International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Tuned Fusion Deep Learning Approach for Unsupervised Cyclone Passive Microwave Rainfall Imagery Prediction

Mamatha N.P [1], Mohmad Umair Bagali[2], Thangadurai N[3]

Submitted: 12/11/2022 **Revised**: 24/01/2023 **Accepted**: 12/02/2023

Abstract: In many cases, tropical cyclones (TCs) can bring about widespread, intense rainfall because of the volume of water vapour they contain. High-resolution, passive microwave (PMW) rainfall assessment of tropical cyclones is essential for disaster forecasting tropical cyclones. Still, it remains a tricky issue because of the low temporal resolution of sensing devices. This research makes an effort to address this issue by directly forecasting PMW rainfall pictures of tropical cyclones. We developed a tuned fusion learning approach to forecast unannotated PMW image classes. The study outcomes demonstrate the system's ability to retrieve essential features from PMW images successfully. The global applicability of this deep learning technique is promising. It offers a fresh angle on satellite-based typhoon rainfall forecasting, which can yield valuable insights for real-time visualization of typhoon precipitation around the world in operational activities.

Keywords: *Deep Learning, Principal Component Analysis (PCA), Tropical Cyclone Detection, Unsupervised Learning*

1. Introduction

Severe weather occurrences such as tropical cyclones are a leading source of catastrophic events in the world, with the potential to bring about widespread destruction through high winds, heavy precipitation, and flooded areas. Since the ocean's water vapour is the primary energy source for tropical cyclones [1], it is crucial to make real-time estimates of TC rainfall to prepare for the potential of widespread flooding [1].

TCs are well-organized meteorological events, with dense cloud cover making it impossible to see through infrared cloud maps and, hence, assess cloud-top features directly connected with rainfall [2]. The thermal radiation of rainfall or the scattering of upwelling radiation induced by rain droplets can be detected by passive microwave (PMW) sensors even through the dense cloud cover of TC [3]. Geostationary satellites frequently carry infrared sensors and continually monitor tropical cyclones around the clock with excellent precision [4]. The long-range signal loss of PMW sensors necessitates their exclusive usage on polar-orbiting spacecraft [5]. Because polar-orbiting satellites only pass the same spot twice a day, PMW is subject to severe geographical and temporal limitations.

[1]*Research Scholar, Department of ECE, Jain University, Bangalore*
[2]*Jain University, Bangalore -562112, ORCID ID : 0000-0001-5342-4067*
[3]*Center for Research and Innovation, Sankalchand Patel University, Gujarat – 384315, ORCID ID : 0000-0002-2149-9440*

Related researchers have used statistical and machine learning methods to translate satellite images into rainfall amounts at the surface; for example, the PERSIANN algorithm established a connection between cloud-top brightness temperature and rainfall amount by employing artificial neural network techniques [6].

There is a scarcity of expert studies on tropical cyclones, which are distinctly prepared and have significantly diverse cloud features compared with cloud-free and thin cloud zones, although extant studies discuss global precipitation directly [7]. Research in this area is crucial due to the unique characteristics of tropical cyclones and the devastating effects they can have [8]. The field of computer vision uses transfer learning extensively for things like picture and video recognition and creation. There are various works based on transfer learning in detecting cyclones too. In this paper, we proposed a tuned transfer learning-based robust clustering model for cluster prediction of tropical cyclones. Considering the anticipated forecast of unsupervised cyclone picture classifications, transfer learning coupled with clustering is an excellent method for resolving this issue

2. Related Works and Challenges

A microwave sensing device is unaffected by the time of day or the presence of clouds. Microwave sensors, both active (such as a scatterometer) and

passive (such as a radiometer), are well-suited to the observation of Polar Regions from space due to their unique characteristics. There are active [9, 10] and passive sensors [11] that have been used to infer the extent of the sea ice cover in Antarctica. In order to consistently track the intensity, inner core morphology, and convective organization of worldwide tropical cyclone life cycles, Yang et al. show that TB adjustments between PMW sensors are crucial [12]. DeepMicroNet was proposed by Wimmers et al., and it has several desirable characteristics that set it apart from other similar methods. They include a probabilistic output, the capacity to function from partial scans, and the ability to tolerate inaccurate cyclone centre corrections [13]. Nimbus-5 F-Electrically Scanning Microwave Radiometer (ESMR-5) data appears to have prospects for use in both tracking and forecasting the strength of these cyclones and can be used to determine their precipitation properties [14]. The strength of tropical cyclones can be estimated with the use of a statistical method presented by Jiang et al. The algorithms are tested using both simulated and real-world data sets, including data collected by aircraft for use in the analysis [15]. For datasets that are lacking in some way, transfer learning can be used to supplement them with information from comparable source datasets. Deo et al. used transfer stacking to investigate cyclones' impacts, testing whether cyclones in various places help with generalization performance. Furthermore, traditional neural networks were employed to assess the impact of cyclone duration on prediction accuracy [16]. A CNN-based methodology was developed by Banerjee et al. to determine if a particular region has been affected by a cyclone and, if so, how severely. The models have been fine-tuned through the use of transfer learning [17].

