Sparrow Search Algorithm Enhanced Deep Learning for Remote Sensing Scene Detection and Classification

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

[1]Research Scholar, Department of Computer and Information Science, Annamalai University,Annamalainagar-608002, Email: rekhamohanraj00@gmail.com
[2]Assistant Professor, PG Department of Computer Science, Government Arts College, Chidambaram – 608102, Email: sivassa77@gmail.com
*Corresponding Author

Received: 15.07.2024 Revised: 11.08.2024 Accepted: 24.09.2024

ABSTRACT
Remote sensing (RS) scene classification using deep learning has become a rapidly advancing field, offering substantial improvements in accuracy and automation. Deep learning (DL) methods excel in extracting intricate features from RS imagery, eliminating the need for manual feature engineering required by traditional machine learning (ML) techniques. Unlike ML, where domain-specific features must be handcrafted, DL algorithms learn relevant features automatically during training. Convolutional neural networks (CNNs) are widely used in RS scene classification for their ability to capture spatial dependencies and local patterns within images. This paper introduces a novel approach, the Sparrow Search Algorithm with Deep Learning Assisted Remote Sensing Scene Detection and Classification (SSADL-RSSDC). The SSADL-RSSDC method automates the identification and classification of multiple scene labels in RS images. It begins with preprocessing using a median filtering (MF) approach, followed by feature extraction using a deep residual network (ResNet) to learn hierarchical representations of input data. The Sparrow Search Algorithm (SSA) is employed to optimize the hyperparameters of the ResNet model. For the final detection and classification, the Extreme Learning Machine (ELM) is used. The SSADL-RSSDC technique's performance is evaluated on a benchmark RS dataset, demonstrating superior results compared to other models.

Keywords: Remote Sensing Image; Scene Detection; Deep Learning; Sparrow Search Algorithm; Computer Vision

1. INTRODUCTION

Remote sensing images (RSIs) are a useful data source for earth monitoring that will support us in measuring and observing comprehensive structures on the Earth's surface [1]. With the developments of earth observation technology, the amount of RSIs is extremely increasing. It will give a specific emergency in the exploration of how to create all usage of ever-increasing RSIs for intelligent earth observation [2]. Therefore, it is tremendously significant to comprehend large and difficult RSIs [3]. For example, a main and challenging problem for efficiently analysing the RSIs, scene classification of RSIs will be an active research domain. RSI scene classification is to properly label specified RSIs with predetermined semantic types [4]. In recent years, wide-ranging research under RSI scene classification has been undertaken driven by its real-time applications.

Developments in computer vision (CV) comprise finding particular features that will be efficient in memory utilization and computational time [5]. The classifiers are needed to have significant ability of generalisation, yet accomplish higher performance rates [6]. RSI properties develop a major exploration domain. The feature-based technique will be added step from data mining approaches to specify high proportion of implementation under analysis of RSI [7]. Image classification is an important application in the CV complex field. This research's main aim is to increase methods for RSI classification employing Machine learning (ML)-based methods. Satellite images can be categorized and features in the images like buildings, landscapes, deserts, and buildings have been analyzed with time-relevant modifications [8]. Among the subdivisions of Artificial Intelligence (AI), machine learning (ML) has gained remarkable accomplishments, and is presently employed in remote detection [9]. Advancing by the usage of deep convolutional features, the approaches rely on deep learning (DL) is achieve higher effectiveness in image classification and the accuracy is continuous existence of progressive with the enhancement of innovative

techniques. DL method can also be familiar with cloud detection in satellite-based images in currently continuing research [10].

This article introduces Sparrow Search Algorithm with Deep Learning Assisted Remote Sensing Scene Detection and Classification (SSADL-RSSDC) technique. The focus of the SSADL-RSSDC technique lies in the automated identification and classification of multiple scene labels in the RS images. For initial preprocessing step, the SSADL-RSSDC technique applies median filtering (MF) approach. Besides, the SSADL-RSSDC technique involves deep residual network (ResNet) model for learning hierarchical representations of the input data. Moreover, the SSA can be applied to selecting the optimum values of the Deep ResNet'shyperparameter. Finally, extreme learning machine (ELM) model is applied for the detection and classification process. The performance assessment of the SSADL-RSSDC technique takes place using benchmark RSI dataset.

2. LITERATURE WORKS

Zhang and Wang [11] introduced a progressive feature fusion (PFF) architecture dependent upon GCN system comprising two components such as a multiscale contextual information fusion (MCIF) and multilayer feature extraction (MFE). The MFE method was employed for extraction, and the MCIF technique was designed to capture significant contextual data. In MCIF, GCN was employed. By using the PFF approach, the graph features in every level have been combined with the subsequent level features. Moreover, collected GCN relies on channel combination was then developed. In [12], an efficient defense model for RSI scene classification, called reconstruction-enabled and distance-optimized adversarial training (RDAT) has been developed. Similarly, a distance-optimized (DO) approach was developed in adversarial training (AT). Additionally, a reconstruction-assisted (RA) block was designed. Particularly, in this block, by SwinT-MSC-UNet (SMUNet) and multiscale convolution (MSC) block, Swin Transformer (SwinT) block was built. Ye et al. [13] presented an innovative incremental learning model called effective channel attention-based multiscale depthwise network (ECA-MSDWNet) technique. Besides, in incremental learning method, the techniques increase the innovative components that rely on a dynamic-structure technique. Lastly, this work focuses on the system for decreasing redundant features and feature sizes via an efficient knowledge distillation approach.

Ahmed et al. [14] projected an IoT-enabled smart surveillance technique. The method offers an AI-based system employing the DL-based segmentation algorithm PSPNet to segment various objects. The method also employed an aerial drone database, executed the data augmentation methods, and leveraged DTL to improve the model's effectiveness. In [15], a modified detection technique was developed that depends upon higher-resolution RSIs and DL. Primarily, the improved simple linear iterative clustering (SLIC) scheme was utilized. Then, annotated databases have been produced employing cropping-with-inpainting, and multiscale extracting techniques. Afterwards, datasets obtained under posttemporal and pretemporal images could be employed. Lastly, an enriched object-oriented CNN method executes classification.

Aljebreen et al. [16] developed an innovative Land Use and Land Cover Classification utilizing the River Formation Dynamics Algorithm with DL (LULCC-RFDADL) system. The dense EfficientNet algorithm can be employed for feature extraction. Likewise, the RFDA method was implemented for tuning. The system utilizes the Multi-Scale Convolutional AE (MSCAE) method for classification. Lastly, the seeker optimization algorithm (SOA) was employed. In [17], an innovative few-shot RSSC called multiple text-task models directed dynamic contrast learning network (MPCL-Net) was introduced. Primarily, two RS-based pretext tasks could be made on the self-supervised learning (SSL) technique. Secondarily, a simplified CNN was established for the deep feature learning subcomponent. Then, the three loss functions have been incorporated and developed into the combined optimization submodules. Lastly, the MPCL-Net must be trained for a meta approach.

