REVIEW

Cyclone detection with end-to-end super resolution and faster R-CNN

Marwa S. Moustafa[1] · Mohamed R. Metwalli[1] · Roy Samshitha[2] · Sayed A. Mohamed[1] · Barma Shovan[2]

Abstract

The accurate identification of extreme weather events (EWEs), particularly cyclones, has become increasingly crucial due to the intensifying impacts of climate change. In the Indian subcontinent, the frequency and severity of cyclones have demonstrably risen, highlighting the need for reliable detection methods to minimize casualties and economic losses. However, the inherent limitations of low-resolution data pose significant challenges to traditional detection methods. Deep learning models offer a promising solution, enabling the precise identification of cyclone boundaries crucial for assessing regional impacts using global climate models data. By leveraging the power of deep learning, we can significantly enhance our capabilities for cyclone detection and contribute to improved risk mitigation strategies in the vulnerable Indian subcontinent. Therefore, this paper introduces an edge-enhanced super-resolution GAN (EESRGAN) leveraging an end-to-end detector network. The proposed approach comprised of a generator network equipped by residual-in-residual dense block (RRDB) and discriminator containing Faster RCNN detector. The precise patterns of cyclone had been effectively extracted to help boundary detection. Extensive experiments have been conducted on Community Atmospheric Model (CAM5.1) data taken into account only seven variables. Four matrices including precision, recall, intersection over union, and mean average precision have been considered to assess the proposed approach. The results have been found very effective while achieving accuracy up to 86.3% and average precision (AP) of 88.63%. Moreover, the proposed method demonstrates its superiority while compared with benchmarks object detectors methods. Thus, the proposed method can be employed in the area of extreme climate detection and could enrich the climate research domain.

Keywords Extreme cyclone detection · Deep learning · GAN · Faster R-CNN · Community atmospheric model (CAM5.1)

Introduction

The intensity of severe climate events, such as heatwaves, torrential rainfall, prolonged droughts, and violent storms, has emerged as a worldwide issue in recent times. This concerning pattern greatly enhances the susceptibility of communities, especially in urban and rural regions, presenting a huge hazard and requiring thorough planning. Numerical climate models provide valuable predictions of shifting weather patterns, but precisely identifying and forecasting extreme occurrences remains a serious challenge (Flaounas et al. 2022; Mezősi 2022; Olaoluwa et al. 2022).

Statistical conventional approaches remain center of analyzing extreme climatic conditions, offering vital understanding of the features and patterns of these occurrences. The methods encompass Extreme Value Theory (EVT), Block Maxima (BM), and Peaks-Over-Threshold (POT) approaches, Generalized Extreme Value (GEV) distribution,

✉ Marwa S. Moustafa
  marwa@narss.sci.eg

  Mohamed R. Metwalli
  m.roshdym@narss.sci.eg

  Sayed A. Mohamed
  sayed.abdo@narss.sci.eg

  Barma Shovan
  shovan@iiitg.ac.in

[1] Data Reception, Analysis and Receiving Station Division, National Authority for Remote Sensing and Space Science, Cairo, Egypt

[2] Indian Institute of Information Technology, Guwahati, Assam, India

Index Threshold Method (ITM), and Partial Duration Series (PDS) approach (Hulme 2014). Despite their robust theoretical basis, simple implementation and understandable outcomes, their limitations include susceptibility to data quality issues, reliance on the assumption of stationarity, subjectivity in the choosing of thresholds, and restrictions associated with parametric models. As climate change continues to intensify, it is becoming increasingly important to develop more robust and sophisticated methods for extreme climate patterns. Hybrid approaches, such as TECA (Rübel et al. 2012), introduced as a possible solution for overcoming these constraints by merging the advantages of conventional with more sophisticated approaches.

Deep learning (DL), which draws inspiration from the brain architectural, has catalyzed a revolution in artificial intelligence. Diverging from conventional approaches, DL leverages extensive datasets to learn intricate patterns, thereby facilitating breakthroughs in domains such as computer vision and natural language processing (Kaur and Singh 2022; Zaidi et al. 2022). Numerous studies have laid the groundwork for exploring the profound influence of deep learning in various domains. For instance, the potential of deep learning in reservoir characterization (Zhang et al. 2020) was demonstrated to integrate seismic and electromagnetic data for improved mapping. Extending beyond the realm of image analysis, (Afzal et al. 2023) delves into the extensive landscape of visualization and visual analytics techniques empowered by deep learning. Additionally, deep learning was extended in environmental monitoring (Hittawe et al. 2019) specifically focusing on anomaly detection in sea surface temperatures. The remarkable performance of hash deep learning model in multi-label remote sensing image retrieval had been investigated (Moustafa et al. 2020). The convergence between DL and statistical methods in optimizing traffic management solutions was explored (Harrou et al. 2021). By drawing inspiration and insights from these prior works, the present research endeavors to contribute to the ever-evolving landscape of DL applications.

This paradigm shift has the potential to reshape the extreme weather analysis (Chen, Zhang et al. 2020). An ensemble of deep learning methods were utilized to detect cyclones (Kumler-Bonfanti et al. 2020) using a twenty-year dataset of simulated data. Another convolutional neural network (CNN) architecture (Kim et al. 2017) was developed to accurately pinpoint severe occurrences, achieving a remarkable accuracy rate of 99.98%. ClimateNet (Kashinath et al. 2021) was created as a baseline dataset for annotating the Community Atmospheric Model (CAM5.1). A deep convolutional neural network (CNN) was specifically designed to categorize the intensity of Tropical Cyclones (TCs) using infrared geostationary satellite data. The Single

Shot MultiBox Detector (SSD) was used to pinpoint Extratropical Cyclones (ETCs) in the northern hemisphere (Shi et al. 2022). A refined Deep Convolutional Neural Network (DCNN) (Tong et al. 2022) was introduced to accurately detect tropical cyclone fingerprints in the northern Pacific basin. These approaches showed similar levels of performance in identifying cyclones. In (Pang et al. 2021), the GAN was combined by transfer learning to detect tropical cyclones from meteorological images. A novel transfer learning model (Wang and Li 2023) was proposed to detect center of TC by harnessing knowledge from a vast image dataset and fine-tuning it for TC-specific features, the model achieves a remarkable 14.1% boost compared to traditional methods. Another innovative CNN model was introduced to pinpoints the centers of low-intensity tropical cyclones (Wang et al. 2024) by incorporating physical and historical data alongside satellite imagery, the model captures crucial evolutionary trends in storm structure, achieving exceptional localization accuracy. The Thermal InfraRed (TIR) Atmospheric Sounding Interferometer (IASI) on the Metop satellite was used to detect TCs in the North Atlantic Basin using YOLOv3 (Lam et al. 2023). The model was evaluated at 0.1 and 0.5 intersection over union (IoU) using the Average Precision (AP) measure. Though promising with an AP of 78.31% at the lower level, precision dropped to 31.05% at the higher barrier.

Nevertheless, the limited resolution of climate data is inadequate for detecting variations in small climatic zones, such as India, which may experience cyclones of varying magnitudes (Dabhade et al. 2021). Single Image Super-Resolution (SISR) may be used to generate artificially enhanced High Resolution (HR) images, which can subsequently be employed to improve the accuracy of object detection systems (Park et al. 2003; Anwar et al. 2020; Liu et al. 2021). Dong et al. introduced the pioneer deep learning Convolutional Neural Network-based Super-Resolution (SRCNN) method. More complex CNN architectures was introduced, including VDSR (Kim, Kwon Lee et al. 2016) and LapSRN (Lai et al. 2018) which resulted in the production of SR images with high Peak Signal-to-Noise Ratio (PNSR) values. On other hand, generative adversarial networks (GANs) has displayed enhancing the perceptual quality and minimize smoothing of reconstructed HR images (Lei et al. 2019; Moustafa and Sayed 2021). Single Image Super-Resolution Generative Adversarial Networks (SRGANs) leverage the collaborative power of two subnetworks: a generator and a discriminator (Ledig et al. 2017). The generator network aimed to reconstruct HR images from their Low Resolution (LR) input counterparts. On other hand, the discriminator network anticipates whether the obtained image is ground truth HR or not. After enough training, the generator creates HR images that mimic ground truth.