Our goal is to predict cyclone stages using unsupervised data. Singh et al. carried out an unsupervised ISODATA clustering categorization to identify sea ice zones near Antarctica [18].

The biggest challenge in this kind of work is getting the appropriate dataset for cyclone prediction. Most of the time, publicly available datasets are hard to access due to several restrictions and registration procedures. Sometimes, the acquisition process needs permission, which can take days to get. Another thing is the type of the stored data file issue. Most of the time, the datasets were stored in a hierarchical data format. Extracting the data from the storage files was often challenging due to being corrupted. A large image dataset is essential for a more efficient model prediction.

3. Methodology

The entire work methodology is described in this section. The workflow, dataset details, data pre-processing and the tropical cyclone forecasting model design, are discussed in the following subsections.

3.1 Workflow

The workflow of this experimental study is depicted in Fig. 1. Initially, we downloaded the dataset from the corresponding website and then saved it into cloud storage for further use.

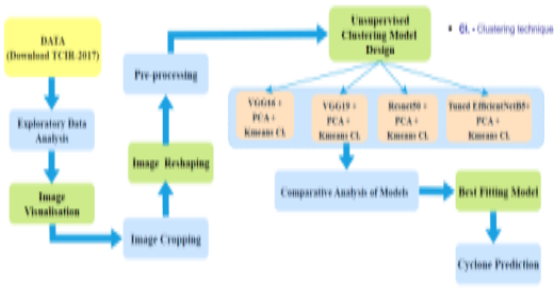

Figure 1. The Workflow of this Clustering Model Design

Exploratory data analysis was conducted to visualize the dataset details. After that, the input dataset was pre-processed as required. Then, this pre-processed data was fed into different predictive models to forecast the cyclone. This particular experimental study was done on unsupervised data. We utilized clustering techniques to predict the target classes. Different feature extraction methods were utilized in this study. They are VGG16, VGG19, Resnet50, and Tuned EfficientNet models. After this stage, the most significant features were selected based on the principal component analysis technique. The best clustering model was selected based on four different evaluation metrics. The most efficient model classified the large input dataset into 5 distinct categories of tropical cyclones.

3.2 Dataset

The Tropical Cyclone Image-to-intensity Regression (TCIR) system compiles information about these storms from four different types of satellite imagery. The goal of the TCIR dataset is to provide scholars with a neutral standard against which to assess the reliability of TC intensity

prediction methods. We used this dataset in our experimental study [19].

The dataset used here was developed to obtain and process Infrared (IR) and PMW channels from the Dataset of Tropical Cyclone for the Image-to-intensity Regression (TCIR) dataset. TCIR is intended to make it convenient for scientists to obtain direct exposure to satellite pictures. The TCIR data incorporates satellite images from two publicly available sources. These sources are known as 'GridSat' [20] and 'CMORPH' [21].

NASA-developed GridSat is a long-term dataset of worldwide infrared window illumination temperatures, including the IR channels. It gathered information from most geostationary meteorological satellites every three hours.

CPC MORPHing technique (CMORPH) offers global precipitation study at comparatively high temporal and spatial resolutions. It accomplishes this by utilizing precipitation projecting primarily from low earth orbit (LEO) microwave satellite observational data. The attributes were conveyed thru the topographically perpetuated information obtained predominately from geostationary satellite infrared data. CMORPHing technique offers global precipitation study at comparatively high spatial and temporal resolution.

