HurriNet: A Framework For Forecasting Tropical Cyclones Using Reanalysis Maps With Vision Transformers

Aaron Brendan Huang
University of Toronto Mississauga
falcom.huang@mail.utoronto.ca

Areeb Imran
University of Toronto Mississauga
areeb.imran@mail.utoronto.ca

Kyle Robert Hobeck
University of Toronto Mississauga
kyle.hobeck@mail.utoronto.ca

Quincy Richard Antonio Garcia
University of Toronto Mississauga
quincy.garcia@mail.utoronto.ca

Abstract

This project aims to refine tropical cyclone forecasting accuracy by incorporating deep learning practices to foresee tropical cyclone trajectories and intensities within the first 24 hours of disaster. Tropical cyclones can inflict substantial harm to communities, furthering the need for precise forecasting which is vital for preparedness and mitigation efforts. Using recent and historical tropical cyclone data with spatial-temporal features in the IBTrACS dataset, we will develop Embedding and Transformer based deep learning models. This project applies spatial patching (using spatial data) to a forecasting problem and observes the final model performance by using the predictions of other forecasting agencies as a benchmark. The target is to achieve a model with a forecast error similar to or less than the forecasting solution used by NHC.

1 Introduction

Hurricanes, typhoons, and cyclones all describe the same phenomenon and can be collectively referred to as Tropical Cyclones (TC). TCs are well known for their adverse consequences and destructive nature, which have continued to increase in severity due to climate change. Coastal areas have suffered dramatically due to TCs, with high life loss and extensive infrastructure damage [1]. For example, the Atlantic hurricane weather experienced in 2024 as of Oct 11 saw the development of the record-breaking Hurricane Milton that reached Category 5 with wind speeds of 180 mph and caused an estimated economic loss between $160 billion and $180 billion. This far exceeds the damages incurred by any one of the TCs during the 2023 Atlantic hurricane season and highlights the rising urgency of hurricanes, furthering the need for efficient TC forecasting with enough lead time to ensure public safety [2, 3]. Deep learning approaches have broken new ground among many tropical cyclone forecasting systems used in practice today, with Recurrent Neural Networks (RNN) and Transformers standing out in particular [4].

Deep learning models derived in previous research have excelled at processing complex data maps efficiently to accurately identify relationships that traditional methods may not be able to discern [5]. The traditional approach relies on numerical forecasting models and fed observational data to arrive at an interpretable prediction. However, these models are computationally taxing and require a very high spatial resolution to capture finer features, so their attention to localized spatial data is otherwise limited. Most forecasting models in operation carry this limitation [6].

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

Our deep-learning approach aims to address these issues through an encoder-decoder model. The architecture encodes statistical data and segmented reanalysis map data together through a transformer to form fused embeddings of the differing data that are equipped to assimilate the intricate details of the input data images. The embedding transformer pass output is aggregated for all time steps in the first 24 hours of a hurricane and fed into a Transformer for refinement and then through a batching layer, mean-pooling layer, and a fully connected layer whose output corresponds to the hurricane track and intensity predictions as a probability distribution [4].

The architecture used in our method seeks to assimilate the input data described by the ERA5 and IBTrACS datasets while remaining computationally efficient and accurate, using fused embeddings and a Transformer, respectively.

2 Background and Related Work

In previous TC research, reanalysis maps are used to visualize important geospatial data that influences TC directionality and intensity [4]. These geospatial features include wind speeds, humidities, and other tropical-cyclone-specific factors (e.g., eye of the storm, diameter of the surrounding storm). Similar to a heat map, reanalysis maps are plots that reflect the intensity of a particular feature in the data through deepening colour. This report will examine reanalysis maps with supplementary statistical data to form predictions.

The work presented in this report takes inspiration from the approach used by Boussioux, Zeng, et al. to track TC paths and intensities in both the North Atlantic and the Eastern Pacific using reanalysis map inputs [4]. In the previous research, a convolutional neural network is employed on the entirety of a reanalysis map to derive input embeddings. However, our model instead uses a Transformer that receives spatial patches, fragments of the spatial data, to create these embeddings.

Similar research has been done by Sun et al. who explored the applications of statistical model ensembles to predict TC activity in the Atlantic. The researchers compared the predictive accuracy of individual models, machine-learning-optimized ensembles, and simple-averaging ensembles [7].

