Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

A Scene Detection and Classification Model for Remote Sensing Images Using Deep Learning Technique with Water Flow Optimization

M. Rega[1], Dr. S. Sivakumar[2]

[1]Research Scholar, Department of Computer and Information Science, Annamalai University, India, rekhamohanraj00@gmail.com
[2]Assistant Professor, PG Department of Computer Science, Government Arts College, (Deputed from Annamalai University, Annamalai nagar), India, sivassa77@gmail.com

Scene classification using deep learning (DL) is a common and effective way in RS and geospatial analysis. It is most vital in environmental monitoring, mapping, land planning, and land management. Nevertheless, the current techniques are issues as vulnerability to noise interference, lower classification accuracy, and poor generalization skills. Remote sensing images are frequently used in the description of urban and rural regions, change detection, and other fields. In general, the RSI is high-resolution and covers extensive and diverse data, appropriate analysis of RSIs is most significant. DL systems like Convolutional Neural Networks (CNNs) are exposed significant result in image detection tasks, making them suitable for scene classification in RSIs. So, this study develops a new Water Flow Optimizer with Deep Learning Enabled Scene Detection and Classification (WFODL-SDC) algorithm on RSIs. The main focus of the WFODL-SDC system lies in the optimal detection and classification of various scenes that exist in it. To accomplish this, the WFODL-SDC technique involves an adaptive median filtering (AMF) method for removing the noise that exists in it. Besides, the WFODL-SDC technique uses SE-DenseNet system for the derivation of useful feature vectors. The experimental values inferred that the WFODL-SDC methodology obtains optimal results with other recent approaches. At last, auto encoder (AE) has been executed for the recognition and classification of various kinds of scenes. The simulation result analysis of the WFODL-SDC technique undergoes utilizing benchmark image database.

Keywords: Scene Detection, Scene classification, Remote Sensing Image, Water Flow Optimizer, Deep Learning, Adaptive Median Filtering.

1. Introduction

Remote sensing images (RSI) normally include many objects owing to their wide geographic analysis. In present scenario, remote sensing (RS) has the overview of deep learning (DL)-based performances for data extraction from scene imageries over object recognition and classification [1]. Accurately identifying objects in such images poses tasks for computers [2], but DL-based models have verified extremely effectual and amazing feature extractor abilities. Given the occurrence of higher-resolution RSI with ample details, object recognition models play a vital part in data extraction by recognizing numerous objects within a remote sensing scene [3]. Scene classification is a popular study topic in the domain of computer vision (CV) used to observation of Earth [4]. Yet, many supervised classification models neglect to simplify when restricted labeled information is accessible [5]. However, labeling huge amounts of RS data is time-consuming, needs expert knowledge, and is labor-intensive. Several models are projected to defeat the issue of inadequate labeled RS data, for example, self-supervised or semi-supervised methods [6]. Few-shot learning (FSL) stimulated by the human capability to acquire novel models from one or few samples to simplify to novel examples is a significant technique in this perspective [7].

Nevertheless, these methods frequently struggle to meet different demands, and as an outcome, their application can normally restrict to exact situations [8]. In comparison, DL has extended important traction in object recognition uses owing to the accessibility of massive databases and advances in computing technology [9]. DL-based approaches deliver faster recognition speed and enhanced accuracy when equated to classical hand-crafted object recognition models. Therefore, DL-driven object detection is quickly developing, allowing precise object recognition and removal of scene data from RSI [10].

This study introduces a new Water Flow Optimizer with Deep Learning Enabled Scene Detection and Classification (WFODL-SDC) technique on RSIs. The main focus of the WFODL-SDC system lies in the optimal detection and classification of various scenes that exist in it. Besides, the WFODL-SDC technique uses SE-DenseNet model for the derivation of useful feature vectors. For parameter tuning of the SE-DenseNet system, the WFO approach can be employed.

2. LITERATURE SURVEY

In [11], a novel RS scene understanding method named multi-scale attention R-CNNs (MSA R-CNNs) model has been developed, which includes a great feature extractor network for improved feature extractor from images. An adaptive dynamic inner lateral (ADIL) connection unit is presented in order to tackle data loss in a distributed lightweight attention module (DLAM) and feature pyramid network (FPN) to enhance feature data processing. Khan and Basalamah [12] present a multi-branch DL structure that competently integrates global contextual features with many features to classify difficult land scenes. Normally, the structure contains dual branches. The initial branch removes contextual data from dissimilar areas of the input images, and the next branch experiences a FCNs model to remove many local features. Wang et al. [13] developed a label-free self-distillation contrastive learning

with a transformer architecture (LaST) model.

In [14], a multi-task network is presented. At first, 3-D ResUnet technique was applied. Next, the network of multi-tasks contains dual sub-tasks such as multi-temporal semantic segmentation and binary change recognition. The shared system branch uses 3-D residual blocks for removal. In the specific branch, a 3D GAN has been combined. Alahmari et al. [15] propose a Hybrid Multi-Strategy Aquila Optimizer with DL-Driven Crop Type Classification (HMAODL-CTC) system. Finally, the projected HMAODL-CTC system employs an ELM model for the classification of crop type.

Guan et al. [16] designed a new attention module (EuPea) method, which is calculated to efficiently take inter-element data in mapping feature and produce extra influential feature map for usage in NNs. In EuPea attention device, the projected model includes distance and Pearson correlation co-efficient data among elements from the mapping feature. In [17], a mayfly optimizer with DL based robust RSI scene classification (MFODL-RRSISC) algorithm is projected. For classification of scene, the developed method uses NasNet for MFO-based parameter tuning, feature extractor, and SAE classification algorithm.

3. THE PROPOSED MODEL

In this paper, we have introduced a novel WFODL-SDC methodology on RSIs. The main focus of the WFODL-SDC system lies in the optimal detection and classification of various scenes that exist in it. To accomplish this, the WFODL-SDC technique involves AMF-based preprocessing, SE-DenseNet based feature extractor, WFO-based parameter tuning, and AE-based classification. Fig. 1 portrays the workflow of WFODL-SDC technique.