3. The Proposed Model

In this article, we have introduced a novel SSADL-RSSDC method. The focus of the SSADL-RSSDC method lies in the automated detection and classification of multiple scene labels in the RS images. Fig. 1 depicts the workflow of SSADL-RSSDC technique.

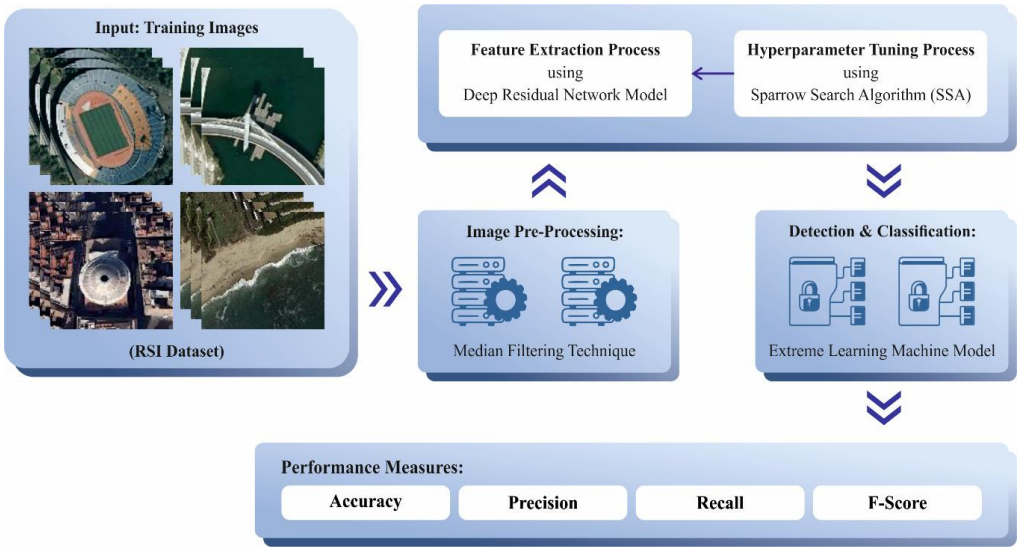

Fig 1. Workflow of SSADL-RSSDC technique

3.1 MF-based Preprocessing

The SSADL-RSSDC technique applies MF approach for initial preprocessing step. MF is a classical image processing method used to remove noise while maintaining edges in digital images [18]. In this method, a sliding window (a rectangular or square neighbourhood) moves across the image, and the pixel values within the window are sorted. The median value is then assigned to the centre pixel of the window. This process efficiently eliminates outliers and decreases the effect of salt-and-pepper noise, making it particularly helpful in scenarios where maintaining fine details is important. MF is robust in handling impulse noise and is computationally efficient, making it a valuable tool for increasing image quality in various applications, such as computer vision and medical imaging.

3.2. Feature Extraction

The SSADL-RSSDC technique involves deep ResNet model for learning hierarchical representations of the input data. F^{set} is the input of deep ResNet [19]. The main benefit of ResNet is it can construct deep networks with huge amounts of weighted layers. From loss function to the dissimilar layers, the gradients discover it very difficult to back transmit without declining to 0 or enlarging as the complexity of a network upsurge. ResNet empowers gradients to move over layers by using a skip connection. In Eq. (1), the residual function $I(P^{set})$ is acquired from the residual block $D(u)$ that conveyed as:

$$D(u) = I(P^{set}) + F^{set} \quad (1)$$

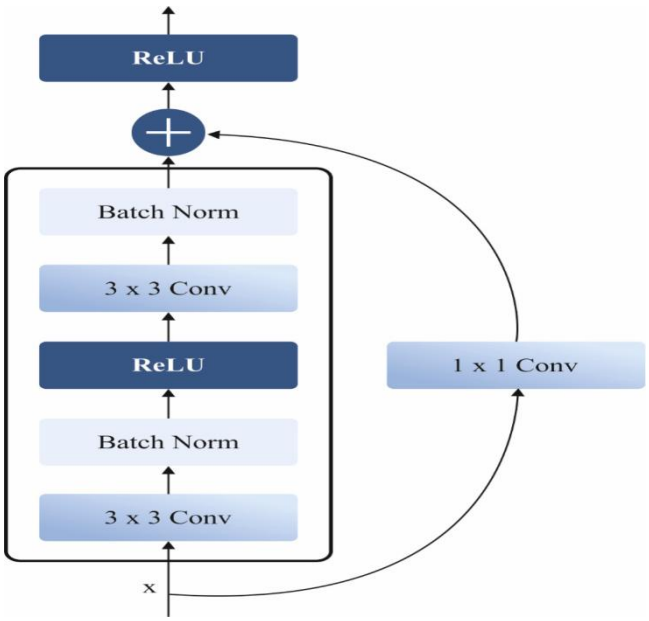

Fig 2. Structure of ResNet

In order to train the weight layers S^* by utilizing labeled data, the residual function $I(P^{set})$ has been learned. Any neural network method like convolutional and fully connected (FC) layers are employed as a weight layer. The residual block permits the forward pass among the system to select to avoid those sections by stating $D(u)$ as zero for specific sections. By employing numerous layers of feature extractors, each can remove a dissimilar possible feature of the input. Now, it is highly probable to build deep networks. Only the system sections are needed for classifying an assumed input case which is initiated for that case. Fig. 2 depicts the framework of ResNet model. The applications of DL regularly experience the problem of overfitting which has a harmful result in forecasting new datasets. In this research, a hybrid classification technique termed CNNs and deep ResNet upsurges the number of passes on the current training data and amount of training data instances, which makes the method lightweight to defeat the problem of overfitting.

3.3. SSA-based Hyper parameter Tuning

At this stage, the SSA can be applied to selecting the optimum values of the Deep ResNet'shyperparameter. SSA is a recent swarm intelligence optimization technique motivated primarily by the anti-predatory and foraging behaviors of natural sparrows [20]. The SSA technique is well-suited for applications in ML and artificial intelligence technologies.

The key concept of SSA algorithm is given below:

(1) Sparrow foraging activity: sparrows migrate in different directions looking for food while foraging. The solution space of the problem is considered a distribution of food, where individuals seek optimum solutions.

(2) Individual sparrows and pops: each individual signifies a solution and looks for the optimum one using their unique action and interaction with others. There are several individuals in the pops that contribute jointly to improve the search performance.

(3) Flight and position updates: Every individual flies based on their position and speed throughout each iteration, and the distance and direction of the flight are directly influenced by global optimum and individual solutions. Based on the flight outcome, the position of individual is adjusted and gradually comes closer to the optimum solution.

(4) Adaptation assessment: The assessment of objective function is used for measuring fitness of the individual. The most suitable solution is the high fitness value and closer to the individual.

(5) Knowledge transfer and upgrades: Each sparrow observes the fitness values and activities of others' locations to upgrade their behaviors for extending their search through upgrade and knowledge transfer.

$$X = (x_1^1 x_1^2 \ \dots \ x_1^D x_2^1 x_2^2 \ \dots \ x_2^D \ \vdots \vdots \vdots \ x_N^1 x_N^2 \ \dots \ x_N^D)$$

Where the location of each sparrow is $X = (x_1, x_2, \dots, x_D)$, N denotes the amount of sparrows in the population, D signifies the size of optimum solution.