3 Data

3.1 Data Overview

Our study uses comprehensive datasets to model and predict cyclone behavior. In particular, we employ the International Best Track Archive for Climate Stewardship (IBTrACS) shapefile data and the ERA5 Reanalysis Maps from the European Centre for Medium-Range Weather Forecasts (ECMWF). The IBTrACS dataset includes records of cyclones that have wind speeds of 35kt or more between the years 1990 and 2023. The ERA5 dataset, which provides high-resolution meteorological data, covers the same temporal range with spatial grids at 0.25-degree intervals. To meet the goal of predicting cyclone intensities and movements within the first 24 hours, our dataset only includes cyclones recorded by United States meteorological agencies that span a period of 24 hours or more which encompasses 2050 unique cyclones. For the purposes of our problem and to avoid input length discrepancies due to varying cyclone lifespans, we clip the data to a uniform shape.

3.2 Data Sources

3.2.1 IBTrACS Shapefile Data

The International Best Track Archive for Climate Stewardship (IBTrACS) provides comprehensive shapefile data on historical TCs worldwide. We will be using a subset of data, focused on cyclones with a wind speed of 35 kt or more between the years 1990 and 2023. IBTrACS has two variations of their data set: point data and line data, but the key features we are interested in are the same across both variations. We will choose the point data variation; the key features we are interested in are showcased in Table 1. Cyclone records can be grouped using the `SID` as IBTrACS creates separate point data entries for cyclones in three-hour periods.

3.2.2 ERA5 Reanalysis Maps

The ECMWF provides the fifth generation atmospheric reanalysis of global climate dataset called ERA5. It contains high-resolution spatial-temporal meteorological data which will be used to capture environmental conditions that influence cyclone behaviour. To gather this data specific to the cyclones in the IBTrACS dataset, we will utilize reanalysis maps specific to the latitude/longitude, time, and area of each TC. This ensures more precision with our model and saves both space and computation time. Moreover, ERA5 renalaysis maps cover block grids of 0.25 degrees squared, so it is effective for looking at tropical storms spanning hundreds of kilometres long. We cover the important meteorological features included in our reanalysis maps in Table 2.

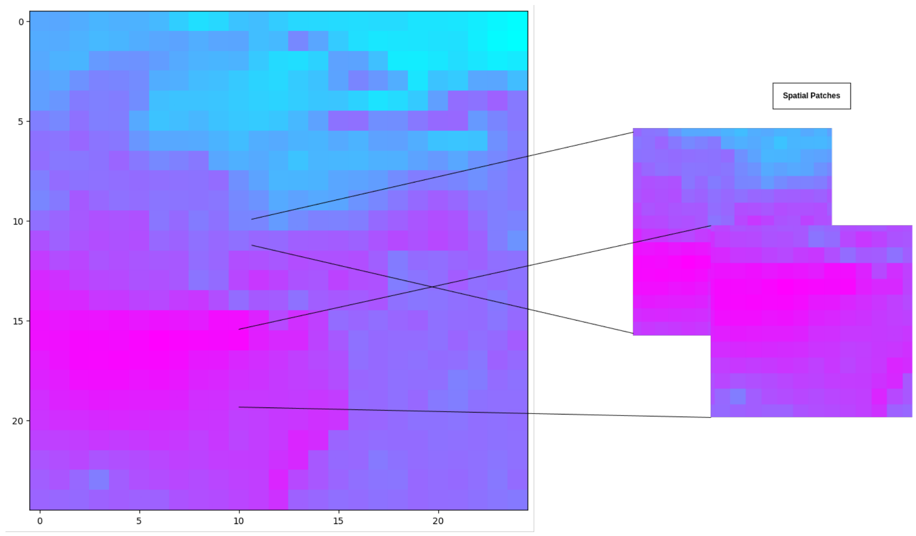

Figure 1: The plot on the left is a plot of an example hurricane from the IBTrACS dataset at a certain time of the day. The screenshots on the right display certain parts of the image, which are the smaller spatial patches. The purple represents very high wind speed areas.

