

Review Article

Remote Sensing Image Classification: A Comprehensive Review and Applications

Maryam Mehmood,^{1,2} Ahsan Shahzad,² Bushra Zafar ,³ Amsa Shabbir,¹ and Nouman Ali ¹

¹Department of Software Engineering, Mirpur University of Science and Technology (MUST), Mirpur AJK 10250, Pakistan

²Department of Computer and Software Engineering, National University of Sciences and Technology, Islamabad, Pakistan

³Department of Computer Science, Government College University, Faisalabad 38000, Pakistan

Correspondence should be addressed to Nouman Ali; nouman.ali@live.com

Received 31 March 2022; Accepted 20 June 2022; Published 2 August 2022

Academic Editor: Afaq Ahmad

Copyright © 2022 Maryam Mehmood et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Remote sensing is mainly used to investigate sites of dams, bridges, and pipelines to locate construction materials and provide detailed geographic information. In remote sensing image analysis, the images captured through satellite and drones are used to observe surface of the Earth. The main aim of any image classification-based system is to assign semantic labels to captured images, and consequently, using these labels, images can be arranged in a semantic order. The semantic arrangement of images is used in various domains of digital image processing and computer vision such as remote sensing, image retrieval, object recognition, image annotation, scene analysis, content-based image analysis, and video analysis. The earlier approaches for remote sensing image analysis are based on low-level and mid-level feature extraction and representation. These techniques have shown good performance by using different feature combinations and machine learning approaches. These earlier approaches have used small-scale image dataset. The recent trends for remote sensing image analysis are shifted to the use of deep learning model. Various hybrid approaches of deep learning have shown much better results than the use of a single deep learning model. In this review article, a detailed overview of the past trends is presented, based on low-level and mid-level feature representation using traditional machine learning concepts. A summary of publicly available image benchmarks for remote sensing image analysis is also presented. A detailed summary is presented at the end of each section. An overview regarding the current trends of deep learning models is presented along with a detailed comparison of various hybrid approaches based on recent trends. The performance evaluation metrics are also discussed. This review article provides a detailed knowledge related to the existing trends in remote sensing image classification and possible future research directions.

1. Introduction

Deep learning and computer vision are used in various applications such as image classification, object detection in industrial production, medical image analysis, action recognition, and remote sensing [1–4]. Satellite images are considered the main source of acquiring geographic information [5], and there are many applications of satellite image analysis in the field of civil engineering such as design, construction, urban planning, and water resource management. The data obtained from satellite sources are huge

and are growing exponentially; to handle these large data, there is a need to have efficient techniques for data extraction purpose. Through image classification, these large number of satellite images can be arranged in semantic orders. The satellite image classification is a multilevel process that starts from extracting features from images to classifying them into categories [6]. Image classification is a step-wise process that starts with designing scheme for classification of desired images. After that, the images are preprocessed which include image clustering, image enhancement, scaling, and so on. At third step, the desired areas of those images are

selected and initial clusters are generated. After that, the algorithm is applied on the images to get the desired classification, and corrective actions are made after that algorithm phase which is also called postprocessing. The final phase is to assess the accuracy of this classification, as shown in Figure 1.

Recent research is focused on the use of mid-level features and deep learning models to build robust decision support systems for smart vehicles, Internet of Things (IoT), and remote sensing images [7–9]. To get the geographical data on large scales, remote sensing plays a significant role, and efficient land use could be achieved through aerial images of Earth [10]. Some are supervised techniques while some of them are unsupervised. Similarly, while keeping in focus the parameters, there are parametric and non-parametric approaches; another type is fuzzy classification [11]; besides this, classification can also be performed on prepixels or subpixels. The latest research in remote image classification is towards hybrid approaches, where two or more techniques are combined to get better classification [3, 12, 13]. The most recent research is focused towards scene-based classification. The whole remote sensing image classification process is divided into three kinds of basic division: supervised learning, unsupervised learning, and deep learning approaches. Supervised learning techniques are further divided into distributed and statistical learning [14–16]. There are many types of distributed learning like logistic regression, decision trees, support vector machine (SVM), ensemble methods, and so on, whereas statistical learning techniques are further divided into parametric and non-parametric approaches. Similarly, different types of unsupervised learning techniques like K-means clustering, spectral clustering, fuzzy C-means, and reinforcement learning are discussed in detail. Moving towards the third division, that is, deep learning approaches, they are further divided into three categories: generative methods, hybrid methods, and discriminative methods. Deep belief network (DBN), network autoencoder, and deep Boltzmann machine (DBM) are discussed in generative methods, whereas deep neural network (DNN) and grey wolf optimization (GWO) are discussed in hybrid methods. In discriminative methods, transfer learning, convolutional neural network (CNN), AlexNet, VGG, GoogLeNet, MobileNet, ResNet, artificial neural network (ANN), and so on are discussed, as shown in Figure 2.

There are basically three types of remote image classifications that are mainly pixel-based classification, object-based classification, and scene-based classification, and the recent research is focused towards scene-based classification [17, 18]. Figure 3 shows that due to the improved research, spatial resolution of images is increased drastically [19]. There is no need to classify remote images on the basis of pixels, and research trend changed towards the object-based classification of images. By objects of remotely sensed images, we mean semantics or scene units [20]. During the last two decades, processing visual features of an image was a time consuming and computationally expensive task, which

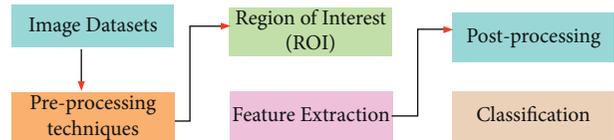


FIGURE 1: Remote sensing image classification process.

required lots of effort and resources. According to the literature in recent years, scale invariant feature transform (SIFT), textual descriptors (TD), color histogram (CH), histogram of oriented gradients (HOG), and global image descriptor (GIST) [21] were proposed by human engineers. After a while, some improvements were made and improved Fisher kernel, spatial pyramid matching (SPM), and bag of visual words (BoVW) were introduced [22]. These encoding techniques were relatively more efficient than the existing techniques [23].

In simple words, we can differentiate between supervised and unsupervised learning as follows: supervised learning algorithms are trained using labeled data, whereas unsupervised learning algorithms are trained using unlabeled data. In case of unsupervised learning, principal component analysis (PCA), sparse coding, and K-means clustering were introduced later on. The benefit of these techniques is that they are able to automatically learn the features. But these unsupervised learning techniques were not doing justice when there are larger datasets [24]. Due to advancement in deep learning techniques and parallel computing, these remote images can be easily classified by initializing weights in training layers so that the prediction of scene could be more accurate in later deep learning layers [25]. There are many deep learning models that exist in literature like AlexNet, GoogLeNet, VGG, and ResNet [26]. AlexNet was proposed in 2012, and it has 60M parameters and is 8 layers deep [27]. GoogLeNet was proposed in 2015, and it has 4M parameters and has 22 layers. It also comes under the category of spatial exploitation [28]. After that, VGGNet was proposed in 2015, and it has 138M parameters and is 16 and 19 layers deeper [26]; it has two types: VGG16 and VGG19. Later ResNet was proposed, and it has various variants like ResNet18, ResNet34, ResNet50, ResNet101, ResNet110, ResNet152, ResNet164, and ResNet1202, and it has 25.6M parameters [29]. The above-mentioned models come under the category of spatial exploitation. Comprehensive reviews relevant to remote sensing image classifications are published in recent years. In [30], we can see a detailed review of multimodel remote sensing image classification. Also, in another article before 2016, all the remote sensing classification techniques were discussed in detail in [31]. In 2017, a detailed review about process of remote sensing image classification was discussed in [32]. The details about the resources for remote sensing research are enlisted in this article [32]. In 2017, a detailed comparison of existing deep learning techniques for hyperspectral classification was given in [33]. In 2017, a review about support vector machine (SVM) techniques relevant to remote sensing was

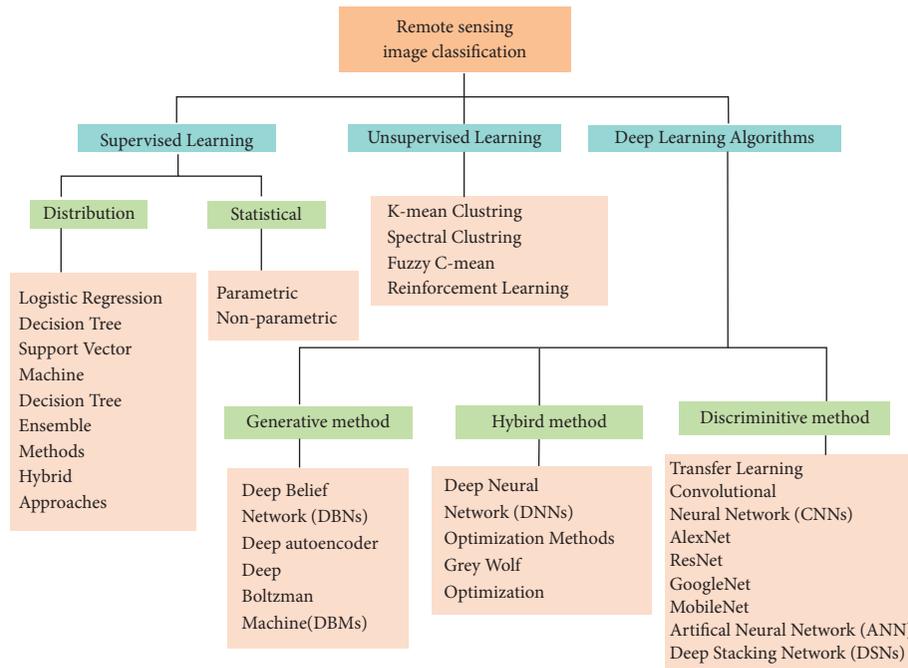


FIGURE 2: Supervised and unsupervised learning techniques.

