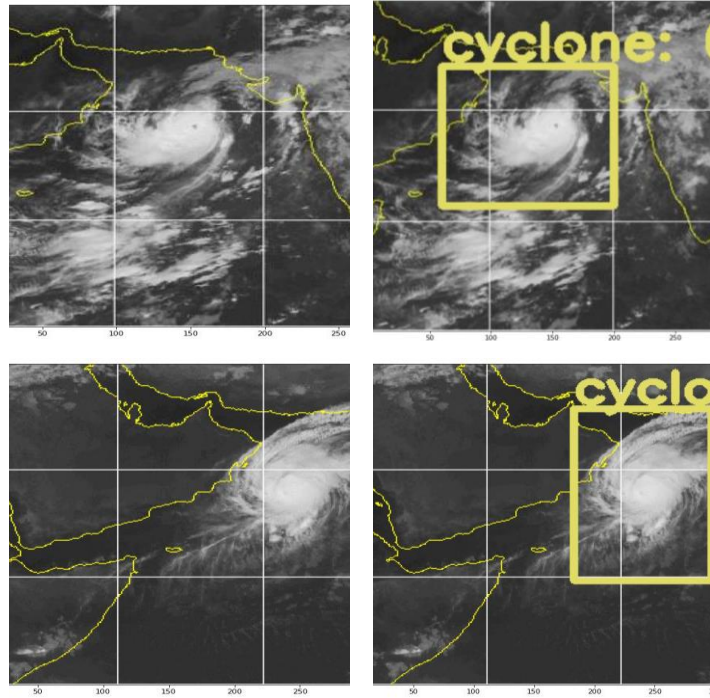| Category	Precision	Recall	F1-score
D (20-25)	0.70	0.64	0.67
DD (30-35)	0.62	0.72	0.67
CS (40-50)	0.54	0.55	0.54
SevereCS (55-65)	0.59	0.59	0.59
VSCS (70+)	0.76	0.60	0.67

Figure 5.8: Original image (left), Detected cyclonic regions (right)

Table 5.6: COMPARISON OF DIFFERENT MODELS ACROSS DIFFERENT TASKS. THE REPORTED VALUES ARE THE MEAN ± STANDARD DEVIATION OF FIVE INDEPENDENT RUNS.

Model	Multiclass Classification Accuracy (%)	Binary Classification Accuracy (%)	Regression RMSE (knots)
Proposed Model	63.83 ± 1.3 [31.2 ± 0.78]	98.4 ± 0.003	16.2 ± 0.9 [26.69 ± 0.88]
VGG-16	39.83 ± 0.86 [30.3 ± 1.04]	96.7 ± 0.007	19.79 ± 0.21 [20.67 ± 0.03]
Inception v3	51.51 ± 1.29 [36.1 ±5.65]	97.5 ± 0.01	12.4 ± 0.29 [17.28 ± 0.53]
Xception Net	40.19 ± 1.07 [30.9 ± 1.78]	97.9 ± 0.0024	22.52 ± 0.8 [21.32 ± 1.43]
ResNet	28.81 ± 0.005 [28.4 ± 0]	98.3 ± 0.0015	13.88 ± 0.28 [22.35 ± 0.47]

* Values in [] are based on Year-wise split (Training: 2000-2019, Testing: 2020-2022)

Chapter 6

ML Approach for Rapid Intensification Prediction

This chapter is based on "An Explainable Machine Learning Approach for Predicting Rapid Intensification in Tropical Cyclones" submitted to Weather and Forecasting (WAF) journal. The work describes the use of SHIPS database which contains various weather, ocean, and climatological features related to TC. The research focuses on predicting RI events, which are defined as significant increase in the strength of TC over a short period. The data from different oceanic basins were used, and a Leave One Year Out (LOYO) approach was employed for training and testing the models, except for the Indian basin, where a different approach was used due to absence of RI events in 2017. The dataset was imbalanced, with a small number of RI cases compared to non-RI cases. To address this, class balancing based on SMOTE was carried out for the training dataset. The study employed SVM for classification and random forest algorithm to rank the importance of features. The optimal set of features was determined based on skill scores such as the POD, FAR, and F1 score. The result showed that certain SHIPS features were important across different basins, and the overall performance of the proposed model was most promising for the Atlantic basin. The study also highlights the challenges of imbalanced data and importance of selecting the appropriate set of features for accurate predictions.

6.1 Methodology

In the original SHIPS dataset, each row indicates a TC record and there are around 500 weather, ocean, and climatological features. Some features in the SHIPS dataset are time-dependent and are available in the original dataset in 6-hour intervals up to 120 hours from the reporting time. There exist records from different oceanic basins in the SHIPS data. This study analysed four oceanic basins: Atlantic, Indian, Western North Pacific, and Eastern North Pacific with a primary focus on the Atlantic and Indian basins.

For the present study, the data from all the available years (1982-2017) was used except for the last year as the training data and the final year data as testing data. For example, for the Atlantic basin, the data from 1982 to 2016 is the training data and 2017 data is the testing data. This is commonly referred to as the LOYO approach and is employed primarily to surpass the sample-size limitation. For the Indian basin however, LOYO was not used as there were no RI events in the year 2017. Thus, records from 2015 to 2017 were instead utilized for testing purposes. All TC records were then labelled as non-RI or RI based on the definition of RI. This essentially turned the problem into a binary classification problem.

To tackle the issue of classification of RI events, different SVM were developed for each basin. The SHIPS dataset is highly imbalanced with respect to RI and non-RI (UNRI) cases. For example, the Atlantic basin which includes a total of 10710 records comprises 10134 UNRI and only 576 RI cases (about 5.6% of the total instances). The SMOTE was employed for the training data to handle such class imbalance [50]. The optimal SVM parameters and sampling strategy (ratio of the majority and minority class) for each basin was determined by the GridsearchCV algorithm and are shown in Table 6.1. The true positives (TP), true negatives (TN), false positives (FP) and false negatives are computed from the classification results of each basin. Based on these values, the POD and FAR are then evaluated which are commonly used metrics for binary classification problems.

In problems relating to meteorological forecasts, certain skill scores are generally calculated for the verification and assessment of the quality of forecast. These metrics can be calculated directly from the confusion matrix of the resulting classification output. Thus, the Peirce Skill Score (PSS), Gilbert Skill Score (GSS), and Heidke Skill Score (HSS) were calculated to assess the performance of the models. The PSS is calculated as the difference of POD and Probability of False Detections (POFD) which tells what fraction of the observed UNRI events were incorrectly forecast as RI. The GSS is a statistical measure which compares the skill of a forecast to a reference forecast that is based on a climatological average. The HSS is

related to GSS and measures the accuracy of the forecast relative to a random chance or a randomly selected forecast.

Table 6.1: SVM PARAMETERS FOR EACH BASIN. THE TABLE LISTS THE SVM CONFIGURATION USED FOR DIFFERENT BASINS, DETAILING THE TRAINING AND TEST BASINS, THE KERNEL TYPE, GAMMA, COST, AND SMOTE RATIO FOR BOTH RARE AND UNREPRESENTED INSTANCES.

Train basin	Test basin	Kernel	Gamma	Cost	RI and UNRI SMOTE ratio
Atlantic	AL (2017)	RBF	0.002	15	1:4
Indian	IO (2015-2017)	RBF	0.005	15	1:1
Western North Pacific	WP (2017)	RBF	0.001	5	1:4
Eastern North Pacific	EP (2017)	RBF	0.001	5	7:20

6.2 Feature Selection

The data pre-processed files used in this study were taken from [32]. They retained mostly time-dependent SHIPS features and based on linear regression experiments, found that using 24-h features instead of 0-24-h averaged features resulted in models with smaller predictive errors. The distance to the nearest landmass (DTL) missing values were interpolated utilizing the track location projection and the remaining missing values were filled by using the mean value of the reanalysis. They retained a total of 121 SHIPS features with a total sample size of 56,223 taken from reanalysis data from 1982 to 2017 of all the basins. To standardize the features, the reanalysis data mean was removed and the data was divided by the reanalysis data standard deviation for each feature.