Lastly, we will further break down these small reanalysis maps into smaller reanalysis maps to be converted into sequences. This technique, called spatial patching [8] shown in Figure 1, analyzes localized properties of each data piece for time series forecasting with RNNs; further details on spatial patching are described below as part of our pipeline for data formatting and cleaning. By dividing up the reanalysis maps into sub-regions, we examine localized spatial variations in the environment that influence TC behaviour.

The pipeline for formatting and cleaning the data is written as follows:

1. **Filtering and Cleaning**: Filter key features from the IBTrACS shapefile data, while also ensuring those important values existing in the fields without artifacts.
   (a) One-hot encoding, standard normalization, and sine/cosine encoding are all employed for these data types.
2. **Integrating ERA5 Data**: Using this cleaned IBTrACS data, make API requests to ERA5 using the spatial attributes to describe a 0.25-by-0.25 degree sub-region of data (or approximately 27.76km-by-27.76km). For every three hour interval of the day, this data is centered at the recorded cyclone latitude and longitude. ERA5 stores reanalysis maps based on rounded pressure and time parameters so API requests are made with parameters closest to the value of IBTrACS spatial attributes.
   (a) The ERA5 reanalysis maps will now correspond to `SID`s from the CWFIS dataset.
3. **Spatial Patching and Sequencing**:
   (a) Spatial patching is implemented by dividing each ERA5 reanalysis map into smaller, localized sub-regions (or patches). This technique allows for the analysis of localized environmental variations that influence cyclone behaviour.

(b) Convert these spatial patches into sequential data suitable for a RNN input. Our method focuses on preserving temporal sequences within each patch where CNNs are typically used for feature extraction.

3.3 Spatial Patching Methodology

Spatial patching is the process of segmenting each ERA5 reanalysis map into smaller, more localized patches to capture smaller patterns across each map. Each patch will represent a sub-region of the cyclones environment so that our model can focus on detailed spatial variations influencing the dynamics of cyclones.

4 Model Architecture

model.png

Figure 2: Architecture of the Fused Embedding Encoder showing the integration of statistical data from IBTrACS and visual data from ERA5 through spatial patching and transformer-based encoding.

The model that will be used in this paper, shown in Figure 2, combines a custom made Vision Transformer with a *Temporal Transformer* to predict the displacement and position of cyclones on the next time step. The model is broken up into two parts: the *Fused Embedding Encoder* and the *Temporal Processing Pipeline*.