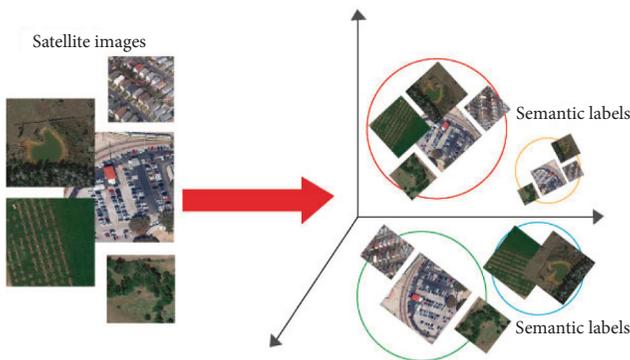


FIGURE 3: Introduction to remote sensing image classification.

discussed in [34]. In [35], AID dataset was proposed, and it also includes remote sensing image classification surveys before 2017. A review of multiple remote sensing techniques and also NWPU-RESISC45 dataset was proposed in [36]. In recent years, comprehensive reviews relevant to hyper-spectral and spatial-spectral images analysis are published [20, 37]. A detailed summary about deep learning in remote sensing applications, current challenges relevant to deep learning methods, benchmarks, and possible future research directions are referred to the following review articles [38–40].

The article is organized as follows: there is a basic introduction about remote sensing image classification in the start. Section 2 is about machine learning. Section 3 contains the detailed description of CNN models and their applications. Section 4 deals with existing deep learning techniques. Section 5 is about the datasets commonly used for

remote sensing image classification which are discussed in detail. Section 6 deals with unsupervised learning techniques. Section 7 is about optimization techniques. Sections 8, 9, and 10 are about feature fusion techniques. Section 11 deals with hybrid approaches. Section 12 is about performance evaluation criterion for classification. In the last section, a conclusion of the proposed research is presented.

2. Machine Learning

Machine learning (ML) is the field of computer science which incorporates both supervised and unsupervised learning techniques [41–43]. It covers both regression and classification problems [44]. In machine learning, a detailed dataset is constructed that covers maximum of system parameters. ML is useful in the scenarios where theoretical knowledge is not sufficient to predict some information out of it [45, 46]. It has a huge number of applications in many areas like land use and cover concerns [47] disaster management, atmosphere changes, and many more [48]. ML is the subdivision of artificial intelligence (AI) [49]. ML basically designs an algorithm to be able to learn from the data to predict something out of it. There are many algorithms present in the field of machine learning that are doing exceptional job like support vector machine, Bayesian network, decision trees, ensemble methods, random forest, neural networks, genetic programming, and many more. ML has a huge impact on remote sensing and geosciences. It automatically extracts features from the data using statistical techniques [50, 51]. At the start, the classification of remote sensing images was considered to be “shallow structures.” To perform remote sensing classification, there exist different

techniques like decision trees, SVM, artificial neural network, bag of visual words, and many more [52–54]. Another important application of ML techniques is to detect the change from the normal scenarios. Images are captured through satellites or drones and then ML techniques are applied to predict the behavior or change [55]. SVM and GA are combined to detect the change. Both supervised learning approaches and unsupervised learning techniques are combined to get the association of adjacent pixels of images, while using SVM, radial basis kernel is used and its parameters like C and Ω are optimized using genetic algorithm (GA); this optimization process increased the efficiency of the process. The authors have performed experimentation while using Mexico dataset and Sardinia image datasets. The results are validated with existing results and the proposed approach outperforms when compared to the existing results [56]. In early years of ML, the accuracy of only high spectral images was high [57]. To overcome this issue, a new 3-D approach was used in combination with spatial and spectral images. The experimentation was performed on Pavia University (PU), Pavia Center (PC), and Kennedy Space Center (KSC) datasets, and the results show that the proposed methods achieve better accuracy with low computational cost [58]. It has many applications in different fields of life like speech recognition systems, search engines, and other AI-based applications like robotics [59]. There are many ML techniques available in literature like K-means clustering and PCA for classification tasks, and to perform regression, there are techniques like SVM, decision trees, ANN, ensemble methods, random forest, and so on [60, 61]. Remote sensing image classification can be performed using existing CNN methods, but they require high computational power and a big labeled dataset for better performance. There are freely available datasets. We can use pretrained networks to get better accuracy. There exist strategies to avoid overfitting and dropouts which also play an important role. The training time of CNN models is quite long, but GPUs help us to solve this issue [62]. Remote images captured from satellite images have a huge importance, but there are some issues in the clarity of images when weather conditions are not so clear which affect the feature selection part of ML process and thus performance degrades [63]. The article described below fills this gap by using a specially designed toolbox. In the first step, gaps between spatial relationships and pixels are filled, while remaining gaps of temporal dynamics of each pixel are filled in the second phase. The experimentation of above algorithm was performed on two datasets Sentinel-3 SLSTR and Terra MODIS. Data were collected in different seasonal conditions. Also, the experimentation was performed on GNU GPL3 which is a public repository [64].

3. Convolutional Neural Network

Convolutional neural networks are useful in many multimedia applications where we need to classify images without human interference. In this article, four different deep learning models: AlexNet, VGG19, GoogLeNet, and ResNet50, were used for feature extraction. The

experimentations were performed on different datasets: SAT4, SAT6, and UCMD, where images for the datasets SAT4 and SAT6 were extracted from NAIP dataset which has around 330000 scene images of all over US. SAT4 and SAT6 have 4 and 6 classes, respectively, and labels are trees, grassland, barren land, building, road, water, and so on, whereas for UCMD dataset, images were extracted from a large dataset named USGS. It has 20 classes in it. Training and testing ratio for SAT datasets is selected as 80:20, respectively, whereas for UCMD, it is 70:30 [65, 66]. Figure 4 shows the basic process of image classification for CNN.

ResNet50 gives better accuracy on all the three above-mentioned datasets. Accuracy on UCM is 98% and that on SAT4 is 95.8%, whereas that on SAT6 is 94.1%. Satellite image classification is a challenging task due to its variability. Due to this issue, existing approaches are not feasible for object detection in satellite images. In this article, a new model DeepSat V2 is proposed which is basically an augmented version of CNN. The first phase is feature extraction phase where 50 features were extracted and then statistical approaches were used for feature ranking to extract the useful features. It has 2 convolutional layers with RELU layer attached. After convolutional layers, there is a max-pooling layer with dropout layer at the end. After that, feature concatenation layer is present followed by fully connected layers. Last layer is softmax layer based on cross entropy loss function. The optimizer used in this model is Adadelta. All the experimentation was performed on SAT4 and SAT6 datasets. This proposed model has achieved accuracy of 99.9% and 99.84% on SAT4 and SAT6, respectively [67].

4. Deep Learning-Based Methods and Approaches

Satellite images have high importance in many fields of life. This article is about the available datasets on remote sensing and the techniques used to classify satellite images. The existing image classification techniques can be divided into four categories: manual feature extraction, unsupervised feature extraction, supervised feature extraction, and object-based classification, as shown in Figure 5.

Dataset used in this article for classification is UCM land use which has 21 classes and 2100 images. Experimentation was performed using AlexNet. Images used for training are about 10%, and after eight iterations, accuracy reached at 94%. By comparing GoogLeNet and CaffeNet, GoogLeNet gives better accuracy, that is, 97%, on UCM dataset. But AlexNet is almost 4 times faster. Deep learning methods perform better in image classification as compared to other feature extraction techniques [68]. The article is about useful methods for feature extraction using deep learning techniques. AlexNet, VGG19, GoogLeNet, and ResNet50 were used here, whereas experimentation is performed on 3 different datasets: SAT4, SAT6, and UC Merced. The accuracy of UCM dataset on multiple deep learning models is summarized in Table 1.

Performance of the proposed ResNet50 on SAT6 is better as compared to previous models, whereas accuracy on SAT4 is degraded. Classification accuracy of the proposed

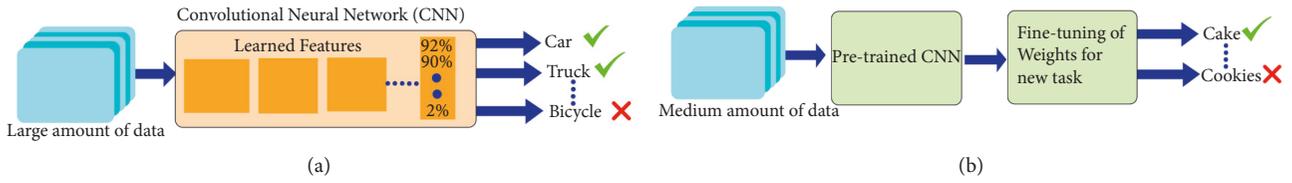


FIGURE 4: Convolutional neural network.

