SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I12P123

Volume 11 Issue 12, 245-257, December 2024

Original Article

Binary vs Multi-class with Gaussian Filter on Typhoon Image Classification for Intensity Prediction

S. Jayasree[1], K. R. Ananthapadmanaban[2]

[1&2] Department of Computer Science, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India.

[1]Corresponding Author : jayasreesrm@gmail.com

Received: 28 October 2024 Revised: 30 November 2024 Accepted: 16 December 2024 Published: 30 December 2024

Abstract *- Strong meteorological events Tropical Cyclones (TCs) pose serious risks to coastal ecosystems and communities. Their strength is usually categorized using a variety of metrics, including wind speed, pressure, and rainfall since it directly corresponds with the possibility of damage and fatalities. An accurate classification of TC severity is essential for disaster preparedness, response plans, and mitigation initiatives. Support vector machines (SVM) {function category}, K-Nearest Neighbors (KNN) {lazy category}, Bayesian networks {Bayes category}, Random forests {Ensemble category}, and decision trees {Tree Category} are among the machine learning classifiers whose performances are compared in this study in binary and multi-class configurations by using Gaussian image processing technique. Performance measures, including time complexity, ROC, PRC, accuracy, precision, recall, and F-measure, were examined. The results indicate that Multi-class with SVM and Multi-class with Random Forest classifiers consistently outperform other models across most metrics, achieving the highest accuracy (0.88) and superior ROC (0.97) and PRC (0.94-0.95) scores. However, SVM models exhibited significantly higher time complexity, particularly in the multi-class with SVM.*

Keywords *- Typhoon images, Random forest, KNN, Binary classification, Multi-class classification*

1. Introduction

Tropical cyclones (TCs) are among the most damaging weather systems, with impacts on coastal areas and the environment nearly always catastrophic. Correct identification of the TC intensity is indispensable in disaster preparedness and response situations and in developing relevant control measures to help alleviate such disasters' impacts. Conventional approaches to classifying the intensity involve interpreting atmospheric variables, including wind speed, pressure, and rainfall. However, these methods are highly time-consuming, error-prone, and inappropriate for large-scale analysis – which might be necessary during a high-stakes incident.

The literature shows that current developments in Machine Learning (ML) and image processing can enhance TC intensity classification and reduce dependency on human input. Authors have also used methods like CNN and other forms of deep learning hybrid models, where intensity predictions are made based on views of satellite imagery and numerical model data. For instance, Juhyun Lee et al. demonstrated a better outlook for forecasting TC intensity by applying deep learning in physical models at equal intervals. Mawatwal et al. obtained a high binary classification accuracy by applying CNN based hybrid systems. Nevertheless, further prospects and problems are observed in the further optimization and systematic evaluation of the yielded ML algorithms, especially within the framework of other preprocessing approaches such as Gaussian filtering.

1.1. Research Gap and Problem Definition

Although prior research proves the usefulness of the ML approach in cyclone intensity prediction, previous findings mostly feature a single classification algorithm or deep learning approach with no comprehensive investigation of their combinations with image preprocessing techniques such as Gaussian filters. These filters are important in improving the quality of the image by removing noises, while the specific details are very important for feature extraction on satellite images. Moreover, relatively limited studies have compared binary classification to the multi-class classification approaches. This was done without considering the type of ML algorithms used.

This research fills these voids by analysing the compounded effects of Gaussian filters on different ML algorithms, including SVM, KNN, Naïve Bayes, Random Forest and Decision Trees within binary and multi-class classification problems. The problem addressed is twofold: implementation of the optimized classification concerning precision and speed and the key evaluation of the trade-off between binary and multidimensional classification.

1.2. Novelty and Contributions

Therefore, the present study is unique because it blends Gaussian filtering and ML algorithms to enhance TC intensity classification. In contrast to previous papers that compare primarily existing approaches based on deep learning, the present research considers a wider range of popular ML algorithms to systematically compare their performance in the contexts of binary and multi-class classification tasks. Key contributions include:

- Innovative Integration: This work utilizes Gaussian filtering to preprocess images as an added feature that enhances ordinary ML classification techniques by providing clearer features by eliminating noise.
- Comprehensive Comparison: An analysis of all the models is done by observing the accuracy, precision, recall and time taken to learn and predict some data sets. This kind of systematic comparison is lacking in the current research and hence fills this gap in determining these models' suitability for cyclone intensity classification.
- Practical Impact: The research results have direct practical implications for improving the efficiency of disaster management by providing more accurate and faster intensity estimates of TCs. It can help to upgrade the systems of early warnings, organize the resources more effectively, and eventually save lives and property.

1.3. Comparison with Existing Research

This study differs from other works in using the following approaches and objectives. Hybrid CNNs that targeted the enhancement of the forecasting accuracy were suggested by Juhyun Lee et al., but the integration of the preprocessing technique is not well explained. Of equal importance, the binary classification in Mawatwal et al.'s hybrid models gave a desirable accuracy but modest information on class disparity in a multi-class scenario. However, in this research, Gaussian filters are compared with the classification performance of a wide range of classification algorithms and offer a general overview of a wide classification range, including binary and multi-class.

Moreover, most of the above-mentioned researchers aim to enhance the accuracy of the data, while most do not take an adequate account of the computational time to which much emphasis is placed, especially when traffic occurs in the real-time environment. This is advantageous since it is more than just on checking the effectiveness of the results with respect to the classification accuracy; the time and the resources are also taken into consideration, which is real-world oriented.

1.4. Scope and Limitations

Because of this, this study concentrates on classifying the intensity of the tropical cyclone with the help of satellite imagery. The given results are quite encouraging, though some peculiarities should be mentioned. The data is obtained only for the region around the Indian Ocean, and thus, the results may not apply to other meteorological events or geographical locations. Furthermore, the methods are restricted to cases where information on infrared imagery is unavailable. Further research may generalise these observations by including other data types and testing the applicability of the approaches described in the present study across different climatic conditions.

The primary objectives of this study are:

- As part of assessing how Gaussian filtering has promoted the efficacy of ML algorithms in cyclone intensity classification.
- To provide a structured view of the performance of different binary and multi-class classification methods using various ML models.
- The decision complexity was compared by observing the classification precision with the run time and the model complexity.

In meeting these objectives, this study benefits from enhancing theoretical and empirical knowledge and deploying ML in meteorology and disaster response. This work is organized as follows: Section 2 presents a review of related research, Section 3 describes materials and techniques, Section 4 displays results and analysis, and Section 5 concludes the work.