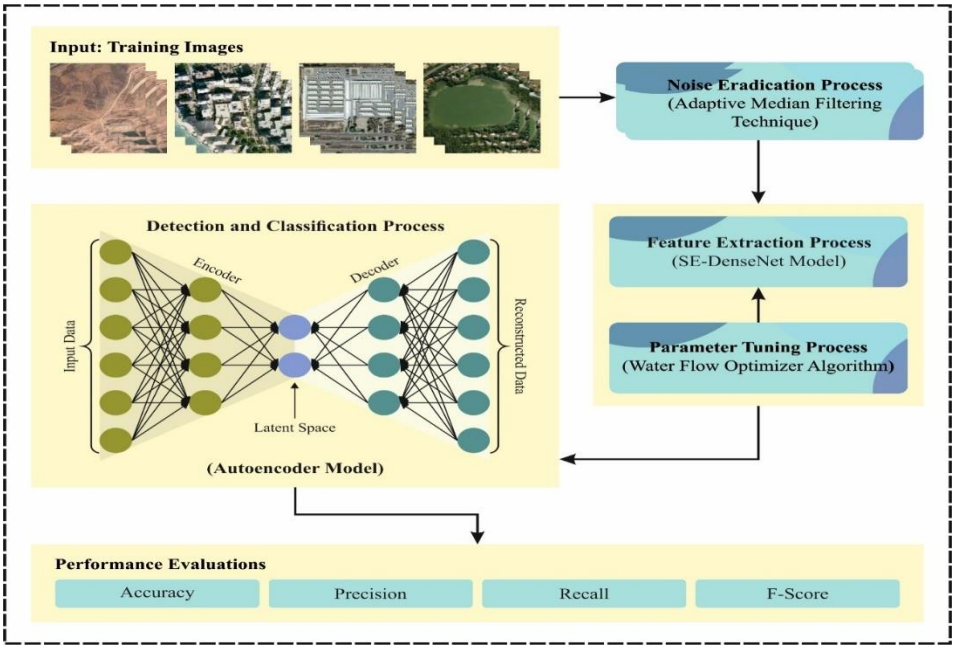


Fig. 1. Workflow of WFODL-SDC technique

3.1. Image Preprocessing

Primarily, the WFODL-SDC technique involves AMF technique to eradicate the noise that exists in it.AMF plays an essential role in improving scene classification by efficiently modifying noise and maintaining crucial particulars within images [18]. In the perspective of scene classification, where images may display fluctuating levels of noise and variations, AMF dynamically alters its filtering parameters depending on the local features of the image. This model excels in upholding the reliability of crucial features while successfully overpowering noise artifacts, donating to amended image quality for following scene analysis. By combining adaptive median filtering into the pre-processing pipeline for scene classification, the method becomes stronger for noise variants, foremost to improve accuracy in classifying scenes across dissimilar visual surroundings.

3.2. Se-Densenet Model

The SE module-based channel attention device permits the system to adaptively rectify the significance of diverse feature channels, permitting it to concentrate on more discriminative features. The WFODL-SDC technique uses SE-DenseNet model for the derivation of useful feature vectors. Due to the restrictions on the simplification performance of rule-based and ML techniques, we accepted a DL-based scene detection method [19]. For feature extraction and identification, we employ the backbone system of a DenseNet169 architecture that joins a scene detection approach and channel attention device. DenseNet enhances the movement of data and gradients among NN layers by linking mapping features from entire previous layers with similar spatial sizes.. Assume that, x_0 is the input image that is delivered over NNs of n layers. Each lth layer in the system contains a non-linear transformer, $H_l(\cdot)$contains a ReLU, pooling, BN, or convolution (Conv). The $(l-1)$th layer output is signified as x_{l-1}, which is linked as an input to lth layer in classical Conv feed-forward network that is exposed in Eq. (1):

$$x_l = H_l(x_{l-1}). \tag{1}$$

The ResNet evades non-linear transformation with identity function by integrating a skip connection. The gradient in ResNet can able to flow straight from final to previous layers over the identity function exposed in Eq. (2):

$$x = H(x) + x_{l-1}. \tag{2}$$

On the other hand, the abstract is employed in order to link the identity function and $H_l$resultant that may delay data flow in the system. To increase data flow among the layers, DenseNet uses straight contacts from any layer to every following layer. As an outcome, the feature map of entire previous layers $x_0, \dots, x_{l-1}$, are delivered to the lth layer as an input that is exposed in Eq. (3):

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]), \tag{3}$$

Whereas, $x_0, x_1, \dots, x_{l-1}$ signifies the sequence of mapping feature created from the $0, \dots, l-1$ layers. In DenseNet, the channel counts produced by mapping feature sequences in deep layer upsurges considerably if the feature map size rests similar. This can outcome in an upsurge in the calculation and memory charges. Therefore, to find out this problem, downsampling layers have been employed in DenseNet structure to reduce the feature map

size. The layers of downsampling will separate the DenseNet structure into many dense blocks, where each one contains of feature map of dissimilar dimensions.

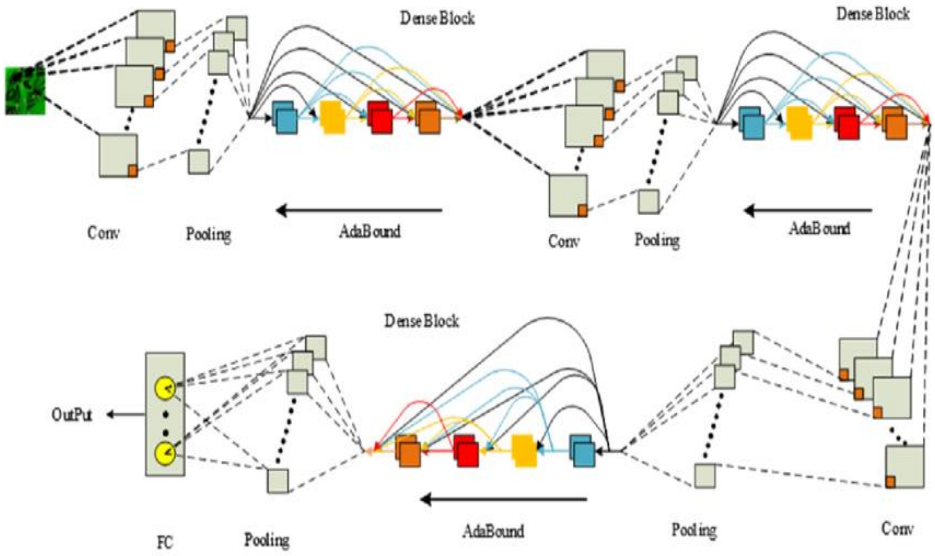
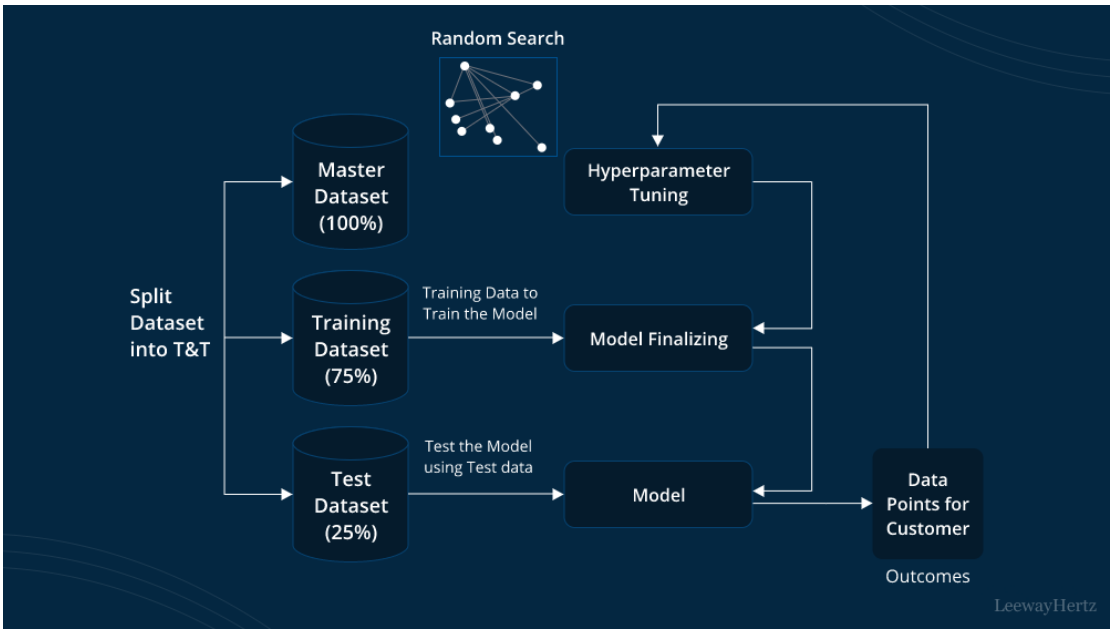


Fig. 2. Architecture of SE-DenseNet model

After the bottleneck layer, the SE unit is applied in order to alter the significance of dissimilar feature map channels. The main intention for employing an SE module is the point that dissimilar networks within mapping feature transmit dissimilar amounts of data. Fig. 2 illustrates the infrastructure of SE-DenseNet model. Standard DNN gives entire channels similarly and individually, so they are incapable to take and emphasize the most selective element of feature map. By clearly demonstrating the channel needs and altering their relative significance, the SE module displays higher solution in differentiating significant features. Initially, the SE squeeze process employs global average pooling and reduces every feature channel spatial size to a single value. The network descriptor is the outcome of pooling process that collects channel specific data.