There was a total of 121 features in the data files to start with. As an attempt to reduce the feature set size to a more reasonable one, the random forest algorithm was used which ranks the features according to their importance. Instead of truncating the number of features from 121 to any random smaller number, separate SVM models were evaluated by adding one feature in each model. The subsequent features were then added based on the random forest

ranking. So, the first model had only one feature with the highest rank and each subsequent model had the next important feature added to the preceding ones. The final model had all the 121 features. This way, it became easier to see what number of features corresponded to the optimal set of skill scores. The list of features used for the respective basin is given in Table 6.2. The feature selection is based on random forest ranking and the features are arranged in descending order of importance score.

The POD and FAR plotted against the number of features for the Atlantic and Indian basin are shown in Fig 6.1. The F1 score which is defined as the harmonic mean of precision and recall scores is a fair metric to assess the model performance. Therefore, the optimal least number of features were determined based on the F1 score and are shown for Atlantic and Indian basins in Fig 6.2. Based on the POD, FAR and F1 score plots, there were a total of 22, 10, 13, and 36 features retained for the Atlantic, Indian, Western
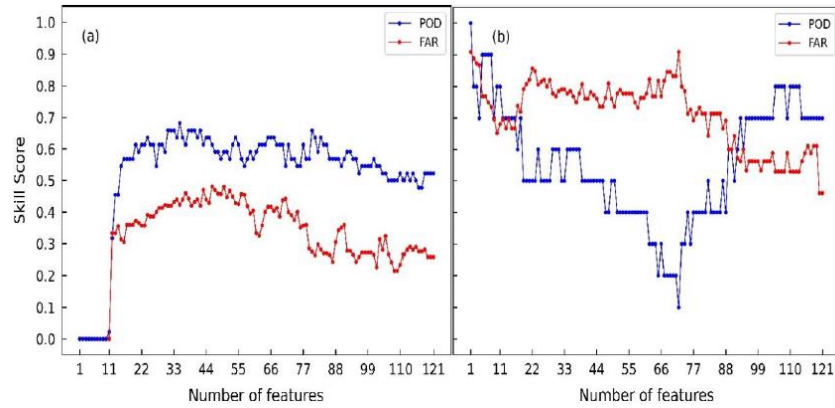

Figure 6.1: POD and FAR for (a) Atlantic and (b) Indian basin with increasing number of features.
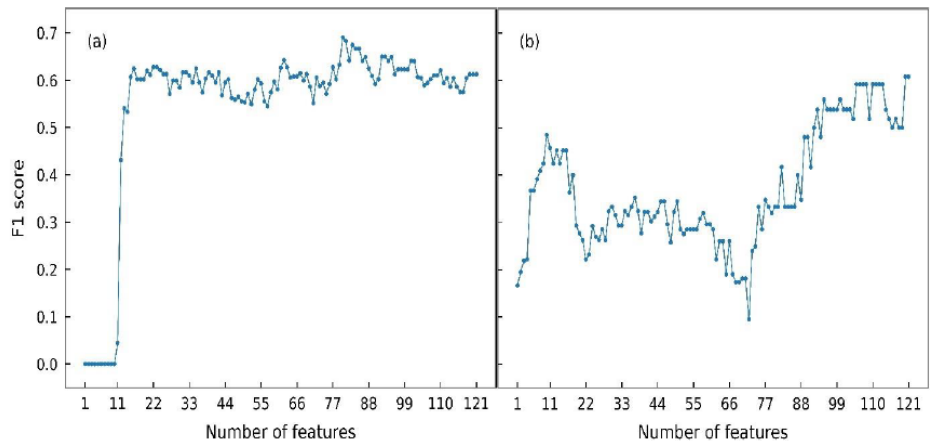

Figure 6.2: F1 scores for (a) Atlantic and (b) Indian basin with increasing number of features. The plots depict how the F1 score, a measure of a model's accuracy that considers both precision and recall, varies with the number of features used in the model for each basin.

North Pacific, and Eastern North Pacific basins, respectively. From the set of features retained for each basin, it was analysed that certain SHIPS features were common among all basins despite the number of features retained varied from 10 up to 36 features. The common features identified were DTL_t24, DELV-12, SHGC_t24, vs0, and TWXC_t24. These are described as the distance to the major landmass, 24-hour TC intensity change, generalized 850 – 200 hPa shear magnitude vs time with vortex removed (and averaged from 0 to 500 km relative to 850 hPa vortex centre), initial maximum 1-minute sustained wind at 10m, and maximum 850 hPa symmetric tangential wind at 850 hPa from NCEP analysis, respectively. Thus, it may be concluded that these features are the most important to the present study and could perhaps be helpful in future RI studies.

6.3 ML Framework Results

For RI detection, the SVMs for each basin were tested on the last one year of TC records. The Indian basin, was tested on data from the years 2015 to 2017. This was done both ways by using all the 121 initial feature sets and the optimal feature set for respective basins. The results of both the cases are summarised in Table 6.3. A comparison of the results in both cases i.e., predictions with optimal number of features and with all the features

Table 6.2: LIST OF ALL FEATURES USED FOR DIFFERENT BASINS

Atlantic	Indian	Western North Pacific	Eastern North Pacific
vs0	vs0	DELV-12	DELV-12
DTL_t24	DTL_t24	vs0	vs0
DELV-12	DELV-12	COHC_t24	ENEG_t24
CFLX_t24	MTPW_v18	DTL_t24	DTL_t24
SHGC_t24	SHGC_t24	TLON_t24	CFLX_t24
SHDC_t24	COHC_t24	TWXC_t24	VVAC_t24
TWXC_t24	SHDC_t24	VMPI_t24	TWXC_t24
TWAC_t24	TWXC_t24	CSST_t24	SHGC_t24
V500_t24	V300_t24	V500_t24	SHDC_t24
TLON_t24	CFLX_t24	CD20_t24	IR00_v5
CD20_t24		SHGC_t24	VMPI_t24
VMPI_t24		TWAC_t24	ENSS_t24
IR00_v3		RSST_t24	U200_t24
V000_t24			TWAC_t24
TLAT_t24			SHRG_t24
CD26_t24			TLAT_t24
VVAC_t24			SHRD_t24
SHRG_t24			PSLV_v6
EPOS_t24			V300_t24
MTPW_v3			U20C_t24
COHC_t24			RSST_t24
V300_t24			VMFX_t24
			D200_t24
			PSLV_v8
			CD26_t24
			V850_t24
			V500_t24
			V000_t24
			E000_t24
			PSLV_v5
			PSLV_v16
			TLON_t24
			MTPW_v11
			CD20_t24
			Z850_t24

Table 6.3: RI DETECTION RESULTS OF SVMS USING THE LEAST NUMBER OF OPTIMAL FEATURES (LEFT ENTRIES IN EACH COLUMN) AND ALL THE 121 FEATURES (RIGHT ENTRIES IN EACH COLUMN) FOR EACH BASIN.

Basin	Hits		POD		FAR	
ATL	44(27)	44(23)	0.61	0.52	0.35	0.26
IND	10(8)	10(7)	0.80	0.70	0.65	0.46
WPAC	34(12)	34(16)	0.35	0.47	0.62	0.58
EPAC	28(17)	28(15)	0.57	0.53	0.51	0.50

Basin	F1		PSS		GSS		HSS	
ATL	0.63	0.61	0.572	0.500	0.412	0.404	0.584	0.575
IND	0.48	0.61	0.732	0.672	0.291	0.416	0.451	0.588
WPAC	0.36	0.44	0.305	0.416	0.186	0.242	0.314	0.389
EPAC	0.52	0.52	0.516	0.486	0.313	0.308	0.477	0.471

*Numbers in () indicate the correctly detected number of RI cases.

suggested that the results in the latter case were better for the Western North pacific and Indian basins. There was a 34.3% improvement (increment) in POD and a 6.4% reduction in FAR for the Western North Pacific basin. The same for the Indian basin was a 12.5% loss in POD but a 14.7% reduction in FAR. For the Atlantic and Eastern North Pacific basins however, it was favourable to retain the optimal least features. The POD for the Atlantic basin came out to be 0.61 with a FAR of 0.35 with reduced feature set used. The POD and FAR for the same basin was 0.52 and 0.26, respectively with all the 121 features. This indicated a 19.2% improvement in POD and a 25.7% reduction in FAR, whereas, there was a 7.5% improvement in POD and a 1.9% reduction in FAR for the Eastern North Pacific basin.