TCIRRP dataset collects all tropical cyclone-related data from 2003 to 2017, for which it is likely to achieve a match between the IR channel and the PMW channel, with the centre of the TC in the middle. This match is based on the fact that the IR and PMW channels have the same frequency. The resolution of the PMW channel that CMORPH generated was originally 1/4-degree latitude/longitude. Linear interpolation was utilized to scale the PMW channel to approximately four times its original size in order to combine the Infrared and PMW channels and ensure that each pixel point corresponds to its neighbouring point. 7 degrees is the spatial resolution that can be achieved in latitude and longitude. The total size of the photos is 201 × 201 points, and the distance between each point (s) in real life is around 4 kilometres.

$$s = 14 \deg rees / 200 \approx 4\ Kms. \quad (1)$$

The data contains a few null values, which were interpolated utilizing 0 to fill in the gaps. The infrared and passive microwave channels were normalized independently using the statistical values, and the result was the generation of over 750,000 pairs of identical images. All of the paired data from 2017, a total of 4020 photos, were included in this experimental study because the researchers only had access to so much processing power**.**

3.3 **Exploratory Data Analysis**

We conducted an exploratory data analysis to determine the image conditions in the TCIR-2017 dataset. The dataset has 4020 images, each containing IR and PMW images side by side. Fig. 2 depicts a sample of PMW image data of tropical cyclones.

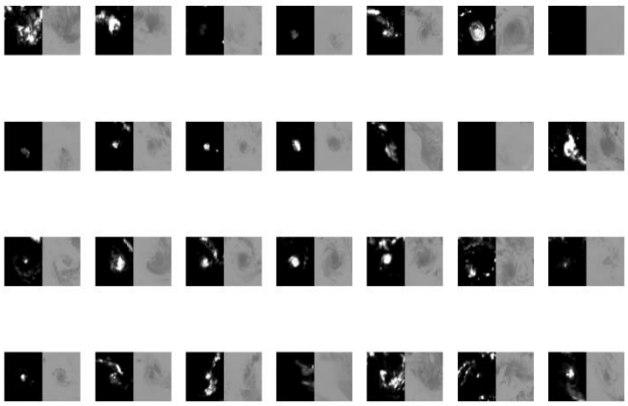

Figure 2. Sample PMW Image Data

3.4. Data Pre-processing

Initially, the images are PMW images, cropped from paired images. Then, 4020 cropped images are saved in a folder for further usage.

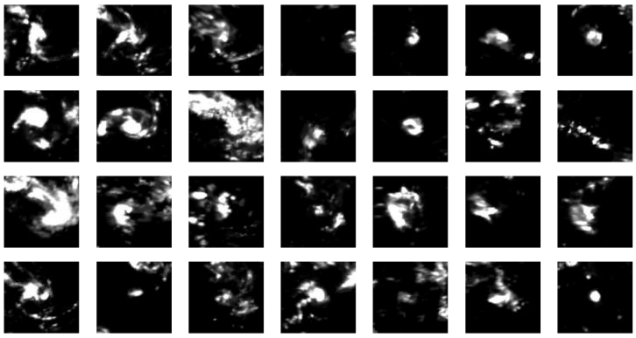

Figure 3. Sample Training Images after Cropping the PMW Images

After that, we loaded the images in a 224×224 array. Then, they are converted into a NumPy array. After the conversion stage, reshaped the images to fit them into the predictive models. Then, the images are pre-processed using Keras's 'preprocess_input' function. It will transform the input photos from RGB to BGR.

After that, it will zero-centre the colour channels in accordance with the ImageNet dataset, avoiding scaling.

3.5 Feature Extraction Model Design

We explored 4 distinct models for extracting features from the PMW image dataset. They are discussed as follows.

3.5.1 VGG16 & VGG19

Visual Geometry Group, or VGG, is a common multi-layer deep Convolutional Neural Network (CNN) architecture. There are two versions of VGG, VGG-16 and VGG-19, with 16 and 19 convolutional layers, respectively, hence the "deep" designation. Models for advanced object recognition are constructed using the VGG framework. The VGGNet, which was trained using a deep neural network, outperforms the benchmark on a broad range of assignments and datasets, not just ImageNet. The VGG network is built with very small convolutional filters as its building blocks. It consists thirteen convolutional layers and three fully connected layers. In the case of VGG19, it has sixteen layers of convolutional processing. Because it is so simple to put into action, the VGGNet is an excellent building block that may be used for educational reasons.
In our experimental study, we removed the last layer from both VGGNet structures to get the features. 4096 features were collected from each model.

3.5.2 ResNet50

ResNet is an example of a convolutional neural network (CNN). He et al. first described it in the paper called "Deep Residual Learning for Image Recognition" [22]. ResNet stands for "Residual Network." Computer vision applications frequently make use of CNNs to fuel their operations. CNN with 50 layers that goes by the name ResNet-50. Among these 50 layers, 48 are convolutional layers, one is a MaxPooling, and one is an average pooling layer. Residual networks are a sub-category of artificial neural networks (ANNs) that are designed by layering residual blocks in the network formation procedure. In comparison to a VGGNet, a ResNet is simpler and contains fewer filters overall. Two primary principles guide the design of ResNet. First, regardless of the size of the final feature map, the number of filters within every layer remains constant. Furthermore, the time complexity of each layer doubles if the feature map's size is reduced by half. This is because of the larger number of filters required to process the smaller map. Fig. 4 represents the basic architecture of a ResNet50 model used for feature extraction in this study. We were able to extract 4096 attributes from the TCIR-2017 dataset using this model.