2. Literature Survey

This current section features newspapers, journals, and articles pertinent to this research project. To forecast the severity of the tropical cyclone, Juhyun Lee and his team built a new model by mixing satellite images with the rest of the numerical model. The models improved the accuracy of the official forecasts, achieving skill score gains of up to 22%, 110%, and 7% for predictions in the 24, 48, and 72 time zones, respectively. These applications showed 62%, 87%, and 50% improvements for the various types of strengthening. This demonstrates that deep learning approaches can greatly increase TC intensity accuracy for both rapid intensification categories.[1] To be more precise, Mawatwal et al. developed a CNN-based hybrid model for computerized cyclone strength prediction. Its primary components are the regression module, the YOLOv3-based cyclone detector, and the binary and multi-class classifiers. The results showed that it performed well in binary classification (0.984; ±0.003), multi-class classification (0.6383, ±1.3), and intensity estimation (RMSE = 16.2; ±0.9) knots. The hybrid approach shows promise in carrying out cyclone assessments, from identification to intensity estimation.[2]

Wang et al. proposed a deep CNN that incorporates physics to control the climates of weak typhoons in satellite infrared photos. Compared to the waves model, the mean distance errors of tropical depression and tropical storm levels have improved to 20.1 km and 19.1 km, respectively, indicating 63.0% and 54.6% improvements. As mentioned, this demonstrates how crucial it is to transform physics-based learning and temporal data into deep learning models to analyze TCs. [3]

So, Nguyen and Kieu proposed that further investigation should be conducted based on many environmental characteristics for the ResNet and UNet model for tropical storm generation. These results indicated that the models performed best when projected lead durations were 12-18 h and were largest, covering more than 50% of the Pacific region despite the absence of exact numbers. This indicates that CNNs might enhance early warning systems by including the appropriate spatial patterns for TC formation. [4]

Therefore, Mu et al. proposed the TC3R model for tropical cyclone rainfall data using a C-band Sentinel-1 SAR image. It is observed that the model has a high level of approximation when looking at the HRR for heavy rain with a bias of -0.69 m/s, an RMSE of 4.08 m/s and R = 0.91 when compared to the actual SFMR values. It calculated the root mean square error between the assessed wind speed and the SMAP measured wind speed as 3.78 m/s. As stated above, these findings demonstrate that deep learning can extract rainfall attributes, including areas affected by TCs, from SAR data.[5] TCIP-Net Cameron Tian et al. provided TC intensity prediction where TCIP-Net seeks specific convective structure information from the infrared satellite data. While the model specifies the Hovmöller diagrams for spatiotemporal dynamics and a subnetwork for acquiring information regarding asymmetric TC construction to predict TC intensity, both these features can be argued to remain relatively innovative. [6] Kitamoto et al. collected the Digital Typhoon dataset of over 40 years of satellite images to validate machine learning benchmarks of consistent spatiotemporal data for tropical cyclones. This dataset is fairly helpful for comparing the performance ratings of the TC analysis models presented in this paper, even though the authors did not give exact quantitative values.[7] Fu et al. have developed an ST Synthesis CNN to estimate the intensity of Typhoons using remote sensing pictures from different sources. Compared to the progressive Dvorak technique, in which RMSE amounted to 12.59 kt, the prognosticating model performed better with an RMSE of 8.89 kt and changed by 29.7%. This significant improvement illustrates how deep learning can enhance traditional approaches towards predicting TC intensity.[8]

Griffin et al. developed two machine learning approaches, D-MINT and D-PRINT, to forecast current and short-term changes in typhoon intensity worldwide. Even when no specific quantification was attempted, the models fared higher in certain aspects of operational instructions for rapid intensification prediction. [9]

Roy et al. used a physical process model to predict the typhoon intensity in the BoB for both 12 and 24 forecast hours. Finally, the system demonstrated the potential of the introduced biologically inspired algorithms in predicting TC intensity by achieving over 90 % accuracy on the held-out data. Zhang et al. introduced SiamTCNet, a typhoon tracking model using DL blended with infrared data. As for TC tracking, even though we did not quantify the longitudinal evolution of the tracking-storm features, this is the first time tracking has incorporated spatiotemporal evolution features. Song et al. proposed MTCIE, a kind of CNN for the typhoon strength probabilistic estimation, using a range of source pictures. [12] Zhao and his colleagues proposed a multi-task learning framework to recognize and estimate typhoon intensity from the pictures of the newly launched FY-4A geostationary satellite. Using the above-defined metrics, the Multi-task learning model demonstrated the possibility of adopting TC analysis by having an overall deviation of 9.50knots as well as an F1 score of 0.64. [13]. Based on the existing Xception network, Ma et al. have developed a typhoon intensity estimation method by applying different advanced methodologies.

The improvements in the TC intensity estimation were best presented by the model’s extreme cyclone wind speed estimation deviations, which were 0.08 and 0.11 less than the models in three different datasets. [14] Raynaud et al. proposed a U-Net-based model to identify the TC wind structure in AROME model outputs. Thus, obtaining a normal intersection-over-union measure of about 8%, the model demonstrated that semantic segmentation techniques could be helpful in TC wind field analysis. According to Jung et al., a DL-based approach to estimating typhoon intensity was proposed, and the Swin Transformer model was utilized to classify typhoon images. The method presented how transfer learning can be effectively used when adjusting to new satellite missions. It cut down the RMSE by 20 per cent more efficiently than relying solely on the latest satellite data and 5.5 per cent for the scientific, operational model. [16]

Cui et al. characterized an understandable machine learning procedure to predict and understand spatial structures of amplitude of typhoon induced sea surface temperature cooling. They, however, did not give specific metrics for their experiments. However, the model that uses XGBoost with SHAP values for interpretation holds much potential towards understanding interactions between TC and the ocean. [17]

To this end, Tian et al. proposed to develop an effective approach known as AWL-Net, which integrates 3D convolutional architectures with adaptive weights learning. Short-term intensity prediction was relatively sound, with the model’s average error of 10.62 kt in a 24-hour typhoon intensity forecast. [18]

Pal et al. proposed Small Skip Net (SSN), a lightweight CNN architecture for classifying TC satellite images. The architectures can also do well in TC classification tasks, from which, based on the final testing, the classification rate of our model reaches 92.35%.

Sharoni et al. developed an artificial neural network technique for determining wind speed from the microwave altimeter data in TC conditions. The simultaneous retrieval of surface properties in the GPR approach showed how machine learning could improve remote sensing data for severe weather conditions and how the method delivered the most precise wind speed estimations up to 35 m/s. [20]

Another paper by Bharathi et al. described using another discriminative model, Support Vector Machines, for typhoon prediction and other kernels, which are polynomial, linear, and radial basis functions. Despite the lack of quantitative metrics, the use of the method shows the sustained importance of traditional machine learning approaches in TC prediction.

He et al. developed deep learning models, and U-Net versions were used. Conditional on the model type, threat scores increased by 12.8% to 22.9% as outcomes confirmed the models' enhancements compared to the GFS model. Using deep learning to improve the numerical prediction of rain related to TC, as expounded in this, demonstrates how it can complement existing results. [22] To improve the efficiency, Tian et al. introduced Easy-RP-R-CNN, which is a convolutional-based cyclone detection method.