3.3. Hyperparameter Tuning Process

Hyperparameter tuning is the process of selecting the optimal values for a machine learning model's hyper parameters. Hyperparameters are settings that control the learning process of the model, such as the learning rate, the number of neurons in a neural network, or the kernel size in a support vector machine.



The goal of hyper parameter tuning is to find the values that lead to the best performance on a given task. At this stage, the WFO approach can be deployed for parameter tuning of the SE-DenseNet algorithm. WFO is a SI method developed by Prof. Kaiping Luo in year 2021, takes motivation from the 2 distinguishing categories of water flows determined by nature turbulent and laminar [20]. Naturally, water movements from higher to lower can be the same as the method of examining for an outcome in an optimizer complexity. This method analyzes the behaviour of turbulent and laminar flows from the water flow technique by applying mathematical expressions and determines the optimum performance by constant iteration. The mathematical definition of turbulent and laminar flow will be given by:

1. Laminar Operator: During the laminar flow, every particle is moved in parallel to one another in the identical direction however, their speed changes because of the settings. The approach of motion will be indicated by applying Eq. (4).

$$y_i(t) = x_i(t) + s * d \; \forall i \in \{1,2,\dots,m\} \tag{4}$$

Whereas,tindicates the existing iteration number, m implies the size of populations, $x_i(t)$ refers the location of the i^{th} elements at the t^{th} iteration, $y_i(t)$ is the potential movement location of the t^{th} individual at the t^{th} rounds, s describes the random number among zero to one, and the d vector denotes the mutual way of movements of each individual at the existing iterations; d will be measured as presented in Eq. (5).

$$d = x_{best} - x_k(t), x_{best} \neq x_{k(t)} \tag{5}$$

Now, x_{besi} indicate the optimum solution acquired by the existing iteration of the population and $x_{k(t)}$ characterizes thearbitrarily chosen particlewithin population.

During the laminar flow operator, each individual of the population utilizes a consistent parallel unidirectional search, wherein the similar direction vector d confirms that explore

will be unidirectional, and arbitrariness of s denotes that various individuals are diverse movement steps.

2. Turbulen Operator: During turbulence, water particles were impacted by alternative difficulties and demonstrated unbalanced rotational movements. The probable movement location y_i was produced by the arbitrary dimensions of the i^{th} individual by employing Eq. (6).

$$y_i = \begin{cases} x_i^{j1}(t) + \left|x_i^{j1}(t) - x_k^{j1}(t)\right| * \theta * \cos(\theta), & ifr < p_e \\ \left(ub^{jl} - lb^{j1}\right) * \dfrac{x_k^{i^2}(t) - lb^{i^2}}{ub^{i^2} - lb^{i^2}} + lb^{j1}, & otherwise \end{cases} \quad (6)$$

The vortex conversion of water particles in similar layer is denoted in upper part of Eq. (6) and the second part of Eq. (6) represents general cross-layer movements of particles. Whereas j1 indicates a dimension arbitrarily chosen in the particles, j2 refers to a size varying from j1 randomly chosen from the particle, and $x_k^{j1}(t)$ describes the significance of the j1th dimensional of k^{th} particle at the t^{th} iteration. θ indicatesthearbitrary number within ranges $-'\pi$ to $'\pi, ub^{j1}$ and lb^{j1} means the up and low boundaries of the chosen dimensional, r defines a randomized number from zero to one, and $p_e \in (0,1)$ was a control parameter termed the vortex possibility

In the iterations of the method execution, the technique executes a stochastic analysis of the 2 behaviors, turbulent and laminar, and their individual execution probabilities have been measured by the factor p_l. The fitness choice is asignificantfeaturein managing the efficiencyof WFO methodology. The hyperparameter choice procedure comprises the encoded solution for measuring the performance of candidate results. During this case, the WFOapproachassumes accuracy as the main condition to plan the fitness function (FF) that is expressed as:

$$\text{Fitness} = \max(P) \quad (7)$$

$$P = \frac{TP}{TP + FP} \quad (8)$$

In which, FP and TPsignifies false and true positive values.

3.4. Classification using AE model

The AE is dependent upon a FFNN [21]. Its output and input layers are generally similar, so the system can able to rebuild the output to create it as near to the input At last, AE can be applied for the detection and classification of various kinds of scenes.. In AEs, H is said to be hidden layer(HL); X denotes the input layer; $\tilde{X}$ refers to the output layer; f and gare the encoder and decoder, respectively. The encoder f and decoder g methods have been presented in Eqs. (9) and (10), correspondingly.

$$f: x \to h: h_j = \sigma(W_{ei}x_i + b_{ei}), \quad (9)$$

$$g: h \to \tilde{x}: \tilde{x}_i = \sigma\left(W_{dj}h_j + b_{dj}\right), \quad (10)$$

Here, x_i denotes the feature of original. x_i is changed to h_j after encrypted by the encoding. h_j is transformed to X_i after being decrypted by the decoding. b_e and W_e signifies the bias and weight of the encoder, correspondingly. b_d and W_d indicates the bias and weight of decoder, correspondingly. σ states the activation function, which can create the new feature, so enhances the model's skill. This research clarifies that the method can absorb the deep semantic features of abnormal and normal traffic. So, it has been create the size of H lesser than X. Once the AE rebuilds an output feature $\tilde{x}$ similar to the input feature x. The HL takes most top feature h between the new feature at AE. The learning model of AE is defined as reducing the loss function. Generally, the loss function is said to be the MSE as below.

$$\mathrm{MSE} = \frac{1}{2n}\sum_{i=1}^{n}(x_i - g(f(x_i))^2, = \frac{1}{2n}\sum_{i=1}^{n}(x_i - \tilde{x}_i)^2. \quad (11)$$

The SAE training method contains training every AE layer separately, beginning from the input layer. After the layer is trained, its encoded output becomes input for the next layer. This stacking procedure permits the SAE to learn gradually abstract and hierarchical representations of the input data.

4. RESULT ANALYSIS

The experimentation validation of the WFODL-SDC methodology is examined using the UCM and AID databases. The UCM [22] dataset includes 2100 samples with 21 classes as exposed in Table 1.

