Research article

Enhancing land cover classification in remote sensing imagery using an optimal deep learning model

Abdelwahed Motwake[1,2,*], Aisha Hassan Abdalla Hashim[1], Marwa Obayya[3] and Majdy M. Eltahir[4]

[1] Department of Electrical and Computer Engineering, International Islamic University Malaysia 53100 Kuala Lumpur, Malaysia
[2] Department of Information Systems, College of Business Administration in Hawtat bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia
[3] Department of Biomedical Engineering, College of Engineering, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[4] Department of Information Systems, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia

* **Correspondence:** Email: am.ismaeil@psau.edu.sa.

Abstract: The land cover classification process, accomplished through Remote Sensing Imagery (RSI), exploits advanced Machine Learning (ML) approaches to classify different types of land cover within the geographical area, captured by the RS method. The model distinguishes various types of land cover under different classes, such as agricultural fields, water bodies, urban areas, forests, etc. based on the patterns present in these images. The application of Deep Learning (DL)-based land cover classification technique in RSI revolutionizes the accuracy and efficiency of land cover mapping. By leveraging the abilities of Deep Neural Networks (DNNs) namely, Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN), the technology can autonomously learn spatial and spectral features inherent to the RSI. The current study presents an Improved Sand Cat Swarm Optimization with Deep Learning-based Land Cover Classification (ISCSODL-LCC) approach on the RSIs. The main objective of the proposed method is to efficiently classify the dissimilar land cover types within the geographical area, pictured by remote sensing models. The ISCSODL-LCC technique utilizes advanced machine learning methods by employing the Squeeze-Excitation ResNet (SE-ResNet) model for feature extraction and the Stacked Gated Recurrent Unit (SGRU) mechanism for land cover classification. Since 'manual hyperparameter tuning' is an erroneous and laborious task, the

hyperparameter selection is accomplished with the help of the Reptile Search Algorithm (RSA). The simulation analysis was conducted upon the ISCSODL-LCC model using two benchmark datasets and the results established the superior performance of the proposed model. The simulation values infer better outcomes of the ISCSODL-LCC method over other techniques with the maximum accuracy values such as 97.92% and 99.14% under India Pines and Pavia University datasets, respectively.

Keywords: land cover classification; artificial intelligence; remote sensing images; computer vision; parameter tuning
Mathematics Subject Classification: 11Y40

1. Introduction

With the ever-growing advancements being made in Remote Sensing Imaging (RSI) techniques, the RSIs are often used to describe the rural and urban regions and detect the modifications occurring in other areas. Since the predominant part of the RSI technique is based on producing high-resolution images, it comprises a wide range of data. Thus, accurate analysis of RSIs is especially prominent [1]. In the application and analysis of RSI, the exact per-pixel classification remains a highly significant measure of Land Cover (LC) type detection [2]. The evolution occurring in the RSI technology enables the development to occur in terms of high-resolution and extended range of space and time [3]. With high-resolution images becoming a norm, the types of LC on the surface get continuously complex. On the other hand, different objects frequently intervene with one another to work on the land [4]. Forest land is covered by trees and orchards whereas different trees are misclassified as forest land [5]. Problems like ships and bridges introduce errors in ultimate segmentation outcomes. In other terms, manual feature extraction techniques cannot be relied upon to achieve accuracy and attain efficiency outcomes.

The conventional techniques employed in the land cover classification process are segregated into thresholding, Support Vector Machines (SVM), clustering and so on [6]. These traditional approaches for land type classification primarily focus on feature-based classification in which the features are mostly developed by human skills and with the help of Machine Learning (ML) or possibility techniques [7]. Apart from the distinct classical approaches, the Deep Learning (DL) technique gains knowledge on higher-level semantic features and represents a considerable interest among the LC community. Among the existing DL algorithms, the Deep Convolutional Neural Networks (DCNNs) are mostly used since it has a well-structured deep multi-layer framework. Therefore, different DCNNs are leveraged in the process of LC classification and achieved substantial developments [8]. In this scenario, the pixel-based DCNNs are developed to attain end-to-end LC classification outcomes. This approach depends on the abstract features, created from the highest layer of the DCNNs. However, spatial context data is misplaced owing to continuous down-sampling processes [9]. This contextual data (i.e., object correlation, spatial location, and object scale) highlights the features required for the classification while it suppresses an undesirable change needed for the accurate classification of the tasks, particularly under numerous scales and orientations of similar LC objects [10].

The current research article develops the Improved Sand Cat Swarm Optimization with Deep Learning based Land Cover Classification (ISCSODL-LCC) algorithm on the RSIs. The specific objective is to develop a model that can autonomously learn about spectral and spatial characteristics

present in the RSIs, thus leading to highly accurate and efficient classification results. The ISCSODL-LCC technique uses the squeeze-excitation ResNet (SE-ResNet) approach for the purpose of feature extraction. Besides, the ISCSO technology is employed for enhanced hyperparameter selection of the SE-ResNet approach. For land cover classification process, the ISCSODL-LCC system uses the Stacked Gated Recurrent Unit (SGRU) algorithm. Eventually, the Reptile Search Algorithm (RSA) is deployed for better hyperparameter selection of the SGRU technique that helps in accomplishing the enhanced performance. The performance of the ISCSODL-LCC model was validated through simulation using two benchmark databases.

2. Related works

Temenos et al. [11] presented an intelligible DL technology named SHAPs algorithm for Land Use and Land Cover (LULC) detection from the RS images. The study used a compact CNN method for the classification of the satellite images and further provided outcomes to a SHAP deep explainer for enhancing the classification performance. The authors [12] proposed a Multi-level LC Contextual (MLCC) method that could adaptably incorporate efficient global context with a local context for LCC. In this study, the MLCC approach had two modules, namely the Multi-level Context Integration Module (MCIM) and the DCNN-based LC Classification Network (DLCN). Moreover, the MCIM allows the adaptive integration of both local and global contexts, following the guidance of uncertainty maps in an effective manner. Ekim and Sertel [13] employed three various DNN-Ensemble (DNNE) techniques and compared it for LCLU classification. The DNNE technique enables the DNNs to be effective by making sure a variety of approaches are integrated.

In [14], the authors developed the Optimum Guidance-Whale Optimizer Algorithm (OG-WOA) for selecting the important features and mitigating the over-fitting issues. For this study, the input images were normalized and used upon AlexNet–ResNet50 algorithm for feature extraction. The proposed OG-WOA method was employed to extract the features so as to select the related features. Lastly, the chosen features were processed for classification by employing the Bi-LSTM approach. Luo and Ji [15] introduced a new two-phase Domain Adaptation algorithm for Cross-Spatio-Temporal classification, in the name of DACST technique, with unlabeled targets and labelled source data as inputs. Zhou et al. [16] used a recent transformer-based multi-modal DL technique for extraction and integrating of the image features in satellite images with textual features of Point-Of-Interest (POI) data.

The authors [17] suggested an Ensemble of DL-Based Multi-modal LC Classification (EDL-MMLCC) method. In this method, the DL algorithm VGG-19, MobileNet, and the Capsule Network (CapsNet) were exploited for the purpose of feature extraction. Next, the Hosted Cuckoo Optimization (HCO) algorithm was utilized to train the development process. Eventually, the SSA with a Regularized ELM (RELM) model was exploited for the purpose of classification. In the study conducted earlier [18], a Multi-Scale FCN (MSFCN) method was developed with a multiscale convolution kernel, Channel Attention Block (CAB), and Global Pooling Module (GPM) technique to use discriminative models in 2D satellite images. The MSFCN method can be extended to 3D i.e., 3D-CNN which is an efficient method to be utilized in every LC type's time-series interaction.

Tariq and Mumtaz [19] aimed at prediction and evaluation of the urban growth and effect on Land Surface Temperature (LST) of Lahore and LULC with cellular automata Markov chain (CA-Markov chain). The aim of the study conducted by Tariq et al. [20] was to predict and assess the urban growth

of Peshawar city and LULC with CA-Markov-Chain. Lastly, the overall accuracy and kappa coefficient values were used to validate the models and the accuracies of the LULC maps. The aim of the study conducted by Tariq et al. [21] was to assess the urban growth and its effect on the LST of Lahore, the second largest city in Pakistan. The development of integrated application of RS and GIS, combined with cellular automata–Markov models, provided new means of evaluating the changes that occur in LULC. Further, it has also empowered the projection of trajectories into the future. Tariq and Shu [22] aimed at evaluating the impact of urban growth on Faisalabad. The aim of this study was to predict the seasonal LST and LULC with CA-Markov-Chain. A CA-Markov-Chain was introduced for simulating long-term landscape changes at 10-year time steps from 2018 to 2048.

3. The proposed model

In the current study, the ISCSODL-LCC algorithm has been proposed for land cover classification from the RSIs. The major intention of the proposed ISCSODL-LCC method is to detect and classify various kinds of land cover present in the RSIs. To accomplish this objective, the ISCSODL-LCC method comprises of four major processes such as the SE-ResNet feature extractor, ISCSO-based hyperparameter tuning, SGRU-based classification, and the RSA-based parameter optimization. Figure 1 portrays the overall working flow of the proposed ISCSODL-LCC system.
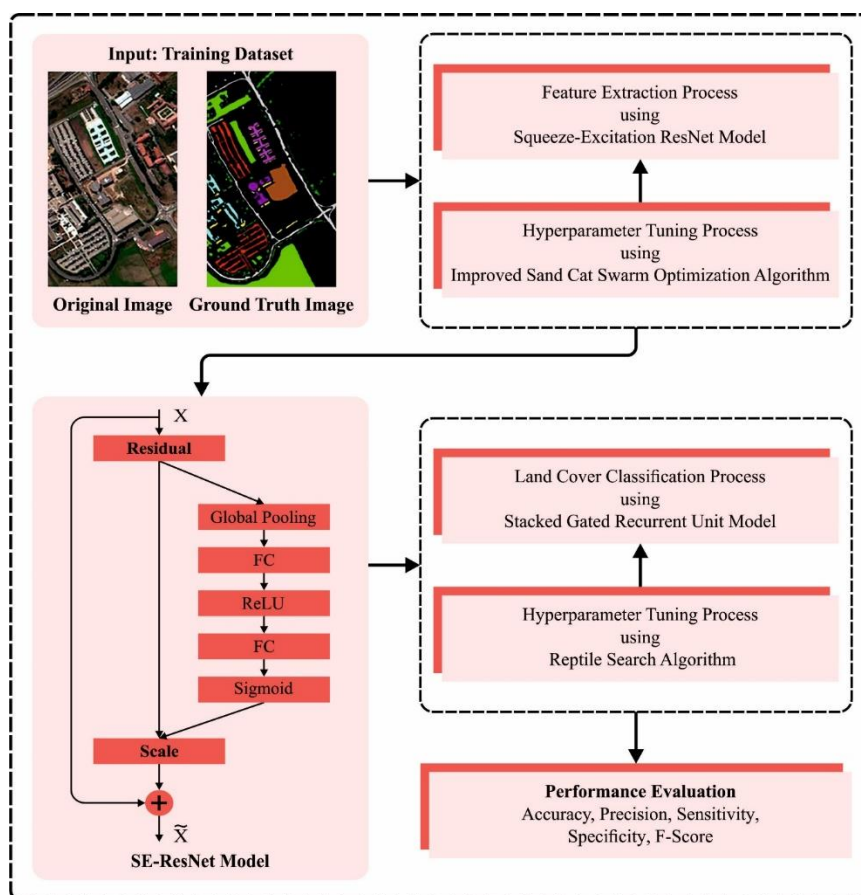

Figure 1. Overall flow of the ISCSODL-LCC algorithm.

3.1. Feature extraction: SE-ResNet model

In the current study, the SE-ResNet algorithm is used for feature extraction. Here, the ResNet model is added with a shortcut connection branch, outside the convolution layer, to carry out endless mapping and processing of the basic components of Residual Learning (RL). Then, the degradation problems of the network are solved as it is more challenging to train the model. Once the CNN model reaches the in-depth layers by successively loading the RL unit, it is possible to train the DCNN network [23]. The basic components of the RL either enhance the computing complexity or establish a new set of variables.

The principles of the ResNet model are given below.

$$x_l = h(x_l) + F(x_l, W_L), \tag{1}$$

$$x_{l+1} = f(y_1). \tag{2}$$

Here $f()$ shows the activation function and $h()$ represents the direct mapping.
The residual block is expressed as follows.

$$x_{l+1} = x_l + F(x_l, W_1). \tag{3}$$

The connection between the lth layer and the deeper L layer is shown below.

$$x_L = x_l + \sum_{i=1}^{L-1} F(x_i, W_i). \tag{4}$$

The loss function gradient ε, in terms of x_l, is defined according to the chain rules for derivatives, utilized in the BP.

$$\frac{\partial \varepsilon}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L} \left(1 + \frac{\partial}{\partial xl} \sum_{i=1}^{L-1} F(x_i, W_i) \right) = \frac{\partial \varepsilon}{\partial x_L} + \frac{\partial \varepsilon}{\partial x_L} \frac{\partial}{\partial x_l} \sum_{i=1}^{L-1} F(x_i, W_i). \tag{5}$$

In the trained model, $\frac{\partial}{\partial x_l} \sum_{i=1}^{L-1} F(x_i, W_i)$ cannot be -1 always. So, there is no existence of gradient vanishing problem from the ResNet method. $\frac{\partial \varepsilon}{\partial x_l}$ mentions that the gradients of the L layer, taken directly from the l layer, are less.

Feature extraction using CNN over a stacked convolution layer is a great dimension feature. Though some of the cases get lost, the residual blocks of the ResNet model tend to bounce the features extracted via the convolutional layer and combine the features before n layers containing convolution feature. After the n layer, either low-dimensional features or high-dimensional features can be retained, and so the efficacy of the network gets improved. Further, GAP is used to replace the FC layers from the standard CNN. GAP assists the correspondence between the classifications and mapping features on the FC layer that is better for the convolution structure. In addition to these, no parameter has been enhanced from the GAP that prevents the over-fitting issue. Furthermore, GAP is robust in nature in terms of spatial transformation of the input and integration of the spatial information.

The SE-ResNet model concentrates on the interdependency amongst the convolved feature networks by employing 1D convolutional layer. The SE blocks, obtained by the squeeze function,

summarize the overall data of the mapping features. The excitation function scales the significance of the feature map. In this work, the squeeze operation extracts the crucial data from all the channels whereas the excitation function evaluates the need amongst the networks with the help of the FC layer with a nonlinear function.

3.2. Hyperparameter Tuning: ISCSA Technique

In this phase, the ISCSA technique is employed for the purpose of hyperparameter tuning of the SE-ResNet model. The SCSO algorithm is inspired from the nature that the sand cats have low-frequency recognition capability and incredible hunting capability [24]. The searching method of SCSO is related to other SI optimization techniques. In this technique, the searching process occurs in two stages such as local and the global search stages. A proper benefit of this method is that there exist several parameter settings as well as an exact balanced conversion method between the global and the local searches.

Based on the size of the problem and the amount of self-defined population, an initialization population matrix is generated while the set expression of all the solutions is as follows; $X_i = (X_{i,1}, X_{i,2}, \dots X_{i,j})$. SCSO technique is simulated by the auditory capabilities of the sand cats at low frequencies, and its accurate formula is shown in Eq (6).

$$L_G = LSC - \left(\frac{2 \times LSC \times k}{2 \times k_{\max}}\right). \tag{6}$$

Here, LSC implies the auditory features and the stability to a range of 2. The variable k represents the present iteration number whereas k_{ax} stands for the maximal number of permitted iterations. In the primary search for the optimum method, a global search model is required to be applied in a large-scale search so that the estimated range of solutions is controlled. The particular implementation formula is provided by Eq (7).

$$X_{i(k+1)} = r \times \left(X_{bs(k)} - rand \times X_{i(k)}\right), \tag{7}$$

$$r = LG \times rand. \tag{8}$$

Here, $X_{bs(k)}$ represents the optimum performance at k^{th} loop during the kth iteration, $X_{i(k)}$ corresponds to a symbol for the present solution, while $rand$ represents an arbitrarily created number in the range of $(0,1)$. A local search procedure is implemented as follows.

$$X_{i(k+1)} = X_{bs(k+1)} - X_{md} \times \cos(\theta) \times r. \tag{9}$$

Here, X_{rnd} defines the arbitrary solution. The SCSO technique differs from the entire exploration step to a partial exploration step, controlled by the parameter, R. If the magnitude of R exceeds 1, then the global search step is executed and if not, the local search step is accomplished.

$$R = 2 \times LG \times rand - LG. \tag{10}$$

The SCSO technique cannot define a better optimum solution, if the convergence accuracy is low. This is attributed to the fact that the scope of the algorithm is confined to the global search step, which makes it unable to avoid the local optimum results. So, the current work presents three

enhancement approaches to increase the widespread outcomes of the SCSO technique.

In the ISCSO algorithm, a Tent map-based chaotic approach has been established to extend the searching space in the primary stage of this model. This approach has a main purpose to serve i.e., to avoid the probability of absent possible solutions.

$$X_{n+1} = f(x_n) = \begin{cases} x_n/a, x_n \in [0, a) \\ (1 - x_n)/(1 - a), x_n \in [a, 1] \end{cases}. \tag{11}$$

The function is nothing but a chaotic mapping in the parameter value and takes an even distribution function and optimum link. When the range is a, this method lies in a chaotic state or if a is 0.5, then it lies in the shorter duration state.

3.3. Image classification: Optimal SGRU model

In this work, the SGRU approach is applied for an effectual classification and land cover detection. Similar to LSTM, the GRU network intends to prevent the gradient exploding and vanishing problems that occur in RNN [25]. In most of the cases, the GRU exhibits a comparable performance with that of the LSTM method. However, the GRU performs considerably faster due to its computational simplicity. The LSTM model controls the data stream in the Hidden Layer (HL) through output and forget gates. On the other hand, the GRU model tracks the state of the sequence without using a single memory cell. The GRU model comprises of reset and update gates. The input and forget gates of the LSTM method are merged into the update gate z_t, which decides the amount of HL that gets updated. The reset gate r_t decides the amount of data that should be passed from the prior HL over the existing HL. Both reset and update gates control the manner in which the data is updated. The GRU model does not have a cell layer memory whereas it can share the entire network state at t timestep. The GRU model exploits the HL to transmit the data. The GRU is evaluated using the subsequent equations.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z), \tag{12}$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r), \tag{13}$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h(h_{t-1} \odot r_t)W_h), \tag{14}$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t. \tag{15}$$

Here, the update gate works similar to the f_t forget gate from the LSTM method, where the update gate z_t decides the amount of information to be retained in the HL outcome h_{t-1} of the preceding memory unit. Both $\tilde{h}_t$ and h_t correspond to the candidate and HL at timestep, t. σ and $\odot$ indicate the logistic sigmoid functions and component-wise multiplication, correspondingly. W and U show the weight matrices to be learnt. Even though the GRU and LSTM models attain comparable outcomes, both the models obtain better results with various tasks. The GRU model is a great selection for small databases while the LSTM technique can frequently be utilized with massive quantities of information. SGRU is an architecture of GRU in which multiple GRU layers are stacked one above the other to make a DNN structure for analyzing the sequential data. All the layers receive input from the prior layer and the output is generated and passed on to the following layer. Stacking multiple layers can assist the network in learning abstract and complex representations of the input dataset, which might result in better efficiency on specific tasks. Figure 2 displays the infrastructure
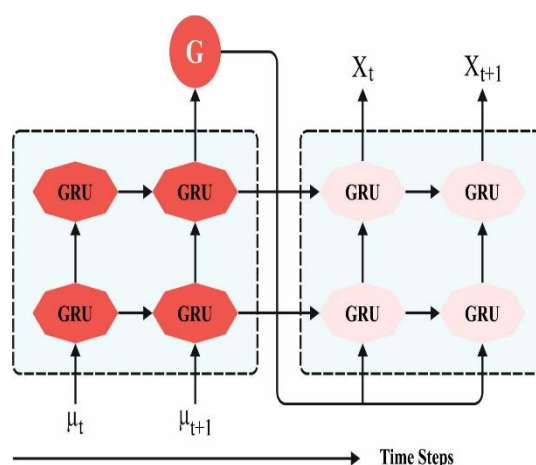
of SGRU.

Figure 2. Structure of SGRU.

At last, the RSA adjusts the hyperparameter values of the GRU method. In RSA, the optimizer model begins with a group of solution candidates. During all the iterations, the optimum performance attained is regarded as a nearby optimum value [26]. Particularly, X denotes the arbitrarily created group of solution candidates, as written in Eq (16).

$$X_{i,j} = rand \cdot (UB - LB) + LB, i = 1,2,3, \dots, Nj = 1,2,3,4, \dots, n. \tag{16}$$

In this formula, $X_{i,j}$ defines the location of the ith crocodile individual from jth dimension, N represents the number of candidate solutions, n implies the size of the provided problems, and goes to an arbitrary function in the range of 0 and 1; UB and LB signify the upper and lower bounds of the problems.

If $t \leq T/4$, then the crocodile population implements a high-altitude walking approach. If $t \leq T/2$, this method is at the primary phase of the iteration. But the crocodile population searches worldwide and arrives at the bounding stage. If $T/4 < t \leq T/2$, the crocodile population executes a belly walking method. During the encirclement exploration stage, the location upgrade formula for the crocodile population is depicted in Eq (17).

$$x(i,j)(t+1) = \begin{cases} Best_j(t) - \eta_{(i,j)}(t) \cdot \beta - R_{(i,j)}(t) \cdot rand, t \leq \dfrac{T}{4} \\ Best_j \cdot \mathrm{x}_{(r1,j)} \cdot ES(t) \cdot rand, \quad \dfrac{T}{4} < t \leq \dfrac{T}{2} \end{cases}. \tag{17}$$

Here, $Best_j(t)$ stands for the position of the optimum result at the present moment, t implies the number of present iterations, T represents the maximal iteration counts, and $\eta_{(i,j)}(t)$ signifies the hunting actions of the ith candidate outcome from j dimensional operator; the computation is depicted in Eq. (18). β refers to the sensitive parameter to adjust the exploration accuracy of the encirclement phase during the iterative procedure, and is set at $(0,1)$. $R_{(i,j)}(t)$ indicates the reduction function employed for reducing the searching region value and it is computed by employing Eq. (19). $r1$ refers to the arbitrary integer between 1 and N; $x_{(r1,j)}$ represents the jth dimension position of the $r1$ arbitrary candidate outcome. N defines the number of candidate solutions and the evolution factor

$ES(t)$ is a probability ratio. In the complete iteration method, the value is arbitrarily reduced between $[2, -2]$, and it is computed by employing Eq (20).

$$\eta_{(i,j)} = Best_j(t) \cdot P_{(i,j)}, \tag{18}$$

$$R_{(i,j)} = \frac{Best_j(t) - x_{(r2,j)}}{Best_j(t) + \epsilon}, \tag{19}$$

$$ES(t) = 2r_3 \left(1 - \frac{t}{T}\right). \tag{20}$$

In this formula, e refers to a small positive number, $r2$ denotes an arbitrary integer of 1 and N; r_3 demonstrates the random integer between 1 and 1; and $P_{(i,j)}$ denotes the percentage dissimilarity among better outcomes and j denotes the dimensional location of the present solution and is computed as depicted in Eq (21).

$$P_{(i,j)} = \alpha + \frac{x_{(i,j)} - M_{(x_i)}}{Best_j(t) \cdot \left(UB_{(j)} - LB_{(j)}\right) + \epsilon}. \tag{21}$$

$M_{(x_i)}$ denotes the average location of the ith candidate's performance and its computation is depicted in Eq (12). $UB_{(j)}$ and $LB_{(j)}$ denote the upper and lower bounds of the jth dimension's position, α refers to the sensitive parameter utilized for adjusting the searching accuracy of the hunting co-operation in iteration procedure (difference among candidate performances) that is set at 0.1.

$$M_{(x_i)} = \frac{1}{n} \sum_{j=1}^{n} x_{(i,j)}. \tag{22}$$

If $T/2 < t$, then the population enters into a final phase of iterations and later, it enters the hunting phase. In this method, if $T/2 < t \leq 3T/4$, then the crocodiles carry out hunting co-ordination. If $3T/4 < t \leq T$, then the crocodiles hunt in a cooperative manner. The relevant formula is represented in Eq (23).

$$x(i,j)(t + 1) = \begin{cases} Best_j(t) \cdot P_{(i,j)}(t) \cdot rand, & \frac{T}{2} < t \leq \frac{3T}{4} \\ Best_j(t) - \eta_{(i,j)}(t) \cdot \epsilon - R_{i,j}(t) \cdot rand, & \frac{3T}{4} < t \leq T \end{cases}. \tag{23}$$

Fitness selection is a significant part of the RSA method. The encoder solution is useful for evaluating the goodness of the candidate solution. Here, accuracy value is a primary condition used to develop the FF.

$$Fitness = \max(P), \tag{24}$$

$$P = \frac{TP}{TP + FP}, \tag{25}$$

where TP and FP represent the true and false positive values respectively.

4. Performance validation

The proposed model was simulated using the Python 3.8.5 tool on a PC configured with the specifications as given herewith; i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU.

4.1. Dataset details

In the current study, the performance of the proposed ISCSODL-LCC model was validated through simulation on the Indian Pines (IP) database and Pavia University (PU) database, available at the URL, https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes. The IP database has 10,349 samples in total under 16 classes, as represented in Table 1.

Table 1. Description of the IP database.

Indian Pines Database		
Class	**Labels**	**No. of Instances**
Alfalfa	C1	46
Corn-Notill	C2	1428
Corn-Mintill	C3	830
Corn	C4	237
Grass-Pasture	C5	483
Grass-Trees	C6	730
Grass-Pasture-Mowed	C7	28
Hay-Windrowed	C8	478
Oats	C9	20
Soybean-Notill	C10	972
Soybean-Mintill	C11	2455
Soybean-Clean	C12	693
Wheat	C13	205
Woods	C14	1265
Buildings-Grass-Trees-Drives	C15	386
Stone-Steel-Towers	C16	93
Total No. of Instances		**10349**

This IP dataset is a hyperspectral dataset, captured by the AVIRIS sensor, over the Indian Pines test site at Northwestern Indiana, USA. It includes 224 spectral bands with spectral resolutions in the range of 0.2 to 2.4 micrometers. The spatial resolution of the images is around 20 meters, which provides a comprehensive data about the land cover within the scene. The dataset covers the images captured from an agricultural region and comprises of different classes namely forests, corn, and soybeans. With a size of 145 pixels by 145 lines, the IP dataset acts as a standard benchmark for assessing the algorithms in hyperspectral image analysis, classification, and feature extraction processes. The PU dataset is a hyperspectral dataset gathered over Pavia, Italy, using the ROSIS sensor. It includes a total of 103 spectral bands with a spectral range of 0.43 to 0.86 micrometers.

The spatial resolution of the images is about 1.3 meters, which offers high-detail information that is suitable for urban land cover classification. The dataset covers the images captured from an urban region and contains classes including asphalt, buildings, and trees. With an image size of 610 pixels by 340 lines, the PU dataset is often applied in research works conducted in the domain of urban remote sensing, along with the applications in environmental monitoring, land cover mapping, and change detection.

Table 2 provides a detailed description about the PU database that contains 42,776 samples under nine class labels.

Table 2. Details of the PU database.

Class	Labels	No. of Instances
Asphalt	C1	6631
Meadows	C2	18649
Gravel	C3	2099
Trees	C4	3064
Painted metal sheets	C5	1345
Bare Soil	C6	5029
Binunen	C7	1330
Self-Blocking	C8	3682
Shadows	C9	947
Total No. of Instances		42776

4.2. Results and discussion

Figure 3 represents the classification analysis outcomes achieved by the ISCSODL-LCC algorithm on the IP database. Figure 3a–b shows the confusion matrices generated by the ISCSODL-LCC technology at 70:30 of the TR set/TS set. The outcome values depict that the ISCSODL-LCC system predicted and categorized all the 16 classes precisely. Then, Figure 3c displays the PR curve of the ISCSODL-LCC methodology. The outcome implies that the ISCSODL-LCC method accomplished improved PR outcomes in all the 16 classes. Figure 3d depicts the ROC outcomes of the ISCSODL-LCC algorithm. The simulation value exhibits that the ISCSODL-LCC algorithm attained the maximum solution with increased ROC values on all the 16 classes.

In Table 3, the overall classification results attained by the ISCSODL-LCC model upon the IP database is shown. The simulation values signify the improved outcomes of the ISCSODL-LCC method. On 70% TR set, the ISCSODL-LCC technique reached the average $accu_y$, $prec_n$, $sens_y$, $spec_y$, and F_{score} values such as 97.92%, 66.36%, 56.80%, 98.82%, and 58.30% correspondingly. Besides, on 30% TS set, the ISCSODL-LCC system attained the average $accu_y$, $prec_n$, $sens_y$, $spec_y$ and F_{score} values such as 97.92%, 66.36%, 56.80%, 98.82%, and 58.30% correspondingly.
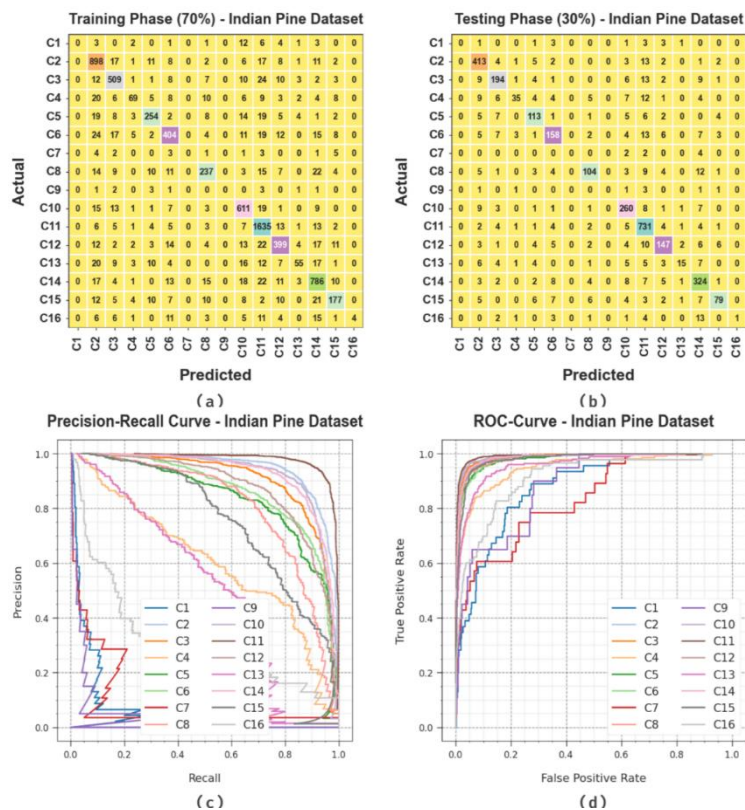
Figure 3. Performances on IP database. (a and b) Confusion matrices, (c) PR_curve, and (d) ROC.

Table 3. Classification outcomes of the ISCSODL-LCC algorithm on the IP database.

Class Labels	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	F_{Score}
TR set (70%)					
C1	99.54	00.00	00.00	100.00	00.00
C2	96.29	82.92	91.45	97.05	86.97
C3	97.43	82.90	86.27	98.42	84.55
C4	98.54	73.40	46.00	99.65	56.56
C5	98.00	80.89	74.93	99.13	77.79
C6	96.96	79.68	77.54	98.47	78.60
C7	99.72	00.00	00.00	100.00	00.00
C8	97.71	76.95	71.39	98.97	74.06
C9	99.83	00.00	00.00	100.00	00.00
C10	97.25	82.46	89.85	98.02	86.00
C11	96.37	88.96	96.46	96.34	92.56
C12	97.25	80.77	79.32	98.59	80.04
C13	98.36	73.33	35.71	99.72	48.03
C14	96.33	83.80	87.33	97.60	85.53
C15	97.98	75.64	66.54	99.18	70.80
C16	99.13	100.00	05.97	100.00	11.27
Average	**97.92**	**66.36**	**56.80**	**98.82**	**58.30**

Continued on next page

Class Labels	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	F_{Score}
TS set (30%)					
C1	99.58	00.00	00.00	100.00	00.00
C2	96.91	86.76	92.60	97.63	89.59
C3	97.20	82.55	80.83	98.57	81.68
C4	98.04	79.55	40.23	99.70	53.44
C5	97.84	75.84	78.47	98.78	77.13
C6	96.97	78.61	75.60	98.52	77.07
C7	99.74	00.00	00.00	100.00	00.00
C8	97.84	80.62	71.23	99.16	75.64
C9	99.74	00.00	00.00	100.00	00.00
C10	97.07	81.50	89.04	97.90	85.11
C11	95.49	86.82	96.18	95.27	91.26
C12	97.49	80.77	77.37	98.80	79.03
C13	98.62	68.18	29.41	99.77	41.10
C14	96.04	79.80	88.77	97.01	84.05
C15	98.04	79.80	65.83	99.33	72.15
C16	99.19	100.00	03.85	100.00	07.41
Average	**97.86**	**66.30**	**55.59**	**98.78**	**57.17**

Figure 4 demonstrates the training accuracy TR_accu_y and VL_accu_y values attained by the ISCSODL-LCC system on the IP database. TL_accu_y is determined by evaluating the ISCSODL-LCC approach on the TR database whereas VL_accu_y is calculated by assessing the model's performance on a distinct testing database. The outcomes display that both TR_accu_y and VL_accu_y values improved with an increase in the number of epochs. Thus, the results of the ISCSODL-LCC model got improved in both TR and TS databases with an increase in the number of epochs.
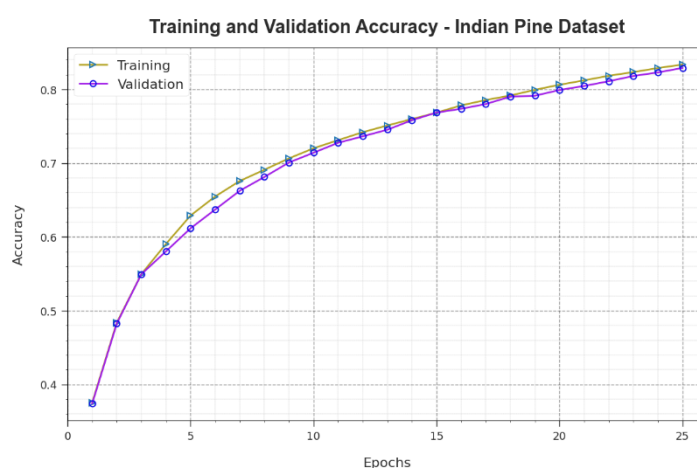

Figure 4. $Accu_y$ curve of the ISCSODL-LCC algorithm on the IP database.

In Figure 5, TR_loss and VR_loss outcomes of the ISCSODL-LCC technique on the IP database are portrayed. TR_loss defines the error between the predictive results and unique values on the TR data. VR_loss represents the extent solution of the ISCSODL-LCC approach on distinct

validation data. The results designate that both TR_loss and VR_loss values tend to reduce with increasing number of epochs. This outcome of the ISCSODL-LCC system shows its prowess in making precise classifications. The reduced TR_loss and VR_loss values establish the superior performance of the ISCSODL-LCC model in taking the patterns and relationships.
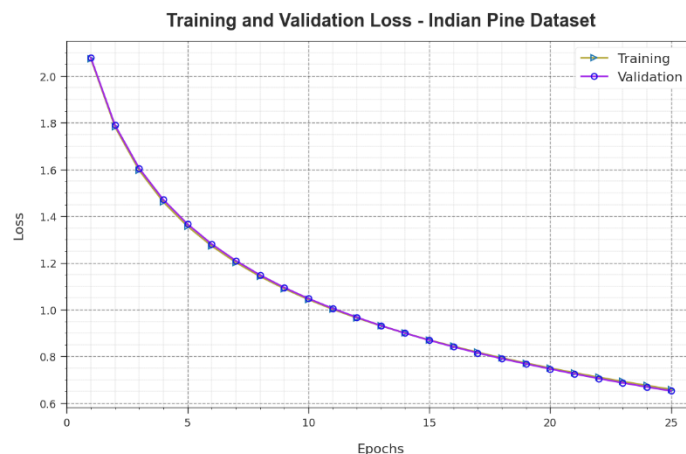

Figure 5. Loss curve of the ISCSODL-LCC algorithm on IP database.

In Table 4, the comparative analysis results attained by the ISCSODL-LCC model and other models upon the IP database are shown [27]. The outcomes imply that the SVMT algorithm achieved poor performance. Furthermore, the SVMEPFT, CoSVMT, and the CoSVMEPFT methods obtained slightly improved results. Moreover, the GFSVMT and the GFSVMEPFT models accomplished a considerable performance. However, the ISCSODL-LCC technique achieved the maximum performance with an $accu_y$ of 97.92%.

Table 4. $Accu_y$ outcomes of the ISCSODL-LCC system and other approaches on the IP database.

Indian Pine Database	
Methods	**Accuracy**
SVMT	81.06
SVMEPFT	89.65
CoSVMT	90.78
CoSVMEPFT	90.98
GFSVMT	94.91
GFSVMEPFT	95.37
ISCSODL-LCC	97.92

Figure 6 shows the classification performance of the ISCSODL-LCC algorithm on the PU database. Figure 6a and b describe the confusion matrices generated by the ISCSODL-LCC model on 70:30 of the TR set/TS set. The simulation values imply that the ISCSODL-LCC technology classified as well as detected all the nine classes correctly. Figure 6c establishes the PR outcomes of the ISCSODL-LCC system. The outcomes define that the ISCSODL-LCC system gained better PR outcomes in all the nine classes. Figure 6d depicts the ROC outcomes of the ISCSODL-LCC algorithm. The outcomes establish that the proposed ISCSODL-LCC method achieved excellent

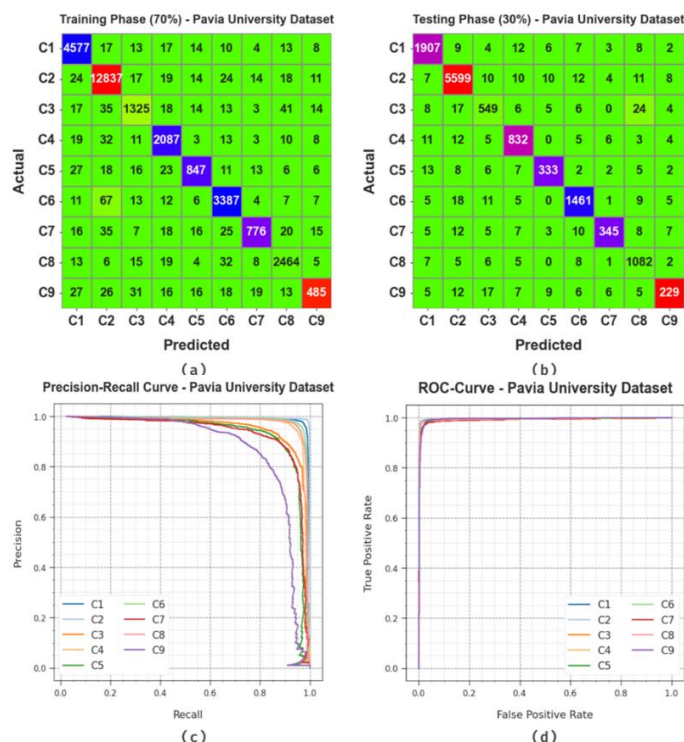performance with greater ROC values on all the nine classes.

Figure 6. Performances on the PU database. (a and b) Confusion matrices, (c) PR_curve, and (d) ROC.

In Table 5, the overall classification outcomes of the ISCSODL-LCC approach upon the PU database are shown. The simulation values infer the improved performances of the ISCSODL-LCC system. On 70% TR set, the ISCSODL-LCC methodology gained the average $accu_y$, $prec_n$, $sens_y$, $spec_y$, and F_{score} values such as 99.14%, 93.38%, 91.11%, 99.47%, and 92.16% correspondingly. Afterwards, on 30% TS set, the ISCSODL-LCC method achieved average $accu_y$, $prec_n$, $sens_y$, $spec_y$, and F_{score} values such as 99.14%, 93.33%, 91.58%, 99.48%, and 92.40% respectively.

Table 5. Classification outcomes of the ISCSODL-LCC algorithm on the PU database.

Class Labels	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	F_{Score}
TR set (70%)					
C1	99.17	96.74	97.95	99.39	97.34
C2	98.74	98.19	98.91	98.61	98.55
C3	99.07	91.51	89.53	99.57	90.51
C4	99.20	93.63	95.47	99.49	94.54
C5	99.31	90.69	87.59	99.70	89.11
C6	99.09	95.87	96.39	99.45	96.13
C7	99.27	91.94	83.62	99.77	87.58
C8	99.23	95.06	96.02	99.53	95.54
C9	99.20	86.76	74.50	99.75	80.17
Average	**99.14**	**93.38**	**91.11**	**99.47**	**92.16**

Continued on next page

Class Labels	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	F_{Score}
TS set (30%)					
C1	99.13	96.90	97.40	99.44	97.15
C2	98.71	98.37	98.73	98.70	98.55
C3	98.96	89.56	88.69	99.48	89.12
C4	99.18	93.38	94.76	99.51	94.06
C5	99.39	90.98	88.10	99.74	89.52
C6	99.14	96.31	96.44	99.51	96.37
C7	99.38	93.75	85.82	99.81	89.61
C8	99.17	93.68	96.95	99.38	95.29
C9	99.21	87.07	77.36	99.73	81.93
Average	**99.14**	**93.33**	**91.58**	**99.48**	**92.40**

Figure 7 illustrates the training accuracy TR_accu_y and VL_accu_y values attained by the ISCSODL-LCC approach on the PU database. TL_accu_y is defined by evaluating the ISCSODL-LCC system on the TR database whereas VL_accu_y is computed by estimating the outcomes of the model on a separate testing database. The results display that both TR_accu_y and VL_accu_y values increase with an increase in the number of epochs. So, the ISCSODL-LCC model has been proved to achieve increased performance on the TR and TS database with an increase in the sum of epochs.
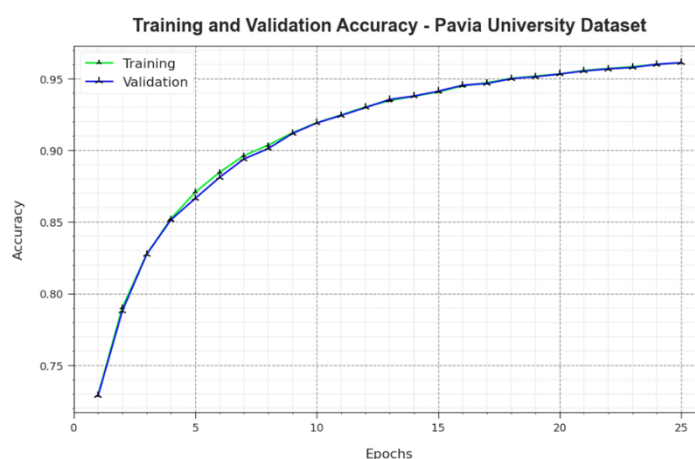

Figure 7. $Accu_y$ curve of the ISCSODL-LCC system on the PU database.

In Figure 8, the TR_loss and VR_loss curves of the ISCSODL-LCC system on the PU database are portrayed. The TR_loss demonstrates the error between the forecasted outcomes and the original TR data values. The VR_loss defines the extent of performance of the ISCSODL-LCC system on individual validation data. The performances show that both TR_loss and VR_loss values tend to be lesser with an increase in the number of epochs. The outcomes revealed the enhanced performance of the ISCSODL-LCC approach and its aptitude upon correct classification. The minimal TR_loss and VR_loss values demonstrate the better outcomes of the ISCSODL-LCC model on arresting the relationships and patterns.
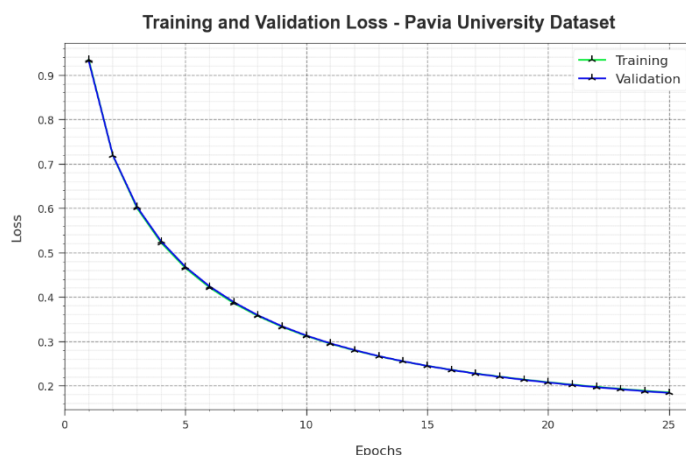
Figure 8. Loss curve of the ISCSODL-LCC algorithm on the PU database.

In Table 6, the comparison analysis outcomes of the ISCSODL-LCC system and other models on the PU database are shown. The outcomes infer that the SVMT method was the worst performer. Afterwards, the SVMEPFT, CoSVMT, and the CoSVMEPFT approaches achieved somewhat improved outcomes. Moreover, the GFSVMT and the GFSVMEPFT algorithms too achieved considerable results. However, the ISCSODL-LCC system outperformed all other models with a maximum $accu_y$ of 99.14%.

Table 6. $Accu_y$ outcomes of the ISCSODL-LCC system and other approaches on the PU database [27].

Pavia University Database	
Methods	**Accuracy**
SVMT	94.26
SVMEPFT	95.57
CoSVMT	96.47
CoSVMEPFT	96.73
GFSVMT	96.56
GFSVMEPFT	97.33
ISCSODL-LCC	99.14

These simulation values confirmed the improved outcome of the ISCSODL-LCC system over other techniques.

5. Conclusions

In the current study, the automated ISCSODL-LCC algorithm has been proposed for land cover classification on the RSIs. The main purpose of the ISCSODL-LCC method is to classify and detect various kinds of land cover present in the RSIs. To achieve this objective, the ISCSODL-LCC technology comprises four major processes such as the SE-ResNet feature extraction, ISCSO-based hyperparameter tuning, SGRU-based classification, and the RSA-based parameter optimization. In this work, the ISCSO method is employed for optimum hyperparameter selection of the SE-ResNet system. For land cover classification, the ISCSODL-LCC algorithm uses RSA for optimal

hyperparameter selection of the SGRU algorithm that supports in accomplishing a better performance. The proposed ISCSODL-LCC methodology was validated for its performance through simulation using two benchmark databases. The simulation values concluded the superior results of the ISCSODL-LCC model over other techniques with the maximum accuracy values such as 97.92% and 99.14% under the IP and PU datasets, respectively. In the future, the ISCSODL-LCC method should be validated upon real-time applications, which explore to different geographical areas. Further, spectral and temporal data sources are to be integrated to enhance the classification performance. Moreover, the incorporation of the emerging remote sensing technology and continual refinement of the hyperparameter optimization strategy can improve the model's versatility and performance.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. T. Kwak, Y. Kim, Semi-supervised land cover classification of remote sensing imagery using CycleGAN and EfficientNet, *KSCE J. Civ. Eng.*, **27** (2023), 1760–1773. https://doi.org/10.1007/s12205-023-2285-0
2. T. He, S. Wang, Multi-spectral remote sensing land-cover classification based on deep learning methods, *J. Supercomput.*, **77** (2021), 2829–2843. https://doi.org/10.1007/s11227-020-03377-w
3. L. Wang, J. Wang, Z. Liu, J. Zhu, F. Qin, Evaluation of a deep-learning model for multispectral remote sensing of land use and crop classification, *Crop. J.*, **10** (2022), 1435–1451. https://doi.org/10.1016/j.cj.2022.01.009
4. A. Tzepkenlis, K. Marthoglou, N. Grammalidis, Efficient deep semantic segmentation for land cover classification using sentinel imagery, *Remote Sens.*, **15** (2023), 2027. https://doi.org/10.3390/rs15082027

5. Y. Li, Y. Zhou, Y. Zhang, L. Zhong, J. Wang, J. Chen, DKDFN: Domain knowledge-guided deep collaborative fusion network for multimodal unitemporal remote sensing land cover classification, *ISPRS J. Photogramm. Remote Sens.*, **186** (2022), 170–189. https://doi.org/10.1016/j.isprsjprs.2022.02.013

6. L. Bergamasco, F. Bovolo, L. Bruzzone, A dual-branch deep learning architecture for multisensor and multitemporal remote sensing semantic segmentation, *IEEE J. STARS*, **16** (2023), 2147–2162. https://doi.org/10.1109/JSTARS.2023.3243396

7. X. Yuan, Z. Chen, N. Chen, J. Gong, Land cover classification based on the PSPNet and superpixel segmentation methods with high spatial resolution multispectral remote sensing imagery, *J. Appl. Remote Sens.*, **15** (2021), 034511. https://doi.org/10.1117/1.JRS.15.034511

8. J. Yan, J. Liu, L. Wang, D. Liang, Q. Cao, W. Zhang, et al., Land-cover classification with time-series remote sensing images by complete extraction of multiscale timing dependence, *IEEE J. STARS*, **15** (2022), 1953–1967. https://doi.org/10.1109/JSTARS.2022.3150430

9. J. Kim, Y. Song, W. K. Lee, Accuracy analysis of multi-series phenological landcover classification using U-Net-based deep learning model–Focusing on the Seoul, Republic of Korea–, *Korean J. Remote Sens.*, **37** (2021), 409–418. https://doi.org/10.7780/kjrs.2020.37.3.4

10. V. Yaloveha, A. Podorozhniak, H. Kuchuk, Convolutional neural network hyperparameter optimization applied to land cover classification, *Radioelectron. Comput. Syst.*, 2022, 115–128. https://doi.org/10.32620/reks.2022.1.09

11. A. Temenos, N. Temenos, M. Kaselimi, A. Doulamis, N. Doulamis, Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP, *IEEE Geosci. Remote Sens. Lett.*, **20** (2023), 8500105. https://doi.org/10.1109/LGRS.2023.3251652

12. X. Cheng, X. He, M. Qiao, P. Li, S. Hu, P. Chang, et al., Enhanced contextual representation with deep neural networks for land cover classification based on remote sensing images, *Int. J. Appl. Earth Obse.*, **107** (2022), 102706. https://doi.org/10.1016/j.jag.2022.102706

13. B. Ekim, E. Sertel, Deep neural network ensembles for remote sensing land cover and land use classification, *Int. J. Digit. Earth*, **14** (2021), 1868–1881. https://doi.org/10.1080/17538947.2021.1980125

14. V. N. Vinaykumar, J. A. Babu, J. Frnda, Optimal guidance whale optimization algorithm and hybrid deep learning networks for land use land cover classification, *EURASIP J. Adv. Signal Process.*, **2023** (2023), 13. https://doi.org/10.1186/s13634-023-00980-w

15. M. Luo, S. Ji, Cross-spatiotemporal land-cover classification from VHR remote sensing images with deep learning based domain adaptation, *ISPRS J. Photogramm. Remote Sens.*, **191** (2022), 105–128. https://doi.org/10.1016/j.isprsjprs.2022.07.011

16. W. Zhou, C. Persello, A. Stein, Building usage classification using a transformer-based multimodal deep learning method, *2023 Joint Urban Remote Sensing Event (JURSE)*, 2023. https://doi.org/10.1109/JURSE57346.2023.10144168

17. G. P. Joshi, F. Alenezi, G. Thirumoorthy, A. K. Dutta, J. You, Ensemble of deep learning-based multimodal remote sensing image classification model on unmanned aerial vehicle networks, *Mathematics*, **9** (2021), 2984. https://doi.org/10.3390/math9222984

18. R. Li, S. Zheng, C. Duan, L. Wang, C. Zhang, Land cover classification from remote sensing images based on multi-scale fully convolutional network, *Geo-Spat. Inf Sci.*, **25** (2022), 278–294. https://doi.org/10.1080/10095020.2021.2017237

19. A. Tariq, F. Mumtaz, Modeling spatio-temporal assessment of land use land cover of Lahore and its impact on land surface temperature using multi-spectral remote sensing data, *Environ. Sci. Pollut. Res.*, **30** (2022), 23908–23924. https://doi.org/10.1007/s11356-022-23928-3

20. A. Tariq, J. Yan, F. Mumtaz, Land change modeler and CA-Markov chain analysis for land use land cover change using satellite data of Peshawar, Pakistan, *Phys. Chem. Earth Parts A/B/C*, **128** (2022), 103286.

21. A. Tariq, F. Mumtaz, M. Majeed, X. Zeng, Spatio-temporal assessment of land use land cover based on trajectories and cellular automata Markov modelling and its impact on land surface temperature of Lahore district Pakistan, *Environ. Monit. Assess.*, **195** (2023), 114. https://doi.org/10.1007/s10661-022-10738-w

22. A. Tariq, H. Shu, CA-Markov chain analysis of seasonal land surface temperature and land use land cover change using optical multi-temporal satellite data of Faisalabad, Pakistan, *Remote Sens.*, **12** (2020), 3402. https://doi.org/10.3390/rs12203402

23. T. Chen, H. Qin, X. Li, W. Wan, W. Yan, A Non-Intrusive Load Monitoring Method Based on Feature Fusion and SE-ResNet, *Electronics*, **12** (2023), 1909. https://doi.org/10.3390/electronics12081909

24. H. Long, Y. He, Y. Xu, C. You, D. Zeng, H. Lu, Optimal allocation research of distribution network with DGs and SCs by improved sand cat swarm optimization algorithm, *IAENG Int. J. Comput. Sci.*, 2023.

25. A. Al Hamoud, A. Hoenig, K. Roy, Sentence subjectivity analysis of a political and ideological debate dataset using LSTM and BiLSTM with attention and GRU models, *J. King Saud Univ.-Com.*, **34** (2022), 7974–7987. https://doi.org/10.1016/j.jksuci.2022.07.014

26. L. Kong, H. Liang, G. Liu, S. Liu, Research on wind turbine fault detection based on the fusion of ASL-CatBoost and TtRSA, *Sensors*, **23** (2023), 6741. https://doi.org/10.3390/s23156741

27. S. Rajalakshmi, S. Nalini, A. Alkhayyat, R. Q. Malik, Hyperspectral remote sensing image classification using improved metaheuristic with deep learning, *Comput. Syst. Sci. Eng.*, **46** (2023), 1673–1688. https://doi.org/10.32604/csse.2023.034414