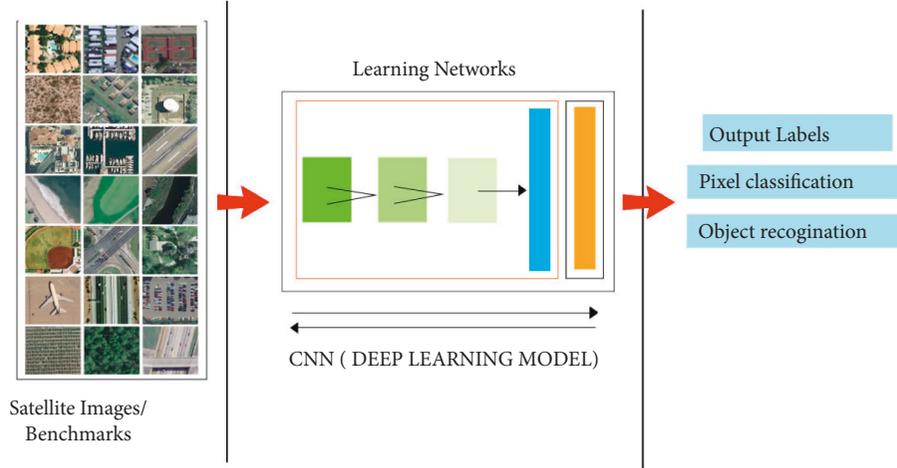


FIGURE 5: Image classification through deep learning.

TABLE 1: Classification accuracy of UCM on different CNN models.

Algorithm	Year	Publication	Accuracy (%)
MCNN [69]	2018	IEEE trans	99.76
ARCNet-VGG16 [19]	2018	IEEE trans	99.12
FACNN [70]	2019	IEEE trans	99.05
SF-CNN [71]	2019	IEEE trans	99.05
CNN-CapsNet [22]	2019	Remote sensing	99.05
D-CNN [72]	2018	IEEE trans	98.93
ADSSM [73]	2018	IEEE trans	98.81
RSFJR [71]	2019	IEEE trans	98.57
GBN [74]	2019	IEEE trans	98.57
MSCP [19]	2018	IEEE trans	98.36
ADFF [75]	2019	IEEE trans	97.53
VGG-VD-16 [76]	2017	IEEE trans	95.21
CaffeNet [76]	2017	IEEE trans	95.02
GBRCN [77]	2015	IEEE trans	94.53
GoogLeNet [76]	2017	IEEE trans	94.31
AlexNet [78]	2021	Journal of Geosciences	94.20
Feature RCGSVM [22]	2019	PLOS ONE	93.81
Inception-V3 [78] [2]	2021	Journal of Geosciences	91.10
LPCNN [79]	2016	IEEE trans	89.90

ResNet50 on SAT4 is 95.8%, that on SAT6 is 94.1%, and that on UCM is 98% [80]. In this article, a new CNN model known as deep convolutional neural network (DCNN) is proposed which works in twofold. In the first phase, multiple filters were introduced to minimize variance, whereas in the 2nd phase, best suited hyperparameters were selected from the pool. Based on these found parameters, a new convolutional neural network (DCNN) model is built and experimentation is performed. The results are validated using

DeepSat model, whereas datasets used are SAT4 and SAT6. Table 2 summarizes the accuracy of RSSCN dataset on different CNN models. Classification accuracy of SAT4 using convolutional neural network (DCNN) is 98.408%, whereas that of SAT6 using convolutional neural network (DCNN) is 96.037% which is better than the model used for validation [83]. In satellite image classification, the process of scale selection is very important task. The remote image datasets are in larger number, and it is very important to

TABLE 2: Classification accuracy of RSSCN on different CNN models.

Algorithm	Year	Publication	Accuracy (%)
CaffeNet [81]	2019	Journal of Geovisualization and Spatial Analysis	88.25
VGG-VD-16 [82]	2018	Applied Sciences	87.18
Inception-V3 [81]	2019	Journal of Geovisualization and Spatial Analysis	87
AlexNet [78]	2021	Journal of Geosciences	85.25
GoogLeNet [82]	2018	IEEE	84.84
POVH [22]	2019	PLOS ONE	84.07
RGSIR [22]	2019	PLOS ONE	81.44

select the relevant techniques for the selection process. In [84], an enhanced technique of CNN model is used and experimentation is performed on WHU-RS, UC Merced, and Brazilian coffee scene datasets. Here classification accuracies of all three datasets are presented. For UCM dataset, the accuracy is better at stage 2, whereas for WHU-RS datasets, the accuracies are measured after stage 1, stage 2, and stage 4 of image scaling. After scales 1 and 2, the accuracy is improved, but after scales 3 and 4, the improvement in accuracy is very small. Land cover land use has a great link between human and nature, and many research studies are available on one-class extraction, but there is a need to focus on multiclass classification. Here in this article, to overcome the issue of low-resolution loss, a new model HR-net is introduced. Comparison was performed on DeepLab and U-Net. The proposed model performs better, test accuracy is 95.7%, mean I/U value is 88.01%, and kappa value is 94.55% [85]. There are datasets available on remote sensing and also techniques used to classify satellite images. The existing image classification techniques can be divided into four categories: manual feature extraction, unsupervised feature extraction, supervised feature extraction, and object-based classification. Dataset used in this article for classification is UCM land use which has 21 classes and 2100 images. Experimentation was performed using AlexNet. 10% images were used for training, and after eight iterations, the accuracy reached at 94%. By comparing GoogLeNet and CaffeNet, GoogLeNet gives better accuracy, that is, 97%, on UCM dataset, but AlexNet is almost 4 times faster. Deep learning methods perform better in image classification as compared to other feature extraction techniques [86]. Xia et al. [76] used google net algorithm and reported a classification accuracy of 94.31% while using UCM image benchmark. Zhang et al. [77] performed scene classification using gradient boosting random convolutional network framework and reported classification accuracy as 94.53% while using UCM image benchmark. Zhong et al. [79] used large patch convolutional neural networks and reported a classification accuracy value as 89.90% for UCM image benchmark. The classification accuracies of AID dataset on different CNN models are summarized in Table 3.

While using the same dataset, the experimentation was performed on CaffeNet, and the accuracy noted in this experimentation is 95.31%. In the third run using same dataset with a different algorithm, i.e., VGG-VD-16, the classification accuracy is 95.21% for UCM dataset. In 2nd run for AID dataset, the experimentation was performed using GoogLeNet and Inception-V3 algorithm, and the accuracies mentioned in

the article are 86.39% and 93%, respectively [76]. Experimentation was performed using ARCNet-VGG16 in scene classification with recurrent attention of VHR remote sensing images reaching the accuracy up to 99.12%. The same experimentation was performed on AID dataset and the accuracy achieved is 93.10% using UCM dataset [19]. In [19], it is stated that they have performed the experimentation using minimum sum coloring problem (MSCP) algorithm and the classification accuracy achieved is 98.36% for UCM dataset. Again the experimentation was performed on AID dataset using three algorithms MSCP, DCNNs, and HW-CNN, whereas the accuracies are 94.42%, 96.89%, and 96.98%, respectively. Zhu et al. [69] reported a value of classification accuracy as 99.76% while using UCM image benchmark. Lu et al. [73] used feature aggregation convolutional neural networks and reported the classification accuracy as 98.81% while using UCM image benchmark. Experimentation performed using feature aggregation convolutional neural network (FACNN) algorithm has achieved accuracy of 99.05%, and the dataset used for experimentation is UCM [70]. Spatial frequency (SF-CNN) has reached the accuracy of 99.05% using UCM dataset. The same algorithm was used for AID dataset and accuracy achieved was 96.66%, whereas using feature aggregation convolutional neural network (FACNN) algorithm on AID dataset, the classification accuracy mentioned in the same article was 95.45% [71].

In [71], it is stated that using robust space-frequency joint representation (RSFJR) algorithm, they have achieved classification accuracy of 98.57% using UCM dataset. In another research, it is stated that they have achieved classification accuracy of 98.57% using GBN algorithm for UCM dataset [74]. ADFE algorithm gave an accuracy of 97.53% in another research using UCM dataset. The same experimentation was performed on AID dataset, and the accuracy achieved is 94.75% [75]. Another research achieved the accuracy of 99.05% using CNN-Caps Net algorithm using UCM dataset. In another article, they achieved accuracy of 93.81% using feature RCGSVM for UCM dataset [22]. AlexNet and inception algorithms gave an accuracy of 94.2% and 911.1%, respectively, using the UCM dataset. Again using AID dataset, the experimentation was performed on VGG-VD-16, and the accuracy achieved is 89.64% [78]. In the article, the author performed experimentation using SCCOV, and the accuracy achieved is 96.10%, and the dataset used for experimentation is AID [89]. Using AID dataset, accuracy achieved is 96.81%, and the algorithm used for experimentation is RSFJR [71]. Using ResNet, another research claimed that they have achieved accuracy of 89.1%

TABLE 3: Classification accuracy of AID on different CNN models.

