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

A Transfer Learning Based Deep Learning Method for Mesoscale Convective Cloud Segmentation

Vidya Patil*[1], Anuradha Phadke[2]

Submitted: 04/11/2023 **Revised**: 26/12/2023 **Accepted**: 03/01/2024

Abstract: Risky weather events associated with Mesoscale Convective Systems (MCS) may end in considerable financial losses and occasionally even fatalities. Owing to the unpredictability of climate scenarios, little is known about the dynamics behind the development and deepening of MCS. Satellite images of MCS Clouds (MCSC) reveal a range of topologies, from open to closed, yet study on MCSC activities remains severely constrained. Through the use of high-resolution mathematical models of the atmosphere and the analysis of remote sensing imagery, high cloud-top temperature gradients, specific spatial shapes of temperature patterns, and other aspects of MCSCs can be investigated. In this study, deep learning (DL) methods are used to segment MCSC images using a transfer learning (TL) strategy. VGG-16 has been improved in the present work by fusing encoder-decoder architecture and taking cues from UNet architecture. The design that is generated is called ENDE-VGG. The proposed approach enables the collection of more pertinent data over a wider region. Images based on the Brightness Temperature in the infrared (K1-IR) channel by the Indian geostationary satellite Kalpana-1 are used in the present research. The TL technique produced a Dice coefficient of 0.935 on the validation data set, an intersection of union (IoU) of 0.875 on the K1-IR data set, and a mean IoU of 0.93. Additionally, using the test data, it obtained a 0.933 Dice coefficient and 0.875 IoU. ENDE-VGG performed better than the most sophisticated cloud image segmentation methods when using IoU as the loss function, based on numerous research studies.

Keywords: *Deep learning, Kalpana-1 infrared images, Mesoscale Convective Systems, Transfer learning*

1. Introduction

The highest temperatures registered in meteorological records have been observed globally over the previous thirty years. In comparison to the worldwide average, temperature levels in Russia are rising twofold quicker [1]. Convective phenomena in the upper atmosphere have intensified as a result of the discovered climate warming, and dangerous convective weather occurrences have become more frequent and intense [2]. These instances significantly harm the country's economy and result in fatalities. The most significant of these is extremely torrential downpours, which includes devastating continuous cyclones in 2010, 2017, and 2021 [3], [4] an intense flood in Krymsk in 2012 [4], severe hurricanes in both the European and Asian areas of the world [5], [6], [7] and high levels of precipitation in general [4]. Mesoscale Convective Systems (MCS) are responsible for producing dangerous convective events, such as huge hail, cyclones, squalls, and strong stormy rains, which often happen during the hottest part of the year [8], [9].

Mesoscale Convective Systems, called the MCS, represent a structured collection of cumulonimbus clouds that generate a zone of precipitation with a width exceeding beyond 100 km no less than a single direction [10]. Although tropical regions are home to the bulk of MCSs, a smaller proportion is often seen there throughout the summer months [11]. Depending upon their geometric and horizontal aspects, Maddox's suggested classification of MCS [12]. Particularly, axisymmetric and regular MCS have been identified. According to the Orlanski identification, the initial categories are further classified into squall contours referred as meso-alpha scale and cumulonimbus peaks called as meso-beta scale [13], [14]. Mesoscale convective complexes (MCCs), that belong to mesoalpha scale frameworks, while cumulonimbus cloud clusters and supercells, resulting in meso-beta scale networks, are examples of axisymmetric MCS.

On the basis of the utilization of remote sensing information, observational and climatological examinations of MCS are carried out. Continuous series of data are specifically given by geostationary satellites like GOES, Meteosat, and Himawari as well as polar satellites in orbit like NOAA or Terra/Aqua. Owing to the contiguous patches of extremely frigid cloud top temperatures that MCS generate, infrared satellite images are used to detect and track them [15]. Additionally, geostationary satellite imagery offers extensive spatial and temporal distribution, alongside any part of the world may employ the techniques used to analyze its information. Laing and Fritsch [11] conducted one of the earliest satellite-based investigations

[1] *Department of Electrical and Electronics Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune-411038, INDIA*
ORCID ID : 0000-0003-0077-3727
[2] *Department of Electrical and Electronics Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune-411038, INDIA*
ORCID ID : 0000-0001-9682-8184
** Corresponding Author Email: vidya.patil@mitwpu.edu.in*

of the climatological properties of MCSs over the period 1986–1997, concentrating on MCCs. Numerous climatological research on MCS have since been conducted globally [16] and macro-regional [17] scales, including the contiguous United States, China, and West Africa. Infrared images from space are employed for evaluating a number of MCS features along with to keep track of them. Intense updrafts and regions featuring a high likelihood of severe weather are particularly indicated by cloud-top signatures including overshooting tops (OTs) [18], cold rings [19], cold U/V features [20], and above-anvil cirrus plumes [21].

In climatological investigations of MCS, meteorological radar measurements are frequently employed in conjunction with satellite images [22]. Such information can be used to assess the properties and progression of MCS because of their great resolution in both time and space. The United States [23] and Eastern Europe each have their own radar-based climatologies of MCSs. Weather radar data has been effectively used to assess the climatological parameters of hailstorms in addition to the MCSs themselves and their morphological properties [24]. Radar data have several spatial coverage gaps in some areas, which is their principal drawback.

Using automated technologies or the expertise of specialists, MCS can be found on satellite images. An increase in the use of computerized data-processing and analyzing methodologies has been brought on by a growing volume of experimental data, particularly those concerning the temporal and spatial variation of different meteorological characteristics [25], [26]. However, it still takes a lot of time to visually locate and analyze MCSs in satellite imagery. Alternately, specialists can employ sophisticated machine learning techniques that can spot associations and patterns in the large amounts of available data to better comprehend the dynamics of MCSs [27], [28].

The remainder of the paper is divided into the following sections: A summary of the research work presented in the published literature for the MCSC segmentation and classification is provided in Section 2. A detailed description of the technique used for the proposed ENDE-VGG model is provided in Section 3. Section 4 describes the experimental findings and the analysis that goes along with them, and section 5 finishes the work provided along with any shortcomings and suggestions for further research.

2. Literature Review

Computer vision techniques are widely used in conjunction with the numerical simulation techniques for the study of two dimensional data that may be visualized. The identification of related components, key point localization, clustering, and image processing methods tailored for the recognition of visual patterns are a few examples. Numerous techniques, such as analyzing images using machine learning algorithms [29], are applicable in addition to image-processing techniques that assess color and brightness [30].

There have been numerous machine learning-based methods created so far to identify different MCC cloud forms. Recent studies have demonstrated the successful use of machine learning techniques for recognizing extreme weather occurrences [31], [32], [33]. Due to the clear depiction of these occurrences in the majority of empirical data and modeled atmospheric dynamics knowledge, the majority of studies concentrate on the recognition

synoptic-scale atmospheric occurrences such as tropical cyclones [34] and atmospheric rivers [35]. Investigations on the detection of mesoscale geophysical phenomena using DL methods have been conducted. Huang et al. [36] discovered sub-mesoscale oceanic vortices in SAR satellite data, and Krinitskiy et al. [37] showed that the convolutional neural network (CNN) capacity to classify polar lows in infrared and microwave satellite mosaics. CNN techniques are also used by Krinitskiy et al. [38] to address the identification of polar lows in satellite mosaics of the Southern

Ocean challenge.