The discoverer, follower, and alert are the three components of SSA algorithm. The sparrow ratio of discoverer and followers is upgraded by the mean square error among true and prediction values as the fitness value. Next, the iteration upgrades are carried out based on the location upgrade:

Discoverer

The individuals who find superior food are answerable for delivering direction to their followers. In each generation, the top PN sparrows with best fitness value are selected as the discoverers. The formulation for location upgrading is given below:

$$X_{i,j}^{t+1} = \{X_{i,j}^t \times exp\left(-\frac{i}{\alpha \times T}\right) if\ R_2 < ST X_{i,j}^t + Q \times L if\ R_2 \geq ST \quad (2)$$

From above-mentioned equation, x denotes the location of the $it\hbar$ sparrow in dimension j at iteration t, i signifies the sparrow count, T signifies the maximal amount of iterations about random number $\alpha \in (0,1]$, $ST \in [0.5,1]$ and warning values $R_2 \in [0,1]$ for security values; Q is the arbitrary value; and L is a matrix of size $1 \times d$. If $R_2 \geq ST$, scouts identify the existence of predators, and the cluster travels quickly near a secure region, If $R_2 < ST$, there are no predators, and finders do wide searches in the region.

Follower

Excepting for the discoverers, all lasting N-PN individuals aid as followers and their location upgrade formula is as follows:

$$X_{i,j}^{t+1} = \{Q \times exp\left(\frac{X_{worst}^t - x_{i,j}^t}{t^2}\right) if\ i > 0.5n X_p^{t+1} + \left|X_{i,j}^t - X_p^{t+1}\right| \times L \times A + if\ i \leq 0.5n \quad (3)$$

$X_{i,j}^{t+1}$ denotes the position with optimum alteration organized through the discoverer at the $t+1st$ iteration, X_{worst}^t signifies the global worst location, A^+ refers to random integer within [1,-1] and

$A^+ = (AA^T) - 1$. If $i \leq 0.5n$, the i^{th} accession will follow the discoverer's foraging centre and arbitrarily feed near the middle location; if $i > 0.5n$, then i^{th}accession flops to attain food and has lower energy level to go to other regions to feed.

Alert

Picking an assured amount of signals to do investigation and initial cautionary task and give food when risks are complex, supposing every generation arbitrarily picks SD sparrows for reconnaissance and cautionary. The formulation for upgrading its location is given below:

$$X_{i,j}^{t+1} = \{X_{best}^t + \beta \times |x_{i,j}^t - X_{best}^t| \; if \; f_i > f_g X_{i,j}^t + K \times \left(\frac{x_{i,j}^t - X_{worst}^t}{(f_i - f_w)\varepsilon}\right) \; if \; f_i = f_g \quad (4)$$

X_{best}^t denotes the global optimum location; β is a step switch parameter, $K \in [-1,1 f_i$ is the fitness value of the it^{th} individual sparrow; f_w and $f_g$refers to the worst and best fitness values; and ε denotes the minimum constant. When $f_i > f_g$, sparrows are at the edge of the population and weak to hunters; when $f_i = f_g$, sparrows are at the midpoint of the population and arbitrarily fall near other sparrows; and when $f_i < f_g$, a spy is not moving.

The SSA derives an FF to achieve higher classification outcomes. It determines a positive integer to represent the superior performance of the candidate solution. Here, the reduction of classifier error is assumed as the FF.

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{No.of\ misclassified\ samples}{Total\ No.of\ samples} * 100 \quad (5)$$

3.4. Classification using ELM technique

Finally, the ELM model is applied to the recognition and classification process. The network of ELM generates voltage degradation features acquired by DBN in order to execute a more precise degradation forecast [21]. The ensemble of output and hidden layers (HL) (H1, H2) are considered a classic ELM system. The attained parameters of RBM2 $\{W, a, b\}$ have been employed to set the output weight vector β and the feature space H of ELM is attained. So, entire parameter of the DBN-ELM system is obtained.

Assume $x_j (j = 1, \dots, m)$ signifies ELM input of network. t_j denotes the training output, which is equivalent to x_t in the entire DBN-ELM system. The activation function of sigmoid is implemented and signified as (W, b, x). The output function $f_l(x)$ is defined by:

$$f_l(x) = \sum_{i=1}^{l} \beta_i G(W_i, b_i, x_j) = t_j, (j = 1, \dots, m) \quad (6)$$

Here, the parameters $\{(W_i, b_i), i = 1,2, \dots, l\}$ have been acquired from the DBN (H2). β_i signifies the output weight linking the it^{th} output and HL. Eq. (6) is stated in a matrix form.

$$H\beta = T$$

$$H = [G(W_1, a_1, b_1) \cdots G(W_l, a_l, b_l) \cdots \cdots \cdots G(W_1, a_1, b_m) \cdots G(W_l, a_l, b_m)] \quad (7)$$

$$\beta = [\beta_1^T \cdots \beta_l^T], T = [t_1^T \cdots t_l^T] \quad (8)$$

Whereas, H represents the HL output matrix that is called ELM feature space. The training procedure is to discover the optimum output weight β^*, which is attained by minimalizing the estimated error as follows.

$$min \; ||H\beta - T||^2 \quad (9)$$

If H is said to be non-singular, then the best output weight β^* is considered as below:

$$\beta^* = H^+ T \quad (10)$$

While, H^+ denotes the Moore-Penrose general reverse matrix of HL output. If H is complete column rank, then H^+ can be stated as below:

$$H^+ = \left(\frac{I}{C} + H^T H\right)^{-1} H^T \quad (11)$$

Where C signifies the regularization co-efficient, which is presented to make the method more steady. To acquire a superior regularization co-efficient, it gets few values for C at first. And then, these regularization co-efficient values were later switched to Eq. (9) to discover the optimal value of C which can diminish the estimated error of ELM.

4. Performance Validation

The experimental analysis of the SSADL-RSSDC technique takes place using UCM [22] and AID [23] datasets. Fig. 3 demonstrates the sample images.