Algorithm	Year	Publication	Accuracy (%)
CNNS-WD [72]	2019	IEEE trans	97.24
HW-CNN [87]	2018	IEEE trans	96.98
DCNNS [87]	2018	IEEE trans	96.89
RSFJR [71]	2019	IEEE trans	96.81
SF-CNN [70]	2019	IEEE trans	96.66
CNN-CAPSNET [88]	2019	Remote sensing	96.32
SCCOV [89]	2019	IEEE trans	96.10
GBN [70]	2019	IEEE trans	95.48
FACNN [70]	2019	IEEE trans	95.45
ADFF [74]	2019	IEEE trans	94.75
MSCP [87]	2018	IEEE trans	94.42
ARCNet-VGG16 [19]	2017	IEEE trans	93.10
Inception-V3 [76]	2017	IEEE trans	93
MCNN [16]	2018	IEEE trans	91.80
VGG-VD-16 [78]	2021	Journal of Geosciences	89.64
CaffeNet [81]	2019	Journal of Geovisualization and Spatial Analysis	89.10
GoogLeNet [76]	2017	IEEE trans	86.39

TABLE 4: Classification accuracy of SIRI-WHU on different CNN models.

Algorithm	Year	Publication	Accuracy (%)
ResNet50 [22]	2019	PLOS ONE	94.03
AlexNet [81]	2019	Journal of Geovisualization and Spatial Analysis	90.20
GoogLeNet [78]	2021	Journal of Geosciences	89.30
Inception-V3 [78]	2021	Journal of Geosciences	89
VGGNet [78]	2021	Journal of Geosciences	86.60
POVH [81]	2019	Journal of Geovisualization and Spatial Analysis	80.14

using AID dataset. From the literature, we have extracted accuracies of different datasets on different CNN models. In Table 4, the datasets with their respective accuracies are mentioned. Similarly, the classification accuracies of SIRI-WHU dataset on different CNN models are also summarized in Table 4.

5. Datasets

The details about different remote sensing datasets are described below:

5.1. SAT4 and SAT6. National Agriculture Imagery Program (NAIP) dataset was used to extract the images to the dataset. SAT4 consists of total of 500,000 image patches while SAT6 consists of 405,000 image patches, as shown in Figure 6.

5.2. Brazilian Coffee Scenes. Dataset is taken from four countries with the size of 64×64 pixels. There are 600 images in 4 different kinds of dataset while the fifth kind has 476 images. Table below summarizes details regarding the total number of classes, images per class, number of images per class and total number of images in the benchmark, image spatial resolution, and dimensions. Figure 7 summarizes the details about coffee dataset images and other dimensions.



FIGURE 6: Picture gallery taken from all classes of SAT dataset.

5.3. RSSCN. The remote sensing image classification dataset comprises images gathered from Google Earth Engine and covers widespread areas. RSSCN consists of 7 classes of quintessential scene images having a size of 400×400 pixels. Further description about this image benchmark is discussed in the dataset description table. Figure 8 shows the picture gallery of all the classes of RSSCN dataset.

5.4. SIRI-WHU. The description such as image size, total number of images, images per class, and date of creation is referred to the following research article [22]. The images



FIGURE 7: Picture gallery taken from all classes of coffee dataset.

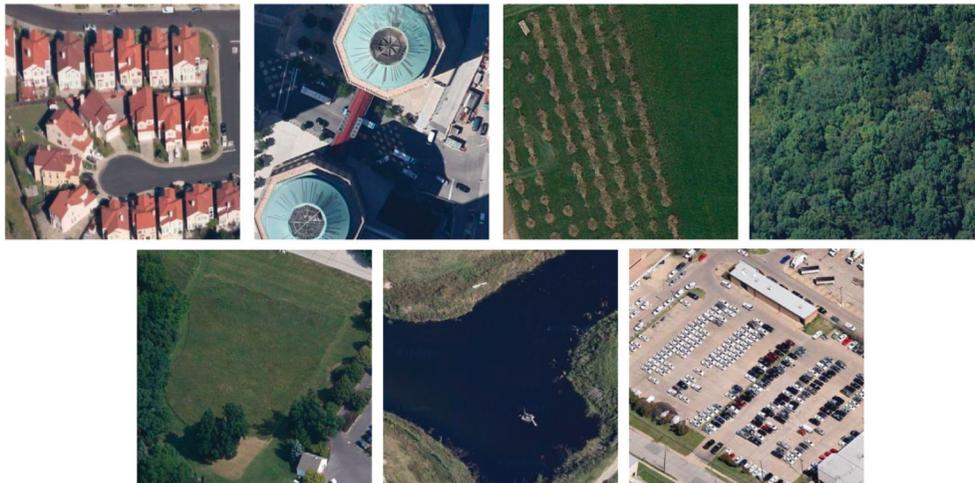


FIGURE 8: Picture gallery taken from all classes of RSSCN dataset.

have a spatial resolution of 2 m with image size of 200×200 pixels. Figure 9 shows randomly selected images taken from each class of SIRI-WHU dataset.

5.5. UC Merced Land Use. The description such as image size, total number of images, images per class, and date of creation is referred to [65]. There are a total of 21 distinctive scene categories with 100 images per class and dimensions of 256×256 pixels, as shown in dataset description table. Figure 10 shows the indiscriminately selected examples of each category included in the dataset (Table 5).

5.6. AID Dataset. AID dataset has 10000 images with 30 different classes. Figure 11 shows the photo gallery of AID dataset.

5.7. DIOR Dataset. The DIOR dataset includes 23,463 images and 192,472 object. Figure 12 shows the photo gallery of

DIOR dataset. Table 5 shows some of the existing datasets with image quantity and other descriptions.

6. Unsupervised Learning Approaches

Due to the advancement of space and satellite technologies, remote sensing has reached a new height [90]. Due to these high-resolution satellites, it has become easier to perform land use land cover surveys, to detect change, to recognize objects, and so on [91]. It has become easier to automatically interpret the image acquaintances due to the advancement in image classification techniques. Using these satellite data efficiently and in effective manner is still challenging. CNNs plays an important role in this image classification process. The article discussed below presents a framework called unsupervised restricted de-convolution neural network (URDNN). The main idea behind this framework is to get unsupervised restricted de-convolution using neural networks. It learns the pixel to pixel and end to end classification and then passes it to CNN model for assigning labels. Due to this, the issue of over and underfitting has been



FIGURE 9: Picture gallery taken from all classes of SIRI-WHU dataset.

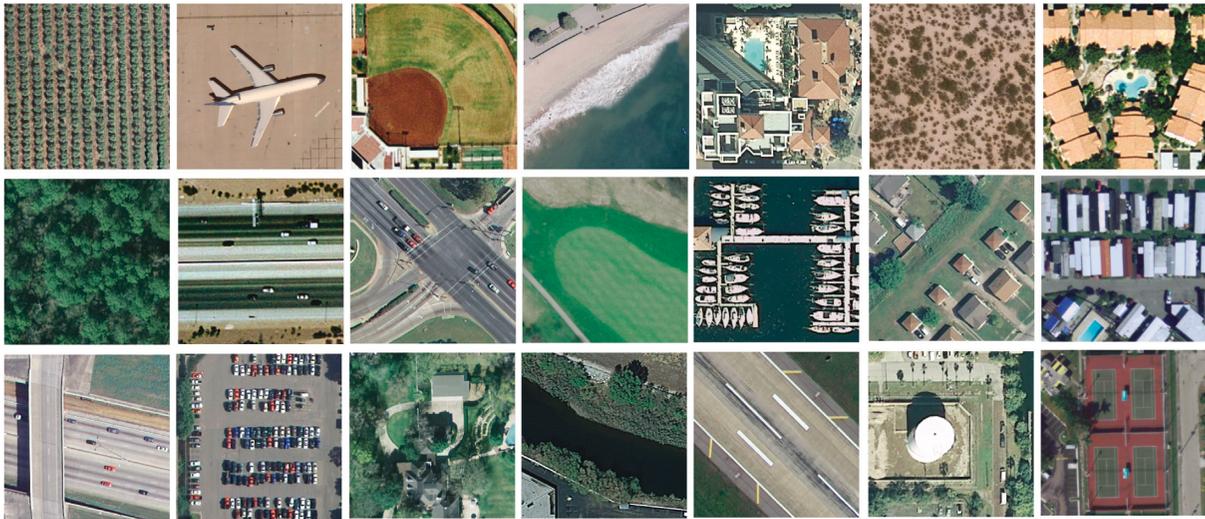


FIGURE 10: Picture gallery taken from all classes of UCM dataset.

TABLE 5: Dataset description.