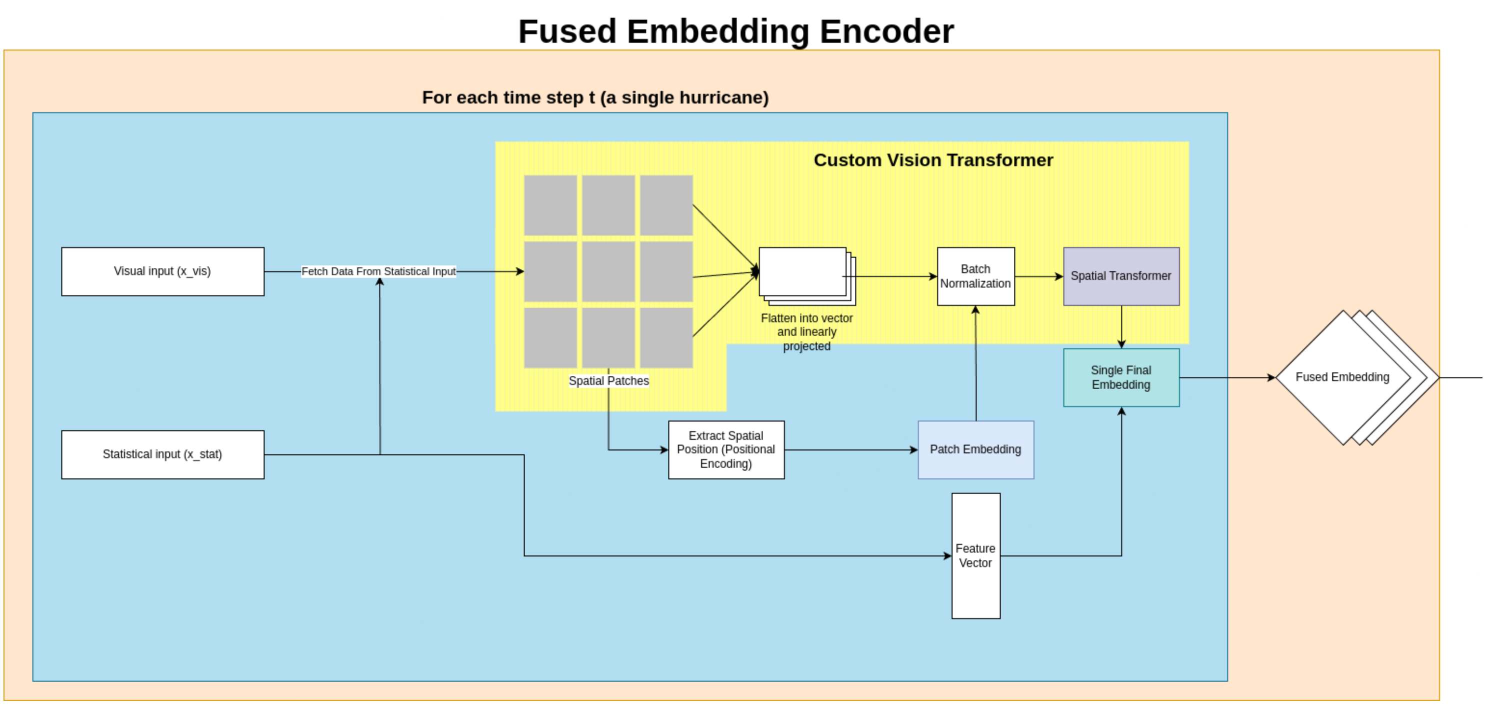

Figure 3: Architecture of the Fused Embedding Encoder showing the integration of statistical data from IBTrACS and visual data from ERA5 through spatial patching and transformer-based encoding.

4.1 Fused Embedding Encoder

In the Fused Embedding Encoder, shown in Figure 3, statistical data from IBTrACS is used to fetch the visual input from the ERA5 dataset about the cyclones. These are reanalysis maps that are 1°-by-1° in both height and width. From here, these reanalysis maps are broken up into spatial patches: each of which is a 5×5 grid of smaller images from the original image. Each of these images is encoded with position data from the Positional Encoding class, then they are flattened, linearly projected, and transformed into a *Patch Embedding*. The Patch Embedding is combined with the flattened vector, batch normalized (to stabilize the learning process), and fed through a *Spatial Transformer*. The Spatial Transformer is a type of transformer designed to capture spatial relationships within the data. The output is integrated with the statistical data to achieve a single final embedding for a single time step t for a singular cyclone. Adding up all the time steps yields 8 *Fused Embeddings*, as they fuse together the statistical and visual encoded data with one embedding for every three hour interval in the IBTrACS data. These embeddings are now a robust representation of cyclone dynamics over time and ready to be used in the next step.