Of course, it is not easy to accurately quantify various ROI subgroups that could not have been categorized in such a way through any other method, and that makes an important part of the model's approach to cyclone detection a creative solution to identification. [23] Typhoon strength was predicted using the enhanced cross CNN-Bi-LSTM model designed by Alijoyo et al. As mentioned before, operating settings should assess such high accuracies well, and it was excellent that the overall accuracy of the system was 99.4%.

Liu et al. proposed TCRainNet, a rainfall nowcasting model for TCs. The deviations of the model's nowcasts were below 2.6 mm, and relative measures were about 0.27 for the probability of detection and 0.20 for the critical success index. These have demonstrated that the model can estimate short-term rainfall for TCs with modest accuracy. [25]

In that study, Rahman et al. attached three deep learning systems, namely the GRU, LSTM and RNN, to a cyclone path prediction experiment. LSTM did the best of the three, even if they had no specific measures, making it clear how valuable long-term memory is when studying how TCs evolve over time. [26]

3. Materials and Methods

3.1. Dataset Description

The satellite images utilized in this study were obtained from the Kaggle database using the INSAT3D Infrared Cyclone Imagery (2012-2021) data set[27]. In this study, we have 136 infrared images and 140 raw images of TCs that took place over the Indian Ocean between 2012 and 2021.

3.2. Image Preprocessing

To ensure consistency and usability for machine learning models, the following preprocessing steps were applied:

- Image Resizing: All the images were then scaled to 224 x 224 pixels to ensure uniformity in inputs over the dataset.
- Grayscale Conversion: Some images were then transformed to a gray scale to minimize the effects of intensity variation unimportant to cyclone classification and to minimize the computational load.
- Normalization: Pixel values were scaled to the range of [0 – 1], and the images were made uniformly to facilitate optimal model convergence during training.
- Gaussian Filtering: A Gaussian filter was then used for image smoothing with less effect on image noise and edges, which are important in classification. The 2D Gaussian function: $G(x1,y1) = (1/(2\pi\sigma^2)) * \exp(-(x1^2+y1^2)/(2\sigma^2))$, where (x1,y1) is the pixel position and σ is the standard deviation.
- Data Augmentation: Thus, using the flipping, rotation (90∘, 180∘, 270∘) and cropping techniques, the dataset was enlarged to about 3,000 new images. This was made to reduce overfitting and increase the model's generalizability.

Akin to feature extraction, these preprocessing steps guaranteed high input data quality in the next classification stages.

3.3. Criteria for Choosing Machine Learning Classifiers

Five machine learning classifiers were chosen for this study:

SVM, KNN, Naive Bayes, Random Forest, and Dec Training Set: The Training set is used in the training analysis to determine the condition to deliver for each input data. The selection was based on the following criteria:

Model Diversity: The highlighted algorithms cover different ML categories:

- SVM: Performs well on high-dimensional data and is relatively immune from over-fitting when the right kernel is chosen.
- KNN: An easy-to-understand and applicable method that is effective for low-dimensional problems and delivers results without parameter estimation.
- Naive Bayes: Its applicability can, therefore, be analyzed as effective for data calibrated in binary or categorical terms according to probabilistic notions.

- Random Forest: An ensemble method used for accurate modeling and suitable for non-linear data.
- Decision Tree: An easily interpretable model with intuitive derivatives that can perform fast for small data sets.

Comparative Analysis: These classifiers were chosen to facilitate a comparison of their performance on binary and multi-class problems.

Computational Efficiency: KNN and Naive Bayes were incorporated since they have less computational requirements when compared to other algorithms, such as Naive Bayes, with basic Random Forest and SVM computational complexity and enhanced accuracy.

3.4. Evaluation Metrics and Cross-Validation

The dataset was split into 80% training and 20% testing sets to ensure robust evaluation. A 5-fold cross-validation was performed to minimize the risk of overfitting and validate the models' generalizability. The following metrics were used for evaluation:

- Classification Metrics: Accuracy, Precision, Recall, F1-score, Receiver Operating Characteristic (ROC) curve, and Precision-Recall Curve (PRC).
- Regression Metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Relative Absolute Error (RAE), and Root Relative Squared Error (RRSE).
- Computational Metrics: Training and prediction times were measured to assess computational efficiency.

3.5. Statistical Analyses and Visual Representations

To support the classification results:

- Cross-Validation Results: Average metrics from cross-validation were reported, including standard deviations to quantify variability.
- Statistical Significance: Paired t-tests were conducted to compare model performance across different classifiers and preprocessing techniques.
- Visual Representations: ROC and PRC curves were generated to illustrate the trade-offs between sensitivity and specificity. Comparative bar charts highlighted performance metrics like accuracy, precision, and recall for binary and multi-class classification.

The following TICMFE-EML algorithm can implemented as a novelty in this research work.

Algorithm: Typhoon Intensity Classification using Multi-Feature Extraction and Ensemble ML (TICMFE-EML)

Input: Set of infrared satellite images IR = $\{IR_1, IR_2, ..., IR_{136}\}$

Output: Typhoon intensity classification model M

Step 1: Image Preprocessing: For each IR, resize the input image IR_i to standard 224x224 pixels. The min and max functions are applied to the resized image. The normalization is performed using these min and max values.

Step 2: Image Augmentation: For each IR_i^n, generate: a. Flip: $IR_i^f = F(IR_i^n)$, where F is the flip function b. Rotate: $IR_i^{rot} = Rot(I_i^n, \theta)$,
where Rot is the rotation function and $\theta \in \{90°, 180°, 270°\}$
Augmented set: $A = \{IR_i^n, IR_i^f, IR_i^{rot}\}$ for all i, resulting in 3000 images

Step 3: Feature Extraction: For each image $J \in A$, compute:
a. Mean: $\mu = (1/N) \sum_{ij} J(i,j)$
b. Standard Deviation: $\sigma = \sqrt{(1/N) \sum_{ij} (J(i,j) - \mu)^2}$
c. Skewness: $\gamma_1 = E[((X-\mu)/\sigma)^3]$
d. Kurtosis: $\gamma_2 = E[((X-\mu)/\sigma)^4] - 3$
e. Entropy: $H = -\sum p(x) \log_2(p(x))$
f. Eye Area: $A = \sum_{ij} E(i,j)$, where E is the binary eye region
g. Gradient Mean: $\mu\_grad = (1/N) \sum_{ij} |\nabla J(i,j)|$
h. FFT Mean: $\mu\_fft = (1/N) \sum_{uv} |F(u,v)|$, where F is the Fourier transform and
i. FFT STD: $\sigma\_fft = \sqrt{(1/N) \sum_{uv} (|F(u,v)| - \mu\_fft)^2}$