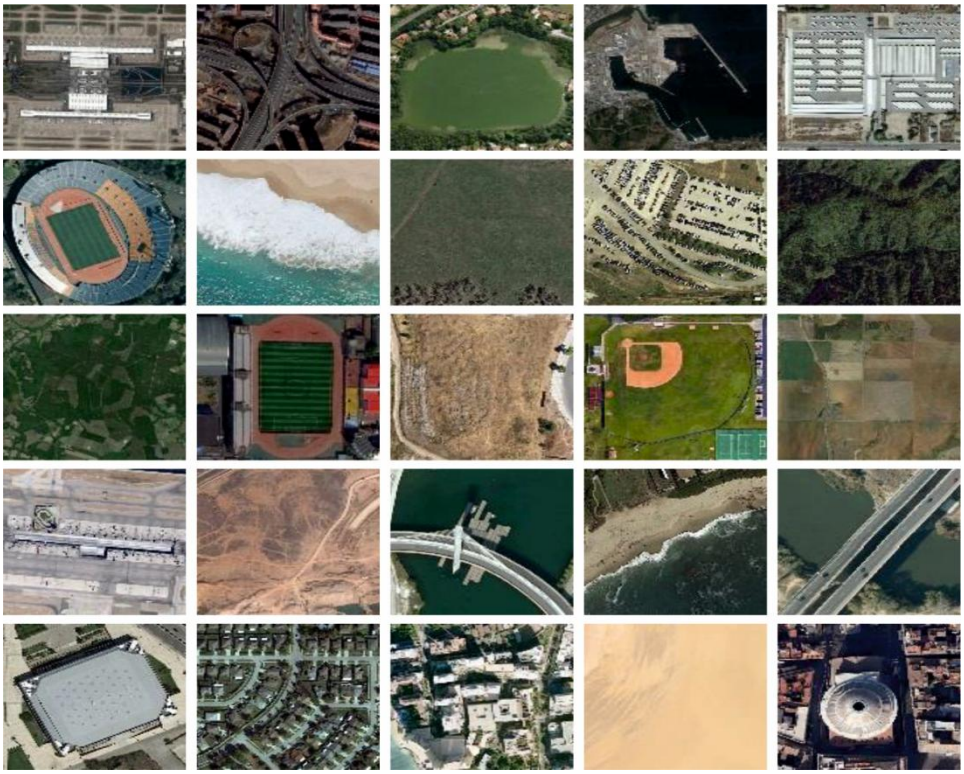

Fig 3. Sample images

Fig. 4 illustrates the classifier results of the SSADL-RSSDC algorithm at UCM dataset. Figs. 4a-4b displays the confusion matrices obtained by the SSADL-RSSDC technique with 70%:30% of TRAPH/TESPH. This figure indicated that the SSADL-RSSDC technique can be identified and categorized with 21 class labels appropriately. Meanwhile, Fig. 4c indicates the PR result of the SSADL-RSSDC method. The figure described that the SSADL-RSSDC algorithm gives greater PR effectiveness with every class. In conclusion, Fig. 4d showcases the ROC result of the SSADL-RSSDC technique. This figure represented that the SSADL-RSSDC method offers efficient outcomes with increased ROC values with diverse classes.

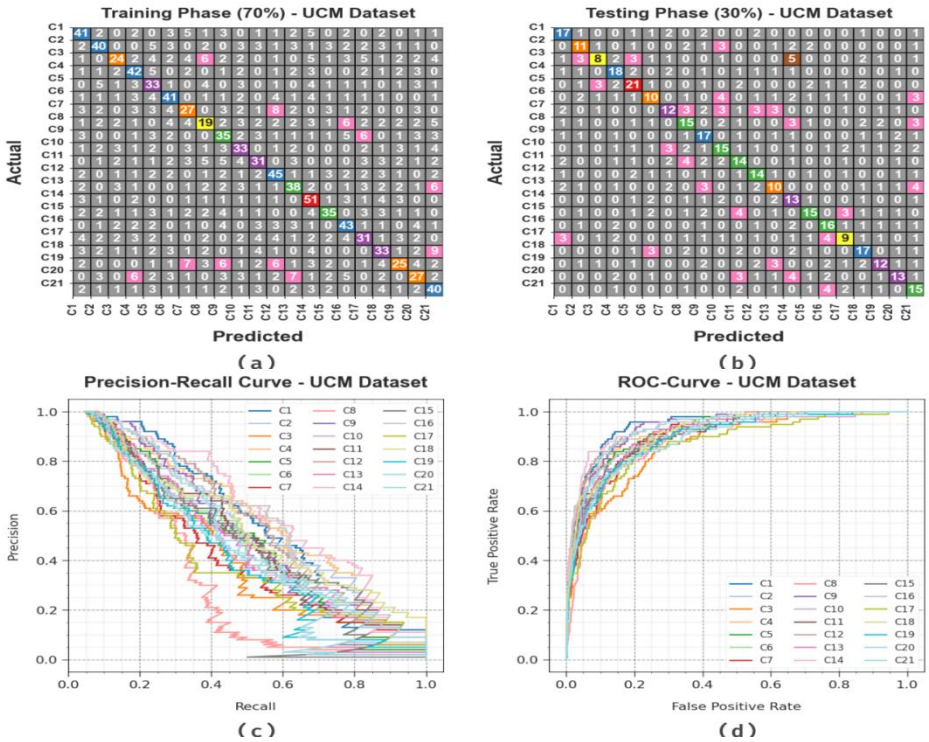

Fig 4. UCM dataset (a-b) Confusion matrices (c-d) PR-curve and ROC-curve

Table 1 highlights the overall scene classification results offered by the SSADL-RSSDC technique on 70% of TRAPH under UCM dataset. These experimentation outcomes underline that the SSADL-RSSDC algorithm properly identified all types of scenes exist in the UCM dataset. It is also noticed that the SSADL-RSSDC technique has the ability to recognize the samples with maximum classifier results.

Table 1. Scene classification of SSADL-RSSDC model with UCM dataset under 70% of TRAPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TRAPH (70%)				
C1	95.65	55.41	56.94	56.16
C2	95.71	58.82	53.33	55.94
C3	95.31	52.17	33.80	41.03
C4	95.71	54.55	60.00	57.14
C5	94.97	42.86	52.38	47.14
C6	95.85	55.41	59.42	57.34
C7	94.83	42.86	40.30	41.54
C8	94.42	35.85	28.36	31.67
C9	94.56	44.30	49.30	46.67
C10	95.44	47.14	52.38	49.62
C11	95.24	49.21	44.93	46.97
C12	95.37	53.57	60.81	56.96
C13	95.51	53.52	53.52	53.52
C14	95.17	52.04	68.00	58.96
C15	96.26	60.34	52.24	56.00
C16	95.58	54.43	59.72	56.95
C17	94.69	45.59	43.06	44.29
C18	94.90	46.48	47.14	46.81
C19	95.10	49.02	35.21	40.98
C20	94.63	42.86	38.57	40.60
C21	94.97	48.19	56.34	51.95
Average	**95.23**	**49.74**	**49.80**	**49.44**

The overall scene classification results offered by the SSADL-RSSDC method at 30% of TESPH at UCM datasets as reported in Table 2. These accomplished outcomes emphasize that the SSADL-RSSDC method appropriately recognized all categories of scenes present in the UCM dataset. It is also perceived that the SSADL-RSSDC algorithm can be the ability to identify the samples with increased classifier outcomes.