Dataset	Total number of images	Total number of classes	Image size	Year
UCM [1]	2100	21	256×256	2010
SIRI-WHU	2400	12	200×200	2016
RSSCN7	2800	7	400×400	2015
Brazilian coffee scene	2876	2	64×64	2015
SAT4/SAT6	500000/405000	4 and 6	28×28	2015
AID	10000	30	600×600	2017
NWPU-RESISC45	31500	45	256×256	2017
OPTIMAL	1860	31	256×256	2018
WHU-RS19	1005	19	600×600	2012
NWPU VHR-10 dataset	800	10	VHR	2014
DIOR and DIOR-R datasets	23,463	20	VHR	2020
IKONOS	1000	10	64×64	2008
In-house dataset	606	6	4500×6000	2011
SPOT image dataset	912	3	512×512	2013
ORNL dataset	690	5	512×512	2014

reduced which occurs due to the large number of labeled data. The experimentation was performed on two datasets Geoeye and Quick-bird sensors [92]. The results are better than the previous models. The accuracy achieved is 97% and

98.9%, respectively [92]. Remote sensing image classification using unsupervised deep learning techniques is introduced here. In the first step, CNN extracts features using unsupervised techniques. After that, parameters of the network



FIGURE 11: Picture gallery taken from all classes of AID dataset.

are trained, which are then passed to the classifier. The cost of computation decreased due to unsupervised learning. SVM classifier was used in the process while spatio-spectral information was efficiently extracted with this technique. Adding new layers in the network improves efficiency, but the problem of overfitting is introduced [93]. While discussing unsupervised remote sensing image classification, the concepts of scale invariant feature transform and histogram-oriented gradient are very important [94]. Image is converted into feature vector by encoding; as compared to hand-engineered image representation, unsupervised learning techniques have achieved a new height [95]. We can get image features right from the start of raw pixels of an image. Gabor filters can be applied to those image patches to get the image features out of those pixels. Bags of word (BOW) is another concept of image classification and image retrieval [96]. To get the best results in terms of accuracy, we need to add SVM with non-linear kernel. While keeping

remote sensing in mind, both color feature and intensity are important while classification. But most of the existing algorithms cannot handle this at a time. The article discussed below discussed this issue. The authors have considered the quaternion of color features and then proposed an unsupervised learning technique with the help of this quaternion concept, and they have jointly considered the color and intensity. The experimentation was performed on UCM and Brazilian coffee datasets. The proposed model has given better accuracy than the existing techniques [97]. With the enhancement of deep learning techniques, we are able to classify remotely sensed images using unsupervised learning techniques, more accurately. When available labeled data samples are limited, it becomes difficult to perform image classification using supervised learning techniques [98]. Scene image classification is a hot topic these days; with these classification and analysis techniques, we are able to perform land cover and land use surveys, urban area planning,



FIGURE 12: Picture gallery taken from all classes of DIOR dataset.

disaster management and planning, crop analysis, weather prediction, and so on much easily and with high accuracy [99]. Previously BOW was the unsupervised learning technique that was used for remote sensing image classification. The article discussed below mentions a new technique. To overcome the issue of less labeled data, a new technique is proposed, and it is multilayered feature matching technique. The model uses both discriminative and generative models for learning unlabeled data. The experimentation was performed on two datasets: UCM and coffee, and as compared to other existing techniques, this proposed model MARTA GAI_{ns} outperforms with the classification accuracy of 94.86% and 89.86%, respectively [100].

6.1. Reinforcement Learning. Reinforcement learning is the concept of training a model for classification purpose where we reward the correct behavior and punish the undesired behavior. Reinforcement learning is the subbranch of machine learning which is quite similar to unsupervised learning where there are no labels assigned to the image. In reinforcement learning, agents learn the parameters and predict the outcomes [101]. On that prediction, there is a reward and punishment, and this process carries on till the game ends. Mostly reinforcement learning is used in gaming

and in AI and robotics where you need to teach a robot some new tricks. There are subelements of reinforcement learning that include policy, reward, value function, and environment as a model [102]. The reinforcement learning has achieved a new height as it is really helpful in minimizing the gap between training loss and matrix evaluation [103]. Captioning image is a challenging and most needed task of remote sensing. Most of the existing ML models suffer from the problem of overfitting. Below mentioned article has overcome this issue by proposing a two-stage model, one stage is for autoencoding variations while in stage 2, reinforcement learning is introduced. CNN is fine-tuned in stage 1, and in stage 2, it generates image captions. Reinforcement learning is then applied to improve the accuracy of the model. The experimentation was performed on NWPU-RESISC45 dataset. The results are far better than the previously mentioned results. But there exists a problem of overfitting, which should be addressed in future [104]. Fully polarized radar has the advantage to capture images throughout the time regardless of weather conditions. They are useful for land cover land use type applications, crop management, forest estimation, disaster prevention, recognition of targets, and many more. The article discussed below proposed a new model called deep Q network (DQN) that is basically a deep neural network model for polarized SAR image classification. The data are first preprocessed to

reduce noise and extract features. The data are then fed into deep neural network for classification purpose where the concept of reinforcement learning is introduced. The experimentation was performed on two PolSAR image datasets and the researchers claim that their model outperforms on many existing models [105].

7. Optimization Techniques for Remote Sensing Image Classification

Optimization is the process of finding those input values that best find the output, and there should be a well-defined objective function. This is a very critical task for which there exist multiple machine learning algorithms. Optimization means to minimize operational cost and improve accuracy. An optimization algorithm tries different solutions until it gets the best suitable result which gives the most optimal solution to that problem [106]. In remote sensing image classification, selection of features is basically the most crucial task and it depends on number of available labeled samples. Feature selection is the process of selecting more important features out of the pool of features and excluding correlated features. In the article discussed below, there is a solution proposed for this feature selection task. For this purpose, the stochastic method is introduced for selection of relevant and important features.

The experimentation was performed on two datasets: AVIRIS and ROSIS, and the results show that the proposed method gives better accuracy than the existing approaches [107]. Feature selection is one of the most important tasks in remote sensing image classification. Due to huge amount of data and correlated features, it becomes very tricky. To overcome this issue, a new methodology is introduced, and it has added the concept of wavelet analysis. Threefold strategy is introduced in this framework: in the first phase, the resolution technique is modified, and in the second phase, 3-D discrete wavelet transformation is introduced. In the last phase, CNN is introduced, and the performance of this new model is tested on three different datasets: Indian Pines, University of Pavia, and Salinas datasets. The accuracy achieved using this model is 99.4%, 99.85%, and 99.8%, respectively [108].

When dealing with hyperspectral remote sensing, we usually have limited samples for training. Using the conventional techniques, it has become difficult to achieve high accuracy. SVM gives the better accuracy as it has good generalization and least structural risk, and it overcomes the issue of high time consumption and less optimized parameters. The article discussed below uses EO-1 Hyperion to optimize the parameters. The proposed model is tested, and the authors claimed that they have reached the accuracy of 91.3% which is quite higher than that of the existing approaches [109]. Remote sensing image classification has a huge benefit for land use land cover cases which is a latest research area. Existing classification methods have the issue of low efficiency and they usually have larger datasets. To overcome this issue, a new method is proposed where the concept of extreme programming is introduced. Ensemble methods along with full use of features and deep learning

methods are introduced into the proposed model. All the three methods give better accuracy in terms of classification and efficiency, and the experimentation was performed on multiple datasets; classification accuracy depends upon the type of dataset. The optimization technique combined with deep learning outperforms as compared to other methods [110]. There are many optimization techniques covering different aspects of image classification task: the summary of the techniques is described in later section.

7.1. Grey Wolf Optimization. Grey wolf optimization (GWO) is a new metaheuristic technique, and it mimics the leadership quality of grey wolves [111]. There are four types of grey wolves, namely, alpha, beta, omega, and delta. The fittest solution is called as alpha, second best is called beta, 3rd one is called delta, and the last one is omega. There are three steps of hunting: searching for prey, encircling prey, and attacking prey; these three steps are implemented to get the optimized performance. This technique is basically a feature selection-based technique [112, 113]. In HSI, there are many consecutive and narrow spectral bands that give information about various land covers. Due to number of features, the time complexity increases. Selecting the best features out of the pool of features is a difficult and challenging task. The article discussed below proposed a new technique for feature selection of HSI, and it reduces redundant features. Fuzzy C-means algorithm is used for the decomposition of feature subset, whereas wolf optimization and max entropy are used for feature selection. The experimentation was performed on three known datasets: Indian Pines, Pavia University, and Salinas. The proposed methods outperform in terms of classification accuracy of existing techniques [114]. Image processing and analysis is an emerging field of computer vision. It has many different applications like image classification, segmentation, medical imaging, compression of image, and many more. There exist multiple algorithms to solve these issues like GA, GP, grey wolf algorithm, bat algorithm, and so on. The article discussed below is a review of multiple optimization techniques, their usage, and their real-world applications [115]. Grey wolf is one of the recent trends which comes under the umbrella of swarm intelligence. It has better performance than swarm intelligence and hence is used more effectively than swarm intelligence, and it is simpler to implement and easy to understand. The article discussed below is a review of multiple applications of grey wolf techniques and its applications [116]. We can summarize these optimization algorithms as follows:

- (i) Grey wolf algorithm can handle large data efficiently, but it ignores smaller details which need to be addressed.
- (ii) Grey wolf divides the features into four groups; in future, there should be a direction where more or less than four groups are formed.
- (iii) Effectiveness of grey wolf should be checked in combination of different optimization algorithms.
- (iv) There should be a focus on solving dynamic problems using GWO.

- (v) Parameter tuning of GWO could also be focused in future.

Table 6 shows the summary of some of the optimization techniques described in literature.

8. Fusion of Deep Learning with Spectral Features

Classification accuracy of hyperspectral images (HSI) has increased drastically when using in combination with CNNs. To perform better, it is needed to have denser network which as a result causes overfitting, degradation of accuracy, and also gradient vanishing. To overcome these issues, a new framework hierarchical feature fusion network (HFFN) is proposed. The main idea behind this model is to fuse the output of all the layers which results in increase of accuracy. The experimentation was performed on three real HSI datasets: AVIRIS Indian Pines image, ROSIS-03 University of Pavia image, and AVIRIS Salinas image. The experimental results were compared with DCNN, SVM, and DRN. The results showed that the proposed method outperforms as compared to existing DL methods [117].

