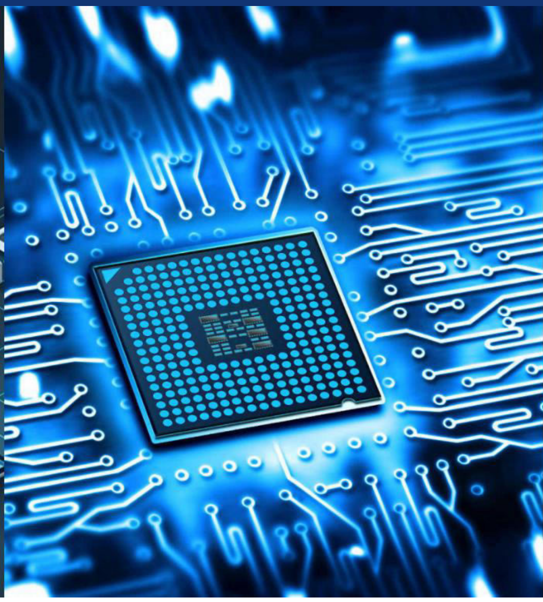
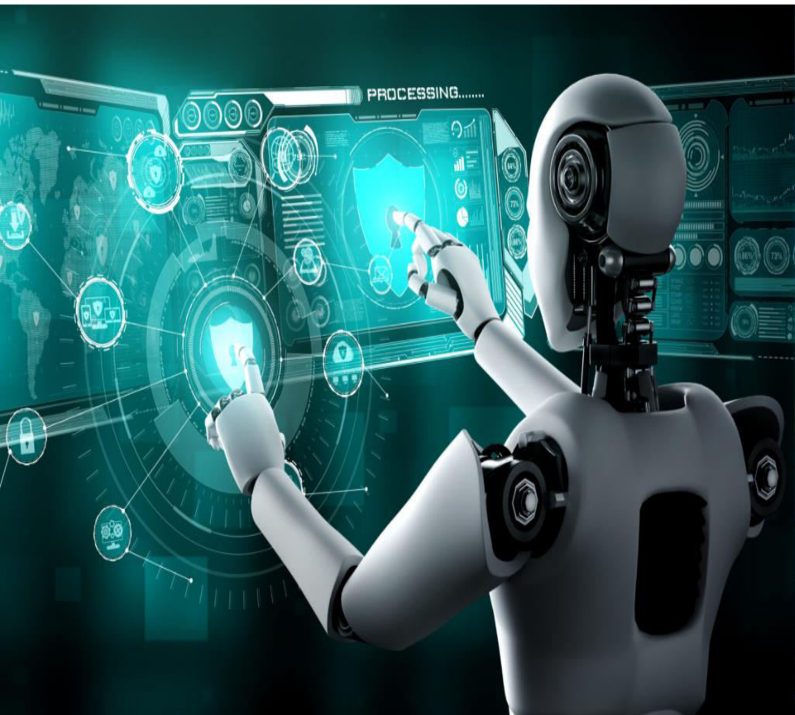
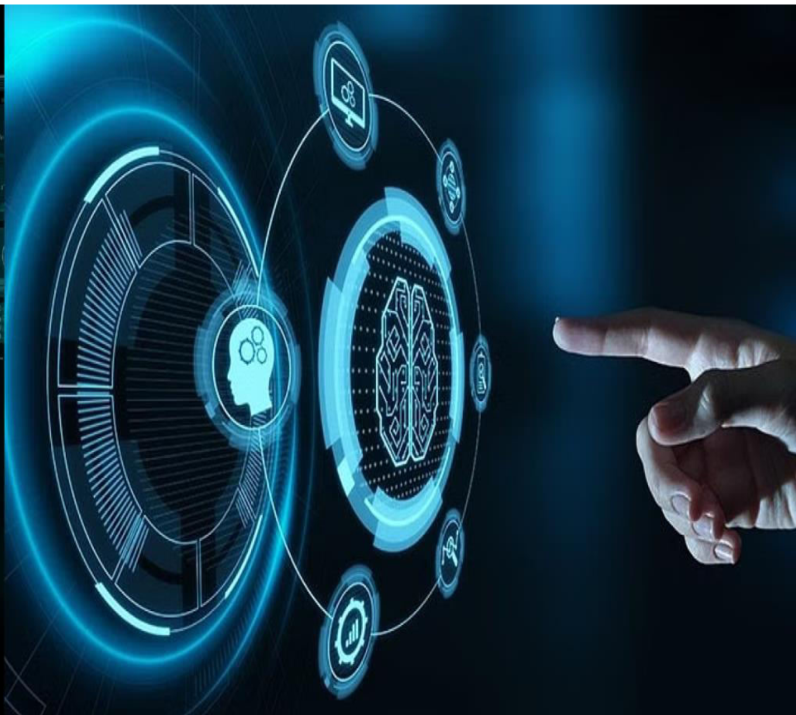
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.625

Volume 13, Issue 1, January 2025

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301104

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Typhoon Prediction System using Historical Data and Live Satellite Imagery

Nitya, Dr. R. Ruhin Kauser

Department of Computer Science and Engineering, Presidency University, Bengaluru, India

Assistant Professor, School of Computer Science and Engineering, Presidency University, Bengaluru, India

ABSTRACT: Typhoons are catastrophic weather events that cause severe damage to lives, infrastructure, and economies, especially in coastal and tropical regions. Timely and accurate forecasting of typhoon behavior, including their formation, intensity, trajectory, and landfall zones, is essential for disaster preparedness and response. This study introduces a Typhoon Prediction System that integrates historical meteorological data with real-time satellite imagery to achieve highly accurate predictions. By leveraging advanced machine learning algorithms and image processing techniques, the system addresses the limitations of traditional forecasting methods and provides a dynamic, real-time solution. The proposed system utilizes critical meteorological variables such as wind speed, atmospheric pressure, and cloud formation patterns from historical datasets, combined with high-resolution satellite imagery from geostationary satellites like Himawari-8. The framework incorporates preprocessing and feature extraction to harmonize diverse data sources. Sequential models like Long Short-Term Memory (LSTM) networks analyze time-series data, while Convolutional Neural Networks (CNNs) process satellite imagery to detect cloud density, wind field structures, and other typhoon-related patterns. The model achieves real-time calibration using live data inputs, enabling adaptive forecasting capabilities that improve accuracy and responsiveness. By offering actionable insights, the system strengthens disaster management strategies, aiding governments and relief organizations in minimizing loss of life and property. This study addresses challenges in data integration and scalability while aligning with global sustainability goals to enhance climate resilience. The proposed solution demonstrates significant potential to mitigate the impacts of typhoons, laying a foundation for future advancements in extreme weather forecasting.

KEYWORDS: Typhoon prediction, machine learning, satellite imagery, real-time forecasting, LSTM, CNN, meteorological data, disaster management, climate resilience, data integration.

I. INTRODUCION

Typhoons, also known as tropical cyclones or hurricanes, are some of the most destructive natural phenomena on Earth, causing widespread devastation to life, property, and infrastructure. These storms, which form in tropical and subtropical regions, are characterized by their extreme wind speeds, heavy rainfall, and intense storm surges, leading to widespread flooding, damage to buildings and infrastructure, and loss of life. The severe impact of these storms on communities, economies, and ecosystems makes accurate and timely forecasting critical for minimizing casualties and damage. Historically, typhoon prediction systems have relied on meteorological data and weather patterns observed over time to predict storm trajectories, intensities, and potential landfall zones. While these methods have been somewhat effective, they face significant limitations in terms of real-time adaptability and accuracy in forecasting typhoon behavior.

Traditional forecasting techniques generally fall into two categories: statistical models and physical models. Statistical models use historical data, such as past typhoon tracks, intensities, and associated meteorological variables, to build predictions based on observed patterns. These models are relatively simple and computationally efficient but fail to account for real-time variations in weather patterns, leading to lower accuracy in dynamic environments. Physical models, on the other hand, use mathematical equations to simulate the physical processes governing typhoon development and movement. While these models are more sophisticated, they require a vast amount of computational power and still face limitations in providing real-time updates during the rapid evolution of storms. As a result, both types of forecasting systems struggle to offer precise predictions that account for the unpredictable nature of modern typhoons. To address these limitations, this study explores the integration of historical data with real-time satellite

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

imagery and machine learning techniques. The combination of these data sources provides a more comprehensive and dynamic view of the storm system, enabling more accurate predictions. By analyzing both historical meteorological parameters and live satellite imagery, the proposed approach incorporates the strengths of both static and dynamic data, enabling real-time adaptability and improving prediction accuracy. This study aims to create a highly accurate and adaptive typhoon prediction system that can forecast not only the occurrence of a typhoon but also its potential trajectory, intensity, and landfall zones. The use of machine learning models, such as Random Forest Regressors and advanced image processing algorithms, plays a central role in the success of the proposed system. These techniques allow for the modeling of complex relationships within the data, enabling the system to predict storm behavior based on multiple variables, including wind speed, atmospheric pressure, cloud density, and sea surface temperature.

The proposed methodology begins by analyzing historical meteorological data, which includes past records of typhoons, their intensities, and trajectories. This historical data serves as the foundation for training machine learning models that can predict future typhoon events based on observed patterns. By incorporating key meteorological variables, such as wind speed, pressure, and temperature, the system can predict the likelihood of a typhoon forming, as well as its potential path and intensity. The historical data is also used to train models that can identify potential landfall zones, based on past storm behavior and geographic factors. Once the machine learning models are trained using historical data, the system moves to the second stage: processing real-time satellite images. Satellite imagery provides a valuable source of information about current weather conditions, including cloud formations, sea surface temperatures, and other factors critical to storm prediction.

The real-time satellite images are processed using advanced image processing algorithms to detect and analyze key weather patterns associated with typhoon formation and intensification. For example, cloud density and structure are crucial indicators of storm development, and variations in sea surface temperature can influence the intensity and path of a typhoon. Image processing algorithms are used to extract these features from satellite images, enabling the system to detect evolving weather patterns and track the movement of the storm.

Once the data from both historical meteorological records and live satellite imagery is processed, the system integrates these data sources to dynamically update forecasts. The integration of these two data types allows for a more comprehensive understanding of the storm's behavior and provides more precise predictions about its intensity, trajectory, and landfall. Machine learning models, including Random Forest Regressors and Long Short-Term Memory (LSTM) networks, are used to generate predictions based on the integrated data. Random Forest Regressors are a type of ensemble machine learning model that works by constructing multiple decision trees and combining their outputs to make predictions. This approach helps capture complex, nonlinear relationships in the data, which is crucial for accurately forecasting typhoon behavior. LSTM networks, a type of recurrent neural network (RNN), are particularly suited for time-series data and are employed to analyze the temporal patterns in the historical data. LSTMs can capture long-term dependencies in the data, enabling the system to predict future typhoon behavior based on past storm dynamics. In addition to machine learning models, advanced data augmentation techniques are employed to improve model performance. Data augmentation is a process of artificially increasing the size of the training dataset by introducing variations, such as rotating or flipping satellite images or adding noise to meteorological data. This helps the models generalize better and reduces the risk of overfitting to specific patterns in the training data. By incorporating data augmentation, the system can improve its prediction accuracy under varying environmental conditions, ensuring robustness and reliability in real-world applications.

II. METHODOLOGY

Data Gathering The data gathering process for this project involves collecting both historical meteorological data and satellite imagery, as well as real-time data streams. Historical meteorological data includes parameters such as wind speed, atmospheric pressure, temperature, and storm trajectories, sourced from reliable meteorological databases. Additionally, a collection of satellite historical data consisting of typhoon images was obtained. These images provide valuable visual insights into past typhoon formations, cloud patterns, and trajectories, serving as an essential resource for training machine learning models.

Real-time satellite imagery is also gathered from live feeds, capturing ongoing atmospheric changes. Together, these datasets offer a comprehensive understanding of typhoon behaviors over time and in real-time contexts. All data is

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

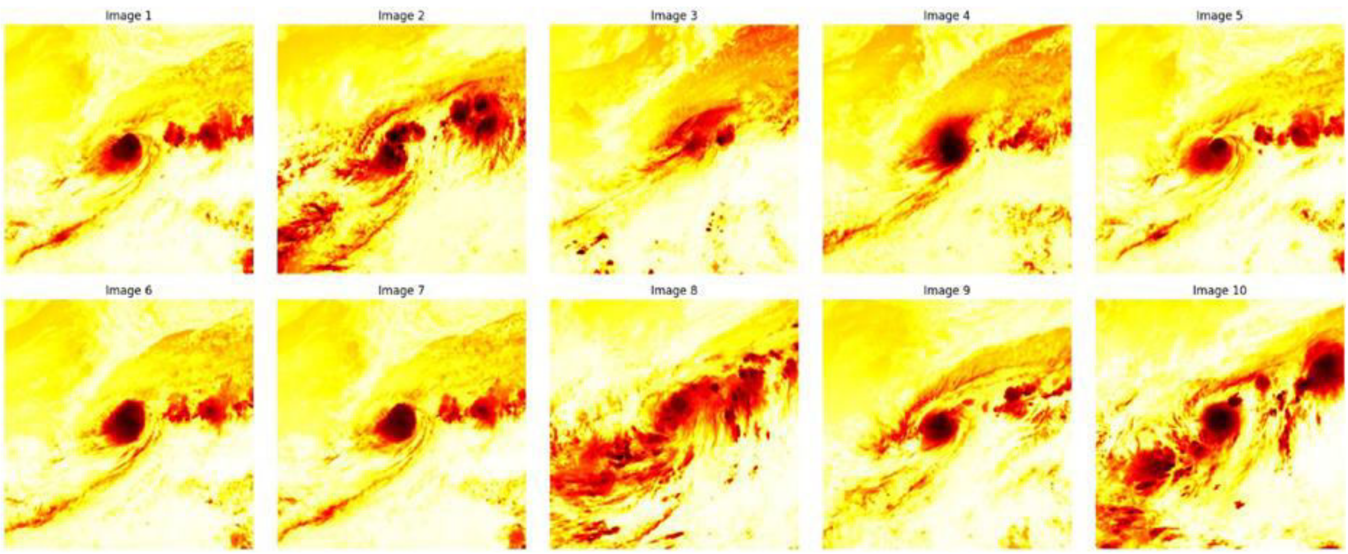
meticulously validated and preprocessed to address issues such as noise, missing values, and inconsistent formats, ensuring high-quality inputs for exploratory analysis and predictive modeling. By integrating historical satellite images with current meteorological data, the project enhances its capability to predict typhoon events dynamically and accurately.The other documents published by The Ministry of Economic Affairs[1] are Introductory Notes and Heat Content of Energy Products, both of which documents have data related to the fuel sources and the units in the Yearly Energy Table. With this, we also gathered data such as Abbreviation and Equivalents of Energy Units for Taiwan, Taiwan Population and also Yearbook 2021 of Taiwan. All of these resources further helped us to visualize and understand the data in the Yearly Energy Table.

The Taiwan Yearly Energy Table has been organized into multiple sheets corresponding to different years, as depicted above. The diverse fuel categories are depicted in the columns, while the rows show the sectors that employ them. From figure 2, we can see that the fuel sources are in their original units. As said earlier, different versions of the table are available with different units.