Table 2. Scene classification of SSADL-RSSDC model with UCM dataset under 30% of TESPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TESPH (30%)				
C1	96.03	54.84	60.71	57.63
C2	95.40	42.31	44.00	43.14
C3	95.40	50.00	27.59	35.56
C4	95.56	52.94	60.00	56.25
C5	94.76	55.26	56.76	56.00
C6	94.44	41.67	32.26	36.36
C7	94.60	48.00	36.36	41.38
C8	94.13	44.12	45.45	44.78
C9	94.76	44.74	58.62	50.75
C10	93.65	45.45	40.54	42.86
C11	94.13	41.18	45.16	43.08
C12	95.71	48.28	53.85	50.91
C13	94.44	38.46	34.48	36.36
C14	94.13	34.21	52.00	41.27
C15	95.24	55.56	45.45	50.00
C16	95.71	51.61	57.14	54.24
C17	94.76	39.13	32.14	35.29
C18	95.24	50.00	56.67	53.12

C19	95.40	50.00	41.38	45.28
C20	95.08	48.15	43.33	45.61
C21	94.13	39.47	51.72	44.78
Average	**94.89**	**46.45**	**46.46**	**45.94**

In Fig. 5, the average classifier analysis of the SSADL-RSSDC technique on the UCM dataset is revealed. The figure shows that the SSADL-RSSDC method accomplished enhanced results. With 70% of TRAPH, the SSADL-RSSDC technique attains average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 95.23%, 49.74%, 49.80%, and 49.44%, respectively. Also, based on 30% of TESPH, the SSADL-RSSDC algorithm provides average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 94.89%, 46.45%, 46.46%, and 45.94%.

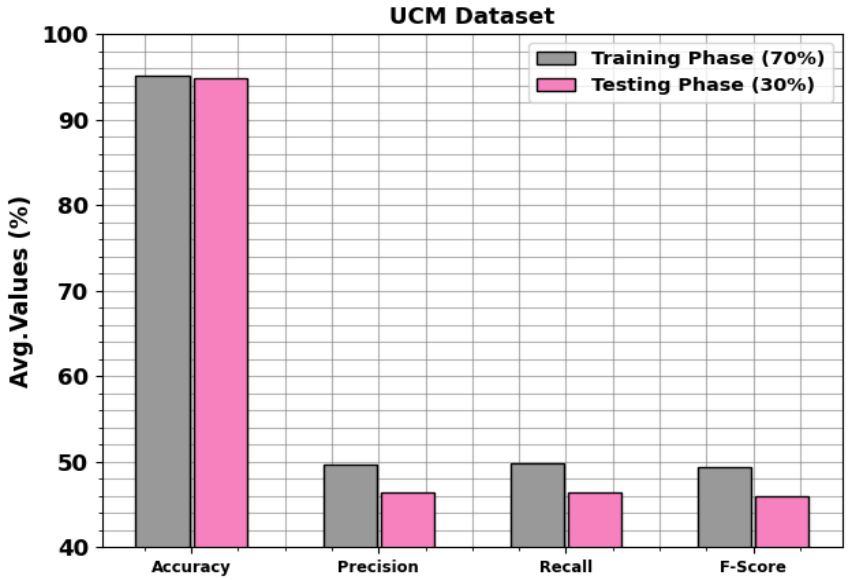

Fig 5. Average of the SSADL-RSSDC system under UCM dataset

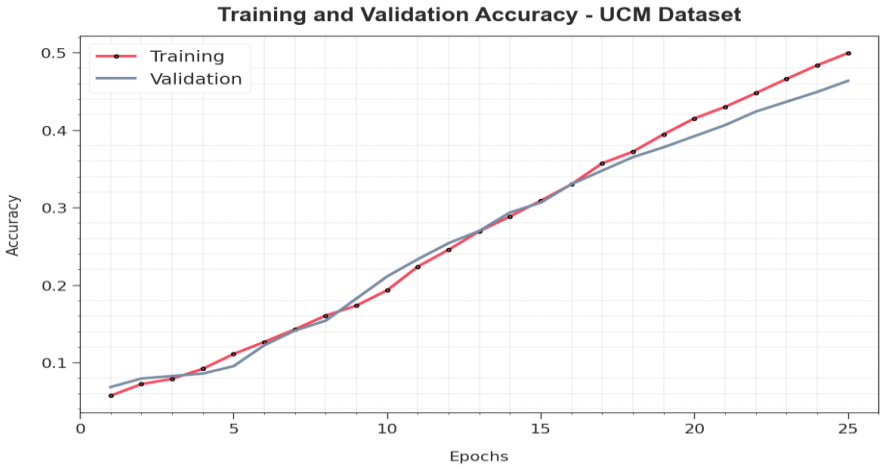

Fig 6. $Accu_y$ curve of the SSADL-RSSDC method at UCM dataset

The effectiveness of the SSADL-RSSDC algorithm with UCM dataset is graphically illustrated in Fig. 6 in the form of training accuracy (TRAA) and validation accuracy (VALA) curves. The figure exhibits useful interpretation of the behaviour of the SSADL-RSSDC method over varying epoch counts, representing its learning process and generalization capabilities. Considerably, the figure infers a constant improvement in the TRAA and VALA with progress in epochs. It makes sure of the adaptive nature of the SSADL-RSSDC technique with pattern recognition process under the TRA and TES data. The rising trend in VALA outlines the capability of the SSADL-RSSDC method to alter the TRA data and also excels in giving exact classification on unnoticed data, pointing out robust generalization abilities.

Fig. 7 shows an extensive representation of the training loss (TRLA) and validation loss (VALL) results of the SSADL-RSSDC method with UCM dataset over distinct epochs. The progressive reduction in TRLA highlights the SSADL-RSSDC technique optimizing the weights and minimizing the classification error on the TRA and TES data. The figure indicates a clear understanding of the SSADL-RSSDC model related to the TRA data, highlighting its proficiency in capturing patterns within both datasets. Remarkably, the SSADL-RSSDC system incessantly improves its parameters in decreasing the differences among the prediction and real TRA class labels.

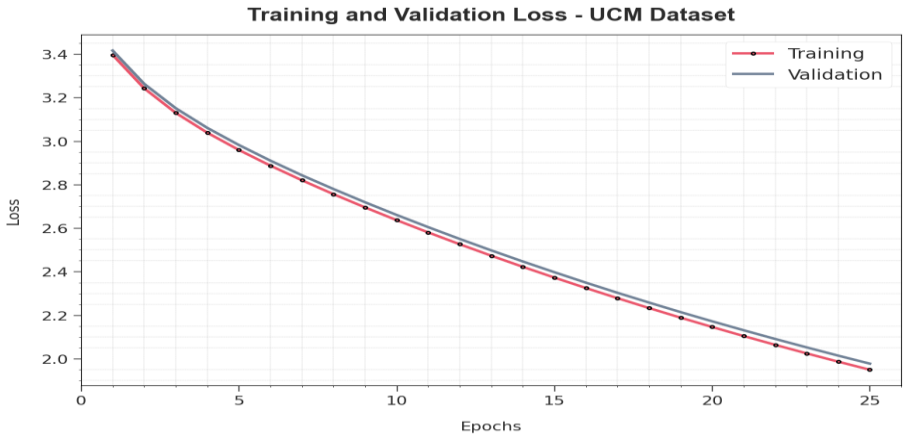

Fig 7. Loss curve of the SSADL-RSSDC model at UCM dataset

The $Accu_y$ comparison study of the SSADL-RSSDC technique on the UCM dataset is reported as displayed in Table 3 and Fig. 8 [24, 25]. These obtained outcomes highlighted that the SC+Pooling, SG+UFL, and CCM-BOVW algorithms have revealed decreased $accu_y$ values of 81.67%, 86.64%, and 86.64%, respectively. Next, the PSR, COPD, and Dirichlet systems have reported closer $accu_y$ values of 89.10%, 91.33%, and 92.80%, correspondingly. Nevertheless, the SSADL-RSSDC technique reaches better results with maximum $accu_y$ of 95.23%.