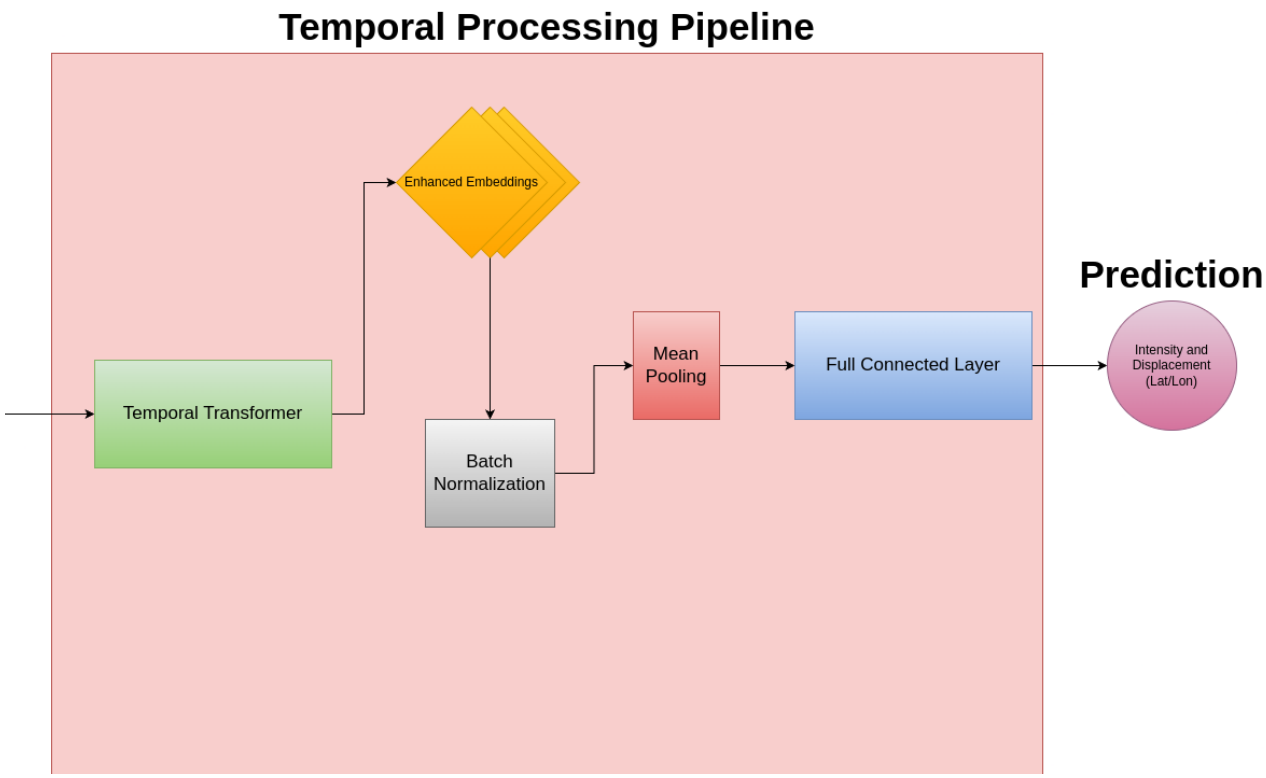

Figure 4: Architecture of the Temporal Processing Pipeline which demonstrates the Temporal Transformer and the steps taken to batch normalize, mean pool, and put these through a full connected layer in order to get our prediction output.

4.2 Temporal Processing Pipeline

As shown in Figure 4, the process is now several sequential stages that transform these *Fused Embeddings* into meaningful predictions.

1. The *Temporal Transformer* is a transformer that captures temporal dependencies and patterns across the time steps; it is fed the sequence of Fused Embeddings for every cyclone.
2. The output of the previous transformer are *Enhanced Embeddings*: embeddings that now model the temporal dynamics of cyclone displacement and intensity. These are batch normalized to stabilize the training process.
3. Mean pooling is applied to these embeddings to reduce the dimensionality while retaining important information about the cyclones.
4. A Fully Connected Layer is deployed to transform these pooled embeddings into the prediction outputs. There is no non-linear activation function applied at this step which is ideal for regression type problems.

The final output of this pipeline consists of two key components: the intensity and the displacement of the cyclones over the next time step (in 3 hours).

Table 1: Intensity Forecast MAE across HurriNet, Hurricast, and Operational Forecast Models

Model name	MAE (kt)	Skill (%)
HurriNet	11.3	-9.7
Hurricast (HUML-(stat/viz, xgb/cnn/transfo))	10.3	0.0
Operational forecast (GFSO)	14.9	-44.6
Operational forecast (HWRF)	10.0	2.9

4.3 Training Settings & Hyperparameters

For training, we form a training set of all cyclones in our dataset from 1990 to 2018, a validation set of cyclones from 2019 to 2021, and a test set of cyclones from 2022 to 2023; the sets are of sizes 1645, 261, and 144, respectively. The choice to partition by years is to avoid the inclusion of cyclones that may have occurred in close proximity or in the same period of time across different sets, reducing the risk of overfitting. The model training procedure utilized the loss function MSE (Mean Squared Error) between the predicted outputs and ground truth labels. This type of loss function is appropriate for regression tasks such as this. Adam optimizer is also used with a specified learning rate and weight decay for regularization. Batches are deployed during this step alongside gradient clipping to prevent exploding gradients during training. The model was trained for 10 epochs and plotted every 10 iterations while shuffling is enabled to diversify the mini-batches.

4.3.1 Hyperparameters

The tuned hyperparameters are as follows: learning rate, batch size, weight decay, and gradient clipping (clip or no clip). Upon looking at several combinations of these hyperparameters, the most optimal hyperparameter values found for our model are:

- Learning Rate = 0.001
- Batch Size = 128
- Weight Decay = 0.0
- Gradient Clipping = None

which yielded a final validation error of 0.046717.