After we gathered data for Taiwan, we step up to gather data for Taichung. First, we start with the industries present in Taichung.
These historical IR satellite images capture the intricate patterns and dynamics of past typhoons, providing critical insights into their development and behavior. The images showcase variations in cloud density, temperature gradients, and storm intensity, which are essential for understanding the lifecycle of these powerful weather systems. Infrared imaging is particularly valuable for detecting heat signatures, allowing meteorologists to study typhoon structures both during the day and night.

The swirling cloud bands, highlighted in these IR visuals, reveal key features like the storm's eye, convective activity, and areas of intense precipitation. By analyzing such images, researchers can identify trends in storm formation, track their trajectories, and estimate wind speeds.
These IR images are integral to developing machine learning models for typhoon prediction, as they provide a rich dataset for training algorithms to recognize patterns associated with storm intensification and movement.

The historical perspective they offer is crucial for improving future forecasting accuracy, making them a cornerstone of modern meteorological studies. While prioritizing inclusivity in our data acquisition efforts, we understand the importance of focused analysis. Following this comprehensive data gathering phase, we will carefully curate and refine the data, retaining only the most relevant elements for our investigation.



Data Organizing The data organization process in this project is a critical step to ensure that the collected information is structured, consistent, and ready for analysis. The dataset comprises two main components: historical meteorological

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

data and satellite imagery. Historical data includes variables such as wind speed, atmospheric pressure, temperature, and storm trajectories, while satellite images capture the visual patterns of typhoons.

To streamline analysis, the data is categorized into distinct groups based on relevance and type. For instance, meteorological data is organized chronologically to facilitate time-series analysis, whereas satellite images are labeled according to typhoon events and their development stages. Metadata, such as timestamps and geographical coordinates, is attached to each entry to maintain context and traceability.

Data cleaning and preprocessing are integral to this phase, addressing missing values, removing redundancies, and standardizing formats for compatibility across various machine learning models. Additionally, satellite imagery is further processed to extract meaningful features like cloud density and storm patterns. This systematic organization ensures seamless integration of diverse data sources, enabling efficient exploratory analysis, modeling, and validation. By evaluating feature importance, the graph highlights the significance of wind speed as a predictive variable for typhoon intensity and trajectory, while atmospheric pressure also plays a crucial complementary role. These insights enable the optimization of the model by focusing on the most relevant parameters, enhancing prediction accuracy and computational efficiency. The comparative visualization underscores the critical role of data-driven feature selection in constructing robust forecasting models.

Data Analyzing and Visualization In this project, the integration of historical meteorological data and satellite imagery provides a comprehensive approach to understanding and forecasting typhoon patterns. Additionally, a novel aspect of this study involves the integration of historical satellite images with live satellite images, specifically focusing on cities in Taiwan. This integration enables a dynamic and detailed analysis of typhoon impacts and urban resilience. This section outlines the processes involved in data analysis, from preprocessing to visualization, to support accurate typhoon predictions.

Year	City	Windspeed(kph)	Rainfall(mm)	Humidity(%)
1960	Hualien	10.993	2.308	66.476
1960	Hualien	13.046	1.829	57.658
1960	Hualien	13.158	6.837	55.305
1960	Hualien	11.085	0.682	55.342
1960	Hualien	10.483	0.000	42.750

Figure 1: preprocessing the data[4].

Before proceeding with analysis, the data must undergo extensive preprocessing to ensure its quality and consistency. Meteorological data is cleaned to address missing values, outliers, and inconsistencies that could affect the accuracy of predictions. In the case of satellite imagery, preprocessing steps include the alignment of images with meteorological data, as well as adjustments for factors like cloud cover removal and resolution enhancement. The integration of live satellite images with historical data required additional preprocessing, such as matching current urban layouts with past event data to accurately analyze the impact on urban areas.

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301104

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

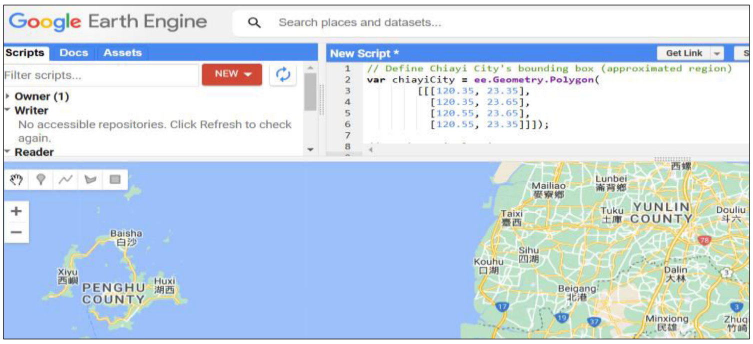


Figure 2: retrieval of real-time satellite images for Taiwanese cities from Google Earth Engine

Data Collection Using Google Earth Engine To gather the necessary satellite imagery, the project employed Google Earth Engine (GEE), a powerful platform for processing geospatial data. GEE's extensive library of satellite datasets and its cloud-based computational capabilities made it an ideal choice for this work. The data collection process involved multiple steps to ensure accuracy and relevance:

Defining the Region of Interest (ROI): Each city's geographic boundaries were precisely outlined using polygonal coordinates. Cities such as Taipei, Kaohsiung, Tainan, and Taichung were selected for their diverse urban structures and geographical settings. The ROI for each city ensured that the analysis focused exclusively on the areas of interest, minimizing unnecessary data processing.