Table 1 Details of the UCM database

UCM Dataset	
Classes	No. of Instances
C1	100
C2	100
C3	100
C4	100
C5	100
C6	100
C7	100
C8	100
C9	100
C10	100
C11	100
C12	100
C13	100
C14	100
C15	100
C16	100
C17	100
C18	100
C19	100
C20	100
C21	100
Total Instances	2100

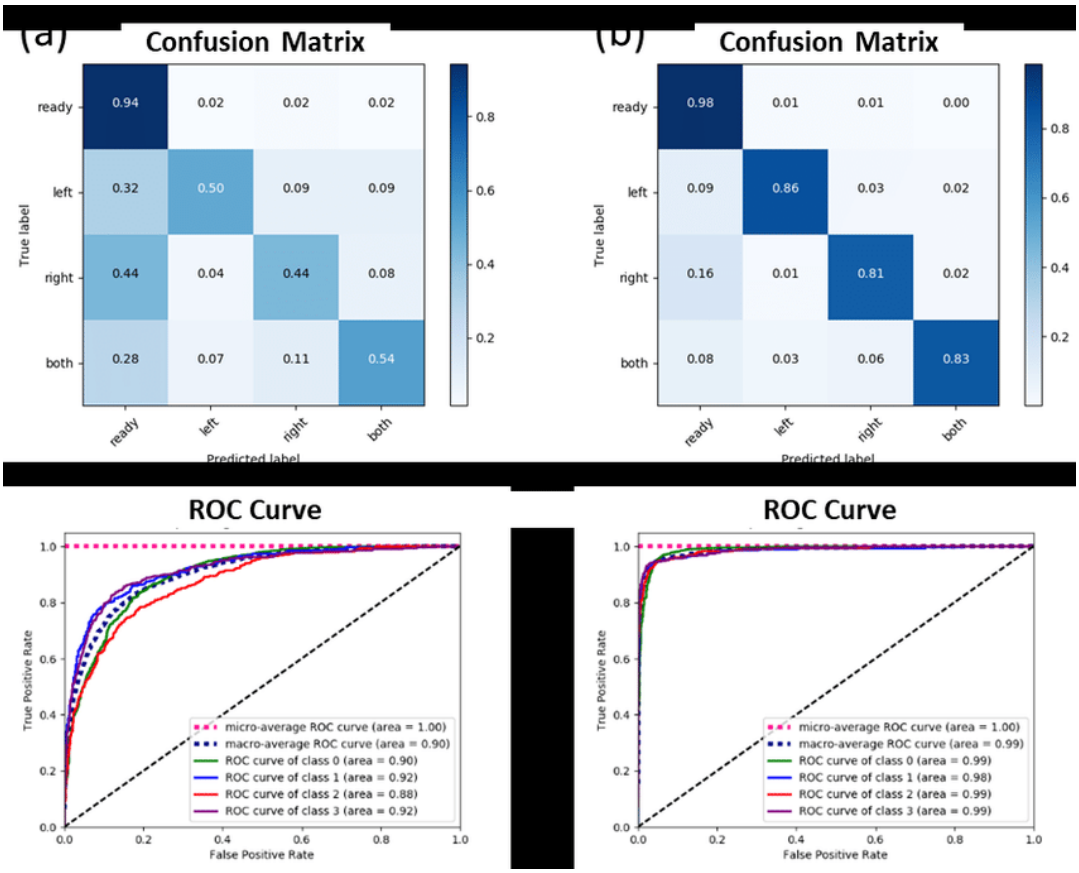


Fig. 3. UCM dataset (a-b) Confusion matrices, (c-d) PR-curve and ROC-curve

Fig. 3 illustrates the classifier results of the WFODL-SDC method at UCM dataset. Figs. 3a-3b showcases the confusion matrices gained by the WFODL-SDC system with 70%:30% of TRAPH/TESPH. This simulation value indicated that the WFODL-SDC model can be identified and categorized with 21 class labels appropriately. Meanwhile, Fig. 3c indicates the PR result of the WFODL-SDC approach. The figure defined that the WFODL-SDC algorithm gives greater PR effectiveness with every class. In conclusion, Fig. 3d showcases the ROC result of the WFODL-SDC algorithm. This figure signified that the WFODL-SDC method offers efficient results with increased ROC values with diverse class labels.

Table 2 reveals the overall scene classifier outcome of the WFODL-SDC algorithm on 70% of TRAPH of UCM dataset. The simulation values stated that the WFODL-SDC technique has proficient effectual recognition of 21 classes.

Table 2 Scene classification of WFODL-SDC model with UCM dataset under 70% of TRAPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TRAPH (70%)				
C1	96.12	62.50	64.94	63.69
C2	96.80	69.09	55.88	61.79
C3	96.39	57.97	62.50	60.15

C4	96.19	59.15	60.87	60.00
C5	96.67	65.62	60.87	63.16
C6	95.92	59.02	50.70	54.55
C7	96.73	63.51	69.12	66.20
C8	96.53	64.00	66.67	65.31
C9	97.21	65.15	70.49	67.72
C10	96.46	59.02	57.14	58.06
C11	96.12	62.16	61.33	61.74
C12	96.26	60.27	62.86	61.54
C13	96.80	66.22	69.01	67.59
C14	95.99	56.10	66.67	60.93
C15	96.46	67.86	52.78	59.38
C16	96.94	66.20	69.12	67.63
C17	96.33	66.27	67.90	67.07
C18	96.73	67.11	68.92	68.00
C19	96.46	60.32	58.46	59.38
C20	95.51	54.29	52.78	53.52
C21	96.94	68.06	69.01	68.53
Average	96.46	62.85	62.76	62.66

A wide-ranging scene classification results were reported by the WFODL-SDC methodology with 30% of TESPH at UCM datasets as illustrated in Table 3. These obtained outcomes underscore that the WFODL-SDC algorithm correctly recognized all categories of scenes present in the UCM dataset. It is also perceived that the WFODL-SDC method can be capable of recognizing the samples with increased classifier outcomes.

Table 3 Scene classification of WFODL-SDC technique with UCM dataset under 30% of TESPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TESPH (30%)				
C1	96.83	57.89	47.83	52.38
C2	95.71	61.90	40.62	49.06
C3	96.51	69.44	69.44	69.44
C4	96.98	73.08	61.29	66.67
C5	96.03	58.33	67.74	62.69
C6	97.46	70.97	75.86	73.33
C7	96.51	64.71	68.75	66.67
C8	96.35	59.26	57.14	58.18
C9	96.19	69.23	69.23	69.23
C10	95.08	57.50	62.16	59.74
C11	96.83	59.26	64.00	61.54
C12	97.14	73.08	63.33	67.86
C13	96.83	64.52	68.97	66.67
C14	97.14	68.57	77.42	72.73
C15	96.67	62.07	64.29	63.16
C16	96.51	64.71	68.75	66.67
C17	97.62	57.69	78.95	66.67
C18	97.62	70.37	73.08	71.70
C19	96.51	67.57	71.43	69.44
C20	97.46	75.00	64.29	69.23
C21	96.83	68.00	58.62	62.96
Average	96.70	65.39	65.39	65.05

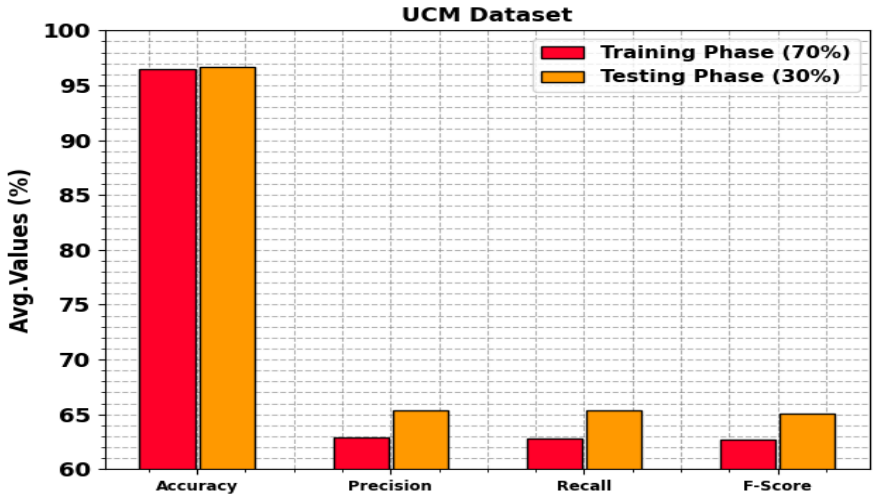


Fig. 4. Average of the WFODL-SDC model at UCM dataset

In Fig. 4, the average scene recognition outcome of the WFODL-SDC technique is reported on UCM dataset. The results imply the proficient capability of the WFODL-SDC technique in the scene classification process. With 70% of TRAPH, the WFODL-SDC technique provides average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 96.46%, 62.85%, 62.76%, and 62.66%, respectively. Also, based on 30% of TESPH, the WFODL-SDC method gives average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 96.70%, 65.39%, 65.39%, and 65.05%.