Table 3. $Accu_y$ outcome of SSADL-RSSDC technique 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
SSADL-RSSDC	95.23

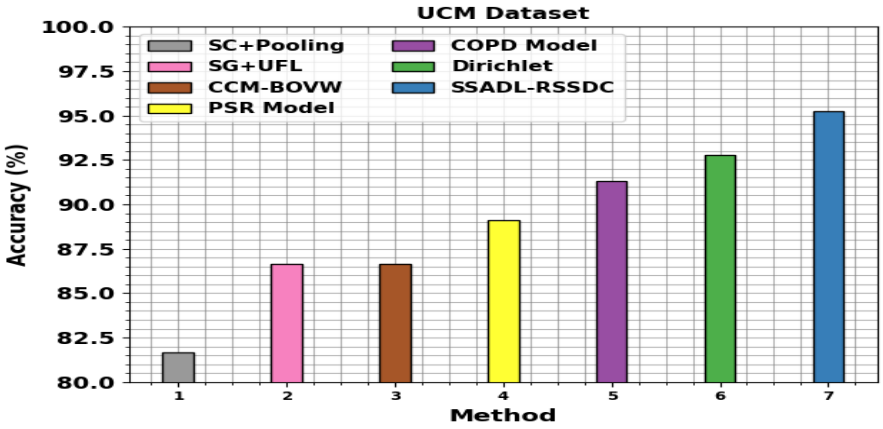

Fig 8. $Accu_y$ outcome of the SSADL-RSSDC method under UCM dataset

Fig. 9 displays the classifier results of the SSADL-RSSDC technique in AID database. Figs. 9a-9b showcases the confusion matrices succeeded by the SSADL-RSSDC system with 70%:30% of TRAPH/TESPH. This figure specified that the SSADL-RSSDC method can be identified and categorized with 30 class labels correctly. Similarly, Fig. 9c specifies the PR result of the SSADL-RSSDC system. The figure displays that the SSADL-RSSDC algorithm provides excellent PR effectiveness with every class. In conclusion, Fig. 9d displays the ROC result of the SSADL-RSSDC technique. This figure represented that the SSADL-RSSDC method gives efficient outcomes with improved ROC values with diverse class labels.

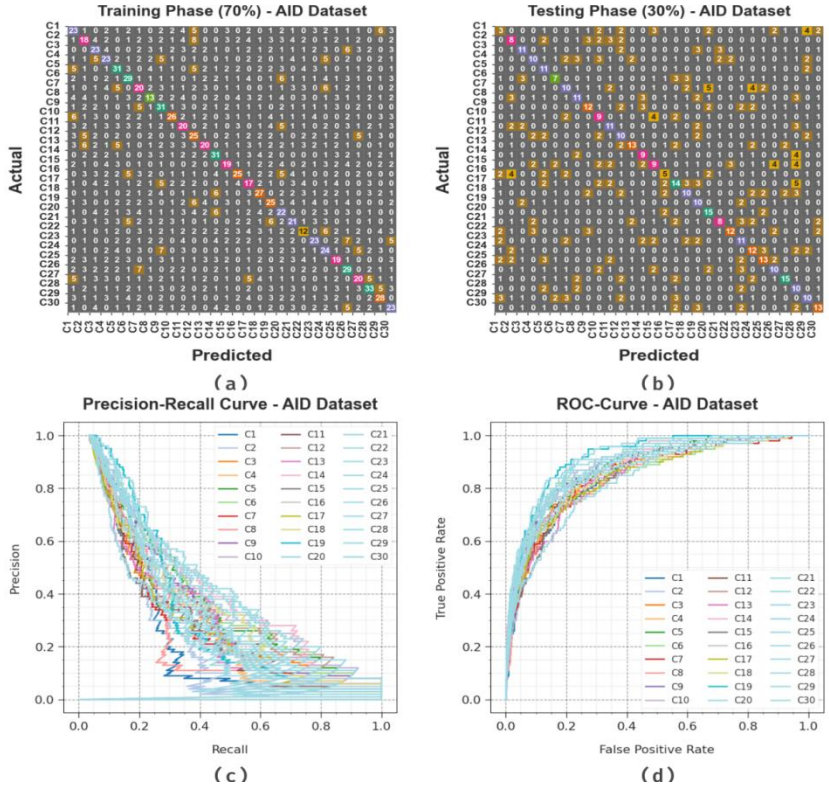

Fig 9. AID dataset (a-b) Confusion matrices (c-d) PR-curve and ROC-curve

Table 4 highlights the overall scene classification outcome offered by the SSADL-RSSDC approach on 70% of TRAPH under AID dataset. These experimentation outcomes underline that the SSADL-RSSDC algorithm accurately identified all types of scenes that exist in the AID dataset. It is also detected that the SSADL-RSSDC technique can recognize the samples with higher classifier outcomes.

Table 4. Scene classification of SSADL-RSSDC model with AID dataset under 70% of TRAPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TRAPH (70%)				
C1	95.43	31.51	33.33	32.39
C2	94.86	26.47	23.68	25.00
C3	95.48	32.86	32.39	32.62
C4	95.62	31.51	35.38	33.33
C5	95.43	39.24	39.24	39.24
C6	95.57	37.18	39.73	38.41
C7	95.24	27.03	30.30	28.57
C8	95.29	22.03	19.70	20.80
C9	95.33	33.70	45.59	38.75
C10	96.29	45.61	35.62	40.00
C11	94.90	26.32	28.17	27.21
C12	95.43	32.89	35.71	34.25
C13	95.95	37.04	28.17	32.00
C14	96.10	43.66	42.47	43.06
C15	96.10	32.20	31.15	31.67

C16	96.14	42.37	34.72	38.17
C17	95.10	22.97	26.98	24.82
C18	96.05	39.71	39.13	39.42
C19	94.95	30.86	33.33	32.05
C20	94.57	26.19	29.73	27.85
C21	95.24	29.58	29.58	29.58
C22	95.29	21.43	17.91	19.51
C23	95.95	38.98	31.94	35.11
C24	95.19	30.38	34.29	32.21
C25	95.76	30.65	29.23	29.92
C26	95.29	33.33	41.43	36.94
C27	95.95	34.48	29.85	32.00
C28	96.24	49.25	42.31	45.52
C29	96.14	41.79	40.00	40.88
C30	95.81	33.33	35.38	34.33
Average	**95.56**	**33.49**	**33.22**	**33.19**

The overall scene classification outcomes succeeded by the SSADL-RSSDC method on the 30% of TESPH under AID dataset as shown in Table 5. These experimental findings emphasize that the SSADL-RSSDC algorithm appropriately recognized all categories of scenes existing in the AID dataset. It is also observed that the SSADL-RSSDC method can be the ability to recognize the samples with boosted classifier results.