One of the oldest is the Wood and Hartmann [39] technique, which uses a Multi-Layer Perceptron model trained using a Moderate Resolution Imaging Spectrometer. Here, the degree and distribution of liquid water paths are used to classify the MCC type on imaging regions of 256×256 km^2. This technique has been found to help identify global and seasonal variations [40], [41]. It gives outcomes for areas with fairly high spatial resolutions. The relevance of the technique for analyzing the dynamics and diurnal cycle of these kinds of systems is nevertheless limited on account of the coarse spatial resolution of the approach and the use of cloud components produced from visible images.

Significant contributions of the present investigation are as follows: i) The conventional VGG-16 architecture's three fully connected layers (FC-Layer), which are used for classification, are adjusted in the present scenario to adopt the traits associated with a symmetric decoder with layers placed in the opposite order from the encoder side, as is done in a traditional U-Net; ii) In order to segment the clouds that can be utilized that represents MCSC after performing the area threshold, DL has been applied for the very first time with Indian satellite IR image data; iii) Considering exclusively two distinct kinds of pixel values in an output image-white for the object, MCSC in this instance and black for the background or other objects in the image-object segmentation, like separating out a MCSC with a certain temperature, can be viewed as a binary classification challenge. Accuracy might not be taken into account as a true performance criterion because the segmented image

contains more black pixels than white ones. An extremely high accuracy value can be produced by an output image with an imbalanced proportion of black and white pixels. Intersection-over-union (IoU) and Dice coefficient are used as performance metrics to evaluate the efficiency of the proposed system; iv) Applying IoU as a loss function has enhanced the accuracy, IoU, and Dice Coefficient performance of the MCS segmentation as overall.

3. Methodology

In this present work, an improved DL model namely ENDE-VGG is utilized to segment MCSC in the MCS environment. In the part that follows, the essential components needed to build a DL-based MCSC segmentation model are outlined. Figure 1 illustrates the suggested DL-based MCC segmentation approach.

3.1. Data Annotation

With the use of thresholding technique, an image is capable of being easily analyzed and the desired features can be extracted by splitting it into two or more sections based on intensity levels. Pixels exhibiting intensities that are higher or lower than the threshold will be categorized differently by specifying a threshold value. This method helps with tasks including recognizing objects.

The thresholded image, T(x, y), is denoted mathematically as,

$$T(x,y) = \begin{cases} 1 \; if \; I(x,y) > T \\ 0 \; if \; I(x,y) \leq T \end{cases} \quad (1)$$

where *I (x, y)* represents the input image.

3.2. Data Augmentation

Data augmentation employs current data to create modified replicas of datasets, which are then used to artificially increase the training set [47]. During augmentation, geometric transformations such as crop, rotation, flipping operations like horizontal flip and vertical flip, grid distortion, and elastic transform are performed to K1-IR images as well as its corresponding mask. The subsections below provide descriptions of each operation.

3.2.1. Crop

The term "random crop process" typically refers to a padding-based random cropping procedure. The input image is first put up with an arbitrary-sized window. The raw image is then clipped out according to the vertical and horizontal lines that traversed a random point within the predetermined window. It was then downsized to the original image size. Here, the ground-truth box in the image can be removed together with the actual image.

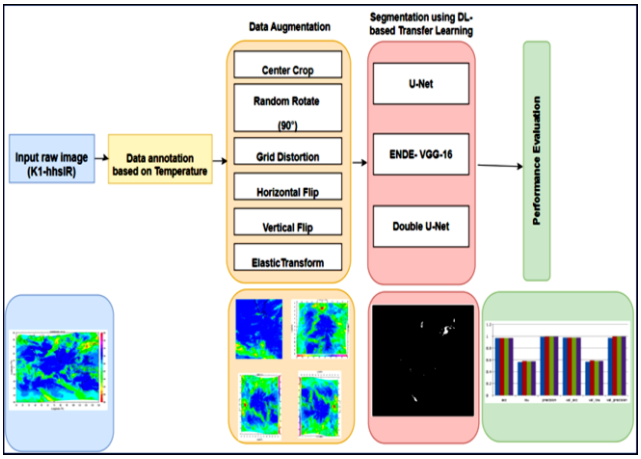

Fig. 1. Overall architecture of TL based DL approach for cold clouds segmentation in MCS environment

3.2.2. Rotation

Rotation augmentations are performed by rotating the image between 1° and 359° to the right or left on an axis. The rotation degree attribute has a significant impact on the effectiveness of rotation augmentations. On identification assessments, slight rotations may be helpful, but as the degree of rotation grows, the marking of the information is not anymore preserved after transformation.

3.2.3. Flipping

Images can be rotated both horizontally and vertically. Transitioning the horizontal axis occurs far more frequently than switching the vertical axis. However, a vertical flip is the same as rotating an image by 180°, then flipping it horizontally. This augmentation is among the simplest to employ and has worked well with datasets like ImageNet.

3.2.4. Grid Distortion

Grid-distortion is an image warping technique which is driven by the mapping between equivalent families of curves, arranged in a grid structure. Here, the aberrant image is broken down into a rectangular grid, and for each grid, each of the current curves are given slightly different shapes, resulting in variances in the MCSC's geometric alteration.

3.2.5. Elastic Transform

In this research, we investigate an interesting method of elastic distortions. The following justifies the use of this kind of data augmentation: The MCSC is a non-rigid object, therefore varied stretching can lead the same object to seem slightly differently in different exams. In order to produce

fresh synthetic samples which are scientifically reasonable, one can artificially induce analogous tiny elastic distortions.

3.3. Segmentation using Deep Learning

A convolutional neural network (CNN) is a particular type of multiple-layered neural network that relies on explicit pattern identification from image data with minimal or no pre-processing [48]. Convolution, pooling, and fully linked layers are the three primary layers that make up a traditional CNN. The convolution layer, which is the base layer of a CNN, performs the majority of the computing effort. Like the term suggests, it filters or performs convolutions on the input before passing the result to the following layer. The input is completely filtered by the applied filter, creating a feature map and serving as a feature marker. The pooling layer is employed to substantially minimize the representation of space and the amount of computation required between successive convolutions. Each sliced input is subjected to the pooling operation by the pooling layer, which lowers the computational burden on the following convolution layer. Features from the input images are extracted and reduced as a result of the use of convolution as well as pooling layers. Applying the completely connected layer results in an end result corresponding to the total number of classes. A CNN is a variation of each of these components layered together to produce a CNN architecture. The architecture is the identical for all CNNs, however there are some differences.

CNNs are a technique for TL that enables the knowledge acquired from a specific task to be applied to another, related task [49]. The fundamental method used in DL to address the issue of limited training data is known as TL. TL might be useful for issues like MCSC segmentation. When employing TL, one aims to transfer from the original domain to the target domain the understanding of the spatial distribution of information in an encoded representation space [49]. TL makes use of the generic properties that CNNs trained on huge datasets, like ImageNet, have discovered and which are pertinent to many perceptual applications. TL uses a pre-trained CNN as its foundation and refines it on the new dataset like K1-IR as opposed to training a CNN from scratch on a fresh dataset. For MCSC segmentation in this work, we used the VGG-16 architecture to perform TL on K1-IR dataset referred as ENDE-VGG. ENDE-VGG is redesigned in the current MCSC classification by integrating encoder-decoder architecture and taking design ideas from UNet architecture.