5 Results

We evaluate the performance of our model by comparing the mean absolute error (MAE) and the mean geographic distance error (MGDE) between the HurriNet model and the results recorded by the relevant research paper [4]. This section will compare our HurriNet model, the best Hurricast model [4] and operational forecast models. The results that are compared will be over different test sets.

The above table is taken from the research paper [4]. The research paper evaluates its results over the Eastern Pacific and North Atlantic basins. The errors for the Hurricast and forecast models are averaged over these two basins, and are then compared to the HurriNet model. The forecast skill is reported relative to the Hurricast model, which is our baseline. The HurriNet model performs about 9.7% worse than the Hurricast model. The GFSO and HWRF operational forecasts, have a forecast skill of -44.6% and 2.9% respectively.

6 Discussion

Based upon the aforementioned results, the proposed model can capture accurate directionality of cyclone movement and generally predict intensity changes in a correct manner.

A notable result of the model is the increased susceptibility in over-predicting longitudinal or latitudinal changes, with the true change in degrees being consistently lower than the produced

Table 2: Track Forecast MAE across HurriNet, Hurricast, and Operational Forecast Models

Model name	MAE (kt)	Skill (%)
HurriNet	28.5	-9.7
Hurricast (HUML-(stat/viz, xgb/cnn/transfo))	90.5	0.0
Operational forecast (FSSE)	62.5	30.9

prediction. However, this series of over-predictions in both the positive and the negative is negligible since the mean geometric distance error of our model is about 28.6 kilometres. This value is less than half of a longitudinal degree or latitudinal degree from the target, affirming that, on average, our model is able to precisely determine the TC track. Comparing this to the Hurricast model (which, when predicting track, has a Mean Averaging Error of 68 kilometres for the most optimal ensemble), we observe notable improvement in our model in terms of track prediction compared to contemporary approaches [4].

The predictive capabilities of our Transformer model with regards to intensity is also competitive with other researched models. Comparing to the CNN-based Hurricast, the proposed model is comparable in terms of the Mean Averaging Error of forecasting hurricane wind intensities with a difference of one knot. Comparing to target values, our model slightly underpredicts the change in wind speed. Unlike the longitude and latitude track predictions which were suceptible to overestimation in both extremes (positive and negative degrees) but captured the direction of cyclones, the model characteristically underpredicted the knot speed of winds throughout testing which is what forms the slightly higher Mean Averaging Error seen.

Overall, the Transformer model with spatial patching makes new strides in path predictive capabilities while still matching previous performance for intensity prediction.

7 Limitations

In this paper, there are several limitations to the architecture. Below, we outline three settings in which this model might underperform.

The model in question heavily relies on ERA5 reanalysis map data and their quality. In certain regions, if this data is not high quality or there is sparse observational data, then the vision transformer may struggle to capture patterns. As a result, suboptimal embeddings are passed through the neural network. Some possible extensions to this would be the use of data imputation techniques to fill in missing values or utilizing unsupervised training on other available data [**?**].

Cyclones in particular are extremely unstable weather phenomena: they produce unusual behavior such as rapid intensification [**?**], weakening from external factors like wind shear, and sudden directional changes. The Temporal Transformers used in this model assume consistent patterns across the time steps. In cases of rare or extreme events deviating from these learned patterns, the predictions will be off. These types of prediction errors could be accounted for by using auxiliary environmental features such as real-time wind shear, ocean heat content, and atmospheric stability indices to gather more context [**?**].

Another limitation of the model involves predicting the next time step: news agencies and other platforms only provide longer-term cyclone trajectories unlike the 3 hour intervals for each of our time steps. This is especially useful for cyclones nearing landmasses and is due to the used dataset being captured in 3 hour intervals across the 24 hours. Future works that address this should consider looking at datasets with longer intervals to use alongside this to make both small time step predictions and larger ones. A model architecture also fitting for this would be Transformers with a memory unit such as LSTMs (Long Short-Term Memory) or Transformers with attention mechanisms [**?**].

8 Ethical Considerations

To ensure the safety of as many people as possible, users of this model should publicize their results so that others are aware of potentially dangerous storms. However, consider that local vendors (and insurance companies) could misuse these results by increasing the costs of their goods/services beforehand and discriminating towards tropical cyclone (TC) refugees.