Selecting Appropriate Data Sources:

- Sentinel-2 Imagery: Known for its high resolution (10 meters) and short revisit period, Sentinel-2 data was used for detailed urban and geographical analysis.
- MODIS Datasets: These were incorporated to capture broader atmospheric conditions, such as cloud cover and temperature variations, complementing the high-resolution imagery.

Filtering and Preprocessing:

- Satellite images were filtered by date to retrieve real-time data reflecting current conditions.
- Preprocessing steps, including cloud masking, atmospheric correction, and cropping, were performed within GEE. This ensured that the imagery was clean and focused solely on relevant features.

Observations from Satellite Imagery The analysis of the collected imagery revealed distinct patterns and characteristics across Taiwan’s cities. These insights were categorized into three primary areas: Urban Layouts: Taipei, as the capital city, showcased a dense and highly developed urban core, with large clusters of highrise buildings and extensive transportation networks.

Smaller cities such as Tainan displayed more dispersed urban layouts, with significant areas dedicated to agricultural or undeveloped land. Geographical Features: Coastal cities, including Kaohsiung, exhibited proximity to natural harbors and waterways, emphasizing their importance as economic and trade hubs. Inland cities like Taichung were characterized by mountainous terrain and green spaces, providing a stark contrast to the urban sprawl of coastal areas.

Real-time weather data can be used to issue timely alerts, mitigating potential damage and loss of life. Urban Development: By analyzing land use patterns, policymakers can make informed decisions on infrastructure development, zoning, and environmental conservation. Climate Resilience:
The integration of meteorological and satellite data contributes to building climate-resilient cities, capable of adapting to extreme weather events. Some of the Advantages of Using Google Earth Engine The use of Google Earth Engine offered several benefits that streamlined the project:
Comprehensive Data Access: GEE provides access to a vast array of datasets, including both historical and real-time satellite imagery, ensuring data diversity and reliability.

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301104

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Cloud-Based Processing: The platform's cloud infrastructure eliminated the need for local computational resources, enabling efficient handling of large datasets.
Real-Time Capabilities: The ability to retrieve and analyze live satellite imagery allowed for up-to-date observations, critical for real-time applications such as typhoon prediction.
Scalability and Customization: GEE's scripting environment enabled the creation of customized workflows, allowing for the seamless integration of data processing, analysis, and visualization. Key Features Observed: Through this project, several unique characteristics of Taiwanese cities were highlighted: Dense Urbanization: Major cities like Taipei and Kaohsiung demonstrated significant urban sprawl, reflecting their economic and cultural significance. Geographical Diversity: The contrast between coastal and inland cities showcased Taiwan's varied topography, from flat coastal plains to rugged mountainous regions. Dynamic Weather Patterns: Real-time observations of atmospheric conditions provided insights into local weather systems and their potential impacts.

Figure 3: Real-Time Satellite Images of Taiwanese Cities

Prediction Using Real-Time Satellite Images with XGBoost The ability to predict the occurrence of typhoons using satellite imagery is a crucial component of early warning systems for disaster management. In this project, we have utilized real-time satellite images and a machine learning model trained using XGBoost to identify potential typhoon events. Below is a detailed summary of the prediction pipeline.

The data Collection and Real-Time Image Processing The first step involves collecting real-time satellite images from various sources, including weather monitoring satellites and satellite imagery APIs. These images provide visual data

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

about cloud formations, atmospheric disturbances, and weather patterns that are indicative of typhoon development. The satellite images used in this study are typically in infrared or visible spectrum formats, which capture temperature variations, cloud structures, and other relevant features.

Preprocessing:

- Resizing: The images are resized to a consistent size, such as 512x512 pixels, to match the input requirements of the prediction model. This ensures that all images are processed uniformly, regardless of their original resolution.
- Normalization: The pixel values are scaled to a range of [0, 1], ensuring that the model receives input data on a consistent scale, improving the accuracy and efficiency of the predictions.
- Feature Extraction: Specific features are extracted from the images, such as patterns of cloud coverage, temperature gradients, and atmospheric anomalies. These features are crucial for identifying patterns associated with typhoon events.

XGBoost Model Training The XGBoost model was trained using a large dataset of historical satellite images of both typhoon and non-typhoon events. This dataset was carefully curated to include various meteorological features and satellite image characteristics that could differentiate typhoons from other weather phenomena.

Training Process:

- Feature Engineering: The historical satellite images were processed to extract various features such as cloud movement, shape, and intensity. These features were then used as input to train the XGBoost model.
- Model Configuration: Hyperparameters such as learning rate, maximum depth, and number of estimators were tuned to optimize the model's performance. The XGBoost algorithm, being a gradient boosting technique, is well-suited for this task because of its ability to handle complex and non-linear relationships in the data.
- Cross-Validation: Cross-validation techniques were employed to validate the model's generalization ability. The dataset was split into training and validation sets to evaluate the model's performance and prevent overfitting.

Making Predictions with Real-Time Data After training, the XGBoost model is used to predict typhoons in new, real-time satellite images. The images are processed and passed through the model, which generates a prediction on whether the image corresponds to a typhoon or not.