3.3.1.Parameters for learning K1-IR features

A domain, D, is made up of the feature space, F, and the associated probability, $P(F)$ [49], where $F = f_1, f_2, \ldots, f_a$ denotes the total number of input images. The region that is relevant can be mathematically stated as,

$$D = [F, P(F)] \quad (2)$$

For two distinct domains, their corresponding feature spaces and marginal probabilities would also vary. A task E in a domain D is made up of two parts: the objective predictive function X(.) and label space S. This is denoted in mathematical concepts as,

$$E = [S, X(.)] \quad (3)$$

The attributes S labelled as l are used during the training procedure to produce the function X(.) in order to predict the testing data. In our proposed method, which uses a pre-trained network of VGG-16, we experience the case when there is just one base domain D_b and one destination domain D_d. Base domain and destination domain are mathematically represented as,

$$D_b = (f_{b1}, l_{b1}), (f_{b2}, l_{b2}), \ldots, (f_{be}, l_{be}) \quad (4)$$

where $f_{bk}$is the instance of the base information that matches label l_{bk}, and k=1, 2, 3,..., e. The desired field information is displayed similarly as,

$$D_d = (f_{d1}, l_{d1}), (f_{d2}, l_{d2}), \ldots, (f_{de}, l_{de}) \quad (5)$$

Our suggested system's two domains, as well as each of their unique characteristic spaces and labels, may not be the same. Transfer learning is the process of learning the predictive function $X(.)$ of a target domain using data from a source domain and tasks. $X(.)$ is used to predict the label of an unidentified instance f, which is mathematically stated as,

$$f(y) = P(l|y) \quad (6)$$

The basic idea of the current study is that each of the innermost layers of CNN are able to extract broad characteristics from images that can be generated in advance for a certain domain (in this case, ImageNet is the source task) and then be utilized for classifying a new task (MCSC classification) D_d.

Additionally, the mathematical depiction of objectives, referred to as loss functions, need to be taken care of, if it possesses the ability of covering even the edge situations in order to comprehend the goal properly and more quickly. Conventional learning methods, from which loss functions have evolved based on label distribution, is where loss functions first appeared. In this work, three types of loss functions are used for the analysis: Binary Cross Entropy (BCE), Dice Coefficient (DC), and Intersection over Union (IoU) [50].

3.4.1.1. Binary Cross Entropy

BCE is the loss function used in this case. BCE loss is a predictive statistic that measures inaccurate data category

labelling through a model and penalizes the model if errors in label classification occur. High accuracy levels correspond to low log loss scores. The goal problem entails two labels, presence of MCSC and absence of MCSC and predictions for two classes: presence and absence of MCSC. As a result, the loss function for the network varies accordingly, and for these binary classification issues, we explore BCE loss (Eq. (7)).

$$Loss_{BCE} = -\frac{1}{N}\sum_{k=1}^{N} Z_i.\left(p(Z_i)\right) + (1 - Z_i).log\ (1 - p(Z_i)) \quad (7)$$

Where Z_i is the actual class, $\log(1 - p(Z_i))$ details the probability of the original class, $p(Z_i)$ represents the probability value of 1 and $1 - p(Z_i)$ represents the probability value of 0.

3.4.1.2. Dice Coefficient

This popular loss is employed to determine the similarity of the two images and it is related to the Intersection-over-Union strategy. As a result, the Dice Coefficient (DC) is transformed into the Dice Loss, a function that represents loss (Eq. (8)).

$$Loss_{DC} = 1 - \frac{2Z_i p + 1}{Z_i + p + 1} \quad (8)$$

where, p represents the predicted value.

3.4.1.3. Intersection over Union

IoU sometimes referred by the term the Jaccard index or Jaccard similarity coefficient, is a popular metric for comparing the resemblance of finite test sets. IoU is usually described as the intersection of two finite test sets, A and B, divided by their union, A and B (Eq. (9)).

$$Loss_{IoU} = \frac{A \cap B}{A \cup B} \quad (9)$$

The identity of indiscernible, non-negativity, symmetry, and triangle inequality are all characteristics of a metric that IoU satisfies. The resemblance between two arbitrary forms A and B is devoid of the scale of their space in particular because IoU is scale-invariant. The IoU has been used in this paper because of the aforementioned qualities.

3.3.2.The Proposed ENDE-VGG Architecture

ENDE- VGG architecture is divided into two sections. VGG16 structure (i.e encoder, EN) is used in one section with five downsampling blocks to capture context. The other U-Net-based component (decoder, DE) [39] employs an expanding path that is symmetrical to enable exact identification and consists of five up-sampling blocks. The design of ENDE- VGG is seen in Figure 2.

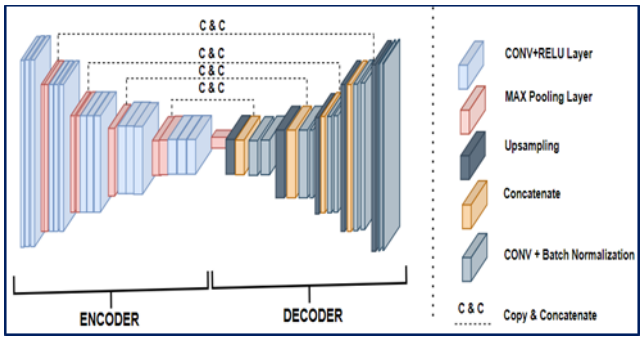

Fig. 2. Overall flow of the proposed ENDE-VGG Architecture

The input images are 512 × 512 pixels containing a total of three channels (Red, Green, and Blue- RGB) are applied in the suggested ENDE-VGG model. Convolutional layers linked to the Rectified Linear Unit (ReLU) layer, MaxPooling layer. Upsampling layers of the ENDE- VGG design bring the component dimensions of the down-sampling block which were reduced to its original dimensions by means of Transposed Convolutional layers. In the meanwhile, Batch Normalization (BN) layer comes after the convolutional layer in the up-sampling block. Certain layers are replicated symmetrically after which they are combined alongside the output of the convolutional layer during the up-sampling block operation. The width and height of the data after every subsequent downsampling block calculation are cut in half. The input applied to the upsampling layer gets doubled in its width and height at the output. The upsampling layer in this case doubles the applied input in both width and height.

Convolutional layers are used in both blocks of up- and down-sampling. The layer employs a set of learnable weight kernels to carry out the convolution operation. MaxPooling layer, which shrinks the overall dimension of the input array and is frequently referred to as downsampling, is another crucial component of ENDE- VGG. It obtains the maximum values using a subarray. If a stride of 2 is taken into account, the input array's width and height are reduced to half once it has passed the MaxPooling layer. The memory shrinks to one-fourth of its initial size once the array has gone through the MaxPooling layer. This increases training efficiency tremendously.