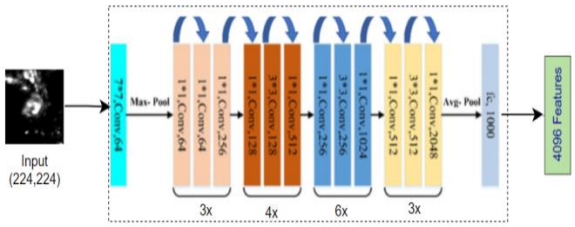

Figure 4. ResNet50 Model for Feature Extraction

3.5.3 Robust Tuned Efficient-NetB5

Google AI presented EfficientNet in work by Tan et al. to develop an approach that is more effective, as the name implies, while enhancing state-of-the-art outcomes. Most of the time, the models are either comprehensive, overly detailed, or excessively high resolution. With EfficientNet, we use a more systematic approach to scaling the features by stea

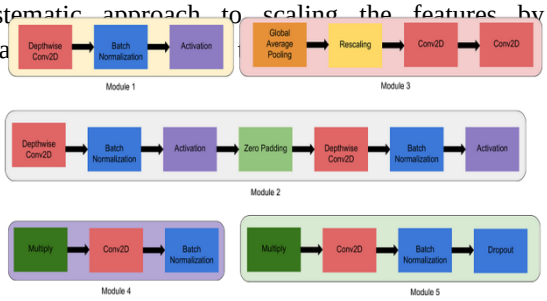

Figure 5. Modules in an EfficientNet Model

We used Keras's EfficientNetB5, a model for categorizing images preloaded with weights learned from ImageNet, for our study [23]. EfficientNetB5's architecture was made possible by incorporating B0's layering into a more complex structure [24]. After establishing that this network's stem is its primary component, the next step in the process is to begin exploring its design. This particular system is comprised of a total of five modules. Each of the modules is presented in Fig 5. This model employed the compound scaling technique to improve both precision and efficacy in larger versions, as needed by the task. Based on careful empirical research, the designers of this network have struck a clever equilibrium between network breadth, complexity,

and resolution by simply scaling all three variables with the same constant ratio.

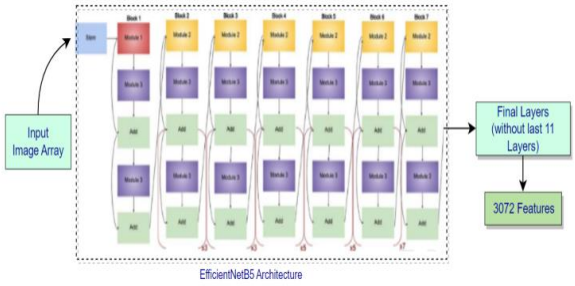

Figure 6. EfficientNetB5 Model for Feature Extraction

This particular model has 577 layers in it. We removed the last 11 layers of this model to extract the features. There were 25 million total parameters present in this network. Among them, 24.9 million were trainable. In this way, we were able to excerpt 3072 attributes out of the input dataset. If,

- α = depth
- β = width
- γ = resolution

$$\text{Then, } FLOPS = (\alpha.\beta^2.\gamma^2)^{\phi} \quad (2)$$

in this Fig. 6 depicts the tuned model utilized experimental work for feature extraction.

3.6 Dimension Reduction and Clustering

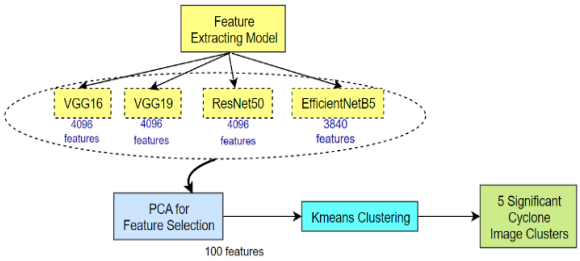

Figure 7. The Full Model Architecture

A model's complexity can be simplified by dimensionality reduction, which also helps prevent the overfitting problem. Regarding feature extraction and dimensionality reduction, Principal Component Analysis (PCA) stands out as one of the most commonly utilized unsupervised linear transformation methods. When applied to high-dimensional data, PCA attempts to locate the directions of the highest variance by mapping the information onto a subspace with the same dimensions as the initial dataset.