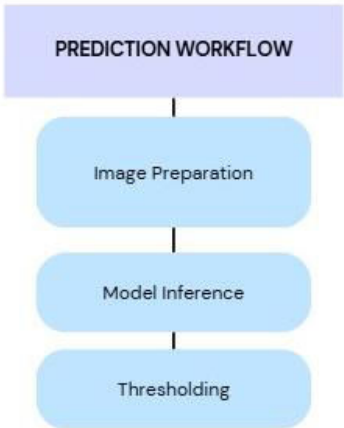


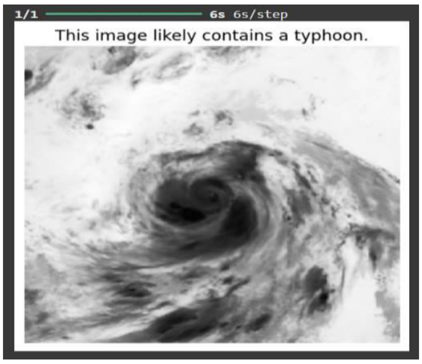
Image Preparation: For each new real-time satellite image, preprocessing steps such as resizing and normalization are performed to align with the input specifications of the XGBoost model. Model Inference: The processed image is passed through the trained XGBoost model, which outputs a probability score indicating the likelihood that the image contains a typhoon. Thresholding: A threshold value, typically set at 0.5, is applied to the model's output. If the probability score is greater than or equal to the threshold, the image is classified as a typhoon event; otherwise, it is classified as non-typhoon.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The predictions made by the model are accompanied by visual representations of the satellite images to aid in interpretation. After the model generates a prediction, the corresponding satellite image is displayed with a title that indicates the predicted event: If the model predicts a typhoon, the title might read, “Prediction: Typhoon Event.” If the model predicts no typhoon, the title might read, “Prediction: No Typhoon.” This visual feedback allows users to quickly assess the outcome and helps validate the model’s prediction by visually inspecting the satellite image in question.



Real-Time Monitoring and Alert System for Typhoon Detection The integration of real-time satellite imagery with machine learning models facilitates continuous monitoring and early detection of typhoons, which is vital for disaster management and public safety. By leveraging this system, authorities can predict the formation of typhoons well in advance, enabling timely responses and minimizing the potential impact on populated regions.

Automated Alerts for Typhoon Detection: The key benefit of this system is the ability to provide automated alerts based on the model’s predictions. Once a typhoon is detected by the model from real-time satellite images, the system triggers an alert mechanism to notify relevant stakeholders, such as disaster management teams, emergency responders, or the general public. This process ensures timely information is delivered for preparation and mitigation efforts, such as evacuations or safety measures. Alerts are generated and sent automatically through various communication channels, including email, SMS, and dashboard notifications. As new satellite data comes in, the system continuously updates predictions, ensuring that the information remains up-to-date and actionable. Twilio Integration for SMS Alerts To ensure effective communication during typhoon emergencies, Twilio API is integrated into the alert system to send SMS notifications. Once a typhoon is detected, the machine learning model triggers the Twilio client to send an SOS alert to recipients. This integration ensures that critical information is quickly disseminated to those at risk.

Alert Message message serves as a vital communication tool to enhance public safety by urging individuals and authorities to act quickly and implement safety protocols. Image Processing Techniques for Typhoon Detection: The following image processing techniques are employed to enhance the detection of typhoons from satellite images,
ensuring accurate identification of key features such as spiral bands and circular patterns, Gaussian Blurring: This technique is applied to smooth the image and reduce noise. By blurring the image, the algorithm can focus on significant structures while filtering out unwanted details. This preprocessing step improves the accuracy of subsequent edge detection, making it easier to identify typhoon-related patterns. Canny Edge Detection: The Canny edge detection algorithm is used to identify the boundaries and edges within the image.

For typhoon detection, this is critical for highlighting the distinct spiral bands that form around the typhoon's center. These bands are key indicators of the typhoon's structure and intensity. Hough Circle Transform: This method detects circular patterns within the image, which are indicative of the spiral band structures commonly associated with typhoons.

The circular shape is an important visual cue in identifying the central eye of the storm, a defining characteristic of typhoons. Grayscale Conversion: Converting the image to grayscale simplifies the image data, reducing it to a single

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301104

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

intensity channel. This simplifies the computational complexity and enhances the efficiency of image analysis while retaining critical features needed for typhoon detection.

	A	B	C	D	E	F	G	H	I	J	K
1	Spiral Band	Eye of the	Cloud Den	Cloud Tem	Rainbands	Wind Field	Symmetry	Color Grac	Brightness	Surroundir	Label
2	0.676431	0.635294	0.796676	0.676431	0.010824	0.592015	0.093834	0	0.22013	0.657746	1
3	0.884072	0.952941	0.982561	0.884072	0.006343	0.347809	0.047792	0	0.108964	0.748208	1
4	0.811975	0.803922	0.924366	0.811975	0.00555	0.759938	0.075385	0	0.182488	0.327953	1
5	0.832473	0.8	0.925223	0.832473	0.008086	0.28432	0.0846	0	0.160974	0.331147	1
6	0.708989	0.047059	0.803253	0.708989	0.01167	0.72533	0.143782	0	0.22035	0.838434	1
7	0.698897	0.427451	0.822824	0.698897	0.008965	0.779361	0.115794	0	0.196709	0.525994	1
8	0.84085	0.894118	0.95727	0.84085	0.005342	0.839381	0.088551	0	0.146201	0.060828	1
9	0.786214	0.909804	0.893794	0.786214	0.011011	0.339926	0.102051	0	0.194687	0.938558	1

Figure 4: Different features on which Typhoon image depends

The CSV file provides a detailed tabular representation of extracted features from satellite images, highlighting critical parameters such as spiral band counts, cloud density, and wind field characteristics. Each row corresponds to an analyzed image, offering insights into typhoon-specific attributes that are vital for understanding storm behavior. These features serve as inputs for training and testing the machine learning model, enabling accurate classification of images as "Typhoon" or "Non-Typhoon."

The organized structure of the dataset not only facilitates robust analysis but also aids in future scalability and application of the system to real-world scenarios in meteorology and disaster management in all the situations.

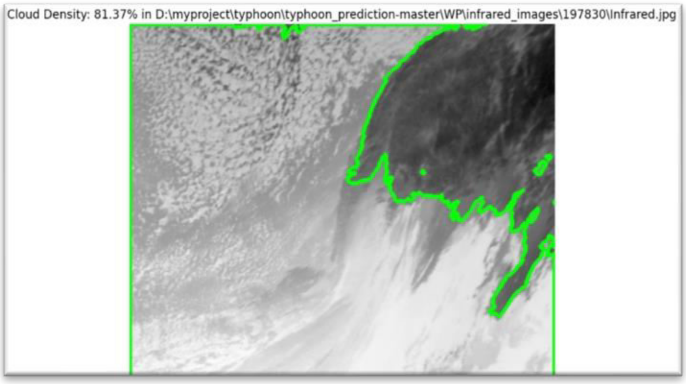
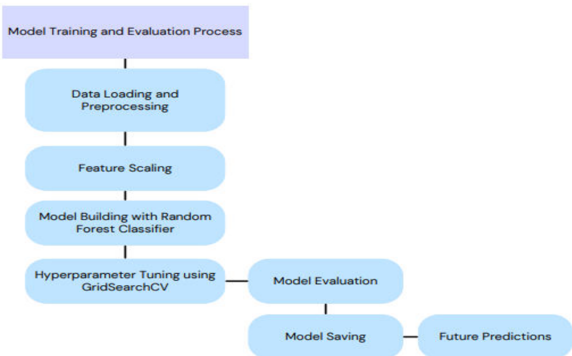


Figure 5: Cloud Density with 81.37% of cloud



www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Data Loading and Preprocessing: The dataset, consisting of features and labels, is loaded from a CSV file using `pandas`. The features (X) represent attributes such as environmental and satellite-based data, while the labels (y) indicate whether the image is of a typhoon (1) or non-typhoon (0). The data is split into training and testing sets using an 80/20 split. The training data is used to build the model, while the testing data serves as an unseen dataset to evaluate model performance.

Feature Scaling: Since Random Forest does not necessarily require feature scaling, the features are standardized using a StandardScaler. This is particularly important for algorithms like neural networks but can also improve model stability and performance in some cases. The scaler is applied to the training set and then used to scale the test set as well. The scaler is saved as a `.pkl` file, ensuring that it can be reused for transforming new data or during future model predictions to maintain consistency in scaling. Model Building with Random Forest Classifier: A Random Forest Classifier is initialized with a fixed random state for reproducibility. The model is an ensemble of decision trees, where each tree is trained on a random subset of data, and predictions are averaged for more stable and accurate results. Hyperparameter tuning plays a crucial role in optimizing machine learning models. For the Random Forest classifier, GridSearchCV provides a systematic approach to fine-tuning hyperparameters by exhaustively searching through a pre-defined parameter grid. This process ensures the model's predictive capabilities are maximized while maintaining generalizability. The model's performance is then rigorously evaluated to ensure that it meets real-world requirements for predicting typhoon events effectively. Hyperparameter Tuning using GridSearchCV. It helps identify the optimal combination of hyperparameters for the Random Forest model. Random Forest is an ensemble learning method known for its robustness, particularly in handling classification problems involving complex datasets. The following hyperparameters were tuned during the grid search:

- GridSearchCV is employed to search through a range of hyperparameters for the Random Forest model to find the best combination. The parameters tuned include:
- n_estimators: The number of trees in the forest (50, 100, or 200).
- max_depth: The maximum depth of each tree, controlling the model's complexity.
- min_samples_split**:** The minimum number of samples required to split an internal node.
- min_samples_leaf: The minimum number of samples required to be at a leaf node.
- Cross-validation (cv=3) is used during the grid search to ensure that the chosen hyperparameters generalize well to unseen data.
- The best combination of hyperparameters is determined through this search, ensuring that the model is not overfitting or underfitting.