Blocks of data are upsampled in order to convert the characteristic vectors to its original pixel values in RGB channels. This process is referred to as image visualization. In the up-sampling blocks (DE blocks), the upsampling layer is frequently utilized. It expands the input array's width and height. Here, any input element would be repeated twice. Following upsampling, resolution would always be 2 × 2 of the input array. The input as well as the output image dimensions are identical since there are precisely the same amount of MaxPooling and Up-Sampling layers.

Layer copying and concatenation are exclusive to U-net and

set it apart from other fully connected networks. For instance, to create annotations in an output image, up-sampling blocks convert the result of convolutional layers, which capture the properties of MCSC into abstract knowledge, or into spatial data. Convolutional and pooling layers, however, will begin to forfeit spatial information, with the pooling process having a greater impact. Both spatial and abstract knowledge are crucial for K1-IR segmentation. As a result, typical fully connected networks lack desirable results in terms of accuracy. Copying and concatenating are suggested as a solution to connect the features in the down-sampling blocks to the equivalent up-sampling blocks in U-net, as each time pooling will drastically lose spatial information.

4. Experimentation Results

4.1. Evaluation Parameters

The metrics that are used to evaluate the models are accuracy, recall, also known as sensitivity, specificity, and false positive rate, which are the most frequently employed metrics. Four key metrics are utilized to determine these measures: (i) correctly identifying cloud pixel cases (true positives, TP); (ii) incorrectly classifying cloud pixel cases (false negatives, FN); (iii) correctly identifying black background cases (true negatives, TN); and (iv) incorrectly classifying black background cases (false positives, FP). 10, 11, 12, and 13 show the mathematical formulas for calculating accuracy, specificity, sensitivity, and false positive rate.

$$SEN = \frac{TP}{TP+FN} \quad (10)$$

$$SPE = \frac{TN}{TN+FP} \quad (11)$$

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

$$FPR = \frac{FP}{TN+FP} \quad (13)$$

Where SEN represents sensitivity, SPE denotes specificity. ACC represents the accuracy, and FPR gives the false positive rate.

4.2. Data Annotation

In order to use supervised learning to obtain the most effective segmentation and classification results, labels need to be collected from K1-IR images. We use a threshold of 220K. This value needs to be calculated by taking into consideration the K1-IR image pixel values that are related to the specific temperature variations. Radiant annotations that have too high of values lose their brilliance, whilst too low of values result in artifacts like less obvious margins around clouds. In order to properly segment MCSC, the following threshold values are selected for this experiment and images are manually annotated:

i) If the temperature is below or equal to 220K, the threshold for the temperature indicates the white pixel values; otherwise,

ii) The values of each pixel are black if the temperature criterion exceeds above 220K.

White pixels serve as the work's representation of objects, MCSC in this scope of investigation. Here, for the purpose of being able to generate manual annotations employing the photo editing program, Photoshop, this threshold is depicted by a color in the color palette.

Figure 3 illustrates the result of the data annotation step. Figure 3 (a) displays the input image from the K1-IR dataset as a sample for illustration. Figure 3 (b) shows the resultant output image at this stage.

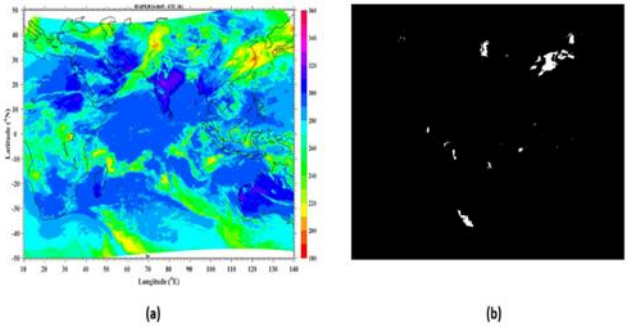

Fig. 3. Result for the data annotation phase

Figure 3 shows that a given RGB K1-IR image is divided into two distinct regions according to a temperature-related threshold value. Therefore, in the output image, pixels with intensity values higher than the specified 220K will be classified as black, or 1, and those with intensity values lower are regarded as white, or 0. This process also helps to know about the ground truth present in the input image.

4.3. Data Augmentation

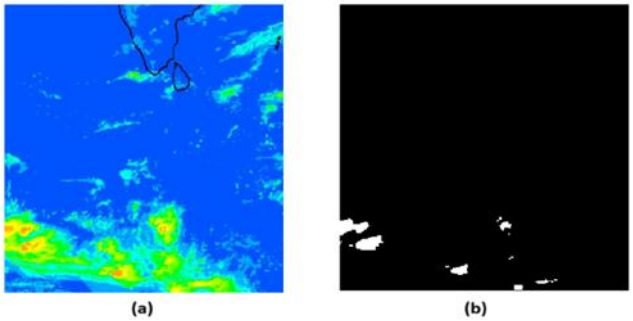

Fig.4. Result of crop operation

An arbitrary-sized window initially loads with the input image. Following that, the raw image is clipped off in accordance with the vertical and horizontal lines that passed through a chosen location inside the predefined window. The image is then shrunk to 512x512 size. In this case, both the image and the ground-truth box can be eliminated. The object is assumed to exist if over 50% of the ground-truth area is left after being chopped off. Additionally, when the

total number of arbitrary points is set to n, the number of original image details are multiplied by n to create new images created by data augmentation. Additionally, we avoided overfitting by enhancing the data by horizontal flipping. Figure 4 displays the result of crop operation.

It is followed by the rotation-based transformations. The image is rotated between 1° and 359° to the right or left on an axis to conduct rotation augmentations. The efficiency of rotation augmentations is significantly influenced by the rotation degree attribute. Slight rotations, such as those between 1° and 20° or -1° to 20°, may be useful for identification evaluations, but as the degree of rotation increases, the imprint of the details is no longer kept following transformation. For the sake of this experiment, the K1-IR images are thus randomly rotated by 90°. The rotation-based transformation result is shown in Figure 5.

Images are rotated in both the horizontal and vertical directions during rotation-based transformation. The horizontal axis changes far quicker than the vertical axis does. The images are rotated by 180° in a vertical flip before being turned horizontally.

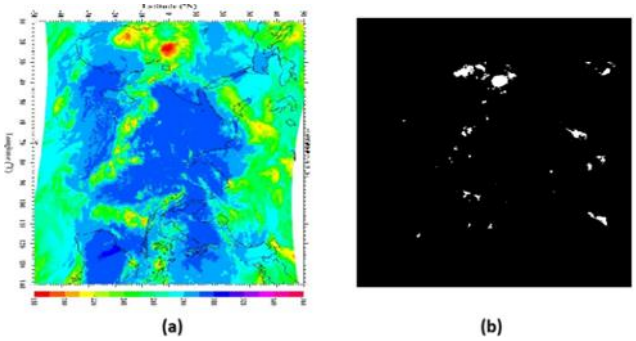

Fig. 5. Result of random rotate by 90° operation

With datasets including ImageNet, this augmentation has shown to be one of the most convenient to use. Figure 6 and Figure 7 displays the result of flip operations, horizontally and vertically, respectively.