Overall, the results for the Atlantic basin were the most promising, followed by the Indian basin which showed a POD of 0.70 and FAR of 0.58. This was also reflected by the higher F1 score and HSS of these basins as compared with the rest. The POD and FAR are conflicting metrics and therefore, other skill scores made it possible to assess the performance better. It is also worth mentioning that the testing data for the Indian basin was very small even though the LOYO approach was not employed. There were 231 TC records out of which there were only 10 cases of RI in total. Thus, the performance metrics for the Indian basin showed a greater degree of imbalance. This can

be seen when only best 10 features were used, the FAR was quite high at 0.65 which dropped to 0.46 when 121 features were used.

6.4 Model Interpretation

To gain deeper insights into the working of ML models, interpretability of ML models is becoming increasingly important. It is helpful to identify the factors that influence the predictions and see the underlying patterns and relationships in the data. There are several techniques commonly used for model explanation and interpretation, including feature importance, partial dependence plots, SHapley Additive exPlanations (SHAP) values, permutation importance, contribution plots, etc. For the present study, three techniques for this purpose were employed: ROC AUC curve, SHAP value feature importance, and SHAP dependence plot. The AUC for the Atlantic basin was 0.943 and 0.940 for the Indian basin as shown in Fig 6.3. This indicated that the model did well in identifying RI cases and minimizing false alarms for both basins.
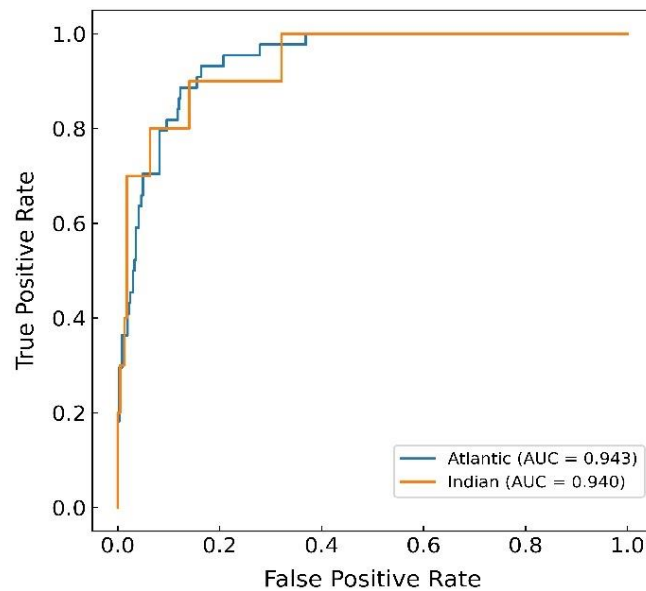

Figure 6.3: The ROC AUC curve of model predictions for both Atlantic and Indian basins. The ROC curves compare the performance of the model in predicting outcomes for the two basins. The blue line represents the ROC curve for the Atlantic basin.

The SHAP dependence plots for the features that have the highest and the lowest mean absolute SHAP value as per Fig 6.4 are displayed in Fig 6.5. It was found that the feature COHC_t24 had the highest SHAP interaction with the feature 'TWXC_t24' in Atlantic basin. 'SHGC_t24' and 'vs0' were the features with the highest and lowest mean absolute SHAP value, respectively, for the Indian basin. The respective features with the most SHAP interaction for these features were 'V300_t24' and 'DTL_t24'. The SHAP values of 'TWXC_t24' are linearly correlated with the value of 'TWXC_t24' implying that increasing 'TWXC_t24' has a positive impact on the model's predictions. 'TLON_t24' has a non-linear and highly complex relation with its SHAP values. The 'TLON_t24' values have significant impact on model's predictions but its relationship with the predictions is not straight forward. Increasing 'SHGC_t24' decreases its SHAP values almost linearly up to some value. vs0 has non-linear correlation with its SHAP
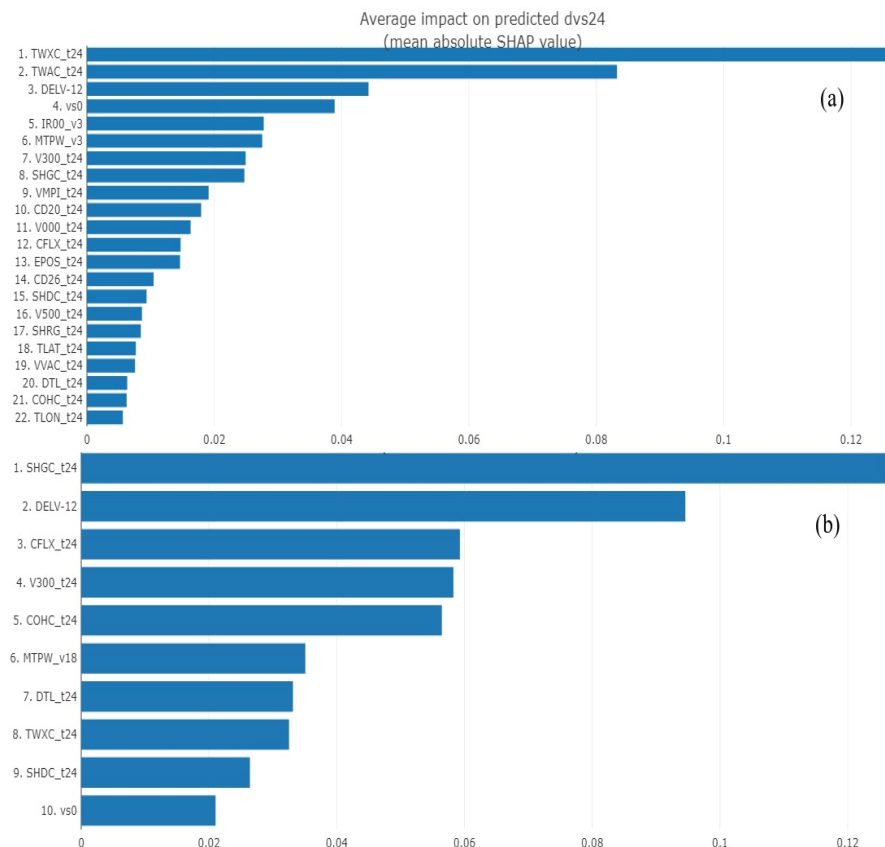

Figure 6.4: Feature importance based on mean absolute SHAP values (average impact on the predicted cyclone intensity, dvs24) of each feature for (a) Atlantic and (b) Indian basin.

values.