III. RESULT AND EVALUATION

Data Pre-processing: Data pre-processing is a crucial step in preparing input data for machine learning models, ensuring that the data aligns with the model's training specifications. In this context, the model receives input satellite images that undergo a series of pre-processing steps to standardize their format, making them suitable for prediction. Image Resizing**:**

Satellite images can vary significantly in their dimensions depending on the source and resolution. To ensure consistency and compatibility with the model, each image is resized to a fixed dimension of 512x512 pixels. This size was chosen as a balance between preserving critical features in the image (such as cloud patterns and structures) and optimizing

Computational efficiency.Rescaling**:** To match the range of pixel values in the model's training data, the image is rescaled. The pixel values, originally in the range of 0 to 255 (typical for 8-bit images), are normalized to a range between 0 and 1**.** Normalization helps the model process the data more effectively, as it reduces potential biases caused by varying pixel intensity scales and ensures that the image data aligns with the input expectations of the trained Random Forest Classifier. Ensuring

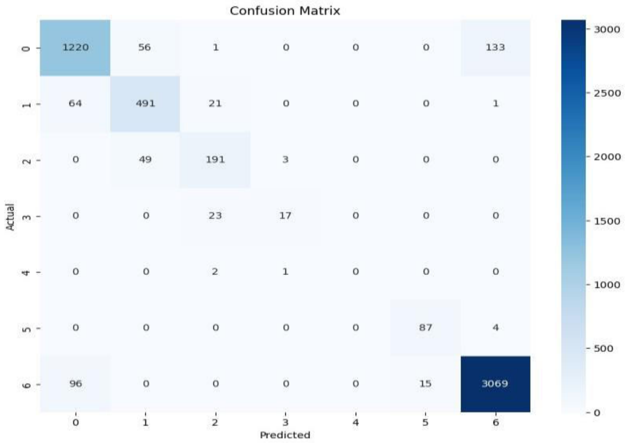
- The image is resized to 512x512 pixels, ensuring a consistent input format.
- The pixel values are scaled to the range of 0 to 1, normalizing the data for the model.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- The pre-processed image is passed through the Random Forest Classifier, which computes a prediction score.
- The score is compared to the threshold of 0.5:
- If the score is 0.7, the model classifies the image as a typhoon (label = 1).
- If the score is 0.3, the model classifies the image as non-typhoon (label = 0).



IV. CONCLUSION

The project aimed to establish a comprehensive and reliable solution for the real-time detection and classification of typhoons using satellite imagery, combining the strengths of advanced machine learning techniques and real-time data processing. By leveraging the capabilities of both deep learning and traditional machine learning models, a robust framework was developed to support disaster management efforts effectively. The system integrated innovative methods such as the Xception deep learning model and the Random Forest machine learning model to ensure high accuracy and efficiency in predicting typhoon events. The amalgamation of computer vision, machine learning, and automated alert systems created a cohesive structure designed to mitigate the impacts of typhoons by enabling early detection and timely response. The Xception model, a sophisticated convolutional neural network (CNN), was selected for its ability to process satellite imagery and identify the intricate features associated with typhoons. The model's architecture, characterized by depthwise separable convolutions, allowed it to excel in detecting the complex visual patterns present in typhoon-related satellite data. Through extensive training on a large dataset of labeled infrared satellite images, the Xception model learned to identify specific typhoon features such as spiraling cloud formations, wind patterns, and other meteorological phenomena. Its state-of-the-art design ensured that the detection process was not only accurate but also efficient, making it highly suitable for the demands of real-time application.

During the development process, several challenges were encountered, particularly in ensuring that the models could handle the diverse quality and characteristics of satellite images. Satellite imagery often varies in resolution, noise levels, and atmospheric conditions, making it essential to develop models that are robust and adaptable. The Xception model, while highly accurate, required extensive data augmentation and additional training to improve its generalization across different scenarios. Similarly, the Random Forest model's reliance on manually extracted features necessitated careful parameter tuning to achieve optimal performance. Despite these challenges, iterative improvements and refinements led to the development of a system that balanced the strengths of deep learning and traditional machine learning approaches, resulting in a high degree of prediction accuracy and reliability. The impact of this project is profound, particularly for regions that are vulnerable to the devastating effects of typhoons. By providing accurate and real-time detection capabilities, the system empowers disaster management authorities to take proactive measures, such as issuing warnings, evacuating at-risk populations, and deploying resources effectively. The project not only demonstrates the potential of machine learning and computer vision in environmental monitoring but also underscores their role in enhancing public safety and disaster preparedness. The integration of predictive modeling with automated alerts further highlights the practical benefits of the system, bridging the gap between technological innovation and real-world application.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

REFERENCES

1. Y.-J. Park, M. Seo, D. Kim, et al., "Long-term typhoon trajectory prediction: A physics-conditioned approach without reanalysis data," OpenReview, 2023.
2. H. Chen, P. Yin, H. Huang, et al., "Typhoon intensity prediction with Vision Transformer," arXiv, 2023
3. H. M. Zahera, M. A. Sherif, and A. Ngonga, "Semantic-based end-to-end learning for typhoon intensity prediction," arXiv, 2020.
4. K. Feng, D. Xi, W. Ma, et al., "Potential paradigm shift in hazard risk management: AI-based weather forecast for tropical cyclone hazards," arXiv, 2024.
5. H.-C. Chen, "An innovative dynamic model for predicting typhoon track deflections over complex terrain," Atmosphere, vol. 15, no. 11, 2022.
6. M. Mainelli, M. DeMaria, L. K. Shay, and G. Goni, "Application of oceanic heat content estimation to operational forecasting of recent Atlantic category 5 hurricanes," Weather and Forecasting, vol. 23, no. 1, pp. 3–20, 2008.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE

SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details