Step 4: Feature Vector: For each image J, create feature vector: $X\_J = [\mu, \sigma, \gamma_1, \gamma_2, H, A, \mu\_grad, \mu\_fft, \sigma\_fft]$

Step 5: Class Preparation: a. Binary: Y_binary = {0 if low/medium(300 images), 1 if high(300 images)} b. Multi-class: Y_multi = {0 if low(300 images), 1 if medium(300 images), 2 if high(300 images)}

Step 6: Filtering: Gabor Filter: $g(x1,y1;\lambda1,\theta1,\psi1,\sigma1,\gamma1) = \exp(-(x1'^2+\gamma1^2y1'^2)/(2\sigma1^2))\cos(2\pi x1'/\lambda\ 1+ \psi1)$
where $x1' = x1\cos\theta + y1\sin\theta$, $y1' = -x1\sin\theta + y1\cos\theta$ $J\_Gab = J * g$

Step 7: Machine Learning Models:
a. SVM: f(x1) = sign(w1^T1 x1 + b1)
b. KNN: y = mode(y_i) for k nearest neighbors
c. Bayes Net : P(ai|bi) = P(bi|ai)P(ai) / P(bi)
d. Random Forest: y = mode(tree_i(x)) for i=1 to n_trees
e. Decision Tree: y = leaf_value(x)

Step 8: Model Evaluation: For each model, m: classification, regression metrics, etc.

Step 9: Model Selection and Fitting: M = argmax_m (F1_m)
Train M on the entire dataset.

3.6. Experimental Setup

The dataset was split into 80% training and 20% testing sets. SVM, KNN, and Naive Bayes were used for binary classification with Gabor filter features. Random Forest and Decision Tree were used with Gaussian filter features for multi-class classification. The models were implemented using a scikit-learn library in Python. Training time was measured for each model.

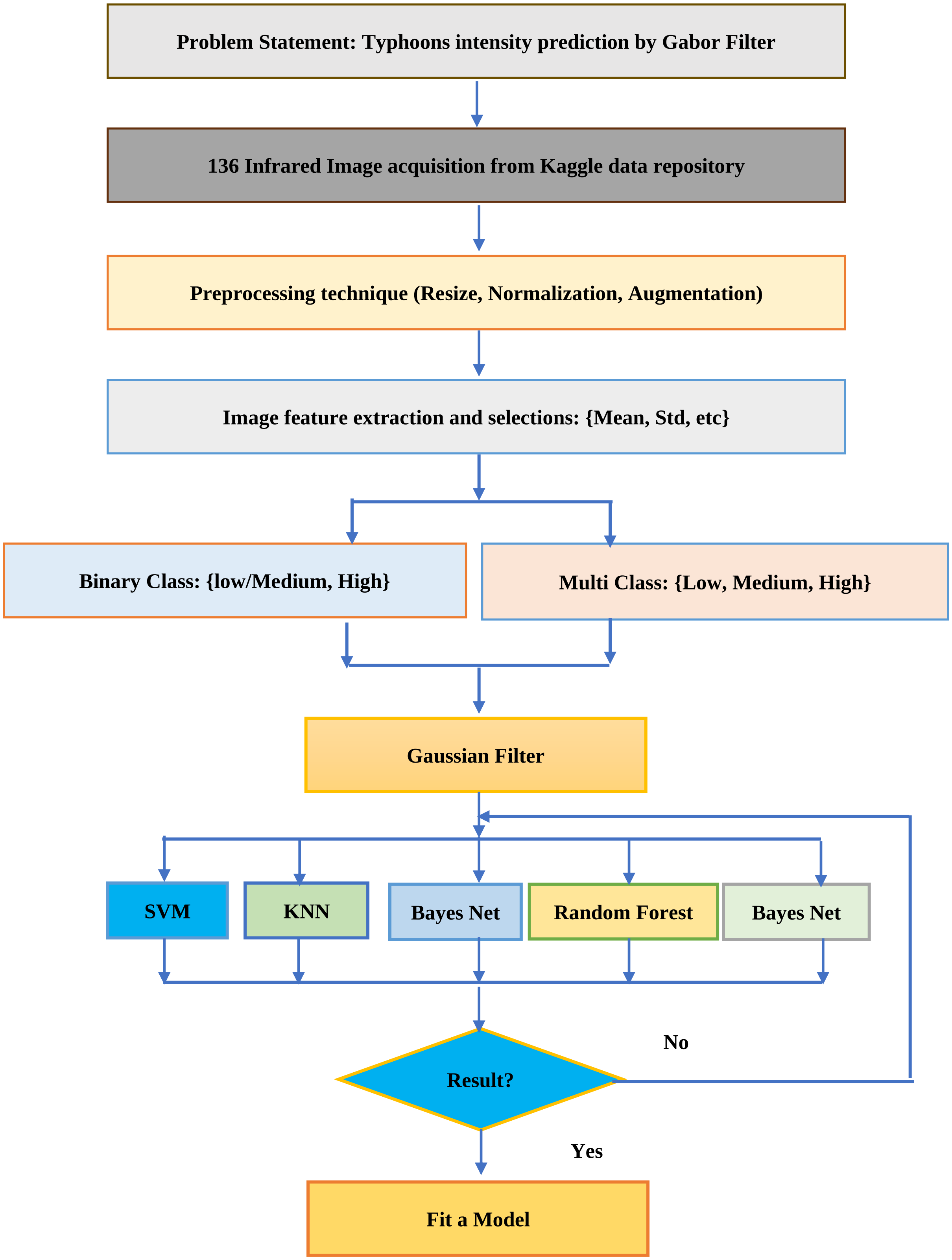

Fig. 1 Schema of the proposed system

Figure 1 shows that using the ML models has the following methods to predict an optimal outcome.

4. Results and Discussion

This section focuses on the results and discussions of the binary class and multi-class classification with Gaussian filters by SVM, KNN, Bayes Net, Decision Tree, and Random Tree for predicting tropical cyclone intensity. Table 1 presents the classification metrics of various classifications for both binary and multi-class with a Gaussian filter.

The SVM and Random Forest had the highest accuracy in the binary class, at 0.88. Bayes Net came in second with 0.81 accuracy, followed by Decision Tree and K-Nearest Neighbors (KNN) with 0.72 accuracy each. When it came to multi-class classification, the best accuracy was maintained by SVM and Random Forest at 0.88, followed by Decision Tree at 0.76, KNN at 0.69, and Bayes Net at 0.68.