Recently, attention-based models or transformers (Lu et al. 2022) could be better feature extraction in local climate zones, these techniques have shown great potential in various computer vision tasks, including super-resolution. Attention mechanisms enable the model to focus on relevant image regions and capture long-range dependencies, which can be beneficial in extracting meaningful features from local climate zones. By attending to relevant spatial or temporal regions, attention-based models can effectively model the complex relationships and patterns within local climate zones, leading to improved performance. Transformers, in particular, have gained significant attention in recent years due to their success in natural language processing and image recognition tasks (Moustafa and Sayed 2021). Despite their booming performance there are some challenges to be considered when utilizing attentions or transformers in case of very large volumes of data: (1) Computational Cost: Transformers heavily rely on attention mechanisms, which involve comparing every element in the input sequence to each other. This leads to quadratic complexity, meaning their computational cost grows with the square of the data size. While techniques like sparse attention and efficient implementations can alleviate this issue, it still remains a hurdle for extremely large datasets. (2) Memory Bottlenecks: Processing entire large datasets at once might not be possible due to memory limitations. Transformers usually need the entire input sequence in memory for attention calculations, making handling massive datasets in a single batch challenging. (3) Training Stability: Training transformers effectively requires careful hyperparameter tuning, especially with large datasets. Learning rate schedules, batch sizes, and optimization algorithms need to be adjusted to ensure convergence and avoid divergence (Khan et al. 2022).

Traditional weather models struggle to accurately identify cyclones due to two key hurdles: (1) their limited resolution, meaning they cannot capture the fine details of cyclones, and (2) the natural variation in cyclone size and structure. These limitations can lead to missed identifications, particularly for smaller or weaker cyclones, impacting weather forecasting and early warning systems. This study tackles these challenges to improve cyclone detection for better weather forecasting and early warning systems. To address this challenge, we propose a novel end-to-end approach that combines edge-enhanced super-resolution (EESRGAN) with a Faster RCNN detector. The proposed framework comprises three subnetworks: a generator, a discriminator, and a Faster RCNN detector. We utilize residual-in-residual dense blocks (RRDB) to extract discriminative features for accurate cyclone detection. We systematically evaluated the proposed approach on Community Atmospheric Model (CAM5.1) image data, considering seven distinct variables.

Extensive experiments were conducted to assess the effectiveness and efficiency of the framework using four metrics: precision, recall, intersection over union, and average precision. The key contributions of this work are:

- The proposed end-to-end framework comprised of a generator network equipped by residual-in-residual dense block (RRDB) and discriminator containing Faster RCNN detector.
- The generator network employs residual-in-residual dense blocks (RRDB) which provides several advantages compared to traditional convolutional blocks allowing extraction of discriminative features. In addition, the skip connections of RRDB enhances gradient flow during training.
- The discriminator network contains Faster RCNN object detection where the gradient of the detection loss function is propagated back to update the parameters of the generator network.
- The proposed EESRGAN with can efficiently detects the tropical cyclone (TC) event which has been verified for India.
- Seven critically important variables for cyclone event analysis from Community Atmospheric Model (CAM5.1) image data have been taken into account for systematically assessment of the proposed network.

The remainder of this paper is structured as follows: Sect. 2 introduces the proposed architecture for the Indian cyclone detection. Experimental setting, and results discussion are presented in Sect. 3. Section 4 concluded the findings.

Methodology

Figure 1 depicts the overall structure of the proposed framework. The proposed framework is composed of two main subnetworks: generator (G), extended discriminator network with object detector network. During training, the gradient of the detection loss function is propagated back to update the parameters of the generator network (G). This backpropagation process guides the generator to refine its image reconstruction, enhancing realism and sharpness in the output images, ultimately improving the performance of the overall framework. On the other hand, the discriminator network (D) aimed to distinguish between ground truth images and estimated SR images whereas, the detector network Leverages the enhanced quality of the SR images created by the generator (G) to perform accurate object detection.
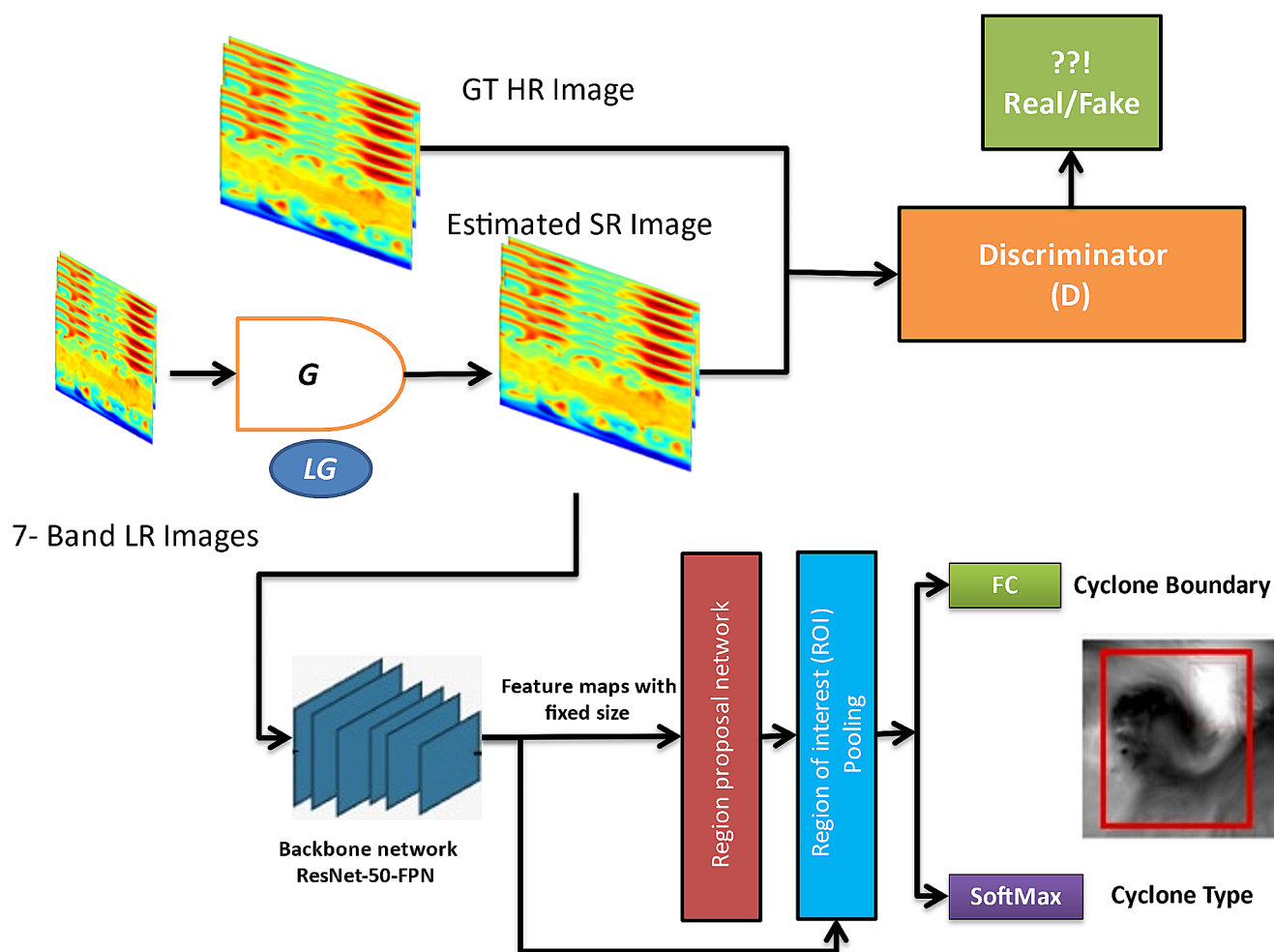
Fig. 1 The overall structure of the proposed end-to-end cyclone detection network