Using principal component analysis (PCA), we can translate a sample vector (y) onto a lower-dimensional feature subspace (j) from the original feature space (n) by constructing a transformation matrix (n x j):

$$y = [y_1, y_2, \ldots y_n], \quad y \in \mathbb{R}^n$$
$$\downarrow yW, \quad W \in \mathbb{R}^{n \times j}$$
$$p = [p_1, p_2, \ldots, p_j], \quad p \in \mathbb{R}^j \quad (3)$$

In our experimental study, we employed PCA in all four application cases for selecting the 100 most significant features out of those created due to the feature extraction models. VGG16, VGG19, and ResNet50 extracted 4096 features, except the EfficientNet model (3072 features). We utilized PCA for each of the four models and reduced the number of essential features to 100.

After all these stages, we used the K-Means clustering method to identify 5 classes of tropical cyclones using the PMW image dataset. The K-means approach in machine learning uses the first set of arbitrarily chosen centroids as the starting points for each class. Then it uses the iteration method to optimize the centroids' placements according to the information being processed. On two occasions, it prevents clusters from being created or optimized. One such instance is when the clustering has been effective, and the values of the centroids have been maintained. Second, it has completed the maximum allowed amount of iterations. The same procedure was implemented in our experimental study.

4. Procedure Implementation

The platform and tools utilized in this experimental study are discussed in this section. We employed Google Colaboratory Pro as a Python coding and execution platform for the research work

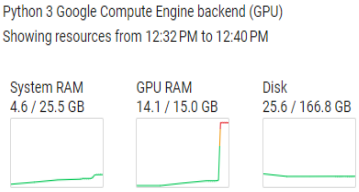

Figure 8. Sample of Hardware Specifications in the Backend

Fig. 8 presents a sample hardware specification during the working procedure of Efficientnet-based feature extraction. The system RAM capacity of the system utilized in this study is 25.5 GB. GPU RAM

capacity was 15.0 GB. Disk support was 166.8 GB. Matplotlib's pyplot library was utilized for graph plotting and image parameter checking. Data pre-processing was done using TensorFlow Keras. One of the novelties of this experimental study lies in feature Extraction. We tuned EfficientnetB5 to extract 3840 significant features from the dataset. Keras applications were employed for this work. We removed several layers of transfer learning-based models to tune the model for feature extraction. As an example, the Robust Tuned EfficientNet model lost eleven layers in this case. We removed the last layer of VGG16, VGG19, and ResNet50 to extract the required features. For model design and execution, TensorFlow Keras and Python programming were used. Principal component analysis (PCA) is a dimension reduction technique which was used here. The k-means clustering technique was employed to group the tropical cyclone passive microwave image classes.

5. Results & Discussion

The research used the TCIR-2017 dataset, which has 4020 pairs of images organized sequentially. We aimed to use the PMW images from this dataset to cluster them into five significant classes of tropical cyclones. The dataset used here is an unsupervised set of data. So, we used different methods for image attribute extraction, feature selection, and image clustering into different classes. We used four distinct evaluation metrics for selecting the best transfer learning approach suitable for extracting features from this dataset. These four metrics are:

- Silhouette Score

Distances between clusters can be visualized and quantified with the help of the Silhouette Score and Silhouette Plot. This visualization shows the distance between each point in a group and the elements in the groups surrounding it. Equation 4 shows the mathematical representation of this metric. If c and d are two elements, then this metric is represented as in equation

$$Silhoutte\ score = \frac{(c-d)}{\max(c,d)} \qquad (4)$$

• The sum of Squared Euclidean Distances (SSE) Value Plot

An SSE is calculated by adding all the squares of all the distances from each place to its nearest centroid. The goal of k-means is to reduce this value because it represents an error metric.

$$SSE = \sum_{i-1}^{n} (y_i - \bar{y})^2 \qquad (5)$$

Here, n is the number of elements and yi is the ith element.

• Calinski-Harabasz Index

The Calinski-Harabasz index is the dispersal ratio within each cluster to the deviation between clusters. Given that it does not necessitate knowledge of the ground truth labels, this score is an excellent measure of a clustering method's effectiveness.

$$s = \frac{tr(B_j)}{tr(W_j)} \times \frac{n_E - j}{j-1} \qquad (6)$$

Here, tr(Bk) is the trace of the between-group dispersal matrix tr(Wk) is the trace of the within-cluster dispersal matrix.