CNNs are known as most powerful methods when talking about hyperspectral image classification. Usually pooling layers and sampling features of CNNs are fixed, so they cannot be used for downsampling of features. A research article proposed a deformable HIS. The proposed method is evaluated on two real HSI datasets: University of Pavia and Houston University, and they have 12 and 15 classes, respectively. 1st experiment was performed on Pavia dataset where training samples (45, 55, and 65) are randomly selected from each class. The results showed that the proposed method accurately classifies pixels in the near edge regions. The 2nd experiment was performed on Houston dataset where training samples (30, 40, and 50) were also randomly selected from each class. It has been observed that the proposed method performs better than other existing methods [118].

A deeper network with 9 layers is proposed called as contextual deep CNN, and the idea behind this research is to have a model that can accurately find local contextual interactions by jointly exploiting local spatio-spectral relationships of neighboring individual pixel vectors, as shown in Figure 13

In the first step, multiscale joint exploitation of the spatio-spectral information is obtained through filter bank which is then combined in a map. The experimentation is performed on three datasets: the Indian Pines dataset, the Salinas dataset, and the University of Pavia dataset. Indian Pines dataset has 12 classes but only 8 were used as there were so many images. Pavia dataset has 16 classes, and all of them were considered for experimentation. The accuracy of Indian Pines using proposed technique is 93.6%, that of Salinas is 95.07%, and that of Pavia is 95.97% [119].

Hyperspectral image (HSI) is a new research area. In this article, a special CNN model is proposed that performs the desired classification by using lesser training and fine-tuning of data. To perform this task, the pixels can be pulled from the same class closer, while pushing the different class pixels farther away. The experimentation is performed on three

TABLE 6: Optimization techniques.

Application	Techniques
Segmentation	Grey wolf optimizer
	Genetic programming
	Genetic algorithm
	Particle swarm optimization
	Cuckoo search algorithm
Classification	Bat algorithm
	Grey wolf optimizer
	Genetic programming
Medical imaging	Particle swarm optimization
	Grey wolf optimizer
	Differential optimization
Compression	Bat algorithm

HSI datasets: Indian Pines, Pavia, and Salinas. The results were validated on AlexNet, VGG-CNN-S, and GoogLeNet. The previous accuracies were 88.45%, 85.5%, and 88.8%, respectively, whereas the proposed model gives accuracies of 96.21%, 86.46%, and 88.48%, respectively [120].

In [121], a new model SAFF is proposed. In the 1st phase, multiple labels were identified by using pretrained CNNs and then a self-attention layer is added for channel-based and spatial-based weight assigning. At the end, SVM was used for classification. The experimentation was performed on three different datasets: (1) UC Merced Land Use Dataset having 2100 images and 21 classes; (2) Aerial Image Dataset having 10000 images and 30 classes; and (3) NWPU-RESISC45 Dataset with 31,500 images and 45 classes. The overall accuracy of UCM dataset is 97.02%, that of AID dataset is 90.25%, and that of NWPU dataset is 84.38%.

9. Feature Fusion

Earlier in the literature, we have seen that for image retrieval, one technique was used; later on, it was observed that fusion of more than one techniques can give better accuracy [122]. In this article, a new model weight feature convolutional neural network (WFCNN) is proposed that performs segmentation and extraction of information from images. The WFCNN model first performs encoding and then classification is performed. The proposed model is trained by using stochastic gradient decent (SGD) algorithm. The experimentation was performed on two datasets: Gaofen 6 images and aerial images. The results are validated using SegNet, U-Net, and RefineNet models. GF-6 datasets give accuracy of 94.13%, and aerial image dataset gives accuracy of 96.9% [123]. Ren et al. [124] proposed a full CNN based on multiscale feature fusion for the class imbalance for remote sensing image classification. The authors named the proposed research model as DeepLab V3+, with loss function based solution of samples imbalance [124]. Experimentation was performed on 2 datasets: sentinel-2 and sentinel-3. When compared with U-Net, PSPNet, and ICNet, the proposed method gives accuracy up to 97% [124]. This article proposed a new technique where large image is divided into small-scale images. To divide the samples into classes, support vector machine (SVM) is used. After this phase, a new module called active learning is added. The proposed model

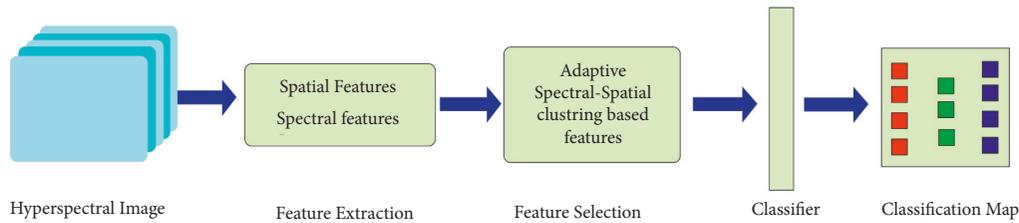


FIGURE 13: Block diagram representing details about spectral feature extraction.

(SSFFSC-AL) performs better in terms of classification accuracy and also gives results in lesser time. The experimentation was performed on two datasets: Indian Pines and Pavia [125]. Feature fusion has two basic methods: local feature fusion and global feature fusion.

Zhu et al. [126] claimed about local and global feature fusion for high-resolution spatial images for scene classification. They have merged two different techniques for feature fusion: local and global fusion. The datasets used for the experimentation are 21-class UC Merced and 12 class Google dataset of SIRI-WHU. Google dataset reached the accuracy of 99.76%, whereas the other one has 96.37%. Future directions identified from this article are

- (1) To use social media data for this training purpose.
- (2) To improve classification accuracies of remote sensing images.
- (3) To implement this research on non-optical data.

Li et al. [127] discussed about scene image classification by fusion strategy to integrate multilayer features using CNN for pretrained data. CNN was used for feature extraction process and then fully connected layers were used for deep feature extraction; then, these extracted features were fused using PCA; after that, classification process was performed. The datasets used for the experimentation are WHU-RS and UCM, and the authors claim that they have achieved better accuracy than previously implemented classification processes. The gap identified in this article is to reduce computational time and to improve classification accuracy [127].

Yuan et al. in [128] discussed scene image classification which was performed by global rearrangement of local features, and the rearrangement of local features helped to get spatial information of the image. The experimentation was performed on four different datasets: UCM, WHU-RS19, Sydney, and AID, and they claimed that the performance was satisfactory. In future, there should be a focus towards improvement of classification accuracy.

In [128], the multilayer covariance pooling technique was used for extraction of features; then, these features were stacked to form a covariance matrix, and finally support vector machine was used for classification. The experimentation was performed on UCM, AID, and NWPU-RESISC45 datasets, and the proposed method outperforms existing methods of classification. In future, there should be an end to end CNN model which is able to classify with better accuracy using lesser features maps at each layer.

A research article discussed feature aggregation to learn about scene classification. This model unites feature

learning, aggregation, and classification into CNN during training process. Fine-tuning is performed to alleviate the training process, and it works for insufficient data as well. The experimentation was performed on three datasets: AID, UCM, and WHU-RS19. The limitation of this research is that there should be a technique that can get semantic information of images without cropping or resizing of images [87].

Figure 14 shows the complete process of how features are extracted and image classification process is completed. Another article is an unsupervised feature fusion technique for training of CNN. Due to this, training becomes easier and more efficient; after that, feature fusion was performed to classify images. The experimentation was performed on UCM and Brazilian coffee datasets, and the proposed model gives better accuracy of 87.83%. There should be a focus on different feature fusion strategies to check their effect [73]. Table 7 summarizes the above-explained research articles.

10. Texture Features

Feature selection and extraction are the most important tasks in content-based image retrieval. There could be two types of features: global and local features. Global features include color, texture, shape, and spatial information, whereas local features have the information about image segmentation, edge detection, corners and blobs, and so on [129].

Texture features are considered to be most powerful features among all. They are the most visible and noticeable patterns in any image. But we cannot use texture features separately. Among low-level visual features, texture of an image is considered as a distinguishable image representation. They are the considered as the visible and noticeable pattern of an image. Different fusions of texture features have shown good results in different application of remote sensing and image retrieval [130].

With these pros, there are also some cons of texture feature extraction. Complexity increases while processing and extracting texture feature [131]. To overcome these issues, different forms of texture features extraction methods are reported in literature such as wavelet transform [132] and Gabor filter [133], and Table 8 presents a detailed summary about texture features.

In this article [134], a new technique is proposed for classification and extraction of features from SAR images. The method is divided into three phases. In the 1st phase,

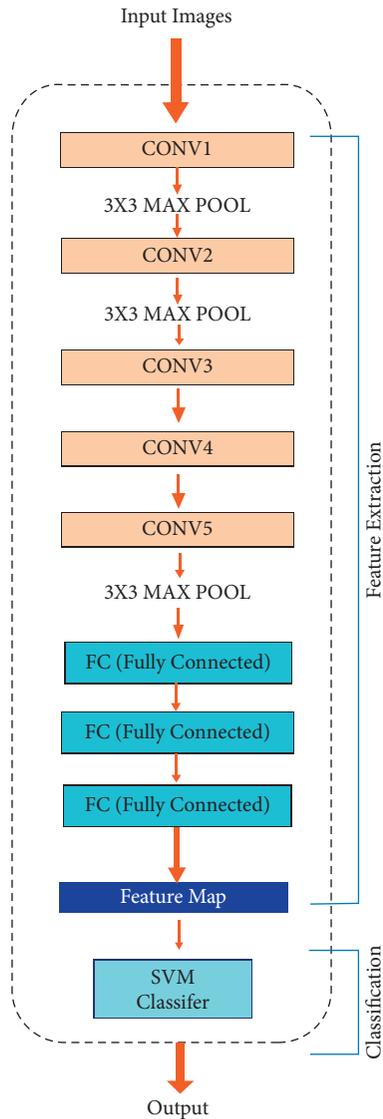


FIGURE 14: Deep learning model.

two types of features were extracted: grey level co-occurrence matrix and Gabor filter. In phase 2, dimensionality is reduced, and at final phase, SVM is introduced for image classification. The experiments showed that this model gives better classification accuracy and is also good for dimensionality reduction. SAR image dataset is used for experimentation, and accuracy is 87.5% [134].