Generator

Building upon the EESRGAN architecture (Jiang et al. 2019), we utilized the generator structure outlined in Fig. 2(a). The key innovation lies in replacing the standard convolution blocks with Residual in Residual Dense Blocks (RRDBs) (Song et al. 2018), as detailed in Fig. 2(b, c), to enhance the generator performance. The inclusion of RRDB in the network offers several advantages over traditional convolutional blocks; (1) Improved feature representation: RRDB architecture enables complicated and discriminative feature extraction and representation. The residual connections in the RRDB let the network capture and convey low-level and high-level information, improving feature learning. (2) Deeper network capacity: RRDB allows for deeper network building without many additional parameters. Densely linking each layer to all subsequent layers in the block achieves this. Thus, the RRDB may take use of deep architectures improved representational capacity and ability to learn abstract features. (3) Efficient gradient flow:

RRDB skip connections improve gradient flow during training. The RRDB combats the vanishing gradient problem by shortening gradient propagation through the network, enabling faster and more stable convergence during training. To mitigate computational complexity, curtail undesirable artifacts, and bolster generalization capabilities in scenarios where training and testing data exhibit substantial statistical disparities, batch normalization layers were judiciously excluded from the architecture (Karras et al. 2019).

We stacked 16 RRDB block with dense connections to increase network capacity. To enhance parameter learning, a Parametric Rectified Linear Unit (PReLU) (El Jaafari et al. 2021) was implemented in conjunction with residual scaling promoting training stability. The PReLU activation function is an extension of the traditional rectified unit, offering improved model fitting without significant additional computational cost or overfitting concerns. By dynamically learning the rectifier parameters, PReLU enhances accuracy without imposing a noticeable burden on computational resources (He et al. 2015). The initial super-resolution (SR)
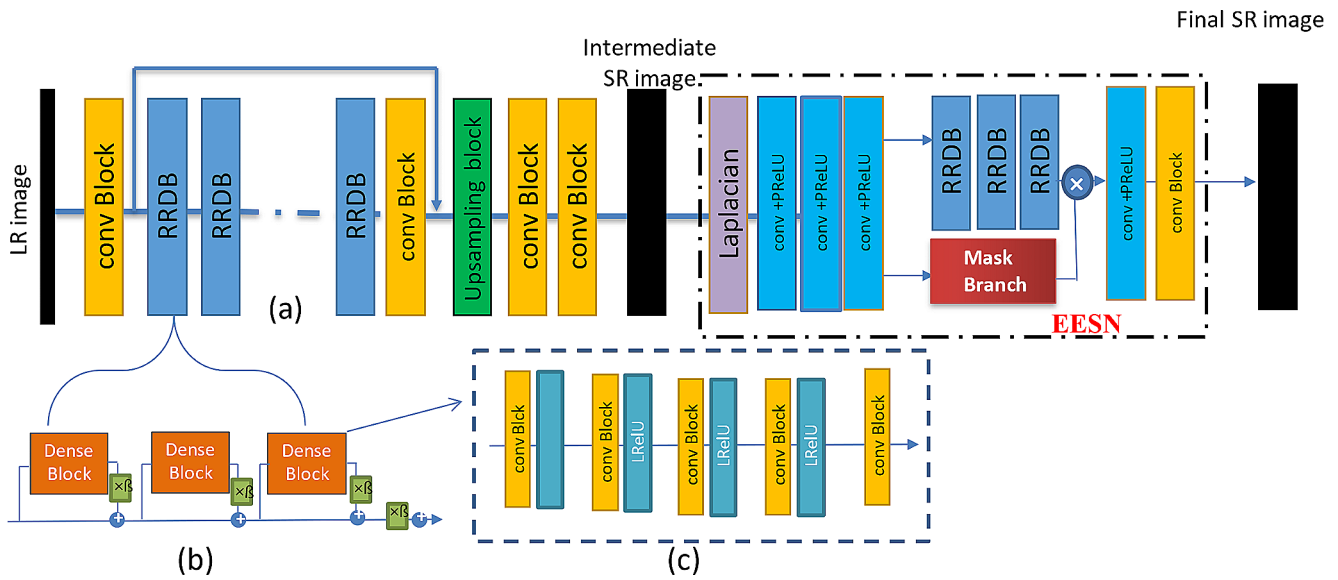
Fig. 2 (a) The generator network architecture with RRDBs and EESN network. (b) Residual in Residual Dense Block (RRDB) where ß is residual scaling parameter. (c) The architecture of the dense block

image generated by the network exhibits undesirable artifacts manifested as noisy edges. The Edge Enhancement Sub-Network (EESN) mitigates these artifacts by replacing the noisy edges with "EESN-purified" edges, yielding the final refined SR image. During training, the generator (G) aims to map the input LR image onto the HR image space, replicating the characteristics of the ground truth HR image. While the intermediate generator output possesses sharp yet jagged edges, the final SR image retains crisply defined contours devoid of spurious artifacts.

The EESN network aims to remove the noise from the initial obtained SR images and sharpen the edges. Laplacian operator is used to extract edges in the image then this edge information is transferred via convolutional, RRDB, and up sampling blocks. Following the architectures in (Jiang et al. 2019), the mask branch equipped by sigmoid activation aimed to eliminate edge noise. Finally, the refined edges are added to the input image. It worth noting that all dense block in EESN were replaced by RRD blocks to improve the performance. The generator network (G) consisted of 16 RRDB while the EEN (Enhanced Encoder Network) employed five blocks. The overall generator (G) cost function (L_G) is defined as in Eq. (1).

$$L_G = \lambda 1 L_{MSE} + \lambda 2 L_{VGG} + \lambda 3 L_{Adversarial} + \lambda 4 L_{EESN} \quad (1)$$

where we prioritized content accuracy ($\lambda 1 = 1$), downplayed perceptual details ($\lambda 2 = 0.001$), used moderate adversarial loss ($\lambda 3 = 0.01$), and emphasized edge preservation ($\lambda 4 = 5$).

The mean square loss L_{MSE} defined in Eq. (2), is the popular in SISR as it is known to increase the PSNR value.

$$L_{MSE} = \frac{1}{r^2 WH} \sum_{w=1}^{rW} \sum_{h=1}^{rH} \left(I_{HR,(w,h)} - G\left(I_{LR}\right)_{w,h}\right)^2 \quad (2)$$

Where r represents the upsampling factor, W and H denoted HR image width and height, respectively. I_{HR}, $G\left(I_{LR}\right)$ stands for the ground truth HR image and SR image.

The L_{VGG} loss, defined in Eq. (1), was originally introduced by (Ledig et al. 2017) to create visually appealing and detailed images. However, their VGG-19 network (Simonyan and Zisserman 2014) was trained on the ImageNet dataset, which differs significantly from the domain of satellite images used in this work. To address this mismatch, we fine-tuned the pre-trained VGG-19 network following the procedure in (Jiang et al. 2019), as shown in Eq. (3). This allows us to calculate the Euclidean distance between the feature maps extracted from the high-resolution (HR) image (I_{HR}) and the super-resolution (SR) image ($G\left(I_{LR}\right)$) using the fine-tuned network.

$$L_{vgg} = \frac{1}{W_{i,j} H_{i,j}} \sum_{w=1}^{W_{i,j}} \sum_{h=1}^{H_{i,j}} \left(\varnothing_{i,j}(I_{HR})_{w,h} - \varnothing_{i,j}(G\left(I_{LR}\right))_{w,h}\right)^2 \quad (3)$$

where $W_{i,j}$ and $H_{i,j}$ indicate the width and height of the corresponding feature map respectively.

The discriminator network loss function can be formulated as in Eq. (4):

$$L_{Adversarial} = -\log\left(D\left(G\left(I_{LR}\right)\right)\right) \quad (4)$$

Finally, the EESN network loss function is formulated as defined in Eq. (5):

$$\boldsymbol{L}_{EESN} = \mathbb{E}_{I_{SR}} \left[\mathcal{P} \left(I_{HR} - I_{SR} \right) \right] + \mathbb{E}_{I_{edge\_HR}} \left[\mathcal{P} \left(I_{edge\_HR} - I_{edge\_SR} \right) \right] \quad (5)$$

where, the first term measures the pixel-wise difference between the generated SR image (I_{SR}) and the ground truth HR image (I_{HR}). $\mathcal{P}$ represents the Charbonnier penalty function. The second term focuses on the preservation of edges in the super-resolved image I_{edge_HR} and I_{edge_SR} denotes the edge maps of the HR and SR images, respectively.

Discriminator

Building on the success of (Jiang et al. 2019), we designed a robust discriminator network crucial for achieving high-quality super-resolution. This network consists of eight convolutional layers with 3×3 filters, progressively increasing in number from 64 to 512, inspired by VGGs architecture. To further enhance discrimination, we incorporate VGG-19 features and leverage Faster R-CNN (Girshick 2015) for object detection within the discriminator, enabling it to effectively differentiate between super-resolved and high-resolution images.

Faster R-CNN (Girshick 2015), developed by Microsoft as a two-stage object detector, has gained significant popularity for its effectiveness in analyzing satellite images. The model comprises of two interconnected subnetworks, namely the region proposal network (RPN) and the detector. The primary task of the RPN is to identify and extract region-specific characteristics associated with objects of interest. Subsequently, these identified regions, along with their corresponding feature maps, are utilized by the detector's classifier and bounding box regressor. To obtain a fixed-size feature map encoding spatial relationships between features, a fully convolutional network known as the backbone is employed. The RPN can accommodate feature maps of any size, leading to the generation of numerous rectangular object proposals. For each sliding position within the feature map, the RPN generates K predictions encompassing diverse sizes and aspect ratios. The regression and classification layers produce four location coordinates and corresponding scores. Consequently, the resulting feature map of size n × n × k represents the regions of interest (ROIs). Through the process of minimizing and refining regional proposals, the RPN contributes to improvements in both speed and accuracy. Several studies (Magdy et al. 2022; Wang and Leelapatra 2022) have demonstrated the superiority of ResNet-50-FPN as the backbone network for this task. This choice stems from its demonstrably higher precision compared to VGG-19 and the baseline ResNet-50 architecture without FPN.

The overall discriminative network (D) which minimizes the cost function is defined in Eq. (6)

$$L_{D\_f} = L_D + L_{OD} \quad (6)$$

The Adversarial decimator network (D) loss function is defined in Eq. (7)

$$L_D = log(D \left(I_{HR} \right) - \log(1 - D(G \left(I_{LR} \right))). \quad (7)$$

where $\boldsymbol{I_{HR}}$ denotes the reference High resolution image, $\boldsymbol{I_{LR}}$ denotes the Low-resolution image

The object detection network Faster RCNN loss function is defined as in Eq. (8).

$$L_{OD} \left(\{ p_i \} , \{ t_i \} \right) = \frac{1}{N_{cls}} \sum_i L_{cls} \left(p_i, \dot{p}_i \right) + \lambda \frac{1}{N_{reg}} \sum_i \dot{p}_i L_{reg} \left(t_i, \right) \quad (8)$$

Where $\boldsymbol{p_i}$ is the predicted probability of anchor, $\boldsymbol{\dot{p}_i}$ is the ground-truth label (1: anchor is positive, 0: anchor is negative), $\boldsymbol{\lambda}$ is balancing parameter, $\boldsymbol{t_i}$ is the predicted box, $\boldsymbol{\dot{t}_i}$ is the ground-truth box.

Training strategy

To better suit climate data characteristics through model training, we depended on data normalization and scaling as an important preprocessing step to ensure that the input seven variable are on a similar scale, which can improve the training process and model performance. We applied Min-Max Scaling to physical climate parameter data which rescaled each variable to a fixed range, typically between 0 and 1. This is achieved by subtracting the minimum value of the variable and dividing by the difference between the maximum and minimum values, as defined in Eq. (9):

$$\boldsymbol{x_{norm}} = (\boldsymbol{x}?\boldsymbol{x_{min}})/(\boldsymbol{x_{max}}?\boldsymbol{x_{min}}) \quad (9)$$

To mitigate computational demands associated with training the proposed model on the entire dataset, we employed a random sampling technique. This resulted in the creation of a smaller, representative subset of data that maintained balanced representation across all four class types, thereby ensuring training efficiency and generalizability.

Instead of training the model from scratch, we benefited from Transfer Learning and adopted the weights from (Jiang et al. 2019) as the initial weights then completed training on climate dataset. This approach leverages the knowledge learned from the pre-training phase and reduces the amount of training required on the target dataset.

Table 1 Class labels for the type of Extreme climate (Cyclone) events

Label	Type of Extreme Climate Event
1	Tropical depression
2	Tropical Cyclone
3	Extratropical Cyclone
4	Atmospheric River

Table 2 Dataset-Splitting for training, validation, and testing

	Training set	Validation set	Testing set
Years	1979–2000	2000–2002	2002–2005
#samples	20,616	11,907	9999

Dataset

The detection task utilized a large-scale Extreme Climate Event dataset (Kashinath et al. 2021) specifically designed for climate analysis. This dataset contains ground truth information for four types of extreme climate events and was generated using the Parallel Toolkit for Extreme Climate Analysis (TECA), which leverages prior knowledge of climate analysis to create accurate labels. The dataset is extensive and stored in a yearly HDF5 file format with a size of 62GB. Each file consists of two variables: "images" and "boxes." The "images" variable has a shape of (1460, 16, 768, 1152), representing 1460 images with 16 channels, a length of 768, and a width of 1152. On the other hand, the "boxes" variable has a shape of (1460, 15, 5), signifying 1460 images with 15 ground-truth boxes per image. The 5 coordinates in each box correspond to x_min, x_max, y_min, y_max, and the associated class label. Table 1 provides a detailed mapping of the class labels for four cyclone classes. For Cyclone detection, the study focused on seven

critically important variables. A sample of the climate dataset is illustrated in Fig. 3. To narrow down the data to a specific region, the dataset was clipped to the extent of the Indian subcontinent. To avoid overfitting and ensure generalizability, we split the data into three different sets: training (50%), validation (27%), and testing (23%) as shown in Table 2. Finally, the generalizability of the model was evaluated on a completely unseen test set (23% of the data), which was never exposed to the model during training or validation.

Experiments setting

The computational environment for all experiments consisted of an Intel Core i7 processor equipped with an NVIDIA Quadro RTX 6000 graphics card (NVIDIA, 2023) and 192 GB of RAM. PyTorch (Paszke et al., 2019) served as the deep learning framework under Windows 10, with CUDA 11.0 and CUDNN 5.1 providing GPU acceleration. Stochastic gradient descent (SGD) with momentum (Ruder, 2017) was employed as the optimizer, utilizing momentum
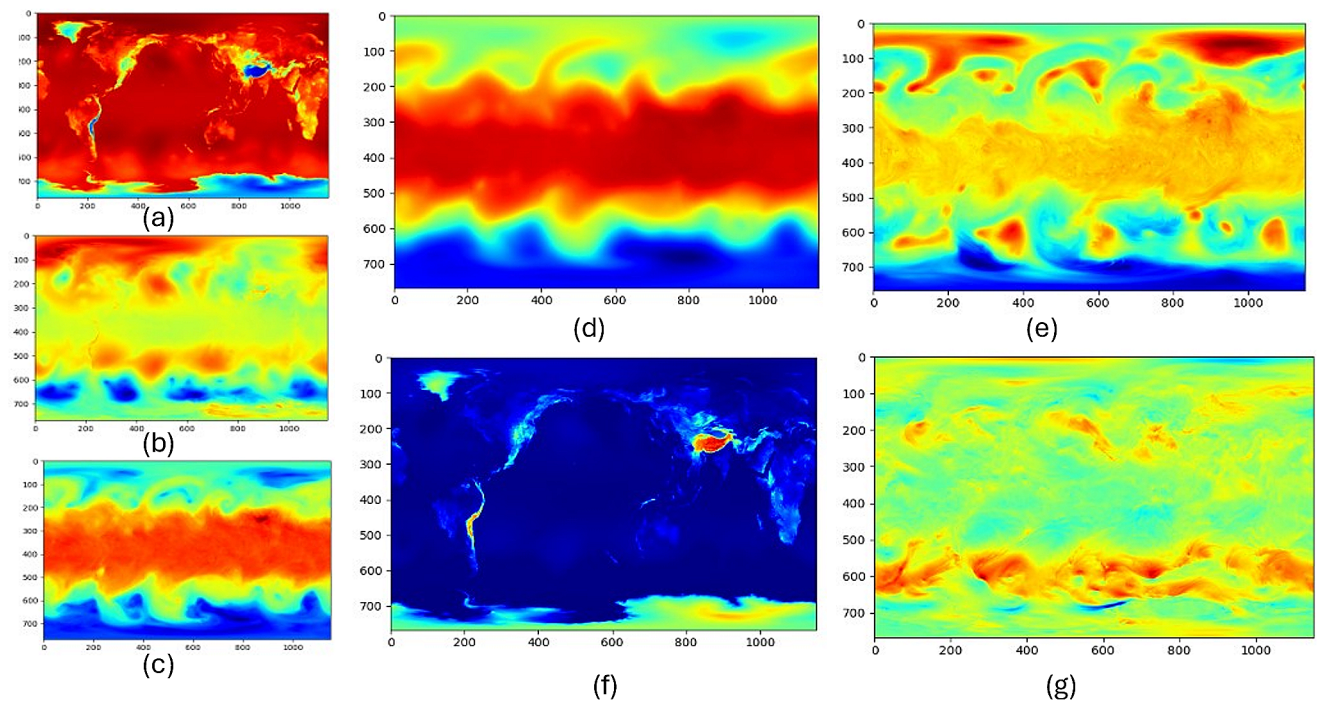

Fig. 3 Worldwide climate parameters generated using CAM5 at 1/6/2001; (a) Sea level pressure (PSL), (b) Temperature at 200 mbar pressure surface (T200), (c) Temperature at 500 mbar pressure sur-face (T500), (d) Zonal wind at 850 mbar pressure surface (U850), (e) Meridional wind at 850 mbar pressure surface (V850), (f) Z100, and (g) Geopotential Z at 200 mbar pressure surface (Z200)

values of 0.9 and 0.999. The learning rate was set to 1×10^{-4}. A batch size of 16 was chosen for training efficiency. The training took 96 hours for 200 epochs. Faster R-CNN infers four images/second. Figure 4 shows the proposed network training and validation loss curves.

Low-resolution (LR) training images were obtained by downsampling ground-truth images using bicubic interpolation to a size of 128×128 pixels. Notably, the experiments were conducted with a 4x scaling factor between the SR outputs and the ground-truth images. During training, both the high-resolution (HR) and low-resolution (LR) images were rescaled to the value ranges of [-1, 1] and [0, 1], respectively. The VGG-19 network (Simonyan and Zisserman 2014) was adapted to accept seven input channels instead of the original three by prepending additional zero channels.

To assess the performance of our proposed architecture, we utilized commonly used metrics for object detection tasks, namely precision, recall, and IoU (Intersection over Union). These metrics are defined as follows:

$$Precision = TP / (TP + FP) \tag{10}$$

$$Recall = TP / (TP + FN) \tag{11}$$

$$IoU = TP / (TP + FN + FP) \tag{12}$$

Where, TP represents true positives, FP represents false positives, and FN represents false negatives. True positives (TP) occur when the predicted cyclone type matches the ground-truth, true negatives occur when the predicted and ground-truth are both negative, false positives occur when the predicted is positive, but the ground-truth is negative, and false negatives occur when the ground-truth is positive, but the predicted ground-truth is negative.

Results

First, we evaluated the SSD and Faster RCNN detectors on both LR and HR images. VGG16 backbone was employed for SSD network while ResNet-50-FPN was employed for Faster R-CNN (FRCNN) detector. For each detector, the training and the testing was conducted on LR and HR images. Table 3 summarizes the obtained detectors results training/testing. Faster R-CNN achieved 79.7% AP when adopting only LR images in training and testing. For both detectors, the obtained accuracy declined when trained on HR images and tested with their LR counterparts.

One can observe that the accuracy of both object detectors excelled in scenario of utilizing HR images in training and testing. The accuracy achieved 74.1% and 81.9% in terms of AP for SSD and Faster RCNN, respectively. This

Fig. 4 The loss curve per epoch for Weather dataset. (a) Generator network, (b) discriminator network

Table 3 The obtained detection results in terms of AP (average precision) on LR and HR images

Model	Image Resolution (Training -Testing)	Climate dataset
SSD	LR-LR	61.8%
	HR-LR	58%
	HR-HR	74.1%
Faster-RCNN	LR-LR	79.7%
	HR-LR	72.5%
	HR-HR	81.9%

Table 4 The detection results in terms of AP on the obtained SR images by the proposed approach, CNNSR, SRGAN networks and bicubic upsampling. Both Detectors are trained separately with both SR and HR images

Model	Image Resolution (Training -Testing)	Climate dataset
Bicubic + SSD	SR-SR	58.9%
	HR-SR	61.3%
Bicubic + FRCNN	SR-SR	62.6%
	HR-SR	66.8%
SRGAN + SSD	SR-SR	80.95%
	HR-SR	83.23%
SRGAN + Faster RCNN	SR-SR	81.68%
	HR-SR	77.92%
CNNSR + SSD	SR-SR	72.83%
	HR-SR	73.78%
CNNSR + Faster RCNN	SR-SR	78.84%
	HR-SR	84.57%
Proposed framework + SSD	SR-SR	84.2%
	HR-SR	86.08%
Proposed framework + Faster RCNN	SR-SR	86.1%
	HR-SR	88.2%

illustrates how image resolution affects object identification quality.

Next, we compared the proposed EESRGAN architecture, CNNSR, SRGAN, and 4× HR estimate from LR image using bicubic upsampling. We trained each network separately. For the assessment, we compared detectors trained on SR images obtained from these approaches versus detectors trained directly on HR images. Table 4 demonstrated that the proposed framework showed the highest results, approaching close to HR-only detection rates. After training, the proposed framework may be immediately used to LR images without HR data and get excellent results. CNNSR and SRGAN have better AP compared with traditional bicubic in prepare LR images. Overall, the proposed framework outperformed the other approaches in climate dataset.

Next, we trained the proposed approach using end-to-end fashion. The discriminator network and Faster RCNN detector were served as the discriminator for the proposed architecture. As a result, the Faster RCNN detector loss being

Table 5 The detection results in terms of AP using end-to-end training for both detectors

Model	Image Resolution (Training -Testing)	Climate dataset
Proposed framework + Faster RCNN	SR-SR	86.3%
Proposed framework + SSD	SR-SR	84.2%

backpropagated into the SR network in order to enhance the network learning during training. The LR-HR images pairs were utilized to train the proposed framework, and the obtained SR images were used to train Faster RCNN detector. In testing, only LR images were feed to the generator to create SR image to be feed to detector network. Table 5 indicates that the proposed approach improved outcomes compared with training the detector network with SR from other SR approaches.

Figure 5 shows the precision-recall curves of the proposed approach, with and without end-to-end training, in comparison to stand-alone Faster-RCNN using LR training/testing images. Precision and recall were determined using IoU = 0.5. One can observe that the proposed framework has superior values in precision and recall than standalone R-CNN models. End-to-end training improved the proposed method performance.

For better comparison and visualization, we plot (1-recall) for X-axes and (1- precision) for Y-axes, as shown in Fig. 6. One can observe that, all detector techniques achieved superb performance for the four categories despite the size of the cyclone. Overall, the proposed method is very effective for detecting extreme climate event in the climate dataset.

The proposed approach yields an SR image with improved visual clarity and detail, thanks to adversarial learning's ability to simultaneously sharpen images and increase detection precision as shown in Fig. 7. In brief, the effectiveness of joint training (detector network and discriminator), improves the obtained SR image images both visually and in detection measures. Also, the proposed approach achieved a considerable improvement compared with other approaches by about 1.5% in terms of average precision (AP).

Discussion

The proposed approach, when tested with SR images generated by itself, improved the detection outcomes compared to training the detector network with SR images from other approaches. It was evaluated using SSD and Faster R-CNN as the detector networks. SSD utilized Vgg16 backbone, while Faster R-CNN employed ResNet-50-FPN.
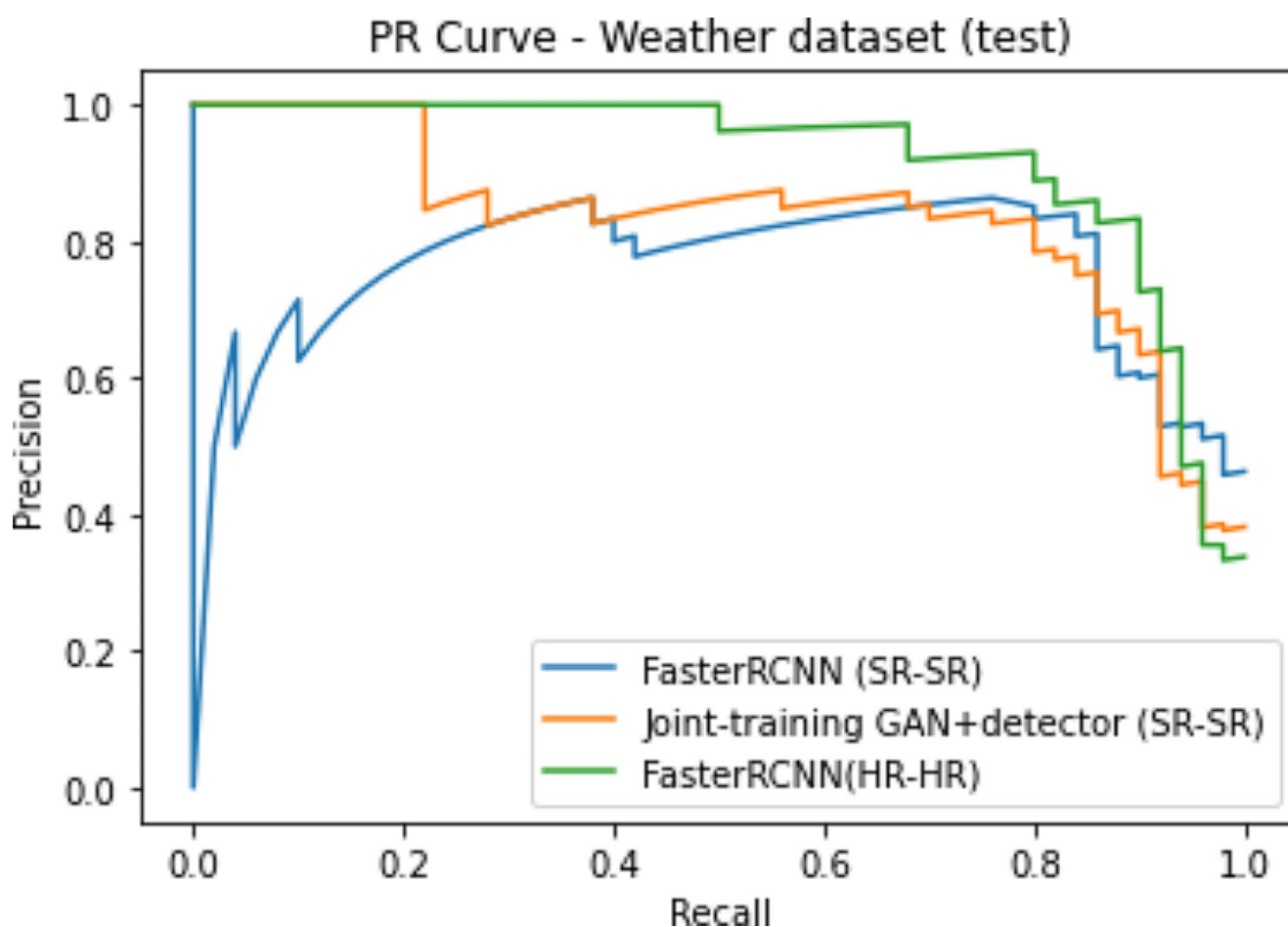
PR Curve - Weather dataset (test)

Fig. 5 The precision-recall curves for the proposed technique, with and without end-to-end training, in comparison to stand-alone Faster-RCNN.

The accuracy of both detectors decreased when tested on LR images. However, the proposed approach utilizing Faster R-CNN and SSD achieved 81.9% and 74.1% AP. A comparison was conducted between the EESRGAN architecture, CNNSR, SRGAN, and bicubic upsampling for training detectors. The proposed approach showed the highest results, approaching the performance of HR-only detection. CNNSR and SRGAN outperformed traditional bicubic upsampling in preparing LR images. Overall, the proposed framework surpassed other approaches in the climate dataset.

Therefore, there are still an open door to integrate recent deep learning-based revelational models to boost the precision of detection in the future. Technically, three main issues had to be addressed in the future. (1) the deep learning-based detection methods mainly used pre-trained, but the nature of climate data is different. Although the large volume of used data in training limited computation affects the model ability to learn from data. (2) The obtained results in Table 5 demonstrate the rather poor performance of the detection utilizing the super resolution images, especially using SSD detectors. The reason for this may be due to limited SSD

ability to extract relevant features especially in local climate zone. The transformer and attention-based model could help in capturing the discriminative features of cyclone events efficiently. (3) Unlike data-driven deep learning-based, the traditional detection techniques employ physics parameters which deep learning-based algorithms disregard. Many studies strive to combine physics into the deep learning for climate forecasting to preserve the benefits of numerical and deep learning-based approaches to enhance deep learning-based TC track detection.

Conclusion

Deep learning can unlock the power of climate data by analyzing its low-resolution obtained from numerical models instead of regional high-resolution counterparts. The proposed approach tackles the challenges of the computational burden and information overload to obtained high-resolution regional data from weather numerical models. The integration between deep learning and numerical data can offer faster analysis, targeted feature extraction, uncovering
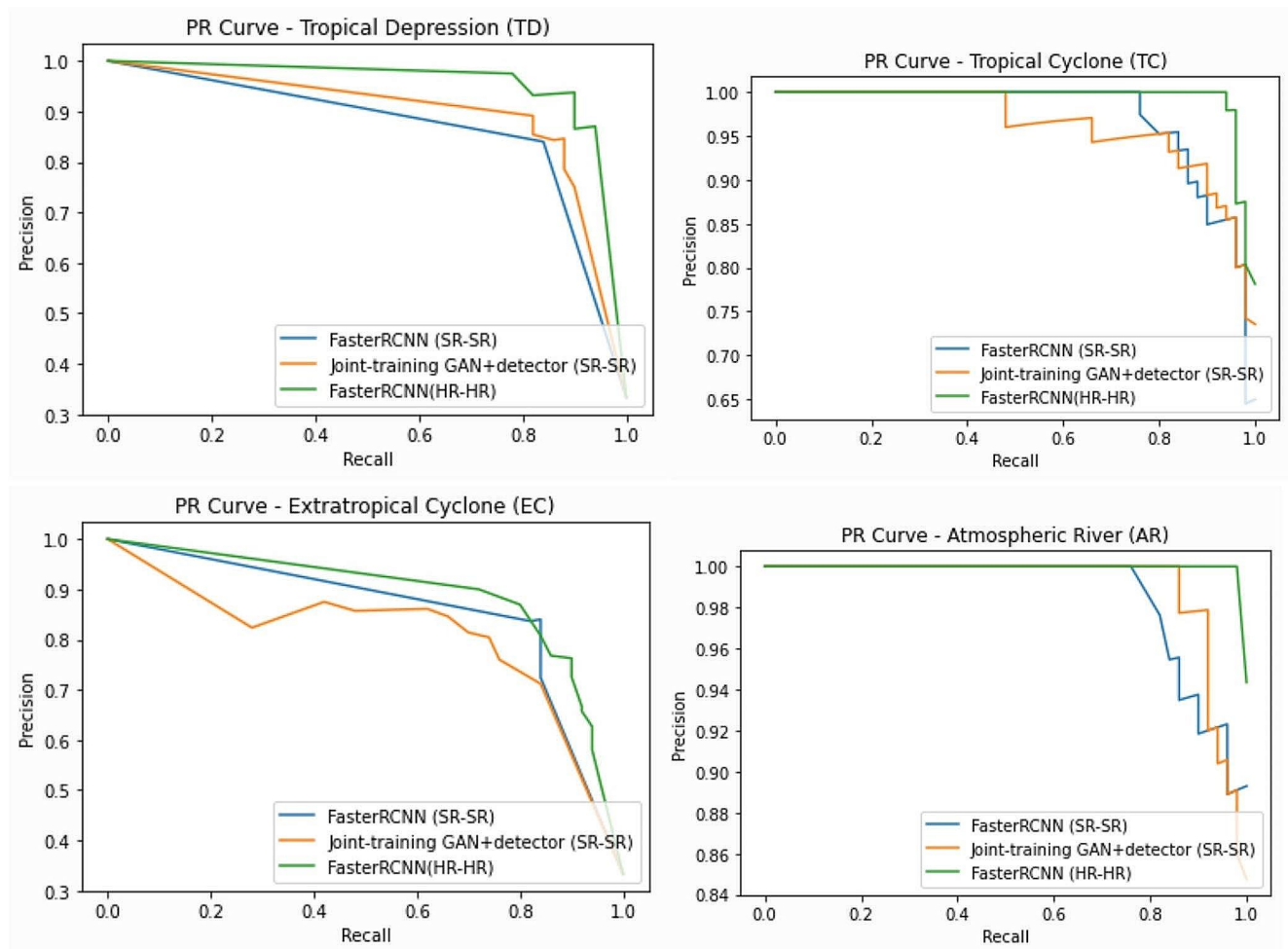
Fig. 6 Precision vs. Recall curves on climate dataset for Tropical Depression (TD), Tropical Cyclone (TC), Extratropical Cyclone (EC), and Atmospheric River (AR), respectively

hidden patterns, broader applicability, and real-time insights. While acknowledging potential information loss and training data challenges, this approach empowers professionals with efficient, scalable, and insightful climate analysis for informed decision-making.

The intensifying impacts of climate change necessitate enhanced detection of extreme weather events (EWEs), particularly cyclones. In the Indian subcontinent, the demonstrably heightened frequency and severity of cyclones necessitate reliable detection methods for mitigating casualties and economic losses. However, traditional detection approaches face significant challenges due to the inherent limitations of low-resolution data. Deep learning models present a promising solution by enabling precise identification of cyclone boundaries crucial for regional impact assessment using global climate model data. By leveraging the power of deep learning, we can significantly improve cyclone detection capabilities and contribute to refined risk mitigation strategies in the vulnerable Indian subcontinent. This paper introduces an edge-enhanced super-resolution

generative adversarial network (EESRGAN) coupled with an end-to-end detector network. The proposed approach comprises a generator, discriminator, and Faster RCNN detector network augmented with residual-in-residual dense blocks (RRDB). This architecture effectively extracts precise cyclone patterns, facilitating accurate boundary detection. Extensive experiments were conducted on Community Atmospheric Model (CAM5.1) data using only seven variables and employed four evaluation metrics: precision, recall, intersection over union, and mean average precision to assess the proposed approach. The results demonstrated remarkable effectiveness, achieving an accuracy of 86.3% and an average precision (AP) of 88.63%. Furthermore, the proposed framework outperformed baseline object detector methods.
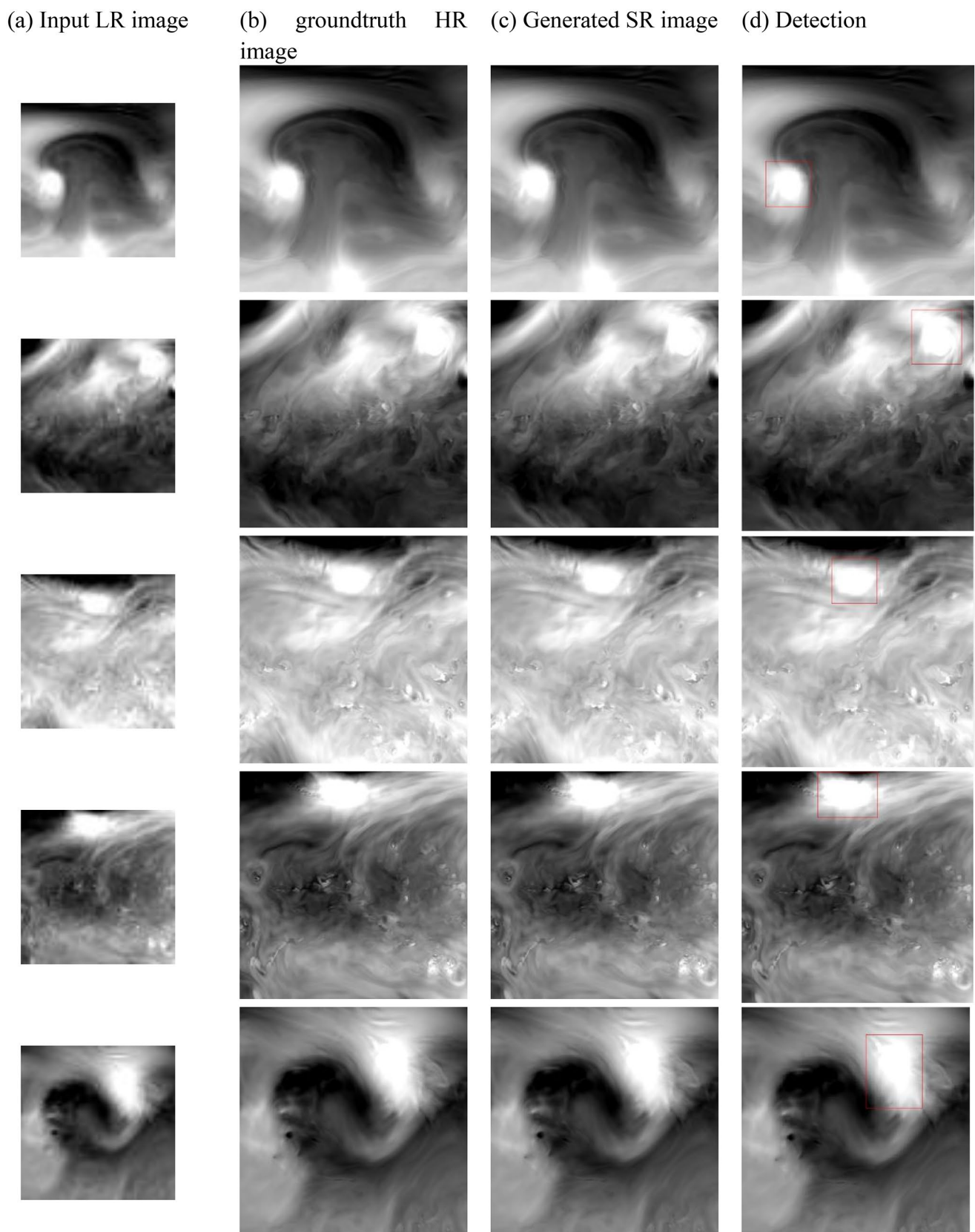
Fig. 7 Examples of the obtained SR images generated from LR images in (a,b). Results of improved edge detection in (c, d)

Declarations

References

Afzal S, Ghani S, Hittawe MM, Rashid SF, Knio OM, Hadwiger M, Hoteit I (2023) Visualization and visual analytics approaches for image and video datasets: a Survey. ACM Trans Interact Intell Syst 13(1):1–41

Anwar S, Khan S, Barnes N (2020) A deep journey into super-resolution: a survey. ACM Comput Surv (CSUR) 53(3):1–34

Chen R, Zhang W, Wang X (2020) Machine learning in tropical cyclone forecast modeling: a review. Atmosphere 11(7):676

Dabhade A, Roy S, Moustafa MS, Mohamed SA, Gendy RE, Barma S (2021) Extreme Weather Event (Cyclone) Detection in India Using Advanced Deep Learning Techniques. 2021 9th International Conference on Orange Technology (ICOT), IEEE

El Jaafari I, Ellahyani A, Charfi S (2021) Parametric rectified non-linear unit (PRenu) for convolution neural networks. SIViP 15(2):241–246

Flaounas E, Davolio S, Raveh-Rubin S, Pantillon F, Miglietta MM, Gaertner MA, Hatzaki M, Homar V, Khodayar S, Korres G (2022) Mediterranean cyclones: current knowledge and open questions on dynamics, prediction, climatology and impacts. Weather Clim Dynamics 3(1):173–208

Girshick R (2015) Fast r-cnn. Proceedings of the IEEE international conference on computer vision

Harrou F, Zeroual A, Hittawe MM, Sun Y (2021) Road Traffic modeling and management. Using Statistical Monitoring and Deep Learning, Elsevier

He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE international conference on computer vision

Hittawe MM, Afzal S, Jamil T, Snoussi H, Hoteit I, Knio O (2019) Abnormal events detection using deep neural networks: application to extreme sea surface temperature detection in the Red Sea. J Electron Imaging 28(2):021012–021012

Hulme M (2014) Attributing weather extremes to 'climate change' a review. Prog Phys Geogr 38(4):499–511

Jiang K, Wang ZY, Yi P, Wang GC, Lu T, Jiang JJ (2019) Edge-enhanced GAN for remote sensing image Superresolution. IEEE Trans Geosci Remote Sens 57(8):5799–5812

Jiang K, Wang Z, Yi P, Wang G, Lu T, Jiang J (2019a) Edge-enhanced GAN for remote sensing image superresolution. IEEE Trans Geosci Remote Sens 57(8):5799–5812

Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition

Kashinath K, Mudigonda M, Kim S, Kapp-Schwoerer L, Graubner A, Karaismailoglu E, Von Kleist L, Kurth T, Greiner A, Mahesh A (2021) ClimateNet: an expert-labeled open dataset and deep learning architecture for enabling high-precision analyses of extreme weather. Geosci Model Dev 14(1):107–124

Kaur R, Singh S (2022) A comprehensive review of object detection with deep learning. Digit Signal Proc : 103812

Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M (2022) Transformers in vision: a survey. ACM Comput Surv (CSUR) 54(10s):1–41

Kim J, Lee JK, Mu Lee K (2016) Accurate image super-resolution using very deep convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition

Kim SK, Ames S, Lee J, Zhang C, Wilson AC, Williams D (2017) Massive scale deep learning for detecting extreme climate events. Climate Informatics

Kumler-Bonfanti C, Stewart J, Hall D, Govett M (2020) Tropical and extratropical cyclone detection using deep learning. J Appl Meteorol Climatology 59(12):1971–1985

Lai W-S, Huang J-B, Ahuja N, Yang M-H (2018) Fast and accurate image super-resolution with deep laplacian pyramid networks. IEEE Trans Pattern Anal Mach Intell 41(11):2599–2613

Lam L, George M, Gardoll S, Safieddine S, Whitburn S, Clerbaux C (2023) Tropical Cyclone detection from the Thermal Infrared Sensor IASI Data using the Deep Learning Model YOLOv3. Atmosphere 14(2):215

Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z (2017) Photo-realistic single image super-resolution using a generative adversarial network. Proceedings of the IEEE conference on computer vision and pattern recognition

Lei S, Shi Z, Zou Z (2019) Coupled adversarial training for remote sensing image super-resolution. IEEE Trans Geosci Remote Sens 58(5):3633–3643

Liu Z-S, Siu W-C, Chan Y-L (2021) Features guided face super-resolution via hybrid model of deep learning and random forests. IEEE Trans Image Process 30:4157–4170

Lu Z, Li J, Liu H, Huang C, Zhang L, Zeng T (2022) Transformer for single image super-resolution. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition

Magdy A, Moustafa MS, Ebied HM, Tolba MF (2022) Backbones-Review: Satellite Object Detection Using Faster-RCNN. International Conference of Remote Sensing and Space Sciences Applications, Springer

Mezősi G (2022) Meteorological Hazards. Natural Hazards and the Mitigation of their Impact, Springer: 97–136

Moustafa MS, Sayed SA (2021) Satellite Imagery Super-resolution using squeeze-and-excitation-based GAN. Int J Aeronaut Space Sci 22(6):1481–1492

Moustafa MS, Ahmed S, Hamed AA (2020) Learning to hash with Convolutional Network for Multi-label Remote sensing image Retrieval. Int J Intell Eng Syst 13(5)

Olaoluwa EE, Durowoju OS, Orimoloye IR, Daramola MT, Ayobami AA, Olorunsaye O (2022) Understanding weather and climate extremes. Climate Impacts on Extreme Weather, Elsevier: 1–17

Pang S, Xie P, Xu D, Meng F, Tao X, Li B, Li Y, Song T (2021) NDFTC: a new detection framework of tropical cyclones from meteorological satellite images with deep transfer learning. Remote Sens 13(9):1860

Park SC, Park MK, Kang MG (2003) Super-resolution image reconstruction: a technical overview. IEEE Signal Process Mag 20(3):21–36

Rübel O, Byna S, Wu K, Li F, Wehner M, Bethel W (2012) TECA: a parallel toolkit for extreme climate analysis. Procedia Comput Sci 9:866–876

Shi M, He P, Shi Y (2022) Detecting extratropical cyclones of the northern hemisphere with single shot detector. Remote Sens 14(2):254

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556

Song T, Song Y, Wang Y, Huang X (2018) Residual network with dense block. J Electron Imaging 27(5):053036–053036

Tong B, Sun X, Fu J, He Y, Chan P (2022) Identification of tropical cyclones via deep convolutional neural network based on satellite cloud images. Atmos Meas Tech 15(6):1829–1848

Wang M, Leelapatra W (2022) A review of object detection based on convolutional neural networks and deep learning. Int Sci J Eng Technol (ISJET) 6(1):1–7

Wang C, Li X (2023) Developing a data-driven transfer learning model to locate Tropical Cyclone centers on Satellite Infrared Imagery. J Atmos Ocean Technol 40(12):1605–1618

Wang H, Xu Q, Yin X, Cheng Y (2024) 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

Zaidi SSA, Ansari MS, Aslam A, Kanwal N, Asghar M, Lee B (2022) A survey of modern deep learning based object detection models. Digit Signal Proc : 103514

Zhang Y, Mazen Hittawe M, Katterbauer K, Marsala AF, Knio OM, Hoteit I (2020) Joint seismic and electromagnetic inversion for reservoir mapping using a deep learning aided feature-oriented approach. SEG Technical Program Expanded Abstracts 2020, Society of Exploration Geophysicists: 2186–2190