6.5 Comparison with existing models

The results of the present work were compared with some of the previous works in this area. The results for the Atlantic basin when optimal features were retained in the present work, showed a 7% improvement in POD and a 61.1% decrease in FAR. They also used the SHIPS database for the Atlantic basin in their study. In comparison to their model with hyperparameter tuning, the present work achieved a 48.7% improvement in POD and a 43.5% reduction in FAR. The comparison is summarised in Fig 6.6.
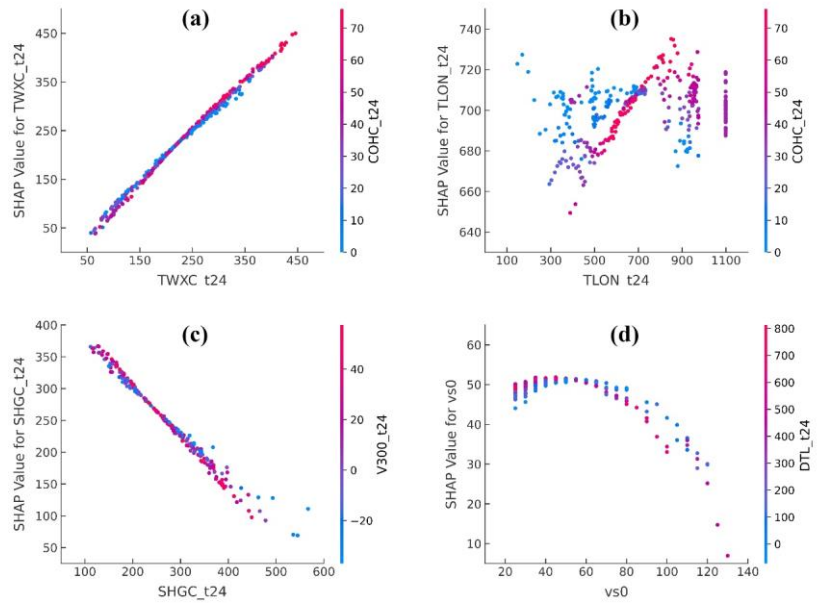
Fig. 6.5. SHAP dependence plots for Atlantic (top) and Indian basin (bottom) for the feature with highest (a and c) and lowest (b and d) mean absolute SHAP value among the optimal feature set. The color of the dot represents the value of feature that has the highest SHAP interaction with the plotted feature.
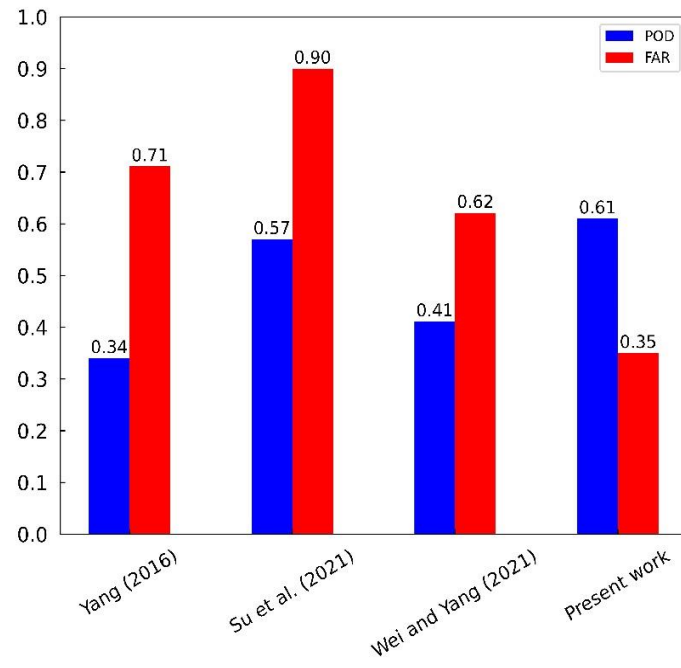
Figure 6.6: Comparison of the Atlantic basin results of present work with previous works in terms of POD and FAR.

Chapter 7

Discussion

TC intensity prediction models can help planners, decision-makers, and inhabitants in coastal regions for emergency preparedness. These models can also help the government departments to develop appropriate strategy for disaster management and mitigation measures.

The study aimed at classifying the NIO cyclones and predict their intensity based on the images. It also aimed at predicting RI events in TCs using ML approaches. It demonstrated the potential of DL and ML in tackling problems related to weather and climate.

Applying a DL technique to a TC could help meteorologists provide more accurate medium-term forecasts and issue timely warnings to communities in the path of these potentially deadly storms. Although NWP models are good and getting better, they are far from perfect. Each prediction may account for a slight variation in the many variables of the weather, such as energy from the ocean and clouds. TCs are generally more difficult to predict because their environment changes many times throughout their lifetime. If we look at the current models, they're imperfect because NWP can't see every molecule of water that it would need and every piece of energy from the sun, and we also know that how we represent some of that information is imperfect. But, when we are facing a hurricane, it's important to know what type of storm we're going to get and when we're going to get it. The models can be deployed on online servers and can be used from anywhere in the world for quick response and analysis.

7.1 Limitation of the present study

One disadvantage of using categories in TC intensity estimation is that values near the boundary value of each category can have a significant impact on the results. Additionally, the scarcity of training and independent testing data for higher intensity TCs limit the conclusiveness of the findings. Furthermore, the proposed model solely relies on the infrared images of cyclones to estimate the TC intensity, neglecting valuable information from

other channels. The infrared and visible channels are two of the most frequently used bands and these provide information about the structure and location of atmospheric systems. The thermal infrared band (8 µm -16 µm) is available during both day and night. It measures radiation emitted from the top of the clouds. But the visible channels (0.35 µm – 0.7 µm) measures the scattered and reflected solar radiation from the top of the clouds. Thus, this data is not available during the night. Information about the atmospheric systems in the lower levels of the atmosphere cannot be provided by these channels because they are often covered by clouds. A TC revolves through irregular shapes at early phases of their development. When direct quantities of environmental variables such as temperature and pressure are not available, the detection of typical circular and curved patterns from remotely sensed data is a possible method to conclude the creation and development of TC. It is, however, also to be noted that the DL model performance is yet to surpass the ADT for regression analysis.

The study also highlights the significance of careful interpretation of the results obtained from data-wise split in DL/ML approach. Based on the study, the application of DL to TC intensity analysis shows tremendous promise for further development with more advanced methodologies and expanded training datasets. Detailed analysis of a particular storm to understand model performance with storm structural changes during rapid intensification is another future work that could be studied. These endeavours will contribute to the refinement and advancement of TC intensity estimation models in the future.

For RI classification, the refined SHIPS data is available only from 1982-2017. The dataset is highly imbalanced which makes the predictions very difficult. The model is not tested on a global scale yet. The model only provides forecast for a 24-hour period, which relies heavily on persistence. However, NHC and JTWC aim to extend intensity forecasts to 5-7 days. This necessitates applying the ML method beyond 24 hours to assess its operational usefulness. Additionally, longer-range forecasts depend more on time-dependent predictors, requiring a re-evaluation of using time-averaging versus single forecast time values, as done in the current ML model.

7.2 Conclusion

In the present thesis, two problems related to TC has been tackled using ML and DL approaches. For the first work, a DL framework has been developed to estimate and classify TC from 10 km spatial resolution Meteosat satellite images over NIO regions for the period 2000-2022. The NIO region has very less dataset compared to other regions, however, is one of the important areas of cyclone genesis. Here we developed an architecture to detect, classify and predict the intensity of the cyclones. Here, for the first time, we tried to categorise the NIO cyclones in five groups. The architecture is statistically evaluated based on a data set spanning 2000-2022, both year-wise and data-wise. The performance of the architecture is found to be better or comparable to other existing studies using both evaluation method. It is to be noted that most of the existing DL/ML based studies either used less data set for performance evaluation or developed for fewer cyclone classes.

7.3 Future Scope

Based on the study, the application of DL on TC intensity analysis holds tremendous promise for further development with more advanced methodologies and expanded training datasets. Detailed analysis of a particular storm to understand model performance with storm structural changes during rapid intensification is another future work that could be studied.

Overall, this study's findings can also contribute to improving the prediction and preparation for RI events of TCs, which can have significant impacts on regions prone to such weather phenomena. Further research can explore other ML algorithms, more advanced feature selection methods, and robustness tests to improve the framework's performance. The research has also generated annotated dataset with cyclone location. The codes and results are publicly available at

https://github.com/manishmawatwal/Cyclone

References

[1] K. Walsh, J. McBride, P. Klotzbach, S. Balachandran, S. Camargo, G. Holland, T. Knutson, J. Kossin, Tsz-Cheung Lee, A. Sobel, M. Sugi, "Tropical cyclones and climate change", WIREs Climate Change, vol. 7, November 2015, https://doi.org/10.1002/wcc.371

[2] N. Mori and T. Takemi, "Impact assessment of coastal hazards due to future changes of tropical cyclones in the North Pacific Ocean", *Weather and Climate Extremes*, vol 11, pp. 53-69, March 2016.

[3] Chen, Rui, Weimin Zhang, and Xiang Wang. 2020. "Machine Learning in Tropical Cyclone Forecast Modeling: A Review" *Atmosphere* 11, no. 7: 676, https://doi.org/10.3390/atmos11070676

[4] Wang, Xin, Wenke Wang, and Bing Yan. 2020. "Tropical Cyclone Intensity Change Prediction Based on Surrounding Environmental Conditions with Deep Learning" *Water* 12, no. 10: 2685, https://doi.org/10.3390/w12102685

[5] K. A. Emanuel and David S. Nolan, "Tropical cyclone activity and global climate", in Proc. 26th Conf. Hurricanes Tropical Meteorology, Miami, FL, USA, vol 10, pp. 240-241, 2004,

https://ams.confex.com/ams/26HURR/techprogram/paper_75463.htm

[6] Zhang, Qiang, Liguang and Qiufeng, "Tropical cyclone damages in China 1983-2006", Bulletin of the American Meteorological Society, vol 90, issue 4, pp. 489, Jan 2009, https://doi.org/10.1175/2008BAMS2631.1

[7] Dvorak Vernon, "Tropical cyclone intensity analysis and forecasting from satellite imagery", Monthly Weather Review, vol. 103, issue 5, p. 420-430, May 1975.

[8] M. F. Pineros, E.A. Ritchie, and J.S. Tyo, "Estimating tropical cyclone intensity from infrared image data", Weather Forecasting, vol 26, no. 5, pp. 690-698, Oct 2011, https://doi.org/10.1175/WAF-D-10-05062.1

[9] Yu-Ju Lee, D. Hall, Q. Liu, Wen-wei Liao and Ming Huang, "Interpretable tropical cyclone intensity estimation using Dvorak-inspired machine learning techniques", *Engineering Applications of Artificial Intelligence*, vol. 101, May 2021.

[10] T.L. Olander and C.S. Velden, "The advanced Dvorak technique (ADT) for estimating tropical cyclone intensity: Update and new capabilities", Weather Forecasting, vol. 34, no. 4, pp. 905-922, 2019, https://doi.org/10.1175/WAF-D-19-0007.1

[11] E.A. Ritchie and K.G. Valliere, "Tropical cyclone intensity estimation in the North Atlantic basin using an improved deviation angle variance technique", Weather Forecasting, vol. 27, no. 5, pp. 1264-1277, Oct 2012, https://doi.org/10.1175/WAF-D-11-00156.1

[12] Lee, C.Y.; Tippett, M.K.; Camargo, S.J.; Sobel, A.H., "Probabilistic Multiple Linear Regression Modeling for Tropical Cyclone Intensity" Mon. Wea. Rev. 2015, 143, 933–954, https://doi.org/10.1175/MWR-D-14-00171.1

[13] T. L. Olander and Christopher S. Velden, "The current status of the UW-CIMSS Advanced Dvorak Technique (ADT)", Conf. on Hurricanes and Tropical Meteorology, Apr. 2016.

[14] Devaraj J, Ganesan S, Elavarasan RM, Subramaniam U, "A Novel Deep Learning Based Model for Tropical Intensity Estimation and Post-Disaster Management of Hurricanes", Applied Sciences, 2021, 11, 4129, https://doi.org/10.3390/app11094129

[15] N. Sharma, M.M. Ali, John A. Knaff and P. Chand, "A soft-computing cyclone intensity prediction scheme for the western North Pacific Ocean", RMetS Atmospheric Science Letters, vol 14, pp. 187-192, May 2013.

[16] S. Chaudhari, D. Dutta, S. Goswami and A. Middey, "Intensity forecast of tropical cyclones over North Indian Ocean using multilayer perceptron model: Skill and performance verification", Nat. Hazards 65, 97–113, 2013.

[17] D. Matsuoka, M. Nakano, D. Sugiyama, and S. Uchida, "Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model," Prog. Earth Planet. Sci., vol. 5, no. 1, pp. 1–16, Dec. 2018.

[18] C. Kumler-Bonfanti, J. Stewart, D. Hall, and M. Govett, "Tropical and extratropical cyclone detection using deep learning," J. Appl. Meteorol. Climatol., vol. 59, no. 12, pp. 1971–1985, Dec. 2020.

[19] W. Tian, W. Huang, L. Yi, L. Wu, and C. Wang, "A CNN-based hybrid model for tropical cyclone intensity estimation in meteorological industry," IEEE Access, vol. 8, pp. 59158–59168, 2020.

[20] J. Liu and B. Feng, "A Neural Network Regression Model for Tropical Cyclone Forecast", Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005.

[21] Moon, I.J.; Ginis, I.; Hara, T.; Thomas, B., "A Physics-Based Parameterization of Air-Sea Momentum Flux at High Wind Speeds and Its Impact on Hurricane Intensity Predictions", Mon. Weather Rev. 2007, 135, 2869–2878, https://doi.org/10.1175/MWR3432.1

[22] S. Kotal, S. Bhattacharya, S.K. Bhowmik, "Estimation of tropical cyclone intensity and location over the north Indian Ocean – a challenge", Meteorol Appl. 2019;26:245–252, DOI: 10.1002/met.1758

[23] J. Combinido, J. Mendoza and J. Aborot, "A Convolutional Neural Network Approach for Estimating Tropical Cyclone Intensity Using Satellite-based Infrared Images", 24th International Conference on Pattern Recognition (ICPR) Beijing, China, August 20-24, 2018.

[24] A. Asif, M. Dawood, B. Jan, J. Khurshid, M. DeMaria and F. Minhas, "PHURIE: hurricane intensity estimation from infrared satellite imagery using machine learning", Neural Computing and Applications, 32:4821-4834, 2020, https://doi.org/10.1007/s00521-018-3874-6

[25] C. Wang, G. Zheng, X. Li, Q. Xu, B. Liu and J. Zhang, "Tropical Cyclone Intensity Estimation from Geostationary Satellite Imagery Using Deep Convolutional Neural Networks," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022, Art no. 4101416, doi: 10.1109/TGRS.2021.3066299

[26] R. Zhang, Q. Liu and R. Hang, "Tropical Cyclone Intensity Estimation Using Two-Branch Convolutional Neural Network from Infrared and Water Vapor Images", IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 1, January 2020, doi:10.1109/TGRS.2019.2938204.

[27] S. Shakya, S. Kumar and M. Goswami, "Deep Learning Algorithm for Satellite Imaging Based Cyclone Detection", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol 13, pp. 827-839, 2020.

[28] M. Maskey, R. Ramachandran, M. Ramasubramanian, I. Gurung, B. Freitag, A. Kaulfus, D. Bollinger, D. Cecil and J. Miller, "Deepti: Deep-Learning-Based Tropical Cyclone Intenisty Estimation System", IEEE

Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, 2020.

[29] K. Biswas, S. Kumar, A. Pandey, "Tropical Cyclone Intensity Estimations Over the Indian Ocean Using Machine Learning", arXiv:2107.05573

[30] C. Kar and S. Banerjee, "Tropical cyclone intensity classification from infrared images of clouds over Bay of Bengal and Arabian Sea using machine learning classifiers", Arabian Journal of Geoscience, 2021, https://doi.org/10.1007/s12517-021-06997-5

[31] T. Olander, A. Wimmers, C. Velden and J. Kossin, "Investigation of Machine Learning using Satellite-based Advanced Dvorak Technique Analysis Parameters to Estimate Tropical Cyclone Intensity", AMS, DOI: 10.1175/WAF-D-20-0234.1

[32] W. Xu, K. Balaguru, A. August, N. Lalo, N. Hodas, M. DeMaria and D. Judi, "Deep Learning Experiments for Tropical Cyclone Intensity Forecasts", *AMS Journal*, vol. 36, pp. 1453-1470, July 2021.