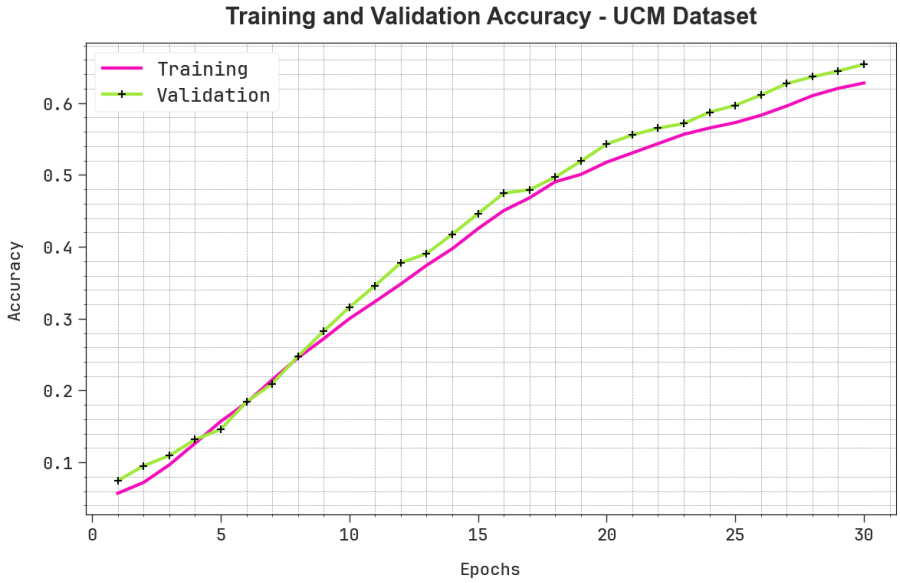


Fig. 5. $Accu_y$ curve of the WFODL-SDC method at UCM dataset

The effectiveness of the WFODL-SDC method with UCM dataset is clearly displayed in Fig. 5 in the usage of training accuracy (TRAA) and validation accuracy (VALA) curves. This

figure signifies useful analysis into the behaviour of the WFODL-SDC technique over diverse epoch counts, signifying its learning process and generalization capabilities. Noticeably, the figure assumes a continuous improvement in the TRAA and VALA with growth in epochs. It ensures the adaptive nature of the WFODL-SDC algorithm with a pattern recognition process under the TRA and TES data. The rising trend in VALA outlines the ability of the WFODL-SDC technique to vary to the TRA data and also surpass in providing correct classification on unnoticed data, providing robust generalization capabilities.

Fig. 6 illustrates an extensive representation of the training loss (TRLA) and validation loss (VALL) results of the WFODL-SDC method with UCM dataset over varying epochs. The progressive lessens in TRLA highpoints the WFODL-SDC model improving the weights and reducing the classification error on the TRA and TES data. The figure specifies a clear understanding of the WFODL-SDC system relevant to the TRA data, highlighting its proficiency in capturing patterns within both datasets. Mainly, the WFODL-SDC system incessantly increases its parameters in lessening the differences among the prediction and real TRA class labels.

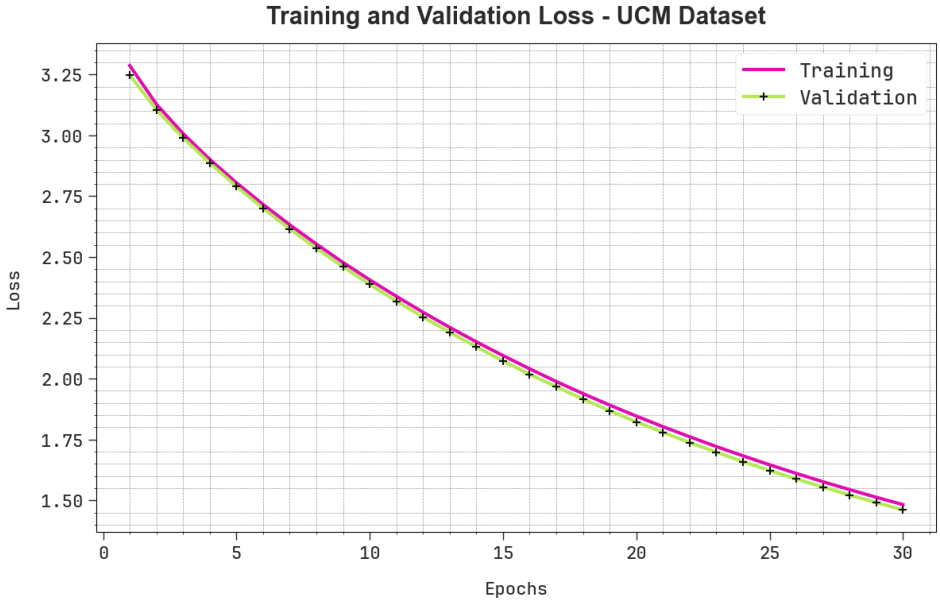


Fig. 6. Loss curve of the WFODL-SDC model with UCM dataset

The performance of the WFODL-SDC methodology is compared with other approaches on UCM database is given in Table 4 and Fig. 7. The results imply that the SC+Pooling, SG+UFL, and CCM-BOVW methods have obtained least performance with $accu_y$ values of 81.67%, 86.64%, and 86.64%, correspondingly. Moreover, the PSR, COPD, and Dirichlet algorithms have attained closer $accu_y$ values of 89.10%, 91.33%, and 92.80%. However, the WFODL-SDC technique accomplishes superior performance with maximum $accu_y$of 96.70%.