The SVM scored 0.88 in the binary class, with Random Forest coming in second with 0.87 precision. K-Nearest Neighbors (KNN) earned 0.8 precision, Decision Tree displayed the lowest precision at 0.72, and Bayes Net achieved 0.81 precision. SVM led the field in multi-class classification with 0.89 precision, Random Forest came in second with 0.88 precision, and KNN came in third with 0.77 precision. With 0.7 and 0.76 precision, respectively, Bayes Net and Decision Tree performed less precisely.

High recall scores of 0.88 and 0.87 were obtained for binary classification using SVM and Random Forest, respectively, while 0.81 was obtained for Bayes Net, 0.73 for KNN, and 0.72 for Decision Tree. SVM and Random Forest kept up their top results in the multi-class classification task, with a recall of 0.88.

Bayes Net and Decision Tree scored 0.68 and 0.76, respectively, while KNN earned 0.69. This shows that while KNN and Bayes Net showed slightly lower recall values, especially in the multi-class classification setting, SVM and Random Forest performed consistently well across both classification types in terms of recall.

SVM and Random Forest produced high ROC values of 0.92 and 0.93 in binary classification, respectively, while KNN, Bayes Net, and Decision Tree produced values of 0.87, 0.82, and 0.72, respectively. With even higher ROC values of 0.97 apiece, SVM and Random Forest outperformed in the multi-class classification, whereas KNN only managed 0.9, Bayes Net 0.8, and Decision Tree 0.82.

Random Forest has the highest PRC value in binary classification (0.92), closely followed by SVM (0.87) and KNN (0.83). Decision Tree and Bayes Net had lower scores, at 0.64 and 0.74, respectively. With a PRC value of 0.95, Random Forest is again in the lead for multi-class classification. SVM is next with a PRC value of 0.94, and KNN lasts at 0.8. The PRC values of 0.59 and 0.65 for Bayes Net and Decision Tree, respectively, indicate moderate performance.

In binary classification, SVM and Random Forest show high MCC scores of 0.75 and 0.73, respectively, indicating good predictive performance. Bayes Net follows with a moderate score of 0.62, while KNN and Decision Tree exhibit lower MCC values of 0.52 and 0.43, respectively.

In multi-class classification, SVM and Random Forest lead with high MCC scores of 0.83 and 0.82, demonstrating their strong ability to handle multi-class problems.

Decision Tree achieves a significantly better MCC in the multi-class setting (0.63) than binary classification. Bayes Net and KNN exhibit moderate MCC values in the multi-class context, with scores of 0.54 MCC and 0.58 MCC, respectively.

In binary classification, SVM and Random Forest demonstrate strong agreement, with 0.75 kappa and 0.73 kappa, respectively, indicating highly reliable classification performance. Bayes Net also performs moderately well, with a Kappa value of 0.62. KNN and Decision Tree show lower reliability in binary classification, with 0.44 kappa and 0.43 kappa, respectively.

For multi-class classification, SVM and Random Forest again stand out with high Kappa values of 0.82, indicating excellent consistency across multiple classes.

Decision Tree exhibits a significantly improved Kappa of 0.63 in the multi-class scenario compared to its binary performance. KNN and Bayes Net show moderate improvements in multi-class settings, achieving Kappa values of 0.53.

Table 1 presents the regression metrics and time complexity for binary and multi-class classifiers with a Gaussian filter on tropical cyclone infrared images.

In binary classification, SVM and Random Forest show minimal prediction error with MAE values of 0.13, indicating highly accurate models. Bayes Net performs moderately well with an MAE of 0.19, while KNN and Decision Tree exhibit higher errors at 0.28.

In multi-class classification, the MAE increases slightly across all classifiers. SVM maintains a low error of 0.13, while Random Forest experiences a small rise to 0.14, remaining among the most accurate classifiers.

KNN and Decision Tree show significant increases in error in the multi-class setting, with MAE values of 0.44 and 0.29, respectively, suggesting challenges in multi-class prediction for these algorithms. Bayes Net also sees an increased error in multi-class classification, with an MAE of 0.38.

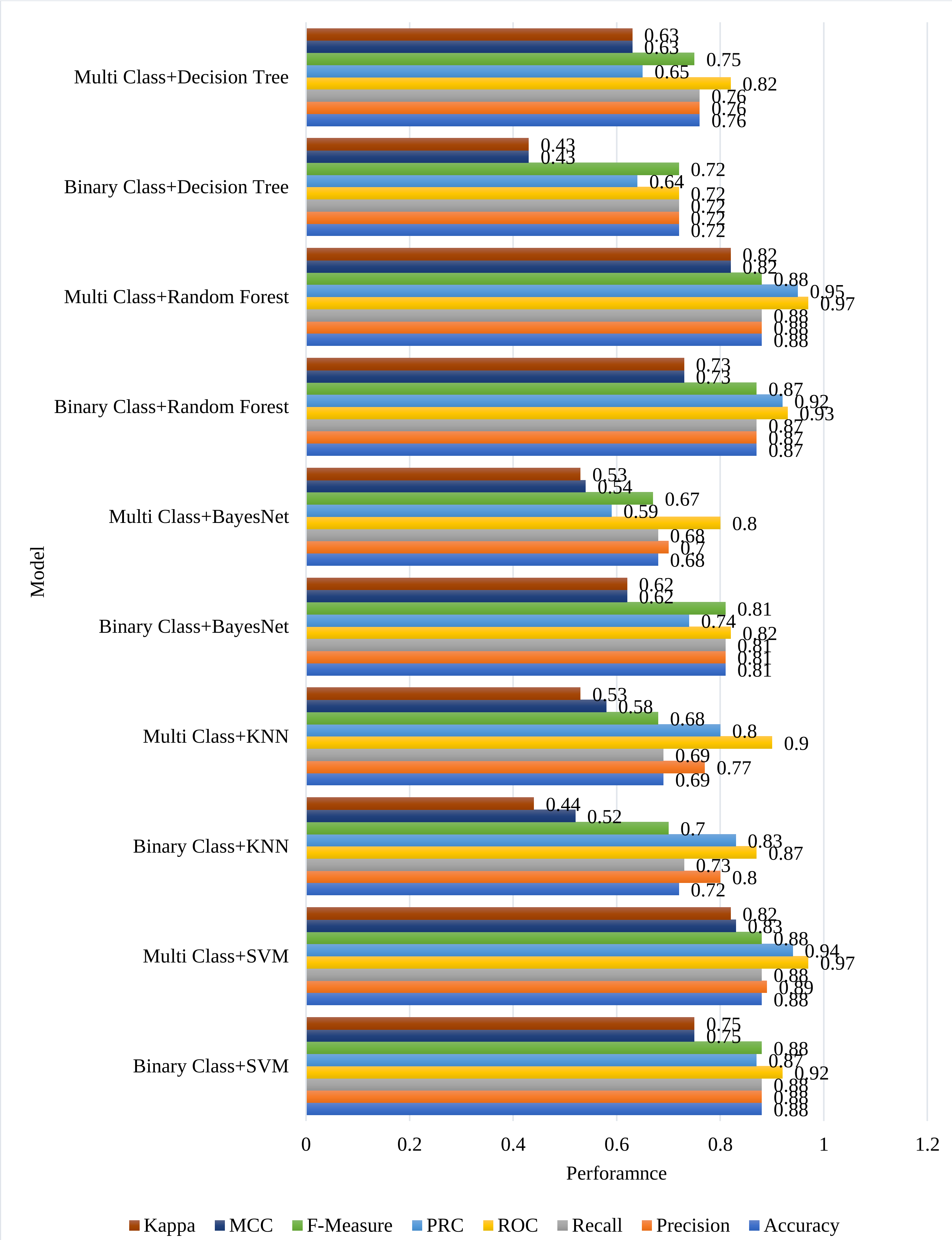

Fig. 2 Gaussian filter with binary & Multi class model vs. Performance

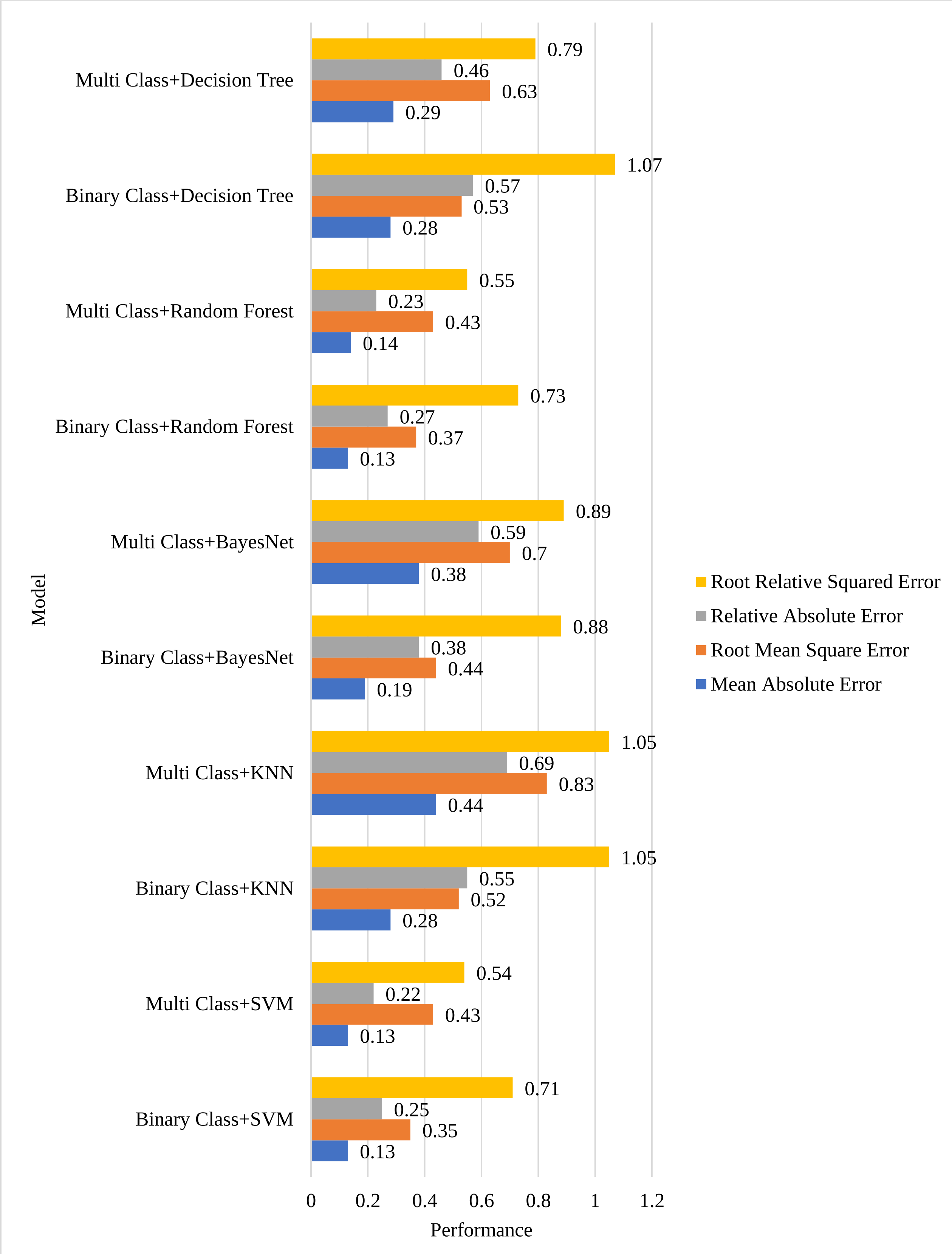

Fig. 3 Gaussian filter with Binary & Multi class model vs. Regression metrics

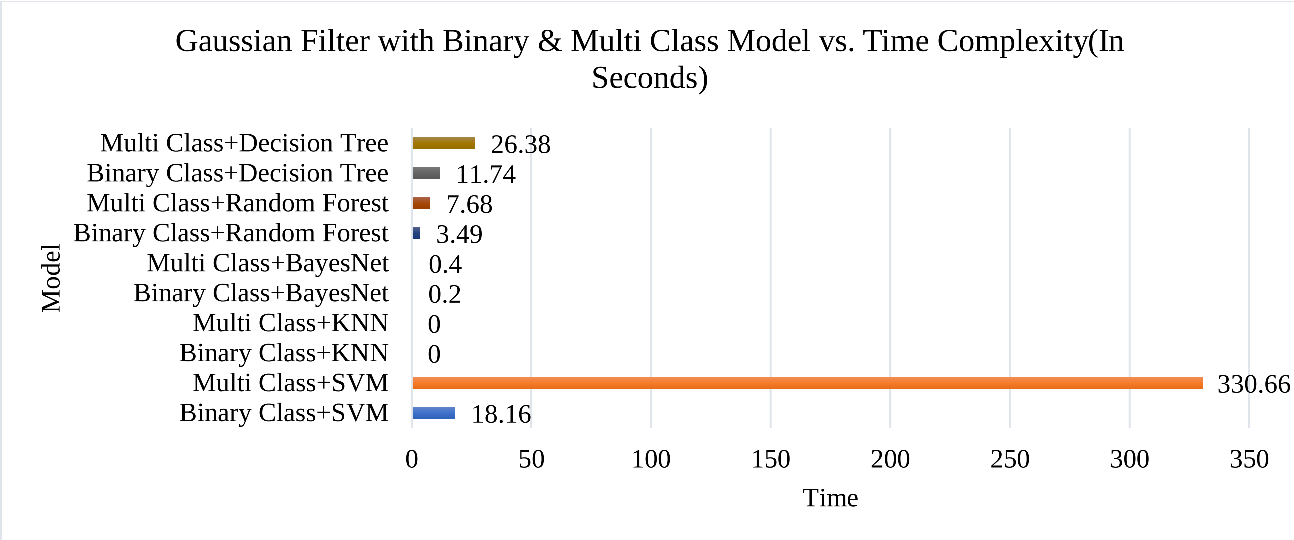

Fig 4 Graphical representation of existing model Vs Proposed model

Table 2. Classification metrics

S.No	Classifier	Accuracy	Precision	Recall	ROC	PRC	Kappa	F1-Score	MCC	MAE	RRSE	RAE	RRSE	Time
1	Binary Class + SVM	0.88	0.88	0.88	0.92	0.87	0.75	0.88	0.75	0.13	0.35	0.25	0.71	18.16
2	Multi-class + SVM	0.88	0.89	0.88	0.97	0.94	0.82	0.88	0.83	0.13	0.43	0.22	0.54	330.66
3	Binary Class + KNN	0.72	0.8	0.73	0.87	0.83	0.44	0.7	0.52	0.28	0.52	0.55	1.05	0
4	Multi-class + KNN	0.69	0.77	0.69	0.9	0.8	0.53	0.68	0.58	0.44	0.83	0.69	1.05	0
5	Binary Class + Bayes Net	0.81	0.81	0.81	0.82	0.74	0.62	0.81	0.62	0.19	0.44	0.38	0.88	0.2
6	Multi-class + Bayes Net	0.68	0.7	0.68	0.8	0.59	0.53	0.67	0.54	0.38	0.7	0.59	0.89	0.4
7	Binary Class + Random Forest	0.87	0.87	0.87	0.93	0.92	0.73	0.87	0.73	0.13	0.37	0.27	0.73	3.49
8	Multi-class + Random Forest	0.88	0.88	0.88	0.97	0.95	0.82	0.88	0.82	0.14	0.43	0.23	0.55	7.68
9	Binary Class + Decision Tree	0.72	0.72	0.72	0.72	0.64	0.43	0.72	0.43	0.28	0.53	0.57	1.07	11.74
10	Multi-class + Decision Tree	0.76	0.76	0.76	0.82	0.65	0.63	0.75	0.63	0.29	0.63	0.46	0.79	26.38

In binary classification, the SVM classifier demonstrates the lowest RMSE at 0.35, indicating high predictive accuracy. Random Forest follows closely with an RMSE of 0.37, reflecting effective performance. KNN and Decision Tree exhibit higher RMSE values at 0.52 and 0.53, respectively, suggesting less reliable predictions. In the multi-class scenario, the RMSE increases across all classifiers. Multi-Class SVM shows an RMSE of 0.43, while Multi-Class Random Forest matches this figure. KNN sees a significant rise to 0.83, indicating a considerable challenge in accurately predicting multiple classes. Multi-Class BayesNet and Multi-Class Decision Tree also display increased RMSE values of 0.70 and 0.63, respectively, suggesting that these models may struggle more with multi-class classification.

In binary classification, the SVM classifier has the lowest RAE at 0.25, indicating a strong predictive performance. The Random Forest classifier also performs well, with an RAE of 0.27. In contrast, the KNN and Decision Tree classifiers exhibit higher RAE values at 0.55 and 0.57, respectively, reflecting less reliability in their predictions. In the multi-class classification scenario, the RAE for Multi-Class SVM improves to 0.22, showing enhanced accuracy compared to its binary counterpart. Multi-Class Random Forest also demonstrates good performance with an RAE of 0.23. However, Multi-Class with KNN experiences a significant increase to 0.69, indicating challenges in managing multiple classes. Similarly, Multi-Class with Bayes Net shows a notable rise to 0.59, while Multi-Class Decision Tree performs better than its binary version at 0.46.

In binary classification, the SVM classifier demonstrates a moderate RRSE of 0.71, indicating relatively reliable performance. The Bayes Net classifier follows closely with an RRSE of 0.88, suggesting a reasonable prediction capability. However, both the KNN and Decision Tree classifiers exhibit significantly higher RRSE values of 1.05 and 1.07, respectively, indicating poorer performance and a higher degree of error in their predictions for the binary class. In the multi-class classification scenario, the Multi-Class with SVM shows an improved RRSE of 0.54, highlighting its effectiveness in handling multiple classes. The Multi-Class Random Forest classifier also performs well with an RRSE of 0.55, suggesting robust predictive power. On the other hand, Multi-Class Bayes Net and Decision Tree classifiers have RRSE values of 0.89 and 0.79, respectively, indicating that they perform adequately but still exhibit notable prediction errors compared to the SVM and Random Forest classifiers.

The Binary Class with SVM has a time complexity of 18.16 seconds, while the Multi-Class with SVM significantly increases to 330.66 seconds, indicating a high computational burden for multi-class classification. In contrast, both the Binary and Multi-Class with KNN classifiers exhibit a time complexity of 0 seconds, highlighting their efficiency due to the instance-based nature of the algorithm. The Binary Class with Bayes Net shows a time complexity of 0.2 seconds, with the Multi-class with Bayes Net slightly increasing to 0.4 seconds, indicating computational efficiency for both classification types. The Binary Class with Random Forest takes 3.49 seconds, and the Multi-Class Random Forest takes 7.68 seconds, demonstrating reasonable efficiency but increased time for multi-class tasks. Decision Trees require a moderate amount of time, with the Binary Class taking 11.74 seconds and the Multi-class 26.38 seconds.

4.1. Gaussian Filtering Impact

Models processed with Gaussian filters consistently outperformed those without, achieving higher accuracy and reduced noise-induced errors.

4.2. Classifier Performance

- SVM and Random Forest achieved the highest accuracy (88%) across binary and multi-class classifications, but SVM exhibited higher computational complexity.
- KNN and Naive Bayes provided faster computation but at the cost of reduced accuracy.
- Decision Tree offered an interpretable solution but struggled with noise in multi-class classification.

4.3. Cross-Validation Findings

Cross-validation confirmed the robustness of SVM and Random Forest, with less than 5% variation in accuracy across folds. Models like KNN exhibited higher variability, indicating sensitivity to data splits.

5. Conclusion

This work postulates a high accuracy precision recall and excellent ROC/PRC score in the Multi-class with SVM and the Multi-class with Random Forest classifiers. These models and their corresponding binary models yield the best accuracy of 0.88 at all levels of evaluation. Therefore, the accuracy improves to 94% when the multi-class versions are complemented by a higher ROC of 0.97 and PRC of 0.94-0.95, implying that the model has better discrimination ability. However, it should also be said that the time testing of both sets of models proved the fact that among all the studied models, the time complexity of the Multi-class with SVM is significantly higher than that of other types for large datasets or for usage in real time projects, it may be critical.

On the other hand, KNN models have the least time complexity but come with the highest level of inefficiency. A general trend observed in this paper is that SVM and Random Forest models perform relatively better regarding different error measures. Interestingly, the classifiers are much better with multi-class data sets. Most multi-class versions have slightly better ROC and PRC scores than the binary data set classifiers. While considering the KNN and

Decision Tree models, the performance ranking is commonly lower than that of the SVM and Random forest algorithms across most evaluation measures. Therefore, whenever computation time is not a main issue, it is advisable to use Multi-class with SVM or Multi-class with Random Forest since they involve powerful algorithms and perform best in the test current. However, the Binary Class with Random Forest gives an equal measure of time space for applications where time operates as a constraint. Therefore, the last choice between these models would be decided by certain application characteristics, such as the need for speed over that of getting the right class.

Limitations and Future Work

This work has been achieved using infrared imagery of the Indian Ocean region. Therefore, the results cannot be generalized to other meteorological phenomena or to other regions of the world. More research can be done on multispectral data or analyze the applicability of these processes in other cyclone affected areas.

References

[1] Juhyun Lee, Jungho Im, and Yeji Shin, "Enhancing Tropical Cyclone Intensity Forecasting with Explainable Deep Learning Integrating Satellite Observations and Numerical Model Outputs," *iScience*, vol. 27, no. 6, pp. 1-18, 2024. [Google Scholar] [Publisher Link]

[2] Manish Kumar Mawatwal, and Saurabh Das, "An End-to-End Deep Learning Framework for Cyclone Intensity Estimation in North Indian Ocean Region Using Satellite Imagery," *Journal of Indian Society Remote Sensing*, vol. 52, pp. 2165-2175, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Han Wang et al, "Determination of Low-Intensity Tropical Cyclone Centers in Geostationary Satellite Images Using a Physics-Enhanced Deep-Learning Model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Quan Nguyen, and Chanh Kieu, "Predicting Tropical Cyclone Formation with Deep Learning," *Weather and Forecasting*, vol. 39, no. 1, pp. 241-258, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Shanshan Mu et al., "High-Resolution Tropical Cyclone Rainfall Detection from C-Band SAR Imagery with Deep Learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Wei Tian et al., "TCIP-Net: Quantifying Radial Structure Evolution for Tropical Cyclone Intensity Prediction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Asanobu Kitamoto et al., "Digital Typhoon: Long-Term Satellite Image Dataset for the Spatio-Temporal Modeling of Tropical Cyclones," *Advances in Neural Information Processing Systems*, 2023. [Google Scholar] [Publisher Link]

[8] Randi Fu et al., "Spatiotemporal Fusion Convolutional Neural Network: Tropical Cyclone Intensity Estimation from Multisource Remote Sensing Images," *Journal of Applied Remote Sensing*, vol. 18, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Sarah M. Griffin, Anthony Wimmers, and Christopher S. Velden, "Predicting Short-Term Intensity Change in Tropical Cyclones Using a Convolutional Neural Network," *Weather and Forecasting*, vol. 39, no. 1, pp. 177-202, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Chandan Roy et al., "Tropical Cyclone Intensity Forecasting in the Bay of Bengal Using a Biologically Inspired Computational Model," *Modeling Earth Systems and Environment*, vol. 10, pp. 523-537, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Chang-Jiang Zhang et al., "Tropical Cyclone Tracking from Geostationary Infrared Satellite Images Using Deep Learning Techniques," *International Journal of Remote Sensing*, vol. 45, no. 18, pp. 6324-6341, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Tao Song et al., "Probabilistic Estimation of Tropical Cyclone Intensity Based on Multi-Source Satellite Remote Sensing Images," *Remote Sensing*, vol. 16, no. 4, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Zhitao Zhao et al., "Multi-Task-Learning-Based Graph Residual Network for Tropical Cyclone Intensity Estimation," *Remote Sensing*, vol. 16, no. 2, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Zhaoyang Ma et al., "A Multiscale and Multilayer Feature Extraction Network with Dual Attention for Tropical Cyclone Intensity Estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] L. Raynaud et al., "A Convolutional Neural Network for Tropical Cyclone Wind Structure Identification in Kilometer-Scale Forecasts," *Artificial Intelligence for the Earth Systems*, vol. 3, no. 2, pp. 1-15, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Hyeyoon Jung et al., "Tropical Cyclone Intensity Estimation through Convolutional Neural Network Transfer Learning Using Two Geostationary Satellite Datasets," *Frontiers Earth Science*, vol. 11, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Hongxing Cui et al., "Modeling Ocean Cooling Induced by Tropical Cyclone Wind Pump Using Explainable Machine Learning Framework," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Wei Tian et al., "Short-Term Intensity Prediction of Tropical Cyclones Based on Multi-Source Data Fusion with Adaptive Weight Learning," *Remote Sensing*, vol. 16, no. 6, pp. 1-21, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Soumyajit Pal, Uma Das, and Oishila Bandyopadhyay, "SSN: A Novel CNN-Based Architecture for Classification of Tropical Cyclone Images From INSAT-3D," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Syarawi M.H. Sharoni, Mohd Nadzri Md Reba, and Hwee San Lim, "Improved Tropical Cyclone Wind Speed Estimation for Microwave Altimeter Using Machine Learning," *Remote Sensing of Environment*, vol. 301, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] K. Bharathi, A. Archita, and S. Gandhimathi Alias Usha, "Predicting Tropical Cyclones: A Supervised Machine Learning Approach," *Predicting Natural Disasters with AI and Machine Learning*, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Lunkai He et al., "Enhanced Tropical Cyclone Precipitation Prediction in the Northwest Pacific Using Deep Learning Models and Ensemble Techniques," *Water*, vol. 16, no. 5, pp. 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Xiaoxian Tian et al., "EasyRP-R-CNN: A Fast Cyclone Detection Model," *The Visual Computer*, vol. 40, pp. 4829-4841, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Franciskus Antonius Alijoyo et al., "Advanced Hybrid CNN-Bi-LSTM Model Augmented with GA and FFO for Enhanced Cyclone Intensity Forecasting," *Alexandria Engineering Journal*, vol. 92, pp. 346-357, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Li Liu et al., "Enhanced Rainfall Now Casting of Tropical Cyclone by an Interpretable Deep Learning Model and its Application in Real-Time Flood Forecasting," *Journal of Hydrology*, vol. 644, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Sabbir Rahman et al., "Tropical Cyclone Track Prediction Harnessing Deep Learning Algorithms: A Comparative Study on the Northern Indian Ocean," *SSRN*, pp. 1-45, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[27] INSAT3D Infrared & Raw Cyclone Imagery (2012-2021), Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/sshubam/insat3d-infrared-raw-cyclone-images-20132021