[33] C. Kar and S. Banerjee, "Tropical Cyclone Intensity Prediction Using Best Track Data Over North Indian Ocean by Machine Learning Classifiers", 2021 IEEE International India Geoscience and Remote Sensing Symposium (InGARSS), DOI: 10.1109/InGARSS51564.2021.9792071

[34] DeMaria, M., & Kaplan, J, A Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic Basin, Weather and Forecasting, vol 9(2), pp 209–220, June 1994, https://doi.org/10.1175/15200434(1994)009%3C0209:ASHIPS%3E2.0.CO;2

[35] E. Rappaport, J.G. Jiing, C. Landsea and S. Murillo, "The Joint Hurricane Test Bed: Its first decade of tropical cyclone research-to-operations activities reviewed", AMS Journal, vol. 93, pp. 371-380, March 2012.

[36] Yang, R., "A Systematic Classification Investigation of Rapid Intensification of Atlantic Tropical Cyclones with the SHIPS Database", Weather and Forecasting, 31(2), 495–513, March 2016 https://doi.org/10.1175/waf-d-15-0029.1

[37] Yu-Ju Lee, "Interpretable tropical cyclone intensity estimation using Dvorak-inspired machine learning techniques", Engineering Applications of
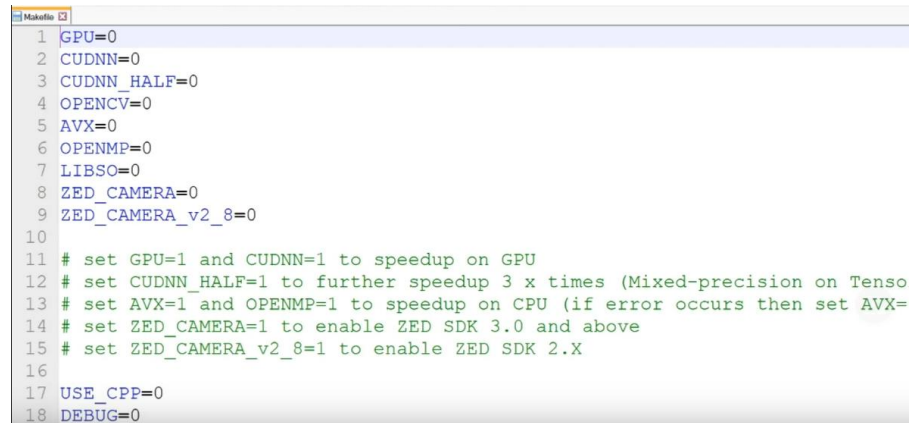
Artificial Intelligence, vol. 101, May 2021, https://doi.org/10.1016/j.engappai.2021.104233

[38] Mercer, A. E., A. D. Grimes, and K. M. Wood, 2021: Application of Unsupervised Learning Techniques to Identify Atlantic Tropical Cyclone Rapid Intensification Environments. *J. Appl. Meteor. Climatol.*, 60, 119–138, https://doi.org/10.1175/JAMC-D-20-0105.1

[39] Shaiba, H., & Hahsler, M. (2016, October 15). Applying Machine Learning Methods for Predicting Tropical Cyclone Rapid Intensification Events. Research Journal of Applied Sciences, Engineering and Technology, 13(8), 638–651, https://doi.org/10.19026/rjaset.13.3050.

[40] Li, Y.; Yang, R.; Su, H.; Yang, C. Discovering Precursors to Tropical Cyclone Rapid Intensification in the Atlantic Basin Using Spatiotemporal Data Mining. *Atmosphere* **2022**, *13*, 882. https://doi.org/10.3390/atmos13060882

[41] Cloud, K. A., Reich, B. J., Rozoff, C. M., Alessandrini, S., Lewis, W. E., & Delle Monache, L. (2019, July 24). A Feed Forward Neural Network Based on Model Output Statistics for Short-Term Hurricane Intensity Prediction. Weather and Forecasting, 34(4), 985–997, https://doi.org/10.1175/waf-d-18-0173.1

[42] Griffin, S. M., A. Wimmers, and C. S. Velden, 2022: Predicting Rapid Intensification in North Atlantic and Eastern North Pacific Tropical Cyclones Using a Convolutional Neural Network. *Wea. Forecasting*, **37**, 1333–1355, https://doi.org/10.1175/WAF-D-21-0194.1

[43] F. J. Tapiador, A. Navarro, R. Martín, S. Hristova-Veleva and Z. S. Haddad, "Predicting Tropical Cyclone Rapid Intensification from Satellite Microwave Data and Neural Networks," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022, Art no. 4205213, doi: 10.1109/TGRS.2021.3128076.

[44] Wei, Y.; Yang, R., "An Advanced Artificial Intelligence System for Investigating Tropical Cyclone Rapid Intensification with the SHIPS Database", Atmosphere 2021, 12, 484, https://doi.org/10.3390/atmos12040484

[45] Wei, Y.; Yang, R.; Sun, D. Investigating Tropical Cyclone Rapid Intensification with an Advanced Artificial Intelligence System and Gridded

Reanalysis Data. *Atmosphere* 2023, *14*, 195. https://doi.org/10.3390/atmos14020195

[46] Hui Su, Longtao Wu, Jonathan H. Jiang, Raksha Pai, Alex Liu, Albert J. Zhai, P. Tavallali, M. DeMaria, "Applying Satellite Observations of Tropical Cyclone Internal Structures to Rapid Intensification Forecast with Machine Learning", Geophysical Research Letter, Aug 2020, https://doi.org/10.1029/2020GL089102

[47] DeMaria, M.; Franklin, J.L.; Onderlinde, M.J.; Kaplan, J. Operational Forecasting of Tropical Cyclone Rapid Intensification at the National Hurricane Center. *Atmosphere* 2021, *12*, 683. https://doi.org/10.3390/atmos12060683

[48] Zhou, G.; Xu, J.; Qian, Q.; Xu, Y.; Xu, Y. Discriminating Technique of Typhoon Rapid Intensification Trend Based on Artificial Intelligence. *Atmosphere* 2022, *13*, 448. https://doi.org/10.3390/atmos13030448

[49] Fischer, M. S., B. H. Tang, and K. L. Corbosiero, 2019: A Climatological Analysis of Tropical Cyclone Rapid Intensification in Environments of Upper-Tropospheric Troughs. *Mon. Wea. Rev.*, 147, 3693–3719, https://doi.org/10.1175/MWR-D-19-0013.1

[50] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002, June 1). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953.
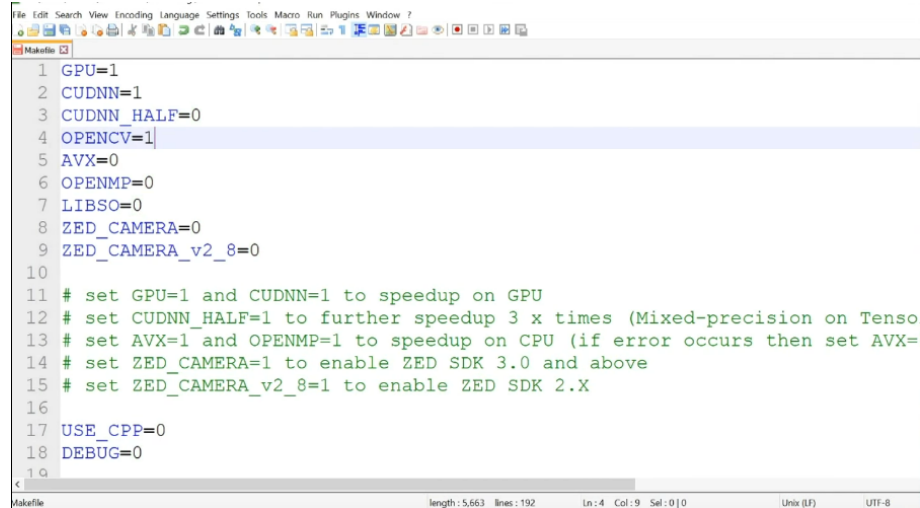
APPENDIX 1

Code segment and configuration changes in YOLOv3

'Make' is command which executes your 'makefile', it is a build script to create/tune the necessary things like environment/folders/... etc. Fig A1 shows the screenshot of the 'make' file. Fig A2 shows the changed configuration of the 'make' file. After this, we will download the make file, to make some changes.

Figure A1: 'Make' file screenshot

Figure A2: Changed configuration of 'Make' file

The 'Make' file is replaced on the original location in Google drive. We will then separate the files into test and train (2:8). We will use transfer learning with 'darknet53.conv.74' weights. We will use this to train on our custom dataset. We will make changes in the configuration file of 'yolov3.cfg'. Our dataset had only one class, i.e., cyclone. Since we had were small number of images to train, we will decrease the batch size. We will decrease the

'max_batches' (number of iterations for yolo training) size to 2000 (No of classes x 2000), 'steps' will be reduced according to the 'max_batches' size, 20% below to 20% above. Fig A3 shows the changes that has to be made in the 'yolov3.cfg' file to train custom dataset of cyclone images. We will make changes to three 'yolo' layers and the preceding 'conv' layer.

Figure A3: Changes in the 'yolov3.cfg' file

APPENDIX 2

Streamlit is an open-source app framework in Python language. It helps us create web apps for data science and ML. It is compatible with major python libraries such as scikit-learn, Keras, PyTorch, NumPy, pandas, Matplotlib etc. Heroku is a container-based cloud Platform as a Service (PaaS), to deploy, manage, and scale modern apps. Fig. A4 shows a screenshot of the local server deployment of Image Classification web app. One of the unique features of Streamlit is its ability to instantly update the app as users make changes to the code. This enables users to quickly iterate and test their code, allowing for a more efficient development process.

Overall, Streamlit is a powerful tool for data scientists and developers who want to create powerful and interactive web apps quickly and easily, with support for a wide range of data science and ML libraries.
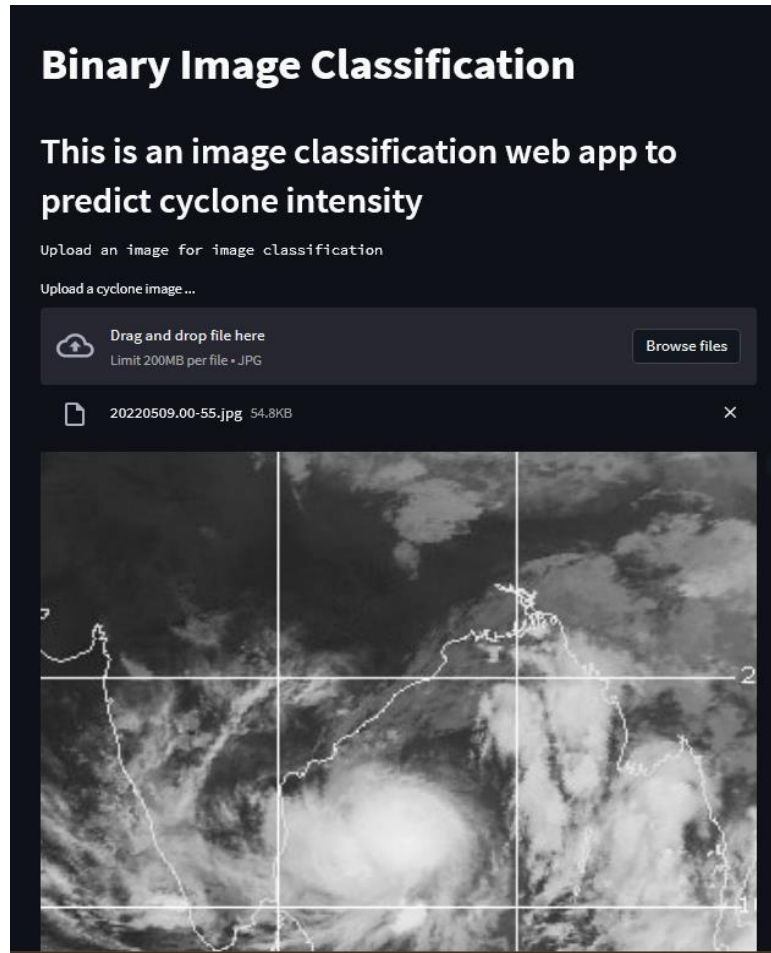
Figure A4: Streamlit (python library) app deployment of proposed model

APPENDIX 3

Table A1. COMPLETE LIST OF ALL 121 ENGINEERED FEATURES USED IN THE 24-H INTENSITY MODEL FOR RI PREDICTION.

Predictors	Long Name
vs0	Initial maximum 1-min sustained wind speed at 10 m (kt)
PSLV_v2	Pressure of the centre of mass (hPa) of the layer where storm motion best matches environmental flow.
PSLV_v3	The observed zonal storm motion component (ms^{-1} x 10)
PSLV_v4	The observed meridional storm motion component (m s^{-1} x 10)
PSLV_v5	As in PSLV_v2, but for the 1000–100-hPa mass weighted deep layer environmental wind (m s^{-1} x 10)
PSLV_v6	As in PSLV_v3, but for the 1000–100-hPa mass weighted deep layer environmental wind (m s^{-1} x 10)
PSLV_v7	As in PSLV_v2, but for the optimally weighted deep layer mean flow (m s^{-1} x 10)
PSLV_v8	As in PSLV_v3, but for the optimally weighted deep layer mean flow (m s^{-1} x 10)

PSLV_v9	The parameter alpha that controls the constraint on the weights from being not too ''far'' from the deep layer mean weights (nondimensional x 100)
PSLV_v10	The optimal vertical weights for p = 100 hPa (nondimensional x 1000)
PSLV_v11	The optimal vertical weights for p = 150 hPa (nondimensional x 1000)
PSLV_v12	The optimal vertical weights for p = 200 hPa (nondimensional x 1000)
PSLV_v13	The optimal vertical weights for p = 250 hPa (nondimensional x 1000)
PSLV_v14	The optimal vertical weights for p = 300 hPa (nondimensional x 1000)
PSLV_v15	The optimal vertical weights for p = 400 hPa (nondimensional x 1000)
PSLV_v16	The optimal vertical weights for p = 500 hPa (nondimensional x 1000)
PSLV_v17	The optimal vertical weights for p = 700 hPa (nondimensional x 1000)
PSLV_v18	The optimal vertical weights for p = 850 hPa (nondimensional x 1000)
PSLV_v19	The optimal vertical weights for p = 1000 hPa (nondimensional x 1000)
MTPW_v2	0–200-km average total precipitable water (TPW) at t = 0 from the GFS analysis (mm x 10)
MTPW_v3	0–200-km TPW standard deviation (mm x 10)
MTPW_v4	200–400-km average TPW (mm x 10)
MTPW_v5	200–400-km TPW standard deviation (mm x 10)
MTPW_v6	400–600-km average TPW (mm x 10)
MTPW_v7	400–600-km TPW standard deviation (mm x 10)
MTPW_v8	600–800-km average TPW (mm x 10)
MTPW_v9	600–800-km TPW standard deviation (mm x 10)
MTPW_v10	800–1000-km average TPW (mm x 10)
MTPW_v11	800–1000-km TPW standard deviation (mm x 10)
MTPW_v12	0–400-km average TPW (mm x 10)
MTPW_v13	0–400-km TPW standard deviation (mm x 10)
MTPW_v14	0–600-km average TPW (mm x 10)
MTPW_v15	0–600-km TPW standard deviation (mm x 10)
MTPW_v16	0–800-km average TPW (mm x 10)
MTPW_v17	0–800-km TPW standard deviation (mm x 10)
MTPW_v18	0–1000-km average TPW (mm x 10)
MTPW_v19	0–1000-km TPW standard deviation (mm x 10)
MTPW_v20	Percent TPW less than 45 mm, r 5 0–500 km in 908 azimuthal quadrants centred on up-shear direction

MTPW_v21	0–500-km averaged TPW (mm 3 10) in 908 up-shear quadrant
MTPW_v22	0–500-km average TPW (mm x 10)
IR00_v2	Time (hhmm) of the GOES image
IR00_v3	Average GOES channel-4 brightness temperature (BT) (C x 10), r = 0–200 km
IR00_v4	Std dev of GOES BT (C x 10), r = 0–200 km
IR00_v5	Average GOES channel 4 brightness temperature (BT) (C x 10), r = 100–300 km
IR00_v6	Std dev of GOES BT (C x 10), r = 100–300 km
IR00_v7	Percent area r = 50–200 km of GOES channel 4 BT< -10 C
IR00_v8	Percent area r = 50–200 km of GOES channel 4 BT < - 20C
IR00_v9	Percent area r 5 50–200 km of GOES channel 4 BT <- 30C
IR00_v10	Percent area r 5 50–200 km of GOES channel 4 BT<- 40C
IR00_v11	Percent area r 5 50–200 km of GOES channel 4 BT<- 50C
IR00_v12	Percent area r 5 50–200 km of GOES channel 4 BT<- 60C
IR00_v13	Max BT from 0- to 30-km radius (°C x 10)
IR00_v14	Avg BT from 0- to 30-km radius (°C x 10)
IR00_v15	Radius of max BT (km)
IR00_v16	Minimum GOES brightness temperature from 20- to 120-km radius (C x 10)
IR00_v17	Avg BT from 20- to 120-km radius (°C x 10)
IR00_v18	Radius of min BT (km)
IR00_v19	Variables No. 1 need for storm size estimation
IR00_v20	Variables No. 2 need for storm size estimation
IR00_v21	Variables No. 3 need for storm size estimation
CSST_24	Climatological SST (°C x 10)
CD20_24	Climatological depth (m) of 20°C isotherm from 2005 to 2010 NCODA analyses
CD26_24	As for CD20, but for the 26°C isotherm
COHC_24	As above, but for ocean heat content (kJ cm^{-2})
DTL_124	Distance to nearest major landmass (km)
RSST_24	Reynolds SST (°C x 10)
U200_t24	200-hPa zonal wind (kt x 10) (r = 200-800 km)
U20C_t24	As in U200_124, but for r = 0-500 km
V20C_t24	As in U20C_124, but for the v component of the wind
E000_t24	1000-hPa theta_e (r = 200-800 km) vs time (K x 10)
EPOS_t24	The average theta_e difference between a parcel lifted from the surface and its environment (200-800-km average) vs time (°C x 10)

ENEG_t24	As in EPOS, but only negative differences are included. The minus sign is not included.
EPSS_t24	As in EPOS, but the parcel theta_e is compared with the saturated theta_e of the environment
ENSS_t24	As in ENEG, but the parcel theta_e is compared with the saturated theta_e of the environment
RHILO_t24	850-700-hPa relative humidity (%) vs time (200-800 km)
RHMD_t24	As in RHILO, but for 700-500 hPa
RHHI_t24	As in RHILO, but for 500-300 hPa
Z850_t24	850-hPa vorticity (ξ x 10^5) vs time (r = 0-1000 km)
D200_t24	200-hPa divergence vs time (r = 0-1000 km)
REFC_t24	Relative eddy momentum flux convergence (m s^{-1} day^{-1}, 100-600-km avg)
PEFC_t24	Planetary eddy momentum flux convergence (m s^{-1} day^{-1}, 100-600-km avg)
T000_t24	1000-hPa temperature (°C x 10) (200-800-km average)
R000_t24	1000-hPa relative humidity (200-800-km average)
Z000_t24	1000-hPa height deviation (m) from the U.S. standard atmosphere
TLAT_t24	Latitude of 850-hPa vortex center in NCEP analysis (°N x 10)
TLON_t24	Longitude of 850-hPa vortex center in NCEP analysis (°W x 10)
TWAC_t24	0-600-km average symmetric tangential wind at 850 hPa from NCEP analysis (m s^{-1} x 10)
TWMX_t24	Maximum 850-hPa symmetric tangential wind at 850 hPa from NCEP analysis (m s^{-1} x 10)
G150_t24	Temperature perturbation at 150 hPa due to the symmetric vortex calculated from the radial thermal wind. Averaged from r = 200 to 800 km centered on input lat/lon (not always the model/analysis vortex position). (°C x 10)
G200_t24	As in G150, but at 200 hPa
G250_t24	As in G150, but at 250 hPa
W000_t24	The tangential wind (m s^{-1} x 10) azimuthally averaged at r = 500 km from (TLAT, TLON). If TLAT, LON are not available, (LAT, LON) are used.
V850_t24	As in W000, but at 850 hPa
V500_t24	As in W000, but at 500 hPa
V300_t24	As in W000, but at 300 hPa
TGRD_t24	The magnitude of the temperature gradient between 850 and 700 hPa averaged from 0 to 500 km estimated from the geostrophic thermal wind (°C m^{-1} x 10^7)
TADV_t24	The temperature advection between 850 and 700 hPa averaged from 0 to 500 km from the geostrophic thermal wind (°C s^{-1} x 10^4)

PENC_t24	Azimuthally averaged surface pressure at outer edge of vortex [(hPa - 1000) x 10]
SHRD_t24	850-200-hPa shear magnitude (kt x 10) vs time (200-800 km)
SHTD_t24	Heading of above shear vector. Westerly shear has a value of
SHRS_t24	850-500-hPa shear magnitude (kt x 10)
SHTS_t24	Heading of above shear vector
SHRG_t24	Generalized 850–200-hPa shear magnitude (kt x 10) vs time (takes into account all levels from 1000 to 100 hPa)
SHDC_t24	As in SHRD, but with vortex removed and averaged from 0 to 500 km relative to 850-hPa vortex center
SDDC_t24	Heading (°) of above shear vector. Westerly shear has a value of 90°
SHGC_t24	As in SHRG, but with vortex removed and averaged from 0 to 500 km relative to 850-hPa vortex center
DIVC_t24	As in D200, but centered at 850-hPa vortex location
T150_t24	200–800-km area average 150-hPa temperature (°C x 10)
T200_t24	As above, but for 200-hPa temperature (°C x 10)
T250_t24	As above, but for 250-hPa temperature (°C x 10)
PENV_t24	200–800-km average surface pressure [(hPa - 1000) x 10]
VMPI_t24	Maximum potential intensity from K. Emanuel equation (kt)
VVAV_t24	Average (0–15 km) vertical velocity (m s^{-1} x 100) of a parcel lifted from the surface where entrainment, the ice phase, and the condensate weight are accounted for. Note: Moisture and temperature biases between the operational and reanalysis files make this variable inconsistent in the 2001–07 sample, compared to 2000 and before.
VMFX_t24	As in VVAV, but a density-weighted vertical average
VVAC_t24	As in VVAV, but with soundings from 0 to 500 km with GFS vortex removed
HE07_t24	Storm motion relative helicity (m^2 s^{-2}) x 10 for p = 1000–700 hPa, r = 200–800 km
HE05_t24	As in HE07, but for p = 1000–500 hPa
O500_t24	Pressure vertical velocity (hPa day^{-1}) at 500 hPa, averaged from r = 0 to 1000 km
O700_t24	As in O500, but at 700 hPa
CFLX_t24	Dry air predictor based on the difference in surface moisture flux between air with the observed (GFS) RH value, and with RH of air mixed from 500 hPa to the surface
DELV-12	Last 12-h intensity change (kt)

APPENDIX 4

The results of the imbalanced dataset were initially not included due to their underwhelming performance. After thorough consideration and analysis, it

was determined that the inclusion of these results would not contribute meaningful insights or enhance the discussion.

The results of the imbalanced dataset are as below: For 20 best Random Forest Atlantic Features

Metrics	Values (unbalanced dataset)	Values (balanced dataset using SMOTE)
POD	0.033	0.662
FAR	0.785	0.819
HSS	0.046	0.216