Speckle effect is a very common issue of PolSAR. To overcome this issue, a special technique is proposed that first extracts the features and then classifies them. Real PolSAR images were used for experimentation process and then validated using existing techniques. This article claims that they have reached the accuracy up to 99.8% [135].

Hyperspectral sensors can collect huge amount of data now. But it is still challenging to classify HSIs accurately. The technique used in previous research was spatio-spectral classification, but these were not able to classify images accurately. In this article, the author proposed a new technique to classify images, and this technique is the process that is carried out in three

phases. In the 1st phase, feature extraction is performed, and in the 2nd phase, images were classified using probabilistic SVM, while in the 3rd and last phase, probabilities were calculated to find the results. The experimentation was performed on two different HSI datasets: Indian Pines and Pavia. The results showed that the classification accuracy of the proposed model is better as compared to previously used techniques [136].

Kai et al. [137] claimed that they extracted texture features using the Gabor method. The datasets used for the experimentation are Corel, Li, and Caltech 101. They managed to improve accuracy. The results showed 83%, 88%, and 70% accuracy of each dataset, respectively. The main limitation identified in this research is the increase in computational cost while feature extraction. In [138], Sajjad et al. reported that texture features could be extracted efficiently using the wavelet method. They have claimed high accuracy of 99%, 56%, and 35% on Corel 1K, Corel 5K, and Corel 10K, respectively. Using wavelet methods of texture feature extraction, we can increase accuracy but computation cost also increases as a result. In another article [139], Sajjad et al. extracted texture features using the histogram method. The experimentation was performed on Corel 1K and Corel 5K datasets. Classification accuracy was 87%. In 2018, it is reported in their research that texture features extracted using the edge detection method give better accuracy, i.e., 98%. The dataset used for the experimentation is NUSWIND. The limitation of this research is the increase in computation cost.

Wang et al. [140] found that texture features extracted using Canny edge detector give better accuracy of 68%. The dataset used for experimentation is Corel 10K. The drawback of this research is increase in running cost as the number of input images was very large.

Nazir et al. [141] stated in their article that texture features extracted through discrete wavelet transform (DWT) and edge histogram descriptor (EDH) have better accuracy than those of other methods. The experimentation was performed on Corel dataset. The accuracy reported in this article is 73.5%. The drawback of this research is that no machine learning methods were used for classification or extraction of features.

In [142], Thusnavis Bella and Vasuki used the ranklet transformation method for texture feature extraction. They claimed that they have increased the accuracy. The datasets used in their experimentation are Corel 5k and Corel 10K. Accuracies measured in the article are 67.4% and 67.9%, respectively. The limitation of this research is that due to many dimensions of texture features, the computation cost increased.

Bella et al. [142] performed texture feature extraction using the grey level co-occurrence matrix (GLCM) method. The dataset used in this experiment was Corel 5K, and they achieved accuracy of 66.9%. The computation accuracy is very high, as there was no algorithm used in their experimentation to reduce computational cost.

In [143], Ashraf et al. [143] claimed that using Gabor filter they have extracted texture features. The accuracy achieved in this experimentation was 79% while dataset was Corel 5K. The limitation of this research is also the increase

TABLE 7: Feature fusion.

Title	Year	Methods	Datasets	Limitations
Integrating multilayer features of convolutional neural networks for remote sensing scene classification [127]	2017	Fisher kernel coding	WHU_RS dataset, UCM dataset	To reduce computational time and to improve classification accuracy.
A deep-local-global feature fusion framework for high spatial resolution imagery scene classification [126]	2018	Deep-local-global feature fusion	21-class UC Merced dataset, 12-class Google dataset of SIRI-WHU	Same process could be implemented on non-optical images.
Remote sensing scene classification using multilayer stacked covariance pooling [129]	2018	Multilayer stacked covariance pooling (MSCP)	UC Merced land use dataset, AID30, NWPU-RESISC45 dataset	There should be an end to end CNN model which is able to classify with better accuracy using lesser features maps at each layer.
Remote sensing image scene classification using rearranged local features [128]	2019	Global and rearranged local features	UC Merced, WHU-RS19, Sydney, AID	To explore more techniques for feature fusion.
A feature aggregation convolutional neural network for remote sensing scene classification [130]	2021	Feature aggregation convolutional neural network (FACNN)	AID, UC Merced, WHU-RS19	There should be a technique that can get semantic information of images without cropping or resizing of images.

in computational accuracy as there are many feature dimensions.

In [144], Alsamadi et al. [144] reported that they have extracted texture features using the DWT method. The dataset used in the experimentation was Corel, and they achieved accuracy of 90%. The limitation of this research is high computational time.

11. Hybrid Approaches

In this article, a hybrid approach is used to accurately classify remotely sensed aerial images. SVM and KNN were combined in this article. First SVM was trained to classify images into different classes. In the testing phase, newly tested samples were entered, and average distance between the test samples for each class was calculated using the distance formula. Lastly, the images are placed to their respected classes where there is minimum average distance. This process is repeated till all the images are sorted into their respective classes. The experimentation was performed on two datasets: the ALOS data of the Yitong River and PMS sensor. The results are quite impressive, i.e., 92.44% and 97.8%, respectively [148].

In an article, both parametric and non-parametric approaches were combined to classify the remote sensing images especially land cover land use data. Also, a new dataset was proposed in this article for this purpose which can also be used in other related research. The data of land were captured for both dry and wet conditions. The proposed model is basically the combination of ISODATA clustering and decision trees. The accuracy achieved for dry conditions is 84.54%, whereas for wet weather conditions, the accuracy is measured as 91.10%, which is better than existing deep learning models [149].

In an article, the authors combined two algorithms: kernel-internal value fuzzy C-means clustering and multi-value C-means clustering; by comparing the results with conventional fuzzy C-means clustering, it was observed that

the proposed methods outperform the existing methods. They have constructed a new dataset: LANDSAT-7 Ba Ria area and Hanoi area. The accuracy noted in this research was 98.2% and 94.13%, respectively [11].

An article explains the phenomenon of sparse code that is used to reduce the calculation time for feature extraction. SC is commonly used for aerial images as it performs better in this particular case. With the help of existing approach, accuracy of local feature extraction is increased as compared to existing techniques. The experimentation was performed on UAS operating system data that are recorded for nearly 2 hours without flight interruptions. The accuracy achieved is 85.7% [150].

An article states that the combination of two techniques: pixel-based multilayer perceptron and CNN. This combined algorithm is applied on a dataset that is obtained through aerial photography and satellite. The dataset contains images of both urban and rural lands of different land uses of Southampton. The proposed method outperforms the existing deep learning methods. The accuracies achieved from this proposed model are 90.93% for urban and 89.64% for rural lands [151].

It states the hybrid approach that combines two techniques, SVM and ANN, for LULC classification of images captured through satellite. The fuzzy hierarchal clustering approach is used for classification purpose as shown in Figure 15.

The dataset "Landsat-8 satellite images" is also proposed in this research. All the data are obtained from lands of Hyderabad and its surroundings. The accuracy achieved in this article are 93.159% for SVM and 89.925% for ANN. The authors claim that the proposed method gave better results than existing methods [153].

Yang et al. [134] proposed an efficient classification technique for agricultural lands that is based on spatial and spectral image features. Here a hybrid approach was used for classification purposes of healthy and non-healthy plants. Unmanned aerial vehicle (UAV) images of rice fields in Chianan Plain and Taibao City, Chiayi County, were collected. The accuracy achieved in this research is 90.67% [154].

TABLE 8: Texture feature-based models.

Title	Methods	Datasets	Accuracy	Limitations
Alexey et al. [145]	Gabor	Corel, Li, Caltech101	83%, 88%, 70%	Increase in computational cost
Song et al. [117]	Wavelet	Corel 1K, Corel 5K, Corel 10k	99%, 56%, 35%	Increase in computational cost
Zhu et al. [118]	Histogram	Corel 1K, Corel 5K	87%	-
Phadikar et al. [146]	MPEG-7 edge detector	NUSWIND	98%	Computational cost increased by using GA
Sharmila and Sharmila [147]	Canny edge detector	Corel10K	68%	Running cost increased due to large image input
Cao et al. [121]	DWT,EDH	Corel	73.50%	No ML used in experimentation
Deselaers et al. [122]	Ranklet transformation	Corel 5K, Corel 10K	67.4%, 67.9%	Computational cost increased due to multiple dimensions of features
Zhang et al. [123]	GLCM	Corel 5K	66.90%	No optimization algorithm is used to reduce computational complexity
Ren et al. [124]	Gabor filter	Corel 5K	79%	Computational cost increased because of feature dimensions
Mu et al. [125]	DWT	Corel	90%	High computational time

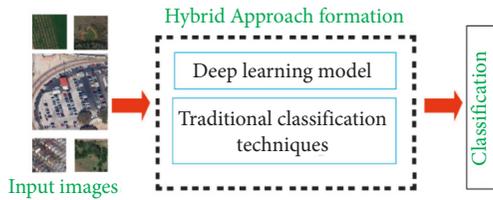


FIGURE 15: Hybrid approaches.