Table 4 $Accu_y$ outcome of WFODL-SDC system with other algorithms under UCM dataset

UCM Dataset	
Method	Accuracy (%)
SC+Pooling	81.67
SG+UFL	86.64
CCM-BOVW	86.64
PSR Model	89.10
COPD Model	91.33
Dirichlet	92.80
WFODL-SDC	96.70

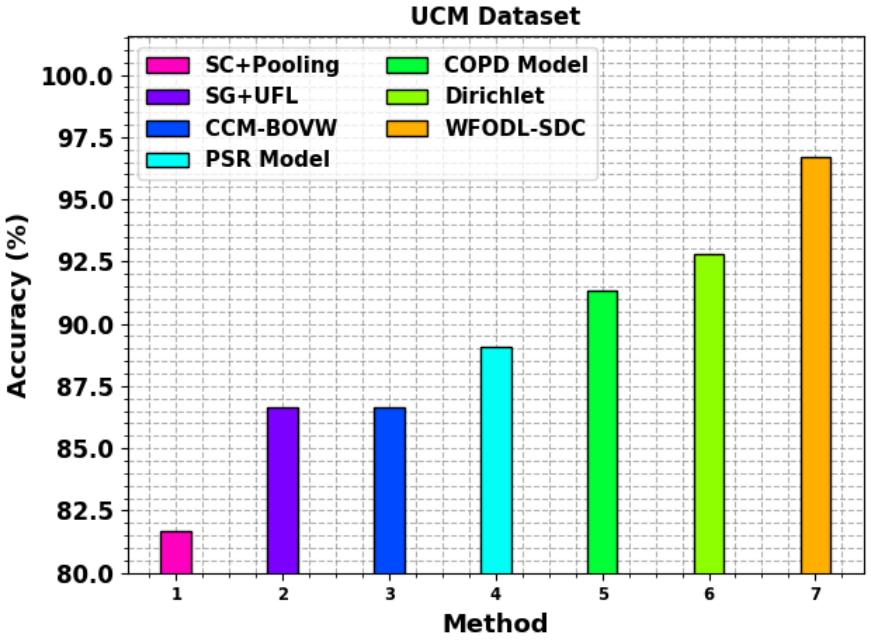


Fig. 7. $Accu_y$outcome of WFODL-SDC model compared under UCM dataset

The AID dataset [23] comprises 3000 instances with 30 classes as shown in Table 5.

Table 5 Details of AID dataset

AID Database	
Class	No. of Instances
C1	100
C2	100
C3	100
C4	100
C5	100
C6	100
C7	100
C8	100

C9	100
C10	100
C11	100
C12	100
C13	100
C14	100
C15	100
C16	100
C17	100
C18	100
C19	100
C20	100
C21	100
C22	100
C23	100
C24	100
C25	100
C26	100
C27	100
C28	100
C29	100
C30	100
Total Instances	3000

Fig. 8 showcases the classifier results of the WFODL-SDC system at AID dataset. Figs. 8a-8b indicates the confusion matrices determined by the WFODL-SDC approach with 70%:30% of TRAPH/TESPH. This figure detailed that the WFODL-SDC algorithm can be recognized and categorized with 30 class labels suitably. Moreover, Fig. 8c denotes the PR result of the WFODL-SDC algorithm. The figure shows that the WFODL-SDC algorithm gains remarkable PR effectiveness with classes. Also, Fig. 8d displays the ROC result of the WFODL-SDC technique. This figure characterized that the WFODL-SDC technique acquires better experimentation results with increased ROC values with diverse class labels.

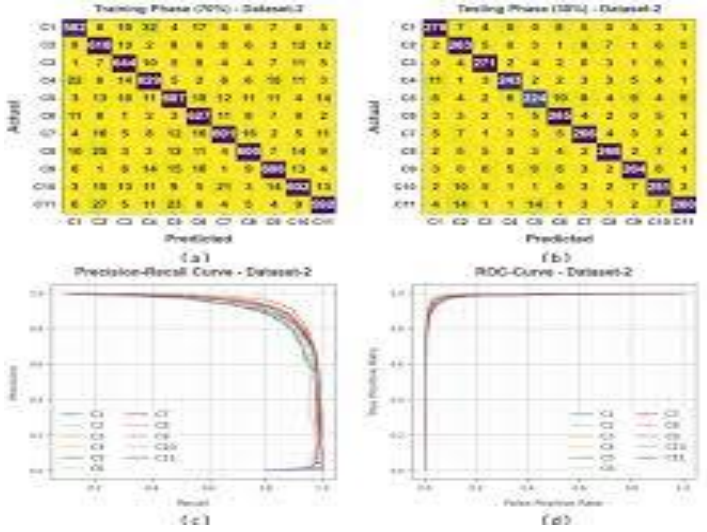


Fig. 8. AID dataset (a-b) Confusion matrices of WFODL-SDC method (c-d) PR-curve and ROC-curve

Table 6 displays the overall scene classifier outcome accomplished by the WFODL-SDC technique on 70% of TRAPH with AID dataset. These experimentation outcomes denoted that the WFODL-SDC algorithm correctly identified all categories of scenes existing in the AID dataset. It is also noticed that the WFODL-SDC algorithm can the capability of recognizing the samples with increased classifier results.

Table 6 Scene classification of WFODL-SDC model with AID dataset under 70% of TRAPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TRAPH (70%)				
C1	96.38	46.03	40.85	43.28
C2	96.10	43.66	42.47	43.06
C3	96.05	40.58	40.00	40.29
C4	96.76	50.00	57.35	53.42
C5	96.48	46.67	40.00	43.08
C6	96.14	42.50	49.28	45.64
C7	96.57	50.72	47.95	49.30
C8	97.00	57.97	54.05	55.94
C9	96.90	53.62	52.86	53.24
C10	96.14	47.67	53.25	50.31
C11	97.14	52.38	52.38	52.38
C12	96.48	50.00	51.35	50.67
C13	96.71	41.38	40.68	41.03
C14	95.90	45.65	53.85	49.41
C15	96.43	44.59	49.25	46.81
C16	96.71	41.43	50.88	45.67
C17	96.67	49.12	40.58	44.44
C18	96.86	53.25	57.75	55.41
C19	96.81	53.12	47.89	50.37
C20	96.24	41.33	46.97	43.97
C21	96.29	45.57	50.70	48.00
C22	96.29	47.37	48.65	48.00
C23	96.05	41.38	32.88	36.64
C24	97.00	53.85	51.47	52.63
C25	96.00	43.42	44.59	44.00
C26	96.76	55.74	45.33	50.00
C27	96.81	53.85	48.61	51.09
C28	96.48	36.36	33.90	35.09
C29	96.24	44.87	49.30	46.98
C30	96.29	46.27	42.47	44.29
Average	96.49	47.35	47.25	47.15

An extensive scene classification outcome succeeded by the WFODL-SDC algorithm with 30% of TESPH under AID dataset as publicized in Table 7. These experimental outcomes emphasize that the WFODL-SDC technique properly recognized all categories of scenes present in the AID dataset. It is also discovered that the WFODL-SDC system can be the ability to recognize the samples with greater classifier results.

Table 7 Scene classification of the WFODL-SDC system with AID dataset under 30% of TESPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TESPH (30%)				
C1	96.89	52.17	41.38	46.15
C2	96.44	39.13	33.33	36.00
C3	96.78	52.38	36.67	43.14
C4	95.78	42.86	56.25	48.65

C5	95.44	32.26	33.33	32.79
C6	97.11	58.62	54.84	56.67
C7	97.44	57.14	59.26	58.18
C8	96.56	40.00	38.46	39.22
C9	96.89	53.57	50.00	51.72
C10	96.44	35.48	47.83	40.74
C11	95.44	44.74	45.95	45.33
C12	96.78	44.83	50.00	47.27
C13	95.89	57.14	39.02	46.38
C14	97.33	46.43	59.09	52.00
C15	96.00	46.34	57.58	51.35
C16	95.67	54.55	55.81	55.17
C17	95.89	40.62	41.94	41.27
C18	97.67	65.38	58.62	61.82
C19	96.78	50.00	44.83	47.27
C20	96.33	51.35	55.88	53.52
C21	96.22	44.19	65.52	52.78
C22	97.56	55.56	76.92	64.52
C23	97.33	56.52	48.15	52.00
C24	97.11	62.50	46.88	53.57
C25	97.22	52.00	50.00	50.98
C26	96.78	42.86	48.00	45.28
C27	97.44	60.00	53.57	56.60
C28	95.67	53.12	41.46	46.58
C29	96.56	45.83	37.93	41.51
C30	96.33	40.00	44.44	42.11
Average	96.59	49.25	49.10	48.69

A comprehensive average classifier result of the WFODL-SDC technique with AID dataset can be shown in Fig. 9. This figure specifies that the WFODL-SDC method gets improved results. According to 70% of TRAPH, the WFODL-SDC system obtained average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 96.49%, 47.35%, 47.25%, and 47.15%. Similarly, with 30% of TESPH, the WFODL-SDC algorithm gives average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 96.59%, 49.25%, 49.10%, and 48.69%.

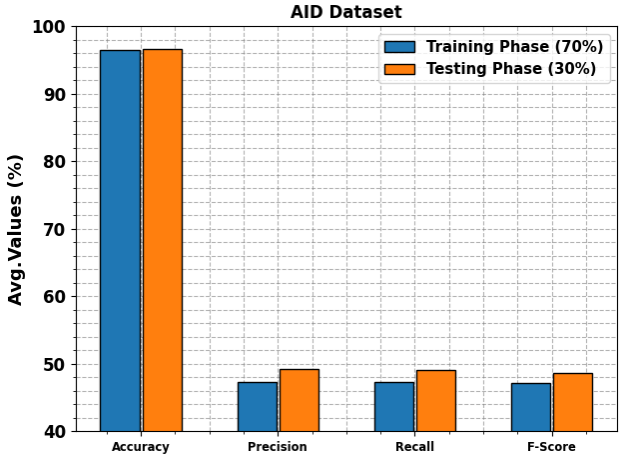


Fig. 9. Average of the WFODL-SDC model with AID dataset

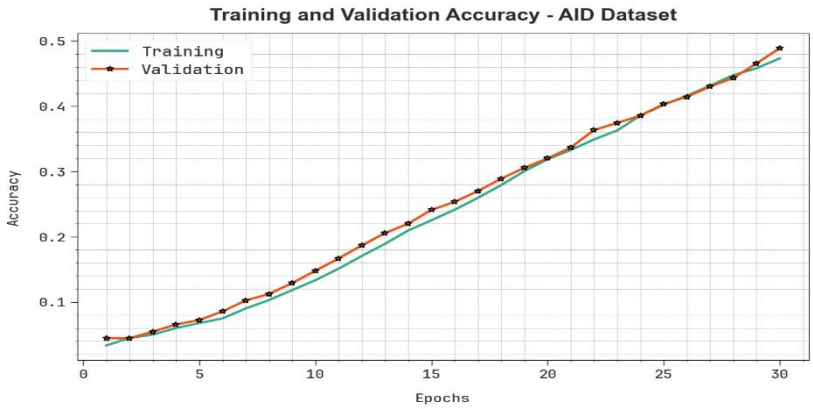


Fig. 10. $Accu_y$ curve of the WFODL-SDC technique at AID dataset

The efficiency of the WFODL-SDC system with AID dataset is graphically illustrated in Fig. 10 in the form of TRAA and VALA curves. This figure exhibits useful analysis of the behaviour of the WFODL-SDC method over multiple epoch counts, demonstrating its learning process and generalization capabilities. Noticeably, the figure infers an incessant improvement in the TRAA and VALA with progress in epochs. It is sure the adaptive nature of the WFODL-SDC method with pattern recognition process at TRA and TES data. The increased trends in VALA outline the capability of the WFODL-SDC technique in altering the TRA data and exceling in providing particular classification on unnoticed data, showing robust generalization abilities.

Fig. 11 shows a wide-ranging view of the TRLA and VALL results of the WFODL-SDC model with AID dataset over multiple epochs. The progressive decreases in TRLA highlight the WFODL-SDC algorithm boosting the weights and lessening the classification error on the TRA and TES data. The figure shows a perfect consideration of the WFODL-SDC model related to the TRA data, highlighting its proficiency in capturing patterns within both datasets. Mainly, the WFODL-SDC method continually increased its parameters in lessening the variances among the prediction and real TRA class labels.

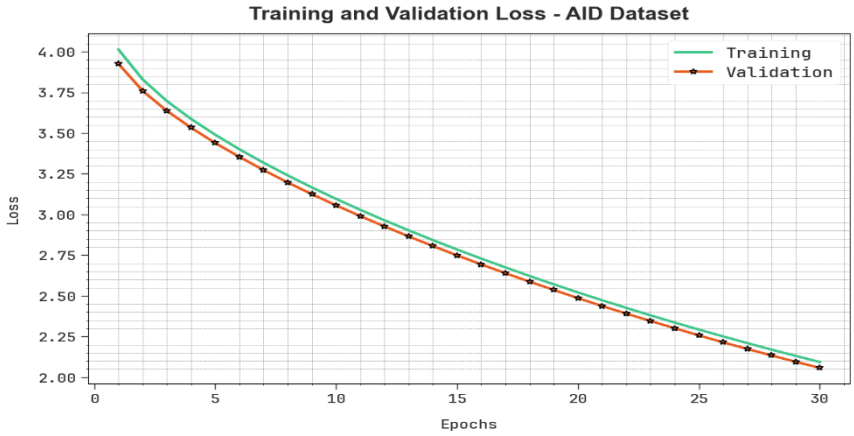


Fig. 11. Loss curve of WFODL-SDC model with AID dataset

A wide-ranging $Accu_y$comparative analysis of the WFODL-SDC method with AID dataset is measured as exhibited in Table 8 and Fig. 12 [24, 25]. These experimentation findings display that the GoogleNet and VGG-VD-16 techniques gain lessened $accu_y$ values of 86.39%, and 89.64%. Similarly, the ResNet50, ResNet-50+EAM, and LCNN-BFF algorithms are attained remarkable $accu_y$ values of 92.57%, 93.64%, and 91.66%. Nevertheless, the WFODL-SDC method gains exceptional outcomes with higher $accu_y$ of 96.59%, correspondingly.

Table 8 $Accu_y$ outcome of the WFODL-SDC model with other algorithms under AID dataset

AID Dataset	
Methods	Accuracy (%)
GoogleNet	86.39
VGG-VD-16	89.64
ResNet50	92.57
ResNet-50+EAM	93.64
LCNN-BFF	91.66
WFODL-SDC	96.59

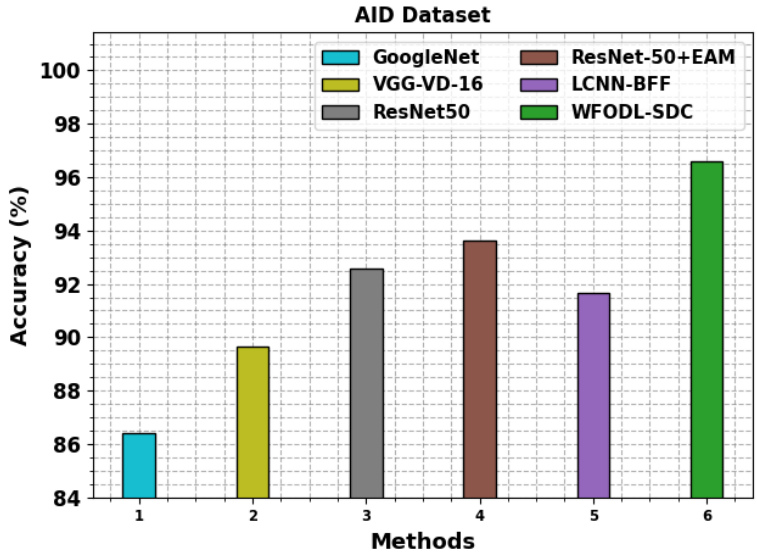


Fig. 12. $Accu_y$ outcome of WFODL-SDC model under AID dataset

Thus, it is apparent that the WFODL-SDC technique has the ability to accurately and efficiently identify distinct scene classes

5. CONCLUSION

In this paper, we have introduced a novel WFODL-SDC methodology on RSIs. The main focus of the WFODL-SDC system lies in the optimal det3ection and classification of various scenes that exist in it. To accomplish this, the WFODL-SDC technique involves AMF-based preprocessing, SE-DenseNet-based feature extractor, WFO-based parameter tuning, and AE-based classification. Primarily, the WFODL-SDC technique involves AMF algorithm for

removing the noise that exists in it. Besides, the WFODL-SDC technique uses the SE-DenseNet model for the derivation of useful feature vectors. Moreover, the WFO system was utilized for parameter tuning of the SE-DenseNet approach. At last, AE has been executed for the recognition and classification of various kinds of scenes. The simulation value of the WFODL-SDC technique occurs using benchmark image database. The simulation results implied that the WFODL-SDC algorithm achieves optimum solution with other recent approaches.

References

1. Huang, H., Mou, Z., Li, Y., Li, Q., Chen, J. and Li, H., 2022. Spatial-temporal invariant contrastive learning for remote sensing scene classification. IEEE Geoscience and Remote Sensing Letters, 19, pp.1-5.
2. Pang, Y., Yu, J., Xi, L., Ge, D., Zhou, P., Hou, C., He, P. and Zhao, L., 2024. Remote Sensing Extraction of Lakes on the Tibetan Plateau Based on the Google Earth Engine and Deep Learning. Remote Sensing, 16(3), p.583.
3. de Roda Husman, S., Lhermitte, S., Bolibar, J., Izeboud, M., Hu, Z., Shukla, S., van der Meer, M., Long, D. and Wouters, B., 2024. A high-resolution record of surface melt on Antarctic ice shelves using multi-source remote sensing data and deep learning. Remote Sensing of Environment, 301, p.113950.
4. Adegun, A.A., Fonou Dombeu, J.V., Viriri, S. and Odindi, J., 2023. State-of-the-Art Deep Learning Methods for Objects Detection in Remote Sensing Satellite Images. Sensors, 23(13), p.5849.
5. Muhammad, U., Hoque, M.Z., Wang, W. and Oussalah, M., 2022. Patch-Based Discriminative Learning for Remote Sensing Scene Classification. Remote Sensing, 14(23), p.5913.
6. Wang, J., Li, W., Zhang, M., Tao, R. and Chanussot, J., 2023. Remote sensing scene classification via multi-stage self-guided separation network. IEEE Transactions on Geoscience and Remote Sensing.
7. Wang, L., Ye, C., Chen, F., Wang, N., Li, C., Zhang, H., Wang, Y. and Yu, B., 2024. CG-CFPANet: a multi-task network for built-up area extraction from SDGSAT-1 and Sentinel-2 remote sensing images. International Journal of Digital Earth, 17(1), p.2310092.
8. Akila, G. and Gayathri, R., 2022. Weighted multi-deep feature extraction for hybrid deep convolutional LSTM-based remote sensing image scene classification model. Geocarto International, 37(27), pp.18217-18253.
9. Jiang, B., An, X., Xu, S. and Chen, Z., 2023. Intelligent image semantic segmentation: a review through deep learning techniques for remote sensing image analysis. Journal of the Indian society of remote sensing, 51(9), pp.1865-1878.
10. Fei, X., Wu, S., Miao, J., Wang, G. and Sun, L., 2024. Lightweight-VGG: A Fast Deep Learning Architecture Based on Dimensionality Reduction and Nonlinear Enhancement for Hyperspectral Image Classification. Remote Sensing, 16(2), p.259.
11. Sagar, A.S., Chen, Y., Xie, Y. and Kim, H.S., 2024. MSA R-CNN: A comprehensive approach to remote sensing object detection and scene understanding. Expert Systems with Applications, 241, p.122788.
12. Khan, S.D. and Basalamah, S., 2023. Multi-branch deep learning framework for land scene classification in satellite imagery. Remote Sensing, 15(13), p.3408.
13. Wang, X., Zhu, J., Yan, Z., Zhang, Z., Zhang, Y., Chen, Y. and Li, H., 2022. LaST: Label-free self-distillation contrastive learning with transformer architecture for remote sensing image scene classification. IEEE Geoscience and Remote Sensing Letters, 19, pp.1-5.
14. Wen, D., Huang, X., Yang, Q. and Tang, J., 2024. Adaptive Self-paced Collaborative and 3D

Adversarial Multi-task Network for Semantic Change Detection using Zhuhai-1 Orbita Hyperspectral Remote Sensing Imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

15. Alahmari, S., Yonbawi, S., Racharla, S., Lydia, E.L., Ishak, M.K., Alkahtani, H.K., Aljarbouh, A. and Mostafa, S.M., 2023. Hybrid Multi-Strategy Aquila Optimization with Deep Learning Driven Crop Type Classification on Hyperspectral Images. Comput. Syst. Sci. Eng., 47(1), pp.375-391.
16. Guan, X., Dong, Y., Tan, W., Su, Y. and Huang, P., 2024. A Parameter-Free Pixel Correlation-Based Attention Module for Remote Sensing Object Detection. Remote Sensing, 16(2), p.312.
17. Ragab, M., 2023. Leveraging mayfly optimization with deep learning for secure remote sensing scene image classification. Computers and Electrical Engineering, 108, p.108672.
18. Ibrahim, H., Kong, N.S.P. and Ng, T.F., 2008. Simple adaptive median filter for the removal of impulse noise from highly corrupted images. IEEE Transactions on Consumer Electronics, 54(4), pp.1920-1927.
19. Devkota, N. and Kim, B.W., 2024. Finger Vein Recognition Using DenseNet with a Channel Attention Mechanism and Hybrid Pooling. Electronics, 13(3), p.501.
20. Zhao, Z. and Luo, S., 2024. A Crisscross-Strategy-Boosted Water Flow Optimizer for Global Optimization and Oil Reservoir Production. Biomimetics, 9(1), p.20.
21. Tu, S., Waqas, M., Badshah, A., Yin, M. and Abbas, G., 2023. Network intrusion detection system (NIDS) based on pseudo-siamese stacked autoencoders in fog computing. IEEE Transactions on Services Computing.
22. http://weegee.vision.ucmerced.edu/datasets/landuse.html
23. https://captain-whu.github.io/AID/
24. Zhao, Y., Liu, J., Yang, J. and Wu, Z., 2022. Remote Sensing Image Scene Classification via Self-Supervised Learning and Knowledge Distillation. Remote Sensing, 14(19), p.4813.
25. Rajagopal, A., Joshi, G.P., Ramachandran, A., Subhalakshmi, R.T., Khari, M., Jha, S., Shankar, K. and You, J., 2020. A deep learning model based on multi-objective particle swarm optimization for scene classification in unmanned aerial vehicles. IEEE Access, 8, pp.135383-135393.