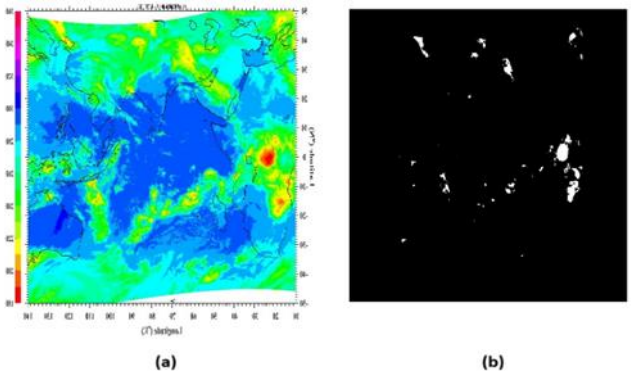

Fig.6 Result of horizontal flip operation

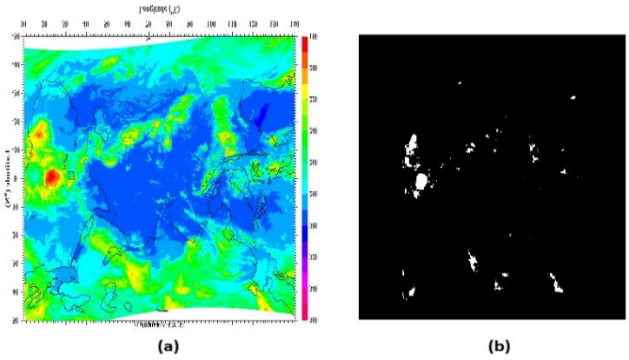

Fig.7. Result of vertical flip operation

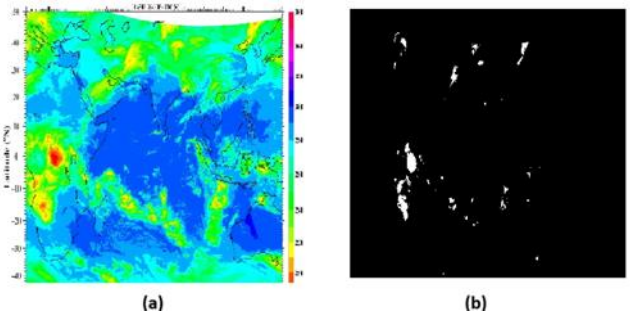

Fig.8. Result of grid distortion

In grid based transformation, the aberrant image is divided into a rectangular grid and each grid's current curves are given slightly varied shapes, leading to variations in the MCSC's geometric transformation. The resultant image is shown in Figure 8.

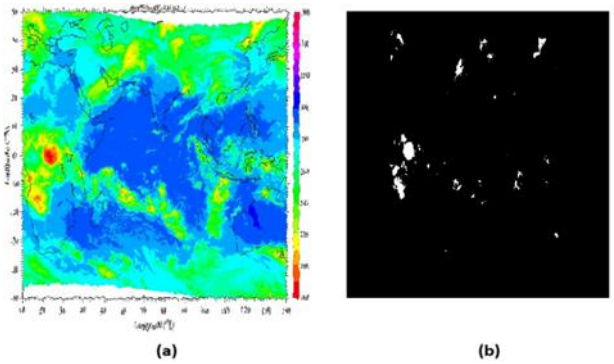

Fig.9. Result of elastic transformation

Additionally, we see that elastic distortions on K1-IR pictures as displayed in Figure 9 substantially enhances MCSC segmentation and classification performance in MCS environments.

4.4. Segmentation using ENDE-VGG Model

Following the data annotation and data augmentation procedure described in Section 4.2 and 4.3, we conducted MCSC segmentation and classification using the proposed ENDE-VGG model. We generated splits of the source K1-IR dataset into training validation and testing subsets employing 70:15:15 ratio. Employing the methodology described in Section 3.3.2, we trained and evaluated our proposed ENDE-VGG model with each split. We present

the resulting measures of quality and their uncertainties in Table 1.

A total of 100 iterations are included in each training phase. The Adam optimizer is used throughout training along with an initial learning rate of 0.0001 as long as it showed little advancement on the validation dataset, then trained using a learning rate of 0.1 through early stopping following another 30 epochs with little improvement on the validation dataset. IoU is the model validation measure used for comparing several models and decide on early halting. A weighted sum of dice loss and BCE that has a 70% weighting in favour of BCE makes up the loss function. By calculating the equilibrium value of precision and recall, dice loss, sometimes referred to as the F1-score, optimizes the incorrect impact of false positives and false negatives. Dice loss and IoU are fairly comparable in terms of numbers. Initial testing revealed that employing both types of loss function offered a greater IoU and dice loss compared to a BCE loss function. This is probably since BCE exhibits more steady gradients and is closer to the actual goal of enhancing IoU than dice loss and IoU loss, which are more resilient to imbalanced classes.

Image investigators frequently examine images employing masks for segmenting and analysing MCSC. It is beneficial to create image processing masks that, like traditional image processing screens, take into account numerical parameters, which includes quantitative assessments of pixel values and nearby pixels, in order to decode basic physical concepts beneath the MCSC image level. Filter sizes of [32, 64, 128, 256, 512] are used in the proposed approach to process labelled objects like MCSC. The validation dataset for filter sizes [16, 32, 48, 64], [16, 32, 64, 128, 256], [64, 128, 256, 512, 1024], [32, 64, 128, 256, 512, 1024], and [32, 64, 128, 512, 1024] does not show any improvement, hence the aforementioned filter size is chosen.

Table 1. Evaluation measures for the proposed ENDE-VGG model

Loss Function	**Training**					**Validation**				
	mean IoU	SEN	SPE	ACC	FPR	mean IoU	SEN	SPE	ACC	FPR
BCE	0.496	0.901	0.999	0.9982	0.00018	0.496	0.897	0.999	0.9982	0.00025
DC	0.926	0.909	0.999	0.9983	0.00016	0.93	0.9	0.999	0.9982	0.00026
IoU	0.931	0.91	0.999	0.9983	0.00014	0.93	0.9	0.999	0.9982	0.00025

Table 1 demonstrates that when the loss function is IoU, mean IoU (mIoU) greatly increases whereas BCE and DC hardly change at all. This is because mIoU elevates the weight of damaged pixels to the identical extent as the weight of non-damaged pixels. This demonstrates that mIoU may enhance the K1-IR dataset's capacity to segment and categorize images, as well as that using a method of assessment that assigns objects more weight is significant especially when such items, like mIoU, only take up a small portion of the total number of pixels in the dataset. Results from every image that offers areas with clear temperature thresholds are better. Since overall K1-IR image is illuminated unevenly, the source image needs to be modified in relation to illumination throughout the image.

Considering the entire K1-IR images, the segmentation performance is shown in Table 1 on the basis of sensitivity, specificity, and accuracy. The effectiveness is still very constant, proving that the variables in the suggested algorithm have the best values. As seen in Table 1, the suggested segmentation method efficiently segments the MCSC using the suggested ENDE-VGG model with an overall accuracy of 0.9983 (training) and 0.9982 (validation). The FPR lies between [0, 1]. In this case, the model performs better when FPR values decrease. For the suggested ENDE-VGG model, this is accurate. From Table 1, it appears that the FPR value, which is within the range, is 0.00014 (for training) and 0.00025 (for validation).

The initial evaluation of the MCSC segmentation method only incorporated ENDE-VGG. Later, two image segmentation models-conventional U-Net [51] and double U-Net [52]-are examined as part of the segmentation algorithms. Then, three objective functions, including BCE, Dice Loss, and IoU, are calculated. Initially, the segmentation method only used 30 to 60 epochs. However, models were under fitting because of an insufficient amount of data and the necessity to train an model with 1.3 million parameters. An average validation Dice Coefficient of 0.91 resulted from overfitting caused by efforts to raise the number of trials. These preliminary findings indicated that the segmentation algorithm on the K1-IR dataset performed around averagely. The comparison of the MCSC segmentation algorithm employing ENDE-VGG ended up being in depth. In this instance, we are able to enhance the optimal Dice Coefficient score from 0.845 to 0.907 considering a total sample of 7868 K1-IR images spread across 100 epochs (Table 2). The Dice Coefficient score increased from 0.915 to 0.951 (Table 2) during the validation phase utilizing 7868 K1-IR images, and from 0.906 to 0.933 during the testing phase.

Table 2. Assessment metrics for objective functions used in ENDE-VGG Model.

Loss Function	Validation		Testing	
	IoU	Dice Coefficient	IoU	Dice Coefficient
BCE	0.845	0.915	0.829	0.906
DC	0.903	0.949	0.874	0.932
IoU	0.907	0.951	0.875	0.933

4.5. Comparison of Performance

Two additional pre-trained models based on DL are employed to assess the effectiveness of the segmentation and classification of the MCSC from the K1-IR dataset in order to gain insight into the impact of the efficiency of the ENDEVGG model. U-Net and double U-Net decoder designs are being put to the test. Adam is employed herein across all the models used for segmentation for uniformity and since the initial evaluation revealed that employing Adam produced decent results, regardless of the various segmentation models utilized in the present investigation. The transferred U-Net model, which is made up of subtle characteristics collected from more than 1,000,000 images, had been pre-trained across the ImageNet dataset. The U-Net architecture's layers are finely tweaked throughout the entire and segmented K1-IR images as the fully connected layers are transferred. The U-Net and double U-Net models retain the low-level attributes from the source domain while transferring the fully connected layers, accelerating learning. A total of 100 epochs are used to record both the training and the validation phases for each image, and the resulting average outcomes of evaluation are provided in Table 3. The U-Net correctly identified MCSC in 0.9982 as part of pixels after training. Table 3 shows that there is a considerable rise in mIoU from 0.911 by U-Net, 0.903 by double U-Net, to 0.931 by ENDE-VGG model. The results showed a sensitivity of 0.906 and a specificity of 0.999. Double U-Net attained a sensitivity of 0.909, specificity of 0.999, and average accuracy of 0.9983. Additionally, noted are the FPRs for U-Net and double U-Net, which are 0.00019 and 0.00015, respectively. In Supplementary Table 1, the results are summarized together with accuracy, sensitivity and specificity. The investigation shows that the suggested ENDE-VGG model improves U-Net and double U-Net topologies, particularly in terms of mIoU, accuracy, and FPR.

Furthermore, we examined the DC and IoU of the recovered ground truth and the binarized forecasted segmentations in both the training and the validation dataset (sections 4.2 and 4.3). When comparing the forecasts based on the training set to the ground truth segmentations, it is as expected that both DC and IoU are relatively large. The projected segmentations nevertheless exhibit a mediocrely good level of agreement with DC and IoU values on unobserved data. Interestingly, we observed performance that is comparable to the outcome of the validation phase. Careful analysis of Table 4 reveals that the trained ENDE-VGG model's performance is mostly unaffected by variations in the ground truth segmentations employed to train the model between different raters. Table 4 displays a summary of various agreement indicators.

Moreover, shown in the present research, we successfully managed to gather the entire MCSC and train an ENDE-VGG segmentation model with a very high DC compared to those formerly reported with only 7868 K1-IR images.

Table 3. Performance comparison criteria for evaluation of the proposed model.

Models	Training					Validation				
	mean IoU	SEN	SPE	ACC	FPR	mean IoU	SEN	SPE	ACC	FPR
U-Net	0.911	0.906	0.999	0.9982	0.00019	0.915	0.889	0.999	0.9981	0.00022
Double U-Net	0.903	0.909	0.999	0.9983	0.00015	0.912	0.9	0.999	0.9982	0.0003
ENDE-VGG	0.931	0.91	0.999	0.9983	0.00014	0.93	0.9	0.999	0.9982	0.0002

Considering more extensive datasets that utilize multi-institutional MCSC batches and more powerful models, the suggested ENDE-VGG segmentation accuracy might probably be increased even more. In this research, we used a dataset of 7868 samples and trained the models utilizing 1.3 million parameters. To maintain success in automatic MCSC segmentation, extensive datasets in MCSC image evaluation shall be required. Using large, heterogeneous records, multi-model techniques to train a single model that is capable of recognizing and segmenting any kind of MCSC should additionally be evaluated.

Table 4. Comparison of objective functions

Models	Training		Validation	
	IoU	Dice Coefficient	IoU	Dice Coefficient
U-Net	0.893	0.943	0.873	0.932
Double U-Net	0.903	0.949	0.879	0.935
ENDE-VGG	0.907	0.951	0.879	0.935

5. Conclusion

MCS are highly dangerous and strong atmospheric phenomena. Due to a lack of labeled data, conducting effective research on MCSC characteristics based on difficult impressions can be difficult. In addition, investigation into the dynamics of the MCSC utilizing DL approaches has become severely constrained due to its multiple topologies, that may be a combination of open or closed in nature. The proposed work uses pre-trained DL algorithms and a TL technique to segment and classify MCSC images employing the K1-IR database obtained from the Indian geostationary satellite. The proposed model, called ENDE-VGG, utilizes the capabilities of the encoder-decoder design. On the validation data set for the K1- IR data set, the TL technique yielded an IoU of 0.875, a mIoU of 0.93, and a Dice coefficient of 0.935. It also cleared the examination with an IoU of 0.875 and a Dice coefficient of 0.933. Numerous investigations indicate that when IoU is used as the loss function, ENDE-VGG outperforms the most modern methods for segmenting cloud images. In the current study, 1.3 million parameters were used to train the ENDE-VGG model using a dataset of 7868 samples. To ensure the achievement of automatic MCSC segmentation, large datasets for MCSC image assessment are required. In the future, studies will examine multi-model techniques for training one predictive model that can recognize and segment any form of MCSC utilizing a variety of huge, heterogeneous datasets.

Acknowledgements

The Kalpana-1 data obtained from the Meteorological and Oceanographic Satellite Data Archival Centre (MOSDAC) website (http://www.mosdac.gov.in) received through Dr. Subrata Kumar Das, Indian Institute of Tropical Meteorology, MoES, Pune, India is gratefully acknowledged.

Author contributions

Vidya Patil: Paper conceptualization, Data curation Methodology, Formal analysis, Visualization, Software Writing original draft and Editing **Anuradha Phadke:** Methodology, Software, Resources and Investigation, Writing-Review and Editing

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] V. M. Kattsov, E. M. Akentieva, O. A. Anisimov, M. Y. Bardin, S. A. Zhuravlev, A. A. Kiselev, M. V. Klyueva, P. I. Konstantinov, V. N. Korotkov, and Kostyanoy, A. G., "Third assessment report on climate change and its consequences on the territory of the Russian federation general summary," Roshydromet Science-Intensive Technologies: St. Petersburg, Russia, 2022.

[2] A. Chernokulsky, A. Eliseev, F. Kozlov, N. Korshunova, M. Kurgansky, I. Mokhov, V. Semenov, N. Shvets', A. Shikhov, and Y. I. Yarinich, "Atmospheric severe convective events in Russia: changes observed from different data," Russ. Meteorol. Hydrol, vol. 47, pp. 343–354, Aug. 22, 2022.

[3] E. P. Meredith, V. A. Semenov, D. Maraun, W. Park, and A. V. Chernokulsky, "Crucial role of Black Sea warming in amplifying the 2012 Krymsk precipitation extreme," Nat. Geosci. vol. 8, pp. 615–619, Jul. 13, 2015.

[4] A. Chernokulsky, M. Kurgansky, I. Mokhov, A. Shikhov, I. Azhigov, E. Selezneva, D. Zakharchenko, B. Antonescu, and T. Kuhne, "Tornadoes in northern Eurasia: From the middle age to the information era," Weather Rev. vol. 148, pp. 3081–3110, Jul. 08, 2020.

[5] A. Chernokulsky, A. Shikhov, A. Bykov, and I. Azhigov, "Satellite- based study and numerical forecasting of two tornado outbreaks in the Ural region in June 2017," Atmosphere. vol. 11, pp. 1146, Oct. 22, 2020.

[6] A. Chernokulsky, M. Kurgansky, I. Mokhov, A. Shikhov, I. Azhigov, E. Selezneva, D. Zakharchenko, B. Antonescu, and T. Kuhne, "Tornadoes in the Russian regions," Russ. Meteorol. Hydrol. vol. 46, pp. 69–82, Feb. 2021.

[7] T. Lister, and J. Masters, "Moscow Storm Kills 16," Injures nearly 170-CNN. Available online: https://edition.cnn.com/2017/05/30/europe/moscow-storm/index.html (accessed on 11 November 2023).

[8] A. Chernokulsky, A. Shikhov, A. Bykov, N. Kalinin,

M. Kurgansky, B. Sherstyukov, and Y. Yarinich, “Diagnosis and modelling of two destructive derecho events in European Russia in the summer of 2010,” Atmos. Res. vol. 267, pp. 105928, Apr. 2022.

[9] A. V. Chernokulsky, A. N. Shikhov, I. O. Azhigov, N. A. Eroshkina, D. P. Korenev, A. V. Bykov, N. A. Kalinin, M. V. Kurgansky, Y. B. Pavlyukov, and A. A. Sprygin, “Squalls and tornadoes over the European territory of Russia on May 15, 2021: diagnosis and modeling,” Russ. Meteorol. Hydrol. Vol. 47, pp. 867–881, Nov. 2022.

[10] R. Houze, W. Schmid, R. Fovell, and H. Schiesser, “Hailstorms in Switzerland: left movers, right movers, and false hooks,” Mon. Weather Rev. Vol. 121, pp. 3345–3370, Dec. 01, 1993.

[11] A. G. Laing, and J. Michael Fritsch, “The global population of mesoscale convective complexes,” Q. J. R. Meteorol. Soc. vol. 123, pp. 389–405, Jan. 1997.

[12] R. A. Maddox, “Mesoscale convective complexes,” Bull. Am. Meteorol. Soc. vol. 61, 11, pp. 1374–1387, Nov. 1980.

[13] I. Orlanski, “A rational subdivision of scales for atmospheric processes,” Bull. Am. Meteorol. Soc. vol. 56, 5, pp. 527–530, May 1975.

[14] S. M. Cheeks, S. Fueglistaler, and S. T. Garner, “A Satellite-based climatology of Central and Southeastern U.S. mesoscale convective systems,” Weather Rev. vol. 148, pp. 2607–2621, Jun. 2020.

[15] Z. Feng, L. R. Leung, N. Liu, J. Wang, R. A. Houze, J. Li, J. C. Hardin, D. Chen, and J. Guo, “A global high-resolution mesoscale convective system database using satellite-derived cloud tops, surface precipitation, and tracking,” J. Geophys. Res. Atmos. Vol. 126, e2020JD034202, Apr. 27, 2021.

[16] D. Chen, J. Guo, D. Yao, Y. Lin, C. Zhao, M. Min, H. Xu, L. Liu, X. Huang, and T. Chen, “Mesoscale convective systems in the Asian monsoon region from advanced Himawari imager: Algorithms and preliminary results,” J. Geophys. Res. Atmos. vol. 124, 4, pp. 2210–2234, Feb. 27, 2019.

[17] X. Yang, J. Fei, X. Huang, X. Cheng, L.M.V. Carvalho, and H. He, “Characteristics of mesoscale convective systems over China and its vicinity using geostationary satellite FY2,” J. Clim. vol. 28, pp. 4890–4907, Jun. 15, 2015.

[18] C. Klein, D. Belusi˘ c, and C.M. Taylor, “Wavelet scale analysis of ´ mesoscale convective systems for detecting deep convection from infrared imagery,” J. Geophys. Res. Atmos. vol. 123, pp. 3035–3050, Mar. 27, 2018.

[19] K. Bedka, J. Brunner, R. Dworak, W. Feltz, J. Otkin, and T. Greenwald, “Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients,” J. Appl. Meteorol. Climatol. vol. 49, pp. 181–202, Feb. 01, 2010.

[20] M. Setvak, D. T. Lindsey, P. Novak, P. K. Wang, M. Radov ´ a, J. Kerkmann, L. Grasso, S. H. Su, R. M. Rabin, and J. S˘˘tastka, “Satellite- observed cold-ring-shaped features atop deep convective clouds,” Atmos. Res. vol. 97, pp. 80–96, Jul. 2010.

[21] J. C. Brunner, S. A. Ackerman, A. S. Bachmeier, and R. M. Rabin, “A quantitative analysis of the enhanced-v feature in relation to severe weather,” Weather Forecast. vol. 22, pp. 853–872, Aug. 01, 2007.

[22] K. Bedka, E.M. Murillo, C.R. Homeyer, B. Scarino, and H. Mersiovsky, “The above-anvil cirrus plume: an important severe weather indicator in visible and infrared satellite imagery,” Weather Forecast. vol. 33, pp. 1159–1181, Oct. 01, 2018.

[23] B. Czernecki, M. Taszarek, M. Marosz, M. Połrolniczak, L. Kolendow´icz, A. Wyszogrodzki, and J. Szturc, “Application of machine learning to large hail prediction—The importance of radar reflectivity, lightning occurrence and convective parameters derived from ERA5,” Atmos. Res. vol. 227, pp. 249–262, Oct. 01, 2019.

[24] A.M. Haberlie, and W.S. Ashley, “Radar-based climatology of mesoscale convective systems in the United States,” J. Clim. vol. 32, pp. 1591–1606, Mar. 01, 2019.

[25] A. Surowiecki, and M. Taszarek, “A 10-year radar-based climatology of mesoscale convective system archetypes and derechos in Poland,” Weather Rev. vol. 148, pp. 3471–3488, Jul. 30, 2020.

[26] L. Nisi, O. Martius, A. Hering, M. Kunz, and U. Germann, “Spatial and temporal distribution of hailstorms in the Alpine region: A longterm, high resolution, radar-based analysis,” Q. J. R. Meteorol. Soc. vol. 142, pp. 1590–1604, Apr. 2016.

[27] J. L. Cintineo, T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, “An objective high-resolution hail climatology of the contiguous United States,” Weather Forecast. vol. 27, pp. 1235–1248, Oct. 01, 2012.

[28] M. A. Krinitskiy, and A. V. Sinitsyn, “Adaptive algorithm for cloud cover estimation from all-sky images over the sea.” Oceanology. vol. 56, pp. 315–319, Jul. 14, 2016.

[29] M. Krinitskiy, M. Aleksandrova, P. Verezemskaya, S. Gulev, A. Sinitsyn, N. Kovaleva, and A. Gavrikov,

"On the generalization ability of data-driven models in the problem of total cloud cover retrieval," Remote Sens. vol. 13, pp. 326, Jan. 19, 2021.

[30] Y. Liu, E. Racah, Prabhat, J. Correa, A. Khosrowshahi, D. Lavers, K. Kunkel, M. Wehner, and W. Collins, "Application of deep convolutional neural networks for detecting extreme weather in climate datasets" arXiv 2016, arXiv:1605.01156.

[31] A. Rupe, K. Kashinath, N. Kumar, V. Lee, J. P. Prabhat, and Crutchfield, "Towards unsupervised segmentation of extreme weather events" arXiv 2019, arXiv:1909.07520.

[32] G. Muszynski, K. Kashinath, V. Kurlin, M. Wehner, and Prabhat, "Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets," Methods Assess. Model. vol. 12, pp. 613–628, Feb. 07, 2019.

[33] D. Matsuoka, M. Nakano, D. Sugiyama, and S. Uchida, "Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model," Prog. Earth Planet. Sci. vol. 5, pp. 80, Dec. 19, 2018.

[34] S. Pang, P. Xie, D. Xu, F. Meng, X. Tao, B. Li, Y. Li, and T. Song, "NDFTC: A new detection framework of tropical cyclones from meteorological satellite images with deep transfer learning," Remote Sens. vol. 13, 9, pp. 1860, May. 10, 2021.

[35] Prabhat, K. Kashinath, M. Mudigonda, S. Kim, L. Kapp-Schwoerer, A. Graubner, E. Karaismailoglu, L. von Kleist, T. Kurth, and A. Greiner, "ClimateNet: An expert-labeled open dataset and deep learning architecture for enabling high-precision analyses of extreme weather," Geosci. Model Dev. vol. 14, pp. 107–124, Jan. 08, 2021.

[36] D. Huang, Y. Du, Q. He, W. Song, and A. Liotta, "DeepEddy: A simple deep architecture for mesoscale oceanic eddy detection in SAR images," in *Proc. ICNSC*, Calabria, Italy, 16–18 May 2017, pp. 673–678.

[37] M. Krinitskiy, P. Verezemskaya, K. Grashchenkov, N. Tilinina, S. Gulev, and M. Lazzara, "Deep convolutional neural networks capabilities for binary classification of Polar mesocyclones in satellite mosaics," Atmosphere. vol. 9, 11, pp. 426, Oct. 31, 2018.

[38] M. Krinitskiy, P. Verezemskaya, S. Elizarov, and S. Gulev, "Machine learning methods for the detection of polar lows in satellite mosaics: Major issues and their solutions," *IOP Conf.* Ser. Earth Environ. Sci. vol. 606, pp. 012025, 2020.

[39] R. Wood, and D. L. Hartmann, "Spatial variability of liquid water path in marine low cloud: the importance of mesoscale cellular convection," J. Clim. vol. 19, pp. 1748–1764, May. 06, 2006.

[40] A. Muhlbauer, I. L. McCoy, and R. Wood, "Climatology of stratocumulus cloud morphologies: Microphysical properties and radiative effects," Atmos. Chem. Phys. vol. 14, pp. 6695–6716, Jul. 03, 2014

[41] I.L. McCoy, R. Wood, and J.K. Fletcher, "Identifying meteorological controls on open and closed mesoscale cellular convection associated with marine cold air outbreaks," JGR Atmos. vol. 122, pp. 11678–11702, Nov. 16, 2017.

[42] T. Yuan, H. Song, R. Wood, J. Mohrmann, K. Meyer, L. Oreopoulos, and S. Platnick, "Applying deep learning to NASA MODIS data to create a community record of marine low-cloud mesoscale morphology," Atmos. Meas. Tech. vol. 13, 12, pp. 6989–6997, Dec. 21, 2020.

[43] A. Vallet, and H. Sakamoto, "A multi-label convolutional neural network for automatic image annotation," J. Inf. Process. vol. 23, 6, pp. 767–775, 2015.

[44] A. Gufan, Y. Lehahn, E. Fredj, C. Price, R. Kurchin, and I. Koren, "Segmentation and tracking of marine cellular clouds observed by geostationary satellites," Int. J. Remote Sens. vol. 37, pp. 1055–1068, Feb. 12, 2016.

[45] D. Watson-Parris, S. A. Sutherland, M. W. Christensen, R. Eastman, and P. Stier, "A large-scale analysis of pockets of open cells and their radiative impact," Geophys. Res. Lett. vol. 48, e2020GL092213, Mar. 28, 2021.

[46] S. Prakash, C. M., and R. M. Gairola, "Large-scale precipitation estimation using Kalpana-1 IR measurements and its validation using GPCP and GPCC data," Theoretical and applied climatology. vol. 106, pp. 283- 293, Apr. 2011.

[47] C. Shorten, and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of big data, vol. 6, no. 1, pp. 1-48, 2019.

[48] Z. Li, H. Shen, Q. Weng, Y. Zhang, P. Dou, and L. Zhang, "Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects," ISPRS Journal of Photogrammetry and Remote Sensing. vol. 188, pp. 89-108, Jun. 2022.

[49] Y. Zhou, A. Ji, L. Zhang, and X. Xue, "Sampling-attention deep learning network with transfer learning for large-scale urban point cloud semantic

segmentation," Engineering Applications of Artificial Intelligence, vol. 117, pp. 105554, Jan. 2023.

[50] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," in *Proc., IWSSIP*, Online, 1–3 July 2020, pp. 237–242.

[51] M. Fawakherji, A. Youssef, D. Bloisi, A. Pretto, and D. Nardi, "Crop and weeds classification for precision agriculture using context-independent pixel-wise segmentation," in *Proc., IRC* Naples, Italy, 25–27 February 2019, pp. 146–152.

[52] L. Jiao, L. Huo, C. Hu, and P. Tang, "Refined UNet: UNet-based refinement network for cloud and shadow precise segmentation," Remote Sensing, vol. 12, no. 12, pp. 2001, Jun. 22, 2020.