Table 5. Scene classification of the SSADL-RSSDC method with AID dataset under 30% of TRAPH

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
TESPH (30%)				
C1	95.11	15.79	09.68	12.00
C2	95.67	25.81	33.33	29.09
C3	96.44	44.00	37.93	40.74
C4	95.67	41.67	28.57	33.90
C5	96.22	31.43	52.38	39.29
C6	96.11	31.82	25.93	28.57
C7	95.00	32.26	29.41	30.77
C8	95.89	44.00	32.35	37.29
C9	95.67	38.71	37.50	38.10
C10	95.44	28.12	33.33	30.51
C11	95.78	35.48	37.93	36.67
C12	94.78	27.03	33.33	29.85
C13	96.78	50.00	44.83	47.27
C14	95.78	31.03	33.33	32.14
C15	94.67	33.33	23.08	27.27
C16	95.89	26.32	17.86	21.28
C17	94.78	36.84	37.84	37.33
C18	95.67	35.71	32.26	33.90
C19	95.67	29.41	40.00	33.90
C20	96.44	41.67	57.69	48.39
C21	96.00	34.78	27.59	30.77
C22	96.11	46.15	36.36	40.68
C23	95.44	31.43	39.29	34.92
C24	95.11	31.58	40.00	35.29
C25	95.56	41.94	37.14	39.39
C26	95.56	33.33	33.33	33.33
C27	96.33	50.00	45.45	47.62
C28	94.89	22.73	45.45	30.30
C29	95.33	31.25	33.33	32.26
C30	95.56	41.94	37.14	39.39
Average	**95.64**	**34.85**	**35.12**	**34.41**

In Fig. 10, the average classifier outcomes of the SSADL-RSSDC technique with AID dataset are revealed. The figure shows that the SSADL-RSSDC algorithm gets improved results. According to 70% of TRAPH, the SSADL-RSSDC technique attains average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 95.56%, 33.49%, 33.22%, and 33.19%. Besides, with 30% of TESPH, the SSADL-RSSDC method provides average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 95.64%, 34.85%, 35.12%, and 34.41%.

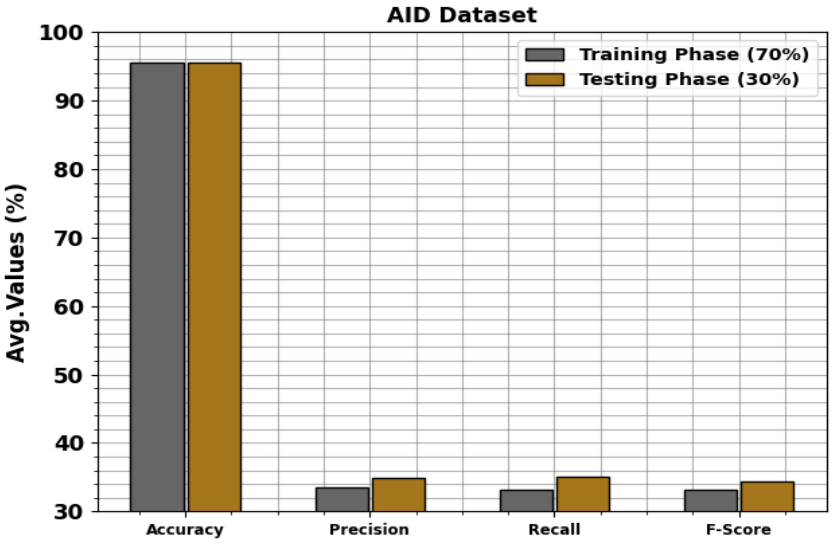

Fig 10. Average of the SSADL-RSSDC method under AID dataset

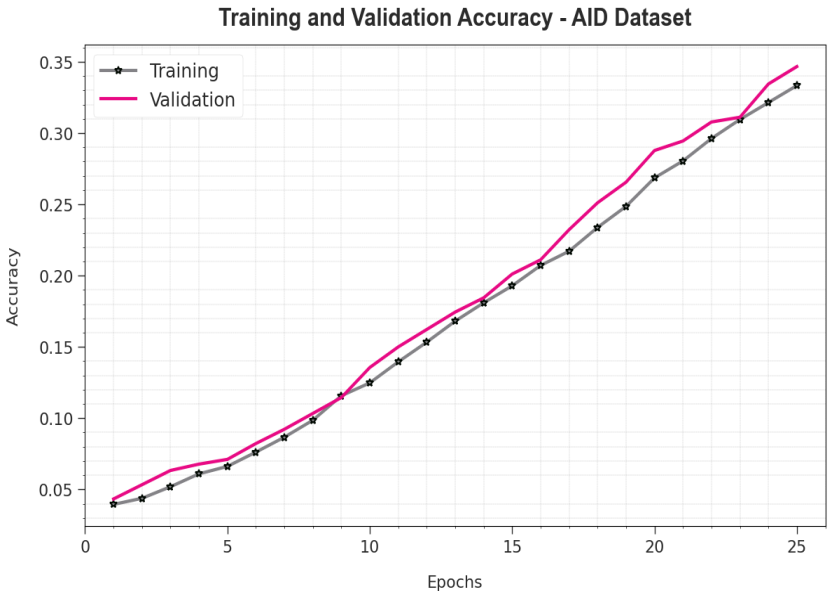

Fig 11. $Accu_y$ curve of the SSADL-RSSDC model at AID dataset

The performance of the SSADL-RSSDC system with AID dataset is graphically demonstrated in Fig. 11 in the form of TRAA and VALA curves. This figure exhibits useful interpretation of the behaviour of the SSADL-RSSDC method over multiple epoch counts, signifying its learning process and generalization capabilities. Noticeably, the figure infers a constant enhancement in the TRAA and VALA with progress in epochs. It makes sure of the adaptive nature of the SSADL-RSSDC technique with pattern recognition process under the TRA and TES data. The rising trend in VALA outlines the capability of the SSADL-RSSDC method to alter the TRA data and excels in giving particular classification on unnoticed data, pointing out the robust generalization abilities.

Fig 12. Loss curve of the SSADL-RSSDC technique with AID dataset

Fig. 12 shows a wide-ranging representation of the TRLA and VALL results of the SSADL-RSSDC system with AID dataset over distinct epochs. The progressive reduction in TRLA highpoints the SSADL-RSSDC system optimizing the weights and decreasing the classification error on the TRA and TES data. The figure indicates a perfect understanding of the SSADL-RSSDC model related to the TRA data, highlighting its proficiency in capturing patterns within both datasets. Significantly, the SSADL-RSSDC algorithm continually improves its parameters in decreasing the differences among the prediction and real TRA class labels.

An extensive $Accu_y$ comparative outcomes of the SSADL-RSSDC algorithm on the AID dataset is informed as shown in Table 6 and Fig. 13. These experimentation outcomes indicate that the GoogleNet and VGG-VD-16 algorithms get diminished $accu_y$ values of 86.39%, and 89.64%, correspondingly. Then, the ResNet50, ResNet-50+EAM, and LCNN-BFF methods are gained closer $accu_y$ values of 92.57%, 93.64%, and 91.66%. However, the SSADL-RSSDC algorithm achieves excellent results with increased $accu_y$ of 95.64%.

Table 6. $Accu_y$ outcome of the SSADL-RSSDC technique 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
SSADL-RSSDC	95.64

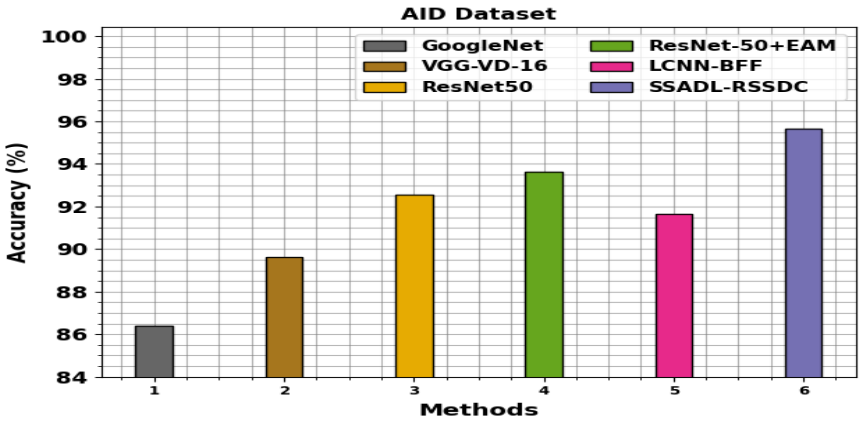

Fig 13. $Accu_y$ outcome of SSADL-RSSDC model under AID dataset

Thus, the SSADL-RSSDC technique can be applied for automated scene classification on the RSI.

5. CONCLUSION

In this article, we have introduced a novel SSADL-RSSDC algorithm. The focus of the SSADL-RSSDC method lies in the automated identification and classification of multiple scene labels in the RS images. For initial preprocessing step, the SSADL-RSSDC technique applies M) approach. Besides, the SSADL-RSSDC technique involves deep ResNet model for learning hierarchical representations of the input data. Moreover, the SSA can be applied to selecting the optimum values of the Deep ResNet'shyperparameter. Finally, the ELM model is applied to the recognition and classification process. The performance assessment of the SSADL-RSSDC algorithm takes place using benchmark RSI dataset. The experimental values inferred that the SSADL-RSSDC technique demonstrates superior performance over other models.

REFERENCES

[1] Li, X., Sayer, A.M., Carroll, I.T., Huang, X. and Wang, J., 2024. MT-HCCAR: Multi-Task Deep Learning with Hierarchical Classification and Attention-based Regression for Cloud Property Retrieval. arXiv preprint arXiv:2401.16520.

[2] Yang, Y., Jiao, L., Liu, F., Liu, X., Li, L., Chen, P. and Yang, S., 2023. An Explainable Spatial-Frequency Multi-Scale Transformer for Remote Sensing Scene Classification. IEEE Transactions on Geoscience and Remote Sensing.

[3] Chen, W., Ouyang, S., Tong, W., Li, X., Zheng, X. and Wang, L., 2022. GCSANet: A global context spatial attention deep learning network for remote sensing scene classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, pp.1150-1162.

[4] 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.

[5] Sun, S., Mu, L., Feng, R., Chen, Y. and Han, W., 2024. Quadtree decomposition-based Deep learning method for multiscale coastline extraction with high-resolution remote sensing imagery. Science of Remote Sensing, 9, p.100112.

[6] Song, H. and Yang, W., 2022. GSCCTL: a general semi-supervised scene classification method for remote sensing images based on clustering and transfer learning. International Journal of Remote Sensing, 43(15-16), pp.5976-6000.

[7] 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.

[8] de RodaHusman, 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.

[9] Hamza, A., Khan, M.A., Ur Rehman, S., Albarakati, H.M., Alroobaea, R., Baqasah, A.M., Alhaisoni, M. and Masood, A., 2023. An Integrated Parallel Inner Deep Learning Models Information Fusion With Bayesian Optimization for Land Scene Classification in Satellite Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[10] Wang, W., Chen, Y. and Ghamisi, P., 2022. Transferring CNN with adaptive learning for remote sensing scene classification. IEEE Transactions on Geoscience and Remote Sensing, 60, pp.1-18.

[11] Zhang, C. and Wang, B., 2024. Progressive Feature Fusion Framework Based on Graph Convolutional Network for Remote Sensing Scene Classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[12] Su, Y., Zhang, G., Mei, S., Lian, J., Wang, Y. and Wan, S., 2023. Reconstruction-Assisted and Distance-Optimized Adversarial Training: A Defense Framework for Remote Sensing Scene Classification. IEEE Transactions on Geoscience and Remote Sensing.

[13] Ye, Z., Zhang, Y., Zhang, J., Li, W. and Bai, L., 2024. A Multiscale Incremental Learning Network for Remote Sensing Scene Classification. IEEE Transactions on Geoscience and Remote Sensing.

[14] Ahmed, I., Ahmad, M., Chehri, A., Hassan, M.M. and Jeon, G., 2022. IoT Enabled Deep Learning Based Framework for Multiple Object Detection in Remote Sensing Images. Remote Sensing, 14(16), p.4107.

[15] Zhang, H., Liu, W., Niu, H., Yin, P., Dong, S., Wu, J., Li, E., Zhang, L. and Zhu, C., 2023. Land Cover Change Detection Based on Vector Polygons and Deep Learning With High Resolution Remote Sensing Images. IEEE Transactions on Geoscience and Remote Sensing.

[16] Aljebreen, M., Mengash, H.A., Alamgeer, M., Alotaibi, S.S., Salama, A.S. and Hamza, M.A., 2024. Land Use and Land Cover Classification using River Formation Dynamics Algorithm with Deep Learning on Remote Sensing Images. IEEE Access.

[17] Ma, J., Lin, W., Tang, X., Zhang, X., Liu, F. and Jiao, L., 2023. Multi-pretext-task prototypes guided dynamic contrastive learning network for few-shot remote sensing scene classification. IEEE Transactions on Geoscience and Remote Sensing.

[18] Seetharaman, R., Tharun, M. and Anandan, K., 2021, September. A novel approach in hybrid median filtering for denoising medical images. In IOP Conference Series. Materials Science and Engineering (Vol. 1187, No. 1). IOP Publishing.

[19] Stateczny, A., Praveena, H.D., Krishnappa, R.H., Chythanya, K.R. and Babysarojam, B.B., 2023. Optimized Deep Learning Model for Flood Detection Using Satellite Images. Remote Sensing, 15(20), p.5037.

[20] Qin, J., Yang, D. and Zhang, W., 2023. A Pork Price Prediction Model Based on a Combined Sparrow Search Algorithm and Classification and Regression Trees Model. Applied Sciences, 13(23), p.12697.

[21] Xie, Y., Zou, J., Li, Z., Gao, F. and Peng, C., 2020. A novel deep belief network and extreme learning machine based performance degradation prediction method for proton exchange membrane fuel cell. IEEE Access, 8, pp.176661-176675.

[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.