The model must make consistently accurate predictions. Otherwise, poor model performance could lead to public mistrust, so people located in TC risk areas could choose not to evacuate while believing that no TC exists or the storm is not dangerous enough, even if the model correctly predicts otherwise.

Note that the model trains on data about past occurrences of TCs to make a prediction for a new TC, so the contents of the datasets could impose bias in the model. If the model is built to predict a TC in one area, but it is trained on data mainly from TCs in other areas, then the model will produce biased and inaccurate predictions. Despite this data incompatibility example being avoidable, there are cases where bias is not as clear as this example.

9 Conclusion

In conclusion, the model created to refine forecasting accuracy using a vision transformer and deep learning techniques is viable for forecasting accuracy in certain conditions. The final prediction error rate is competitive with models associated with other research papers. This is an area that should be explored in further depth by future research as models of these types may pick up on more localized cyclone conditions. These efforts could lead to more precise and context-aware forecasting systems for the future, ultimately benefiting disaster preparedness and response efforts for those most affected by cyclones.

Appendix

Google Colab: Here
Data: Here

Table 3: Key Features of IBTrACS Point Shapefile Data

Feature	Description
`BASIN`	Ocean basin where the cyclone is located, these are values from the list [NA, EP, NI, WP, SP, SI, SA].
`LAT` and `LON`	Geographical coordinates marking the location of each cyclone.
`ISO_TIME`	Timestamp indicating the specific date and time of each cyclone observation in UTC with format YYYY-MM-DD HH:mm:ss.
`USA_WIND`	Wind speed (knots) recorded in the United States dataset.
`USA_PRES`	Atmospheric pressure (hPa) recorded in the United States dataset.
`USA_STATUS`	Cyclone status according to the U.S. categorization.
`NATURE`	Nature of the system, such as tropical or extratropical.
`STORM_SPD`	Translation speed of the cyclone (knots), indicating its movement speed.
`STORM_DIR`	Translation direction (degrees), showing the direction the cyclone is moving.
`SID`	Unique storm identifier that links points across the dataset.

Table 4: Key Features of ERA5 Reanalysis Maps

Feature	Description
`relative_humidity`	Relative humidity at specific pressure levels.
`temperature`	Atmospheric temperature in Kelvin at specific pressure levels, representing the thermal state of the air.
`u_component_of_wind`	Eastward (U) component of wind.
`v_component_of_wind`	Northward (V) component of wind.
`vorticity`	Measure of the local rotation in the atmosphere.
`geopotential`	Height of specific pressure surfaces.

References

[1] Shannon Doocy, Anna Dick, Amy Daniels, and Thomas D Kirsch. The human impact of tropical cyclones: a historical review of events 1980-2009 and systematic literature review. *PLoS currents*, 5, 2013.

[2] 2023 Atlantic Hurricane Season - Center for Disaster Philanthropy — disasterphilanthropy.org. [Accessed 10-11-2024].

[3] 2024 Atlantic Hurricane Season - Center for Disaster Philanthropy — disasterphilanthropy.org. [Accessed 10-11-2024].

[4] Léonard Boussioux, Cynthia Zeng, Théo Guénais, and Dimitris Bertsimas. Hurricane forecasting: A novel multimodal machine learning framework. *Weather and forecasting*, 37(6):817–831, 2022.

[5] Sophie Giffard-Roisin, Mo Yang, Guillaume Charpiat, Christina Kumler Bonfanti, Balázs Kégl, and Claire Monteleoni. Tropical cyclone track forecasting using fused deep learning from aligned reanalysis data. *Frontiers in Big Data*, 3, 2020.

[6] Kerry Emanuel. Will global warming make hurricane forecasting more difficult? *Bulletin of the American Meteorological Society*, 98:495 – 501, 2017.

[7] Xia Sun, Lian Xie, Shahil Umeshkumar Shah, and Xipeng Shen. A machine learning based ensemble forecasting optimization algorithm for preseason prediction of atlantic hurricane activity. *Atmosphere*, 12(4):522, 2021.

[8] Xinhe Liu and Wenmin Wang. Deep time series forecasting models: A comprehensive survey. *Mathematics*, 12(10):1504, 2024.