• Davies-Bouldin Index

The average similarity score between two clusters is what we use to calculate the Davies-Bouldin Index. The measure of similarity is the correlation between the distances inside and outside of a cluster. A higher rating is achieved by having clusters that are separated by a larger distance. If the number of clusters is n, di is the average distance separating the cluster and cluster centre ci, and xij is the distance separating two cluster centres, then the Davies-Bouldin Index is represented as equation 7.

$$Davies - Bouldin\ Index = \frac{1}{n}\sum_{i=1}^{k} \max_{i \neq j}\left(\frac{d_i + d_j}{x_{ij}}\right) \qquad (7)$$

Table 1 compares different evaluation metrics scores for different models used for the tropical cyclone prediction.

No.	Model	Silhouette Score	Calinski-Harabasz Score	Davies-Bouldin Score
1	VGG16 + PCA + Kmeans	0.0532	279.09	2.9752
2	VGG19 + PCA + Kmeans	0.1283	7363.69	2.6923
3	Resnet50 + PCA + Kmeans	0.1294	10102.42	3.03

4.	Tuned EfficientNetB5_1 +PCA+Kmeans	0.3479	1021.27	1.3140

Table 1. Comparison of Different Evaluation Metrics from Different Models

The baseline network of the EfficientNetB5 model was created by utilizing a multi-objective neural architecture search that optimizes both accuracy and FLOPS. This allowed the network to perform at its highest possible level. As we are not concentrating on any one particular piece of hardware with this system, we chose to optimize for FLOPS rather than latency. Because of the higher resolution, the size of the network increases. During the scaling of ConvNets, it is essential to strike a balance between all network dimensions, including its breadth, depth, and resolution. This can allow for improved precision as well as effectiveness. Compound scaling is applied while dealing with this kind of system.

Dimensionality reduction may often be divided into two primary classes: filtering and extracting features. In the process of extracting features, data is gathered from the given feature set in order to build a new feature subspace, but in the process of selecting features, just a fraction of the initial features are selected. In addition to enhancing a learning system's operational and archival effectiveness, extracting selected features can also boost prediction accuracy by alleviating the burden of dimensionality.

Time complexity is a big issue for model execution. Some model fitting and execution takes a lot of time. We compared the transfer learning models we utilized in this study. Fig. 9 depicts the bar plot for the time taken to execute different models to extract important features from the input image. It shows that VGG19 takes the least time to execute the model. VGG16 takes the most time (42.90 minutes) to extract the features. The tuned EfficientNet model took almost 27.19 minutes to do the same.

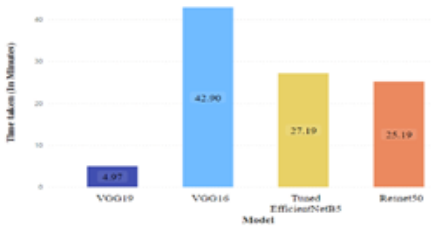

Figure 9. Time Taken for Each Transfer Learning Model for Feature Extraction.

Different approaches reached different number of images in distinct clusters. So, compared the results using the evaluation metrics to find out the most fitting model among them for this type of prediction problems. The ResNet50-based model shows the highest score for Calinski-Harabasz Index. But, the tuned EfficientNetB5-based model reached a silhouette score of almost 0.35. This is a good score for categorizing unsupervised data into clusters. The number of clusters were selected based on the SSE score plot. The number of in our study is five clusters which was evaluated using the SSE score plot. The proposed Efficient-Net-based model for clustering shows that there are less frequency of cluster 5 than the other types of tropical cyclones.

Model	0	1	2	3	4
VGG16 + PCA + Kmeans	994	957	680	563	826
VGG19 + PCA + Kmeans	1639	387	1052	495	447
Resnet50 + PCA + Kmeans	675	1420	803	327	795
Tuned EfficientNetB5_1 +PCA+Kmeans	804	623	804	804	181

Table 2. Count of images in distinct Classes

All of the aforementioned assessments show that the robust tuned EfficientNet-B5 -based fitting model for the classification of TCIR-2017 data. Fig. 10-14 are the sample images from different classes predicted by the EfficientNet-based model.

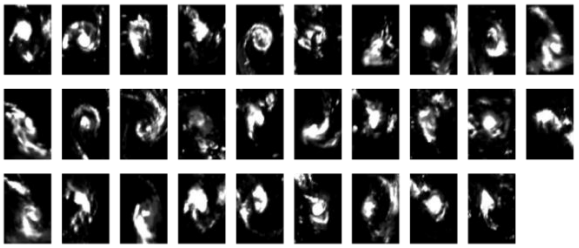

Figure 10. Cluster 1 using Efficient-Net +PCA+K-means Approach

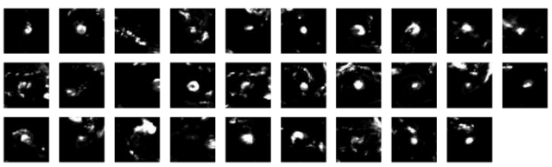

Figure 11. Cluster 2 using Efficient-Net +PCA+K-means Approach

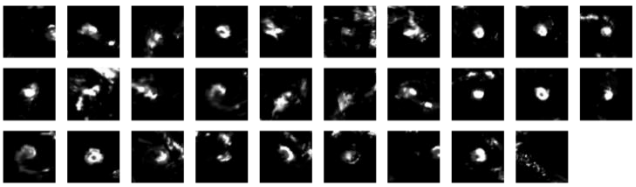

Figure 12. Cluster 3 using Efficient-Net +PCA+K-means Approach

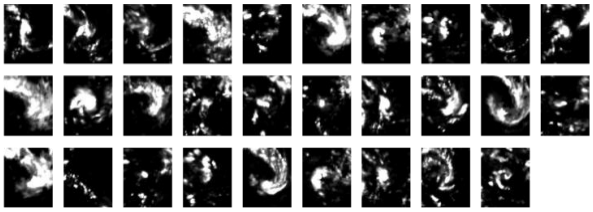

Figure 13. Cluster 4 using Efficient-Net +PCA+K-means Approach

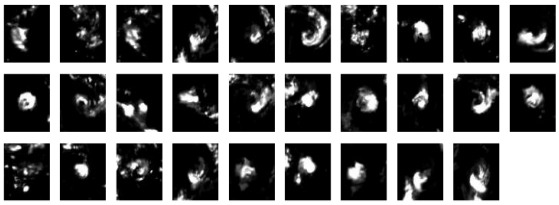

Figure 14. Cluster 5 using Efficient-Net +PCA+K-means Approach

The clusters created by the robust tuned transfer learning-based clustering model can be described as follows:

1. Tropical Cyclone eyes
2. Spiral rain
3. High degree of dispersion 1
4. High degree of dispersion 2
5. Dissipating of Tropical Cyclone

In light of the preceding analysis methods, we may deduce that EfficientNet can robustly and generally extract mapping characteristics of PMW images via training. A more accurate forecast of certain types of tropical cyclones is possible using the suggested model. But the efficiency is not exceptionally precise in the details, and we presume it is because of the restriction of computing and just one year of data has been used for the research. If it can use the complete TCIR data, it can result in a considerable increase in the exact class projections.

6. Conclusion

A novel approach to tropical cyclone unsupervised classification is discussed in this paper. The technique we developed in this research uses EfficientNet to forecast microwave rainfall image clusters from PMW image data. The clustering can help in the monitoring of the tropical cyclone using the PMW images. It can track the type of cyclone intensity based on its visuals. For this, we used the TCIRRP benchmark dataset. To ensure accuracy, the EfficientNet - B5 model is tested on 4020 sets of paired remote sensing photos during training. It was demonstrated that the Efficientnet-based model could successfully and accurately extract the crucial properties of microwave rainfall from PMW images. However, there are limiting conditions in this study: The precision of the automated system is only as good as the PMR products used. Forecasting could benefit from the use of multichannel data, including the water vapour channel and the visible channel, because the intrinsic limitations of PMW sensing devices also lead to erroneous forecasts.

Acknowledgement

This research was supported/partially supported by ISRO. We thank our colleagues from Jain University and Brindavan college of engineering, Bangalore who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper. We thank Dr. Ezhilarasan G, Research Coordinator, Jain University for assistance with clustering model Design and other model design, and Mrs. Ananya, Assistant professor, Brindavan college of engineering for comments that greatly improved the manuscript.

Author contributions

Mohmad Umair Bagali: Conceptualization, Methodology, Software, Field study **Thangadurai. N:** Data curation, Writing-Original draft preparation, Software, Validation., Field study **Mamatha N.P:** Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

All authors declare no conflicts of interest.

References

[1] L. Chen, Y. Li, and Z. Cheng, "An overview of research and forecasting on rainfall associated with landfalling tropical cyclones," *Advances in Atmospheric Sciences*, vol. 27, pp. 967-976, 2010.