Another article explains the research of a hybrid approach used for classification of remote sensing images. SVM and KNN were combined in this research for better results. Two datasets were used for experimentation: dataset-1 contains “ALOS data of the Yitong River in Changchun,” whereas dataset-2 contains “the ortho image of a factory region in Jiangsu Province.” The accuracies achieved are DS-1: 92.4% and DS-2: 97.9% [155].

An article explains a disaster scenario of southern India which was hit by flood. Data captured in this research were 200 flooded and non-flooded images. The approach used for research is the combination of SVM and K-means clustering. The accuracy achieved was 92%. Bitner et al. [157] extracted automatic building footprint by extracting multiresolution remote sensing images using a hybrid approach. The data of World View-2 imagery of Munich, Germany, were collected through satellite. The experimentation is performed by combining approaches, i.e., U-Net on top of the Caffe deep learning framework. The new hybrid technique performs better than existing techniques. The accuracy achieved is 97.4% [157]. The summary of all the hybrid approaches explained in the above section is given in Table 9.

12. Performance Evaluation Criteria

To evaluate the performance of classification, there are many ways that exist in the literature [65, 158]. The selection of performance measure purely depends upon the type of classification we are going to perform and what type of results are

required. The selection of algorithm for classification purpose also plays an important role for the selection of performance metrics. Following are the performance metrics used to check the accuracy of classification:

True positive (TP): number of images that are correctly labeled

True negative (TN): number of images that are incorrectly labeled

False positive (FP): assigning label to an image when it does not belong to that class

False negative (FN): not assigning label to the image when it really belongs to that class

Below mentioned are some of the types to measure performance of content-based image retrieval:

- (i) Precision/predictive value: it is the ratio of relevant output to the total number of output image.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

- (ii) Average precision: it can be defined as the mean of all the related queries.

$$\text{Avg.Precision} = \sum_{k=1}^{nNRI} p(k) \times \frac{R(k)}{NRI} \quad (2)$$

- (iii) Mean average precision: it is defined as the mean of average mean of all the relevant queries.

$$\text{MAP} = \sum_{(q=1)}^S AP \frac{(q)}{S} \quad (3)$$

where S is the no. of queries.

TABLE 9: Hybrid approaches.

Ref.	Approaches	Datasets	Accuracy
Kantakumar et al. [149]	Maximum likelihood supervised, decision tree, and ISODATA clustering technique	Landsat-8's dataset for dry period and wet period	Dry: 84.54%; wet: 91.10%
Nguyen et al. [11]	Fuzzy C-means clustering and multiple kernel interval-valued fuzzy C-means clustering	LANDSAT-7 Ba Ria area and Hanoi area	98.2% and 94.13%
Qayyum et al. [150]	DRT hybrid dictionary with Ricker wavelet function	UAS operating system recorded data for nearly 2 h without flight interruptions	85.70%
Zhang et al. [151]	CNN and MLP (multilayer perceptron)	Urban and rural scenes of aerial imagery of Southampton	90.93% for urban and 89.64% for rural
Nijhawan et al. [152]	Hybrid of CNN integrated with handcrafted (LBP + GIST) features	Satellite images of Uttarakhand, northern part of India	88.43%
Prasad et al. [153]	Landsat-8 satellite images	SVM and ANN	SVM: 93.15%; ANN: 89.92%
Yang et al. [154]	UAV images of rice fields in Chianan Plain and Taibao City, Chiayi County	DSM and texture information	90.67%
Alimjan et al. [148]	SVM and KNN	DS-1: ALOS data of the Yitong River in Changchun DS-2: the ortho image of a factory region in Jiangsu Province	DS-1: 92.4%; DS-2: 97.9%
Akshya et al. [155]	SVM and K-means	A dataset containing 200 flooded and non-flooded images	92%
Hua et al. [156]	Class-wise attention-based convolutional and bidirectional LSTM network	UCM and DFC15 multilabel datasets	
Schuegraf et al. [157]	U-Net on top of the Caffe deep learning framework	World View-2 imagery of Munich, Germany	97.40%

(iv) Precision recall curve: it is the trade-off between precision and recall under different thresholds.

(v) Recall/sensitivity: it is the ratio of relevant output to all the input and output queries.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

(vi) F -measure: it is the harmonic mean between precision and accuracy.

$$F - \text{measure} = 2 \left(\frac{P \times R}{P + R} \right) \quad (5)$$

(vii) Negative predictive value: it can be defined as the ratio between correctly labeled negative images to total number of negatively labeled images.

$$\text{Negative predictive value} = \frac{TN}{TN + FN} \quad (6)$$

(viii) Specificity: it is the ratio between correctly labeled negative images to total number of negative images.

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (7)$$

(ix) Accuracy: it is the ratio of all the results either rightly labeled or falsely labeled to total number of labels that exist.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (8)$$

(x) Overall accuracy: it is defined as the sum of overall accuracy of total correctly labeled images to all the existing images.

$$w = \sum_{i=0}^{nc} e \frac{ii}{NT} \quad (9)$$

$$NT = \sum_{i=1}^{nc} \sum_{j=1}^{nc} eij,$$

where w = overall accuracy, NT = sum of all non-diagonal elements in confusion matrix, and eij = total correct cells.

(xi) Mean square error: the most popular metric used for measuring the error is mean square error. It computes the average of the squared difference between the target value and the value predicted by the model.

$$\text{MSE} = \frac{1}{N} \sum_{j=1}^N (y_j - y'_j)^2 \quad (10)$$

where N = the last iterations, y_j = true value, and y'_j = value predicted by the model.

(xii) Mean absolute error: when we try to compute the average between the actual value and predicated value we use MAE. The mathematical representation of the metric is given below.

$$\text{MAE} = \frac{1}{N} \sum_{j=1}^N |y_j - y'_j|, \quad (11)$$

where N = the last iterations, y_j = true value, and y'_j = value predicted by the model.

- (xiii) Root mean square error: it is easy to compute and gives a better idea of how well the model is performing. We just have to take the square root of average of the squared difference between the target value and the value predicted. Mathematically, it is pictured as

$$D = \sqrt{\frac{1}{N} \sum_{j=1}^N (y_j - y'_j)^2}. \quad (12)$$

- (xiv) Area under the receiver operating characteristic curve (AUROC): this is a very interesting metric and is also known as AUC-ROC score/curves. While computing AUROC, true positive rate (TPR) and false positive rate (FPR) are used. Mathematically, it is represented as follows:

$$\text{TRP} = \frac{TP}{TP + FN}, \quad (13)$$

$$\text{FRP} = \frac{FP}{FP + TN}.$$

13. Conclusion and Future Directions

Remote sensing image analysis is used in various real-time applications such as monitoring of Earth, urban development, town planning, water resources engineering, providing construction requirements, and agriculture planning. Image analysis and classification is an open research problem for the research community working on remote sensing applications. Due to recent development in imaging technology, there is an exponential increase in the number and size of multimedia contents such as number of videos and digital images. Due to this increase in this volume of digital images, the automatic classification of images is an open research problem for computer vision research community. Various research models are proposed in recent years, but there is still a research gap between human understanding and machine perception. Due to this reason, the research community working on remote sensing image analysis is exploring the possible research directions that can bridge this gap. The earlier approaches for remote sensing image analysis are based on low-level feature extraction and mid-level feature representation. These approaches have shown good performance on small-scale image benchmarks with limited training and testing samples. The use of discriminating feature representation with multiscale features can boost the performance of the learning model. These approaches can mainly assign single labels to images, while in existing era, it is a requirement to assign multiple labels to

single image on the basis of contents. One of the main requirements of a deep learning model is to build a large-scale image benchmark that can be helpful to train a complex deep network. The creation of a large-scale image benchmark with all possible classes of remote sensing images is one of the main requirements and an open research problem in this domain. Most of the current research models based on deep learning are mainly using the fine-tuning and data augmentation techniques to enhance learning. If a large-scale image benchmark is available, it will assist the learning model to learn parameters in a more effective way. The available large-scale image benchmarks are used through supervised learning, and this is a time consuming process and such fully supervised learning models are computationally expensive. Exploring the possible learning capabilities based on unsupervised and semi-supervised learning is a possible future research direction. The deep learning models use extensive computational power for training, and mostly, the research models are using GPUs as high-performance computing. Designing a deep learning model with less computations is also a possible research direction, and such model can be used on a device with less computation powers. The use of few-shot/zero-shot learning approaches can be explored in the field of remote sensing image classification.

Data Availability

The details about the data used is mentioned and cited within this manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] C. Shi, X. Zhang, J. Sun, and L. Wang, "Remote sensing scene image classification based on self-compensating convolution neural network," *Remote Sensing*, vol. 14, no. 3, p. 545, 2022.
- [2] S. Karimi Jafarbigloo and H. Danyali, "Nuclear atypia grading in breast cancer histopathological images based on cnn feature extraction and lstm classification," *CAAI Transactions on Intelligence Technology*, vol. 6, no. 4, pp. 426–439, 2021.
- [3] X. Zhang and G. Wang, "Stud pose detection based on photometric stereo and lightweight yolