

## An automated deep learning based satellite imagery analysis for ecology management



Haya Mesfer Alshahrani<sup>a</sup>, Fahd N. Al-Wesabi<sup>b,c,\*</sup>, Mesfer Al Duhayyim<sup>d</sup>, Nadhem Nemri<sup>e</sup>, Seifedine Kadry<sup>f</sup>, Bassam A.Y. Alqaralleh<sup>g</sup>

<sup>a</sup> Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Saudi Arabia

<sup>b</sup> Department of Computer Science, King Khalid University, Muhayel Aseer, Saudi Arabia

<sup>c</sup> Faculty of Computer and IT, Sana'a University, Sana'a, Yemen

<sup>d</sup> Department of Natural and Applied Sciences, College of Community - Aflaj, Prince Sattam bin Abdulaziz University, Saudi Arabia

<sup>e</sup> Department of Information Systems, King Khalid University, Muhayel Aseer, Saudi Arabia

<sup>f</sup> Faculty of Applied Computing and Technology, Noroff University College, Kristiansand, Norway

<sup>g</sup> Computer Science Department - IT Faculty, AL-Hussein Bin Talal University, Jordan

### ARTICLE INFO

**Keywords:**  
 Ecology  
 Biodiversity  
 Satellite imagery  
 Deep learning  
 CNN  
 Environment planning

### ABSTRACT

Ecology is the methodical study of biodiversity which affecting natural life and habitats. Due to the anthropogenic pressure on the atmosphere, there is an upraised threat to wild animals and other habitats of the ecological atmosphere. So, there is a need for efficient ecology management models to map and save nature resources. At the same time, the use of satellite imagery analysis is an effective tool for determining important details on earth resources and the platform. It finds useful for proficient ecology management, such as land use detection, forest fire detection, environment planning, and so on. Earlier satellite imagery classification approaches mainly depend upon feature coding approaches which has limited capabilities and yield mediocre outcomes. The recent developments of deep learning models made the image classification highly effective. In this view, this paper presents a new parameter tuned deep learning based EfficientNet model with Variational Autoencoder (PTDLEN-VAE) model for satellite imagery analysis on ecology management. The presented PTDLEN-VAE model includes a series of operations namely pre-processing, feature extraction, and classification. Primarily, the satellite images are preprocessed to improve the contrast level of the image. Followed by, the PTDLEN based feature extractor is utilized to derive a useful set of feature vectors from the aerial image. Besides, the improved krill herd optimization (IKHO) algorithm is applied for the parameter tuning of the EfficientNet model. Finally, the classification of aerial images using the derived feature vectors takes place by the use of the VAE model.

The efficacy of the PTDLEN-VAE model is validated using a benchmark aerial image dataset and the resultant experimental values highlighted the effectiveness of the PTDLEN-VAE model interms of precision, recall, F1-score, F2-score, and computation time.

### 1. Introduction

Generally, ecological research represents the study of living things and their surrounding environments, include abiotic and biotic entities. Because of the complex nature of biodiversity, it is complicated for simplifying the measure and express biodiversity. Biodiversity must be interrelated to the variance of lifestyles, however ecological multiplexes are part of them. Conservation has to turn into an essential manner of handling the enhanced native ecosystem degradation and conversion

that have considerable negative effects on biodiversity. Biological invasion represents one of the most common risks to ecosystems and local communities (Raffini et al., 2020). Invasion occurs if species are accidentally/intentionally presented from the historic/native range to a novel region, distributed in the new environment effectively and adversely affects it (Hoffmann and Broadhurst, 2016; Hulme et al., 2009). Especially, conspicuous exist the social and ecological effects of alien organism on native species and ecosystems that includes, decreased and availability quality of key natural resources, pollution,

\* Corresponding author at: Department of Computer Science, King Khalid University, Muhayel Aseer, Saudi Arabia.

E-mail addresses: [hmalshahrani@pnu.edu.sa](mailto:hmalshahrani@pnu.edu.sa) (H.M. Alshahrani), [fawesabi@kku.edu.sa](mailto:fawesabi@kku.edu.sa) (F.N. Al-Wesabi), [m.alduhayyim@psau.edu.sa](mailto:m.alduhayyim@psau.edu.sa) (M. Al Duhayyim), [nnemri@kku.edu.sa](mailto:nnemri@kku.edu.sa) (N. Nemri), [seifedine.kadry@noroff.no](mailto:seifedine.kadry@noroff.no) (S. Kadry), [alqaralleh@ahu.edu.jo](mailto:alqaralleh@ahu.edu.jo) (B.A.Y. Alqaralleh).

reduced biodiversity, flooding and increased frequency of wildfires (Charles and Dukes, 2008).

Satellite Image is an image of the entire or portion of the world captured by artificial satellites. It could neither be visible light images, water vapour nor infrared images (Shafaey et al., 2018). The distinct kinds of satellite produce (higher temporal, spatial, and spectral) resolution images which cover the entire World almost every day. The classification and analyses of remote sensing images are highly significant in various real time applications like geospatial object detection, natural hazards, urban planning, precision agriculture, military monitoring, and vegetation mapping (Zhang et al., 2016). Over many years of investigation, the degree of automation for remote sensing image analyses are remained lower (Marmanisad et al., 2016). Satellite Imaging Corporation (SIC) gives satellite image data with the integration of multispectral and panchromatic bands to show up the environmental features which are primary to our projects. GIS maps and Satellite images have increased the chances for modelling, satellite map data integration, satellite map production, and analysis for assessment and monitoring.

As landscapes change, population grows, countries boost their economies, governments have gradually based on current satellite images and other geospatial data for applications like disaster response, agricultural biodiversity, conversation, public health, forestry High resolution satellite imagery, and land registration have assisted systematic study events at regional and landscape scales. Availability of satellite images could give spatial resolution of 0.31 m or enhanced for analyses of transportation development and growth urban for monitoring and assessment. Multispectral sensors could offer improved spectral resolution which could be utilized for additional analyses of change detection and land cover, and how related transportation development and urban growth influence these situations.

### 1.1. Role of DL models for satellite imagery classification

The development of image classification methods finds an application in many areas. The image classification could be separated into 3 major types based on its features. The ‘handcrafted feature based approach’ emphasizes on distinct properties like shape and colour information’s, that is likely properties of sense images (Bian et al., 2017), whereas ‘unsupervised feature learning based approaches’ goal for learning a set of fundamental functions like bag of words module which is utilized to encode the features. The popular encoding technique is named quantization, and highly efficient technique is fisher encoding, whereas the input in Fisher technique is a group of handcrafted features, and the output is a group of learned features (Yuan et al., 2016). Lastly, the ‘deep feature learning based techniques’ are known as Deep Learning (DL). Recently, DL of remote sensing image features are demonstrated a significant ability for classification by selecting suitable feature to the problems of remote sensing image classifier (Cheng et al., 2017). Selecting a suitable DL is a sub region of ML which is depending upon multiple layers of learning. The DL framework expands from the classical NN by including additional layers to the hidden layer portion. It has several frameworks DL, where CNN is one of them. The CNN is extensive and has been utilized recently to handle a different and difficult challenge like image classification and recognition through a series of feed forward layers.

### 1.2. Paper's contribution

This paper presents a new parameter tuned deep learning based EfficientNet model with Variational Autoencoder (PTDLEN-VAE) model for satellite imagery analysis on ecology management. The goal of the PTDLEN-VAE model is to examine the satellite images using DL model for ecological condition monitoring. Initially, the satellite images are pre-processed to improve the contrast level of the image. Besides, the PTDLEN based feature extractor is utilized using EfficientNet model for

the generation of feature vectors from satellite images. In addition, an improved krill herd optimization (IKHO) algorithm is applied for the parameter tuning of the EfficientNet model. The IKHO algorithm is designed based on the integration of conventional KH with free search (FS) operator for avoiding local optima problems. Furthermore, VAE based classification model is used for the allocation of class labels to the satellite images. For validating the improved performance of the PTDLEN-VAE technique, an extensive experimental analysis take place on benchmark aerial image dataset. In short, the key contributions are listed as follows.

- A novel PTDLEN-VAE model encompasses preprocessing, PTDLEN based feature extraction, parameter optimization, and classification is presented. To the best of author's knowledge, the PTDLEN-VAE model has been never presented in the literature.
- The IKHO algorithm is designed by the integration of KH algorithm with free search (FS) operator to avoid the local optimal problem of the KH algorithm.
- A PTDLEN based feature extractor is presented by the use of parameter tuned EfficientNet model using IKHO algorithm, which aids to enhance the prediction outcome of the PTDLEN-VAE model for unseen data.
- Validation of the PTDLEN-VAE model takes place on benchmark dataset and the experimental results are discussed in terms of different aspects.

### 1.3. Paper organization

The rest of the paper is planned as follows. Section 2 elaborates the previous satellite imagery analysis techniques. Besides, Section 3 discusses the proposed PTDLEN-VAE model and Section 4 offers the experimental validation. At last, Section 5 summarizes the paper.

## 2. Prior satellite imagery classification models

Unnikrishnan et al. (2019) employs the red and near infrared (NIR) band data for the publicly available SAT-4 & SAT-6 datasets classification. It is achieved, as NDVI computation needs 2 bands (red and NIR) information's and the class included in the datasets are vegetation kinds. In this study, novel DL frameworks for 3 distinct networks (VGG, AlexNet, ConvNet<sub>x</sub>) have been presented using hyper tuning the network and the input as 2 band data. The adapted architecture using 2 band information's together with decreased amount of filters are tested and trained module handles for classifying the images into distinct kinds. Kadhim and Abed (2019) produced an efficient method for classifying satellite images which are depending upon DL and utilize CNN for extracting the features with the help of VGG19, AlexNet, Resnet50 pretraining, and GoogLeNet modules.

Liu et al. (2020) presented 2 novel high resolution satellite imagery datasets (SAT-4 and SAT-6) and presented a DeepSat architecture for classification is depending upon DBN and “handcrafted” features. The presented study is an extensive form, they presented an end to end architecture leverage an enhanced framework that increases a CNN using handcrafted features (rather than utilizing DBN based framework) for classification. Pelletier et al. (2019) proposed extensive research of TempCNNs, a DL method that employs convolution in the temporal dimension for manually learning the temporal (also spectral) features. The aim of this study is to qualitatively and quantitatively calculate the contributions of TempCNN for classifying SITS, as related to RNN and RF a regular DL method is mainly fitted to temporal data.

Laban et al. (2020) propose an improved method for the classification of satellite images with the help of CNN. It has 2 features of satellite images which makes efficiency problems very critical; firstly, higher data contents within the satellite image, and then, higher computation basics utilize CNN. The development method is based on an effectual selection of appropriate image scales. Xia et al. (2020) proposed a

dilated multi scale cascade forest technique for realizing the classification of satellite cloud images. Multi scale scanning rises the variety of feature vector extractions and enhances feature separability. Additionally, the amount of layers of network cascading doesn't need to be personalized, it could be manually established as per the efficiency of the scheme.

Lunga et al. (2020) presented a new remote sensing data flow (RESFlow) to an advanced ML for computing huge amount of remotely sensed images. The important role is separating huge amount of data for homogeneous distribution to fit easier modules. Nguyen et al. (2020) proposed a new multi temporal high spatial resolution classification technique using an innovative spatio temporal spectral DNN for locating paddy fields at the pixel level for an entire year and temporal instances. Boulila et al. (2021) proposed a new distributed DL based method for classifying the big remote sensing images. Particularly, they proposed an RS-DCNN for managing image classification. The initial phase is to make the training dataset for RS-DCNN. Later, for ensuring a data parallel training on Apache Spark architecture, a pixel based CNN module through the big data cluster is executed by BigDL.

### 3. Materials and methods

#### 3.1. Dataset used

The performance of the PTDLEN-VAE model is examined against UCM dataset and AID dataset (<http://weegee.vision.ucmerced.edu/datasets/landuse.html>, n.d.; <https://captain-whu.github.io/AID/>, n.d.). The UCM dataset comprises a set of 100 images with 21 class labels. Each image measures  $256 \times 256$  pixels. The pixel resolution of this public domain imagery is 1 ft. Fig. 1 showcases the sample images from UCM dataset. The AID multi-label dataset has 3000 RS images from the AID dataset.

#### 3.2. Overall system architecture

The overall system architecture of the PTDLEN-VAE model is



**Fig. 1.** Sample satellite images.

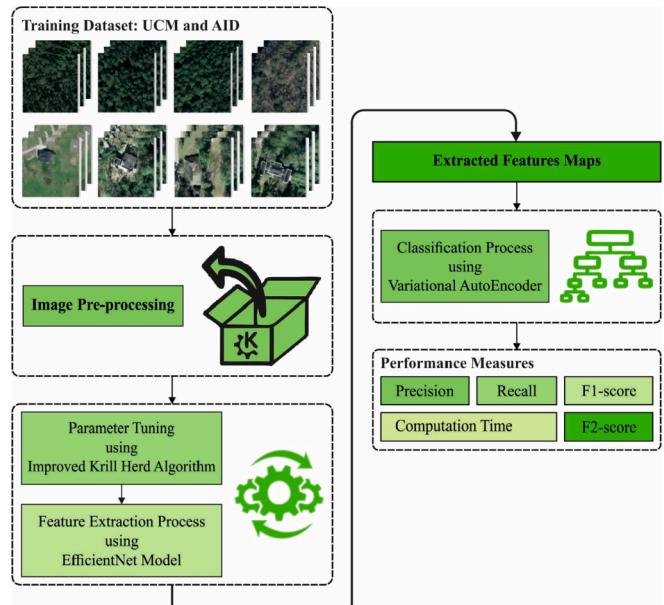
displayed in Fig. 2. The presented PTDLEN-VAE model includes a series of operations such as pre-processing, EfficientNet based feature extraction, IKHO based parameter optimization, and VAE based classification. Once the contrast level of the input satellite images is improved, the features are extracted by the EfficientNet model, and the hyperparameter of the EfficientNet model is optimally adjusted by the IKHO algorithm. Finally, the VAE based classification model is executed to assign the proper class labels to the satellite images.

#### 3.3. Structure of EfficientNet model

The EfficientNet model is employed as a feature extraction module to generate a useful set of feature vectors of the input satellite images. DL is one of the famous structures as DL techniques are learned relevant features from input images at distinct convolution levels same as the function of human brain. The DL is resolving difficult issues mostly well and rapidly with superior classifier accuracy and a minimum error rate. The DL technique is comprised of various modules (convolution, pooling layer, and fully connected (FC) layers, and activation function). A DL structure goal to gain optimum efficiency accuracy and performance with lesser techniques. Different from other recent DL techniques, the EfficientNet framework is compound scaling approach which utilizes the compound coefficient for uniformly scale network width, depth, and resolution. The EfficientNet has 8 various techniques in B0 to B7. EfficientNet utilizes inverted bottleneck convolutional that is initially established in the MobileNetV2 method that has a layer which initially expands the networks and next compresses the channel. This framework decreased calculation by the factor of 2 as related to normal convolutional, in which  $f$  implies the filter size. The researchers in Hassan et al. (2021) outperformed that EfficientNetB0 is the easiest of every 8 techniques and utilizes less parameters. Therefore, it is directly utilized EfficientNetB0 for evaluating efficiency. Fig. 3 illustrates the layered architecture of EfficientNet.

#### 3.4. Design of IKHO algorithm for parameter optimization

In order to optimally adjust the hyperparameter of the EfficientNet model, they are optimally adjusted by the use of IKHO algorithm. KH (Wei and Wang, 2020) is a SI based approach which is based on the swarming behavior of krills based on a particular environmental and biological process. The KH algorithm contains 3 main processes namely



**Fig. 2.** Overall working process of PTDLEN-VAE model.

| Stage | Operator           | Resolution | #channels | #Layers |
|-------|--------------------|------------|-----------|---------|
| 1     | Conv3x3            | 224x224    | 32        | 1       |
| 2     | MBCConv1, k3x3     | 112x112    | 16        | 1       |
| 3     | MBCConv6, k3x3     | 112x112    | 24        | 2       |
| 4     | MBCConv6, k5x5     | 56x56      | 40        | 2       |
| 5     | MBCConv6, k3x3     | 28x28      | 80        | 3       |
| 6     | MBCConv6, k5x5     | 14x14      | 112       | 3       |
| 7     | MBCConv6, k5x5     | 14x14      | 192       | 4       |
| 8     | MBCConv6, k3x3     | 7x7        | 320       | 1       |
| 9     | Conv1x1/Pooling/FC | 7x7        | 1280      | 1       |

Fig. 3. Layered architecture of EfficientNet.

movement influenced by other krills, foraging action, and arbitrary diffusion. The KH algorithm modifies the Lagrangian approach in a  $d$ -dimension decision space using Eq. (1):

$$\frac{dX_i}{dt} = N_i + F_i + D_i \quad (1)$$

Where  $N_i$ ,  $F_i$ , and  $D_i$  denotes the motion supported by other krills, physical diffusion, and the forage motion. In case, the movement gets affected using other kills, the way of movement,  $\alpha_i$ , is determined using local, and target outcomes. For any krill individuals, the movement can be defined using Eq. (2):

$$N_i^{\text{new}} = N_i^{\max} \alpha_i + \omega_n N_i^{\text{old}} \quad (2)$$

and  $N^{\max}$  means higher induce speed,  $\omega_n$  specifies inertia weight of the motion induced in 0 and 1, and  $N_i^{\text{old}}$  characterizes last motion encouraged. The forage motion is determined using two main components. Primarily, it defines the place of the food source, and then it indicates the earlier knowledge about the place of the food. For  $i$ th krill individual, the motion is defined by:

$$F_i = V_f \beta_i + \omega_f F_i^{\text{old}} \quad (3)$$

where

$$\beta_i = \beta_i^{\text{food}} + \beta_i^{\text{best}} \quad (4)$$

Here,  $V_f$  means forage speed,  $\omega_f$  signifies inertia weight of forage motion amongst 0 to 1, and  $F_i^{\text{old}}$  designates last foraging motion. The arbitrary diffusion of krills can be considered as an arbitrary function. It can be represented based on maximum diffusion speed and randomized direction vector, equated in Eq. (5):

$$D_i = D^{\max} \delta \quad (5)$$

where  $D^{\max}$  implies maximal diffusion speed and  $\delta$  exemplifies random direction vector. Next, the position vector of the krills in the duration of  $t$  to  $t + \Delta t$  can be defined by:

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt} \quad (6)$$

It is differentiated that  $\Delta t$  is an important parameter which needs to be optimally chosen for effective outcomes and is represented as a scaling factor of the speed vector. Fig. 4 illustrates the flowchart of KH technique.

For avoiding the local optima problem, the IKHO technique is derived based on the concept of FS. FS (Peney and Littlefair, 2005) model the behaviors of animal and operate on a group of solutions named population. In this method, every animal has original peculiarities named mobility and sense. The animal utilizes its sense to select a position for the future move. Distinct animals could have distinct sensibilities. It also differs at the time of optimization procedure, and individual animals could have distinct sensibilities beforehand distinct

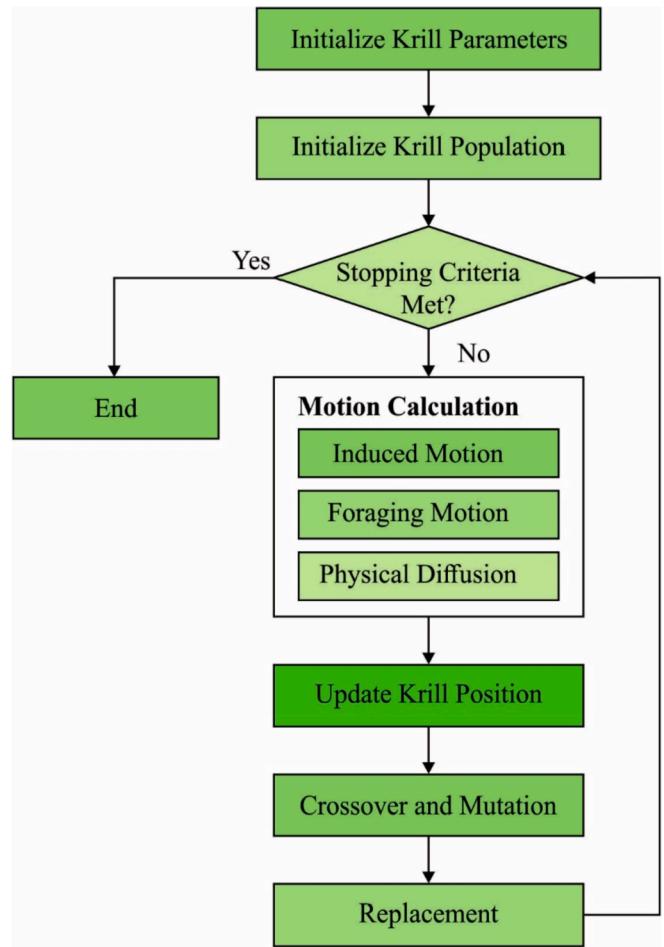


Fig. 4. Flowchart of KH algorithm.

walks. The animal could choose, for starting the exploration walk, some locations are marked with pheromone that fits its sense. Thus, the likelihood of accessing some locations of the search space is null. In the exploration, all krills attains few favors (an objective function solution) and distribute a pheromone in quantity that is related to the quantity of founded favors (the quality of the solution). The pheromone is completely substituted by a novel one afterward every walks.

Specifically, the animals in the process are mobile. Every animal could function whether using smaller accurate steps for local search or using larger steps for global exploration. Also, every animal chooses how to search (using smaller or larger steps) by themselves. The explicit restriction doesn't exists. The prior knowledge could be considered, however it isn't mandatory.

The framework of this method contains 3 main activities: initiation, exploration, and termination.

- (1) Initiation. In this study, the initiation approach is  $x_{0ji} = x_{i \min} + (x_{i \max} - x_{i \min}) * \text{random}_{ji}(0, 1)$ , Whereas haphazard  $(0, 1)$  is an arbitrary value in  $[0, 1]$  and  $x_{i \min}$ ,  $x_{i \max}$  represents the space border.
- (2) Exploration. The exploration walk creates coordinate of a novel position  $x_{tji}$  as

$$X_{tji} = x_{0ji} - \Delta x_{tji} + 2 * \Delta x_{tji} * \text{random}_{tji}(0, 1) \quad (7)$$

The alteration approach is

$$\Delta x_{tji} = R_{ji} * (x_{i \max} - x_{i \min}) * \text{random}_{tji}(0, 1) \quad (8)$$

Whereas  $T$  denotes the step limits for each walk and  $t = 1, \dots, T$  denotes the present step.

The specific behavior, at the time of the walk, is modelled and illustrated as  $f_{ij} = f(x_{ji})$ ,  $f_j = \max(f_{ij})$ , whereas  $f_j$  denotes the position marked by the pheromone from an animal afterward the walk (Li et al., 2014). The pheromone generation can be given by

$$P_j = \frac{f_j}{\max(f_j)} \quad (9)$$

The sensibility generation can be denoted as

$$S_j = S_{\min} + \Delta S_j \quad (10)$$

$$\Delta S_j = (S_{\max} - S_{\min}) * \text{random}_j(0, 1),$$

While  $S_{\min}$  &  $S_{\max}$  denotes the minimum and maximum likely values of sensibility.  $P_{\min}$  &  $P_{\max}$  represents the minimum and maximum likely values of pheromone trails. Also  $P_{\max} = S_{\max}$ ,  $P_{\min} = S_{\min}$ .

Selection and decision making for an initial position  $x'_{0ji}$  for an exploration walk is

$$x'_{0ji} = \begin{cases} x'_{0ji} & (P_k < S_j) \\ x_{ji} & (P_k \geq S_j) \end{cases} \quad (11)$$

whereas  $j = 1, \dots, NK$ ,  $k = 1, \dots, NK$ , and  $k$  denotes the marked location number.

- (3) Termination. In this study, the condition for end is  $\text{iter} > \text{iterMax}$ , whereas  $\text{iterMax}$  denotes the maximal amount of iterations.

### 3.5. VAE based classification

During the classification process, the VAE model gets executed to assign proper class labels to the applied satellite images. Autoencoder (AE) is a kind of NN which is trained to replace the input with the output. It includes a hidden layer  $h$  to define a code employed for input representation. The network comprises of an encoder function  $z = f(x)$  and decoder function  $r = g(z)$ , where  $x$  denotes the input data. The meaningful features from the AE by the constraint that the  $z$  to have lower dimensions compared to  $x$ . The code of the AE which is lesser than the input dimensions are known as under complete. The learning of the undercomplete depiction enforces the AE in capturing the important characteristics of the training data. The VAE is a directed network which makes use of learned approximate inferences and undergoes training using gradient approaches. For generating a sample from the model, the VAE initially draw an instance  $z$  from the code distribution  $p_{model}(z)$ . The sample is executed via a differentiable generator network  $g(z)$ . Lastly,  $x$  undergo sampling from a distribution  $P_{model}(x|g(z)) = p_{model}(x|z)$ . At the time of training process, the estimated inference network (or encoder)  $q(z|x)$  can be employed for obtaining  $z$  and  $P_{model}(x|z)$  is afterward considered as the decoder network (Dai et al., 2019). The main perception of VAE is that it can be trained using the maximization of the variational lower bound  $L(q)$  related to the data point  $x$ :

$$L(q) = E_{z \sim q(z|x)} \log p_{model}(z, x) + H(q(z|x)) \quad (12)$$

$$= E_{z \sim q(z|x)} \log p_{model}(x|z) - D_{KL}(q(z|x)||p_{model}(z)) \quad (13)$$

The initial term in Eq. (12) refers to the joint log likelihood of the visible and hidden parameters. Then, the next term denotes the entropy of the approximate posterior. If  $q$  can be selected as a Gaussian distribution, with noise appended to the estimated average value, maximization of the entropy term inspires the raise of noise standard deviation. The entropy element encourage the variational posterior in placing maximum high probability mass on several  $z$  values which can generate  $x$ , instead of failing to individual point estimation. Besides, the initial term in Eq. (12) refers to the reconstruction log-likelihood in the A. Then, the next term denotes the estimated posterior distribution  $q(z|x)$  and the model prior  $p_{model}(z)$ . Fig. 5 demonstrates the structure of VAE.

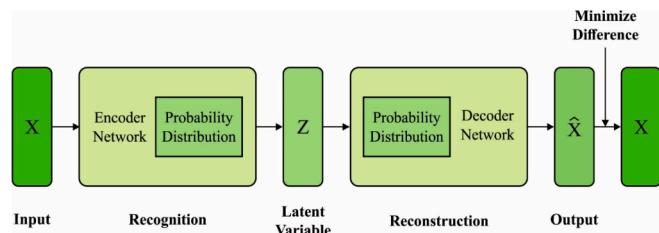


Fig. 5. Structure of VAE.

Stochastic gradient descent (SGD) on BP is managing stochastic input, then not stochastic unit within the networks. The solution is named as “reparameterization trick”, which is to transfer the sampling to input layer. It is easy from  $N(\mu(x), \theta(x))$  by sampling  $\epsilon \sim N(0, I)$ , afterward calculating  $p_{model}z = \mu(x) + \theta^{1/2}(x) * \epsilon$ . Where  $\mu(x)$  and  $\theta(x)$  are the mean and covariance of  $(z|x)$ . So, Eq. (13) is calculated as:

$$L(q) = E_{\epsilon \sim N(0,I)} p_{model}(x|z = \mu(x) + \theta^{1/2}(x) * \epsilon) - D_{KL}(q(z|x)||p_{model}(z)) \quad (14)$$

In VAE is comprised of input layer, various AEs, and output layer. Then, an unsupervised pre-training step, the supervised fine-tuning step is implemented for learning the entire network parameters by employing the BP technique. This technique is comprised of 1 input layer, 5 hidden layers, and 1 output layer.

### 4. Performance validation

The projected method is simulated employing PC i5-8600k processor, GeForce 1050Ti, 4GB RAM, 16GB OS Storage, and 250GB SSD File Storage. The simulation tool utilized is Python 3.6.5 tool together with few packages such as tensorflow, keras, numpy, pickle, matplotlib, sklearn, pillow, and opencv-python. The presented technique is experimented with utilizing a benchmark UCM and AID dataset. Also, the outcomes are studied under folds. The parameter setting of the PTDLEN-VAE method is provided as follows: mini batch size: 200, dropout: 0.5, count of hidden layers: 3, amount of hidden units: 1024, and activation function: softmax.

This section offers a brief experimental results analysis of the PTDLEN-VAE model over the other techniques. Table 1 and Fig. 6 assess the classification outcomes analysis of the PTDLEN-VAE model with other techniques on the applied UCM dataset. On examining the results interms of precision, the experimental results showcased that the PTDLEN-VAE model has accomplished maximum average precision of 96% whereas the DLEN-VAE, DLEN-DNN, and DLEN-LSTM techniques have gained a minimum average precision of 95.70%, 95.68%, and 95.49% respectively. Followed by, on investigative the outcomes with respect to recall, the experimental outcomes demonstrated that the PTDLEN-VAE manner has accomplished maximal average recall of 95.35% whereas the DLEN-VAE, DLEN-DNN, and DLEN-LSTM methods have reached a lesser average recall of 94.90%, 94.93%, and 94.73% correspondingly. Eventually, on exploratory the outcomes interms of F1-score, the experimental results portrayed that the PTDLEN-VAE manner has accomplished superior average F1-score of 95.84% whereas the DLEN-VAE, DLEN-DNN, and DLEN-LSTM algorithms have attained a minimum average F1-score of 95.62%, 95.58%, and 95.37% correspondingly. Meanwhile, on examining the results with respect to F2-score, the experimental outcomes exhibited that the PTDLEN-VAE method has accomplished higher average F2-score of 95.67% whereas the DLEN-VAE, DLEN-DNN, and DLEN-LSTM methodologies have achieved a minimal average F2-score of 95.47%, 95.45%, and 95.35% correspondingly.

Table 2 and Fig. 7 evaluate the classification outcomes analysis of the PTDLEN-VAE method with other approaches on the applied AIM dataset. On investigative the outcomes with respect to precision, the

**Table 1**

Result analysis of different folds on proposed PTDLEN-VAE methods with other methods on UCM aerial dataset.

| No. of Folds  | PTDLEN-VAE   | DLEN-VAE     | DLEN-DNN     | DLEN-LSTM    |
|---------------|--------------|--------------|--------------|--------------|
| Precision (%) |              |              |              |              |
| NF = 6        | 96.70        | 96.39        | 96.35        | 96.21        |
| NF = 7        | 96.22        | 95.91        | 95.84        | 95.72        |
| NF = 8        | 94.19        | 94.01        | 94.03        | 93.79        |
| NF = 9        | 96.75        | 96.47        | 96.46        | 96.26        |
| NF = 10       | 96.15        | 95.70        | 95.71        | 95.45        |
| Average       | <b>96.00</b> | <b>95.70</b> | <b>95.68</b> | <b>95.49</b> |
| Recall (%)    |              |              |              |              |
| NF = 6        | 94.81        | 94.62        | 94.59        | 94.34        |
| NF = 7        | 95.93        | 95.40        | 95.33        | 95.21        |
| NF = 8        | 95.51        | 95.23        | 95.17        | 95.04        |
| NF = 9        | 96.38        | 95.42        | 95.86        | 95.48        |
| NF = 10       | 94.13        | 93.83        | 93.71        | 93.58        |
| Average       | <b>95.35</b> | <b>94.90</b> | <b>94.93</b> | <b>94.73</b> |
| F1-score (%)  |              |              |              |              |
| NF = 6        | 95.93        | 95.65        | 95.63        | 95.36        |
| NF = 7        | 96.43        | 96.37        | 96.26        | 96.18        |
| NF = 8        | 96.29        | 95.95        | 95.97        | 95.70        |
| NF = 9        | 94.86        | 94.68        | 94.71        | 94.37        |
| NF = 10       | 95.67        | 95.47        | 95.35        | 95.25        |
| Average       | <b>95.84</b> | <b>95.62</b> | <b>95.58</b> | <b>95.37</b> |
| F2-score (%)  |              |              |              |              |
| NF = 6        | 95.63        | 95.55        | 95.48        | 95.37        |
| NF = 7        | 96.34        | 96.22        | 96.16        | 96.10        |
| NF = 8        | 96.07        | 95.85        | 95.78        | 95.69        |
| NF = 9        | 94.78        | 94.52        | 94.51        | 94.42        |
| NF = 10       | 95.53        | 95.23        | 95.31        | 95.15        |
| Average       | <b>95.67</b> | <b>95.47</b> | <b>95.45</b> | <b>95.35</b> |

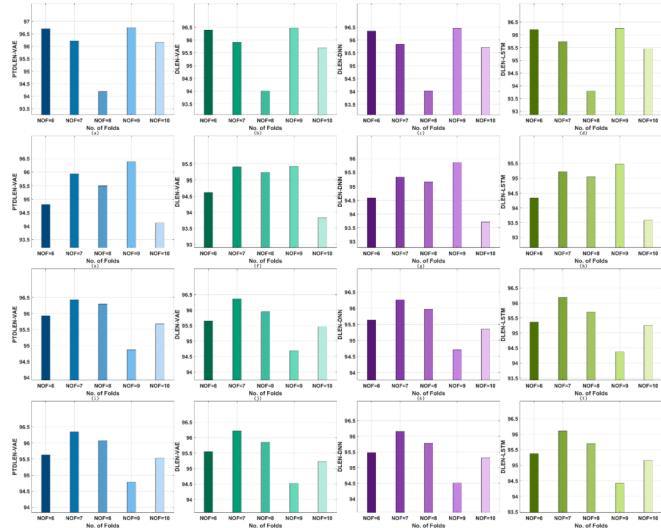


Fig. 6. Results of UCM Dataset (a-d) Precision (e-h) Recall (i-l) F1-score (m-p) F2-score.

experimental outcomes outperformed that the PTDLEN-VAE manner has accomplished enhanced average precision of 95.90% whereas the DLEN-VAE, DLEN-DNN, and DLEN-LSTM methodologies have obtained a minimal average precision of 95.58%, 95.57%, and 95.43% correspondingly. In addition, on examining the outcomes in terms of recall, the experimental results depicted that the PTDLEN-VAE algorithm has accomplished maximal average recall of 95.23% whereas the DLEN-VAE, DLEN-DNN, and DLEN-LSTM techniques have attained a minimal average recall of 94.81%, 94.80%, and 94.69% respectively. Similarly, on determining the results with respect to F1-score, the experimental results demonstrated that the PTDLEN-VAE manner has accomplished maximum average F1-score of 95.75% whereas the DLEN-

**Table 2**

Result analysis of different folds on proposed PTDLEN-VAE methods with other methods on AID multi-label dataset.

| No. of Folds  | PTDLEN-VAE   | DLEN-VAE     | DLEN-DNN     | DLEN-LSTM    |
|---------------|--------------|--------------|--------------|--------------|
| Precision (%) |              |              |              |              |
| NF = 6        | 96.50        | 96.28        | 96.27        | 96.17        |
| NF = 7        | 96.11        | 95.76        | 95.75        | 95.62        |
| NF = 8        | 94.18        | 93.90        | 93.88        | 93.77        |
| NF = 9        | 96.73        | 96.37        | 96.35        | 96.17        |
| NF = 10       | 95.99        | 95.58        | 95.61        | 95.43        |
| Average       | <b>95.90</b> | <b>95.58</b> | <b>95.57</b> | <b>95.43</b> |
| Recall (%)    |              |              |              |              |
| NF = 6        | 94.80        | 94.45        | 94.52        | 94.29        |
| NF = 7        | 95.74        | 95.24        | 95.32        | 95.16        |
| NF = 8        | 95.37        | 95.07        | 95.17        | 95.03        |
| NF = 9        | 96.25        | 95.73        | 95.26        | 95.45        |
| NF = 10       | 94.00        | 93.58        | 93.71        | 93.50        |
| Average       | <b>95.23</b> | <b>94.81</b> | <b>94.80</b> | <b>94.69</b> |
| F1-score (%)  |              |              |              |              |
| NF = 6        | 95.79        | 95.42        | 95.53        | 95.35        |
| NF = 7        | 96.49        | 95.95        | 96.15        | 96.08        |
| NF = 8        | 96.09        | 95.73        | 95.84        | 95.66        |
| NF = 9        | 94.83        | 94.48        | 94.57        | 94.37        |
| NF = 10       | 95.55        | 95.67        | 95.22        | 95.17        |
| Average       | <b>95.75</b> | <b>95.45</b> | <b>95.46</b> | <b>95.33</b> |
| F2-score (%)  |              |              |              |              |
| NF = 6        | 95.81        | 95.64        | 95.48        | 95.30        |
| NF = 7        | 96.51        | 96.17        | 96.23        | 96.05        |
| NF = 8        | 96.07        | 95.64        | 95.84        | 95.71        |
| NF = 9        | 94.85        | 94.51        | 94.58        | 94.39        |
| NF = 10       | 95.56        | 95.28        | 95.23        | 95.10        |
| Average       | <b>95.76</b> | <b>95.46</b> | <b>95.47</b> | <b>95.31</b> |

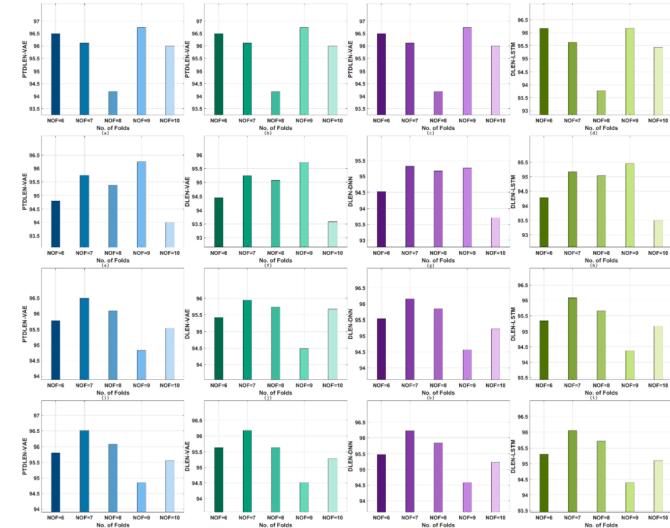


Fig. 7. Results of AID Dataset (a-d) Precision (e-h) Recall (i-l) F1-score (m-p) F2-score.

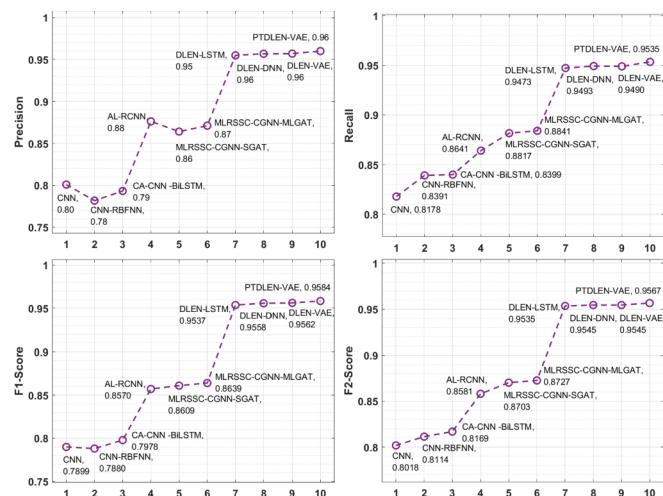
VAE, DLEN-DNN, and DLEN-LSTM techniques have attained a minimum average F1-score of 95.45%, 95.46%, and 95.33% correspondingly. In the meantime, on examining the results with respect to F2-score, the experimental outcomes outperformed that the PTDLEN-VAE technique has accomplished superior average F2-score of 95.76% whereas the DLEN-VAE, DLEN-DNN, and DLEN-LSTM algorithms have achieved a lower average F2-score of 95.46%, 95.47%, and 95.31% correspondingly.

A brief comparative results analysis of the PTDLEN-VAE model takes place with other methods on the UCM dataset in Table 3 and Fig. 8. From the results, it is noticeable that the CNN-RBFNN, CA-CNN-BiLSTM, and CNN models have demonstrated worse performance over the other

**Table 3**

Performances of proposed PTDLEN-VAE with existing methods on the UCM multi-label dataset.

| Methods           | Precision | Recall | F1-Score | F2-Score |
|-------------------|-----------|--------|----------|----------|
| CNN               | 0.8009    | 0.8178 | 0.7899   | 0.8018   |
| CNN-RBFNN         | 0.7818    | 0.8391 | 0.7880   | 0.8114   |
| CA-CNN-BiLSTM     | 0.7933    | 0.8399 | 0.7978   | 0.8169   |
| AL-RCNN           | 0.8762    | 0.8641 | 0.8570   | 0.8581   |
| MLRSSC-CGNN-SGAT  | 0.8641    | 0.8817 | 0.8609   | 0.8703   |
| MLRSSC-CGNN-MLGAT | 0.8711    | 0.8841 | 0.8639   | 0.8727   |
| DLEN-LSTM         | 0.9549    | 0.9473 | 0.9537   | 0.9535   |
| DLEN-DNN          | 0.9568    | 0.9493 | 0.9558   | 0.9545   |
| DLEN-VAE          | 0.9570    | 0.9490 | 0.9562   | 0.9545   |
| PTDLEN-VAE        | 0.9600    | 0.9535 | 0.9584   | 0.9567   |

**Fig. 8.** Result analysis of PTDLEN-VAE model on UCM dataset.

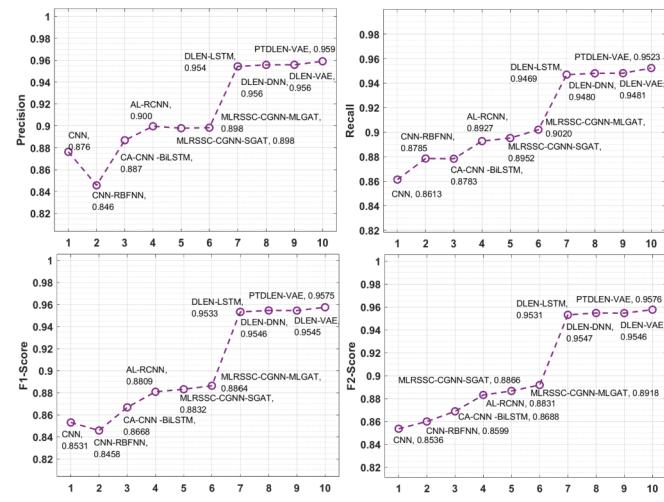
methods. Along with that, MLRSSC-CGNN-SGAT, MLRSSC-CGNN-MLGAT, and AL-RCNN techniques have accomplished moderately closer outcomes. In line with, the DLEN-LSTM, DLEN-DNN, and DLEN-VAE techniques have resulted in competitive outcome in terms of different measures. However, the proposed PTDLEN-VAE model has obtained improved performance with the higher precision of 0.9600, recall of 0.9535, F1-score of 0.9584, F2-score of 0.9567.

A detailed comparative outcomes analysis of the PTDLEN-VAE manner takes place with other techniques on the AIM dataset in Table 4 and Fig. 9 (Li et al., 2020). From the outcomes, it can be obvious that the CNN-RBFNN, CA-CNN-BiLSTM, and CNN manners have outperformed least performance over the other techniques. In line with, MLRSSC-CGNN-SGAT, MLRSSC-CGNN-MLGAT, and AL-RCNN algorithms have accomplished reasonably closer results. Besides, the DLEN-LSTM, DLEN-DNN, and DLEN-VAE approaches have resulted in competitive outcomes with respect to various measures. But, the

**Table 4**

Performances of proposed PTDLEN-VAE with existing methods on the AID multi-label dataset.

| Methods           | Precision | Recall | F1-Score | F2-Score |
|-------------------|-----------|--------|----------|----------|
| CNN               | 0.8762    | 0.8613 | 0.8531   | 0.8536   |
| CNN-RBFNN         | 0.8456    | 0.8785 | 0.8458   | 0.8599   |
| CA-CNN-BiLSTM     | 0.8868    | 0.8783 | 0.8668   | 0.8688   |
| AL-RCNN           | 0.8996    | 0.8927 | 0.8809   | 0.8831   |
| MLRSSC-CGNN-SGAT  | 0.8978    | 0.8952 | 0.8832   | 0.8866   |
| MLRSSC-CGNN-MLGAT | 0.8983    | 0.9020 | 0.8864   | 0.8918   |
| DLEN-LSTM         | 0.9543    | 0.9469 | 0.9533   | 0.9531   |
| DLEN-DNN          | 0.9557    | 0.9480 | 0.9546   | 0.9547   |
| DLEN-VAE          | 0.9558    | 0.9481 | 0.9545   | 0.9546   |
| PTDLEN-VAE        | 0.9590    | 0.9523 | 0.9575   | 0.9576   |

**Fig. 9.** Result analysis of PTDLEN-VAE model on AID dataset.

presented PTDLEN-VAE methodology has gained maximum performance with the superior precision of 0.9590, recall of 0.9523, F1-score of 0.9575, F2-score of 0.9576.

A brief CT analysis of the PTDLEN-VAE model with other techniques on the applied UCM and AID dataset is shown in Table 5. Fig. 10 investigates the CT analysis of the PTDLEN-VAE model on the applied UCM dataset. The figure has shown that the AL-RCNN, CNN-RBFNN, and CA-CNN-BiLSTM techniques have attained poor results with the highest CT of 1.646 s, 1.847 s, and 1.934 s respectively. Followed by, the DLEN-LSTM, MLRSSC-CGNN-MLGAT, MLRSSC-CGNN-SGAT, and CNN approaches have needed a moderate CT of 1.134 s, 1.224 s, 1.432 s, and 1.586 s respectively. Concurrently, the DLEN-VAE and DLEN-DNN models have needed reasonable CT of 1.065 s and 1.097 s respectively. However, the proposed PTDLEN-VAE model has accomplished improved performance with the least CT of 1.021 s.

Fig. 11 examines the CT analysis of the PTDLEN-VAE technique on the applied AID dataset. The figure exhibited that the CA-CNN-BiLSTM, CNN, and CNN-RNFNN approaches have gained worse outcomes with maximum CT of 1.977 s, 1.944 s, and 1.932 s correspondingly. Additionally, the AL-RCNN, MLRSSC-CGNN-SGAT, MLRSSC-CGNN-MLGAT, and DLEN-DNN methodologies have desired a moderate CT of 1.731 s, 1.562 s, 1.209 s, and 1.167 s correspondingly. Simultaneously, the DLEN-DNN and DLEN-VAE algorithms have needed reasonable CT of 1.165 s and 1.134 s correspondingly. Eventually, the presented PTDLEN-VAE technique has accomplished enhanced performance with a minimum CT of 1.129 s.

## 5. Conclusion

This paper has presented an effective PTDLEN-VAE method for satellite imagery analysis on ecology management. The presented PTDLEN-

**Table 5**

Performances of proposed PTDLEN-VAE with existing methods in terms of computation time (min).

| Methods           | UCM Dataset | AID Dataset |
|-------------------|-------------|-------------|
| CNN               | 1.586       | 1.944       |
| CNN-RBFNN         | 1.847       | 1.932       |
| CA-CNN-BiLSTM     | 1.934       | 1.977       |
| AL-RCNN           | 1.646       | 1.731       |
| MLRSSC-CGNN-SGAT  | 1.432       | 1.562       |
| MLRSSC-CGNN-MLGAT | 1.224       | 1.209       |
| DLEN-LSTM         | 1.134       | 1.167       |
| DLEN-DNN          | 1.097       | 1.165       |
| DLEN-VAE          | 1.065       | 1.134       |
| PTDLEN-VAE        | 1.021       | 1.129       |

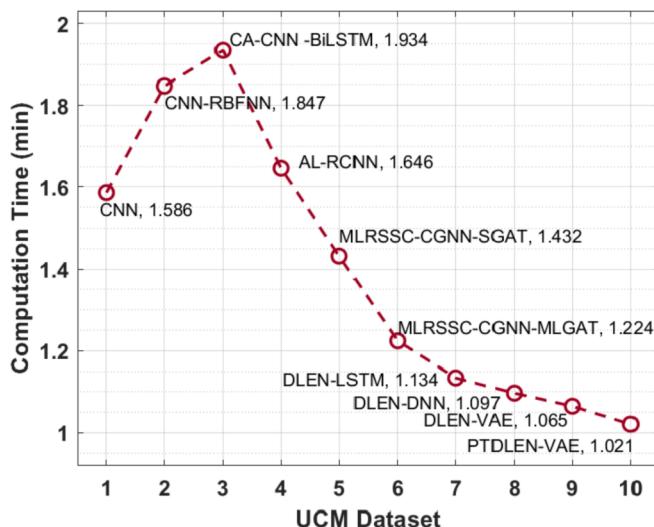


Fig. 10. Computation time analysis of PTDLEN-VAE model on UCM dataset.

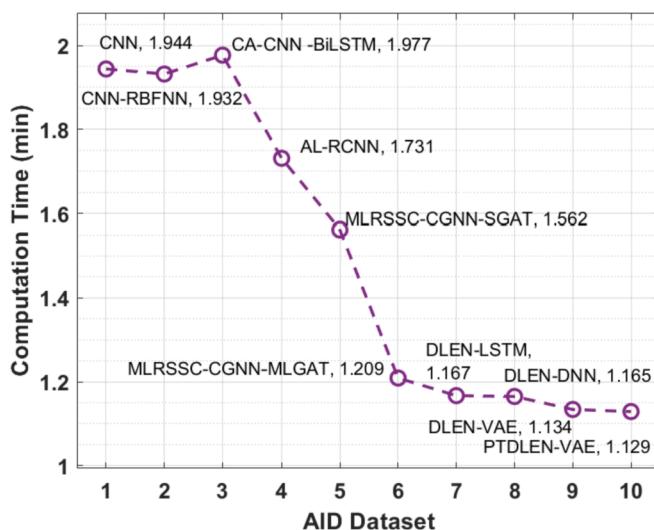


Fig. 11. Computation time analysis of PTDLEN-VAE model on AID dataset.

VAE model includes a series of operations such as pre-processing, EfficientNet based feature extraction, IKHO based parameter optimization, and VAE based classification. Besides, the IKHO algorithm is designed by the integration of KH algorithm with FS operator to avoid the local optimal problem of the KH algorithm. In addition, the PTDLEN based feature extractor by the use of parameter tuned EfficientNet model using IKHO algorithm helps to enhance the prediction outcome of the PTDLEN-VAE model for unseen data. For validating the improved performance of the PTDLEN-VAE method, an extensive experimental analysis take place on benchmark aerial image dataset. The resultant experimental values highlighted the supremacy of the proposed PTDLEN-VAE model over the other techniques. In future, the presented PTDLEN-VAE model can be deployed in unmanned aerial vehicles (UAVs) and lightweight cryptographic algorithms can be designed to accomplish secure UAV communication for satellite imagery analysis.

#### Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

#### Consent to participate

Not applicable.

#### Declaration of Competing Interest

The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

#### Acknowledgment

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under grant number (RGП 2/25/42).

This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Fast-track Research Funding Program.

#### References

- Bian, X., Chen, C., Tian, L., Du, Q., 2017. Fusing local and global features for high-resolution scene classification. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 10 (6), 2889–2901.
- Bouliha, W., Sellami, M., Driss, M., Al-Sarem, M., Safaei, M., Ghaleb, F.A., 2021. RS-DCNN: a novel distributed convolutional-neural-networks based-approach for big remote-sensing image classification. *Comput. Electron. Agric.* 182, 106014.
- Charles, H., Dukes, J.S., 2008. Impacts of invasive species on ecosystem services. In: Nentwig, W. (Ed.), *Biological Invasions*. Ecological Studies; Springer, Berlin/Heidelberg, Germany, pp. 217–237. ISBN 978-3-540-36920-2.
- Cheng, G., Li, Z., Yao, X., Guo, L., Wei, V., 2017. Remote sensing image scene classification using bag of convolutional features. *IEEE Geosci. Remote Sens. Lett.* 14 (10).
- Dai, M., Zheng, D., Na, R., Wang, S., Zhang, S., 2019. EEG classification of motor imagery using a novel deep learning framework. *Sensors* 19 (3), 551.
- Hassan, S.M., Maji, A.K., Jasinski, M., Leonowicz, Z., Jasinska, E., 2021. Identification of plant-leaf diseases using CNN and transfer-learning approach. *Electronics* 10 (12), 1388.
- Hoffmann, B.D., Broadhurst, L.M., 2016. The economic cost of managing invasive species in Australia. *NeoBiota* 31, 1.
- <http://weegee.vision.ucmerced.edu/datasets/landuse.html>.
- <https://captain-whu.github.io/AID/>.
- Hulme, P.E., Pysek, P., Nentwig, W., Vilà, M., 2009. Will threat of biological invasions unite the European Union? *Science* 324, 40–41.
- Kadhim, M.A., Abed, M.H., 2019, April. Convolutional neural network for satellite image classification. In: Asian Conference on Intelligent Information and Database Systems. Springer, Cham, pp. 165–178.
- Laban, N., Abdellatif, B., Ebied, H.M., Shedeed, H.A., Tolba, M.F., 2020. Multiscale satellite image classification using deep learning approach. In: Machine Learning and Data Mining in Aerospace Technology. Springer, Cham, pp. 165–186.
- Li, L., Zhou, Y., Xie, J., 2014. A free search krill herd algorithm for functions optimization. *Math. Probl. Eng.* 2014.
- Li, Y., Chen, R., Zhang, Y., Zhang, M., Chen, L., 2020. Multi-label remote sensing image scene classification by combining a convolutional neural network and a graph neural network. *Remote Sens.* 12 (23), 4003.
- Liu, Q., Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., Nemani, R., 2020. Deepsat v2: feature augmented convolutional neural nets for satellite image classification. *Remote Sens. Lett.* 11 (2), 156–165.
- Lunga, D., Gerrand, J., Yang, L., Layton, C., Stewart, R., 2020. Apache spark accelerated deep learning inference for large scale satellite image analytics. *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.* 13, 271–283.
- Marmanisad, D., Wegnera, J., Galliani, S., Schindlerb, K., Datcuc, M., Stillad, U., 2016. Semantic segmentation of aerial images with an ensemble of CNNs. *ICWG* 3 (4), 1–8.
- Nguyen, T.T., Hoang, T.D., Pham, M.T., Vu, T.T., Nguyen, T.H., Huynh, Q.T., Jo, J., 2020. Monitoring agriculture areas with satellite images and deep learning. *Appl. Soft Comput.* 95, 106565.
- Pelletier, C., Webb, G.I., Petitjean, F., 2019. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sens.* 11 (5), 523.
- Penev, K., Littlefair, G., 2005. Free search—a comparative analysis. *Inf. Sci.* 172 (1–2), 173–193.
- Raffini, F., Bertorelle, G., Biello, R., D'Urso, G., Russo, D., Bosso, L., 2020. From nucleotides to satellite imagery: approaches to identify and manage the invasive pathogen *Xylella fastidiosa* and its insect vectors in Europe. *Sustainability* 12 (11), 4508.
- Shafaey, M.A., Salem, M.A.M., Ebied, H.M., Al-Berry, M.N., Tolba, M.F., 2018, September. Deep learning for satellite image classification. In: International Conference on Advanced Intelligent Systems and Informatics. Springer, Cham, pp. 383–391.

- Unnikrishnan, A., Sowmya, V., Soman, K.P., 2019. Deep learning architectures for land cover classification using red and near-infrared satellite images. *Multimed. Tools Appl.* 78 (13), 18379–18394.
- Wei, C.L., Wang, G.G., 2020. Hybrid annealing krill herd and quantum-behaved particle swarm optimization. *Mathematics* 8 (9), 1403.
- Xia, M., Tian, N., Zhang, Y., Xu, Y., Zhang, X., 2020. Dilated multi-scale cascade forest for satellite image classification. *Int. J. Remote Sens.* 41 (20), 7779–7800.
- Yuan, Y., Wan, J., Wang, Q., 2016. Congested scene classification via efficient unsupervised feature learning and density estimation. *Pattern Recogn.* 56, 159–169.
- Zhang, L., Xia, G., Wu, T., Lin, L., Tai, X., 2016. Deep learning for remote sensing image understanding. *J. Sens.* 2016, 1–2.