[2] J. Shi *et al.*, "Implementation of an aerosol–cloud-microphysics–radiation coupling into the NASA unified WRF: Simulation results for the 6–7 August 2006 AMMA special observing period," *Quarterly Journal of the Royal*

Meteorological Society, vol. 140, no. 684, pp. 2158-2175, 2014.

[3] S. Vahedizade, "Machine Learning for Advancing Spaceborne Passive Microwave Remote Sensing of Snowfall," University of Minnesota, 2022.

[4] P. Minnett *et al.*, "Half a century of satellite remote sensing of sea-surface temperature," *Remote Sensing of Environment*, vol. 233, p. 111366, 2019.

[5] V. Levizzani, "Satellite clouds and precipitation observations for meteorology and climate," *Hydrological modelling the water cycle. Springer, Berlin*, pp. 49-68, 2008.

[6] K.-l. Hsu, X. Gao, S. Sorooshian, and H. V. Gupta, "Precipitation estimation from remotely sensed information using artificial neural networks," *Journal of Applied Meteorology Climatology*, vol. 36, no. 9, pp. 1176-1190, 1997.

[7] S. Yang and J. Cossuth, "Satellite remote sensing of tropical cyclones," in *Recent Developments in Tropical Cyclone Dynamics, Prediction, and Detection*: IntechOpen, 2016, p. 137.

[8] S. Yang, R. Bankert, and J. Cossuth, "Tropical cyclone climatology from satellite passive microwave measurements," *Remote Sensing*, vol. 12, no. 21, p. 3610, 2020.

[9] Q. P. Remund and D. G. Long, "Sea ice extent mapping using Ku band scatterometer data," *Journal of Geophysical Research: Oceans*, vol. 104, no. C5, pp. 11515-11527, 1999.

[10] Q. P. Remund and D. G. Long, "A decade of QuikSCAT scatterometer sea ice extent data," *IEEE Transactions on Geoscience Remote Sensing*, vol. 52, no. 7, pp. 4281-4290, 2013.

[11] J. C. Comiso, "Large-scale characteristics and variability of the global sea ice cover," *Sea ice: an introduction to its physics, chemistry, biology geology*, pp. 112-142, 2003.

[12] S. Yang, J. Hawkins, and K. Richardson, "The improved NRL tropical cyclone monitoring system with a unified microwave brightness temperature calibration scheme," *Remote Sensing*, vol. 6, no. 5, pp. 4563-4581, 2014.

[13] A. Wimmers, C. Velden, and J. H. Cossuth, "Using deep learning to estimate tropical cyclone intensity from satellite passive microwave imagery," *Monthly Weather Review*, vol. 147, no. 6, pp. 2261-2282, 2019.

[14] E. Rodgers and R. Adler, "Tropical cyclone rainfall characteristics as determined from a satellite passive microwave radiometer," *Monthly Weather Review*, vol. 109, no. 3, pp. 506-521, 1981.

[15] H. Jiang, C. Tao, and Y. Pei, "Estimation of tropical cyclone intensity in the North Atlantic and northeastern Pacific basins using TRMM satellite passive microwave observations," *Journal of applied meteorology climatology*, vol. 58, no. 2, pp. 185-197, 2019.

[16] R. V. Deo, R. Chandra, and A. Sharma, "Stacked transfer learning for tropical cyclone intensity prediction," 2017.

[17] S. Banerjee, A. Ghosh, K. Sorkhel, and T. Roy, "Post Cyclone Damage Assessment Using CNN Based Transfer Learning and Grad-CAM," in *2021 IEEE Pune Section International Conference (PuneCon)*, 2021, pp. 1-7: IEEE.

[18] R. K. Singh, K. N. Singh, M. Maisnam, and S. Maity, "Antarctic sea ice extent from ISRO's SCATSAT-1 using PCA and an unsupervised classification," in *Proceedings*, 2018, vol. 2, no. 7, p. 340: MDPI.

[19] B. Chen, B.-F. Chen, and H.-T. Lin, "Rotation-blended CNNs on a new open dataset for tropical cyclone image-to-intensity regression," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 90-99.

[20] NCEI, "The Geostationary IR Channel Brightness Temperature (BT)- GridSat-B1 Climate Data Record (CDR)," *National Centers for Environmental Information*, 2017.

[21] NOAA, "NOAA CPC Morphing Technique ("CMORPH")," *National Weather Service: Climate Prediction Center*, 2023.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[23] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, 2019, pp. 6105-6114: PMLR.

[24] B. Koonce and B. Koonce, "EfficientNet," *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition Dataset Categorization*, pp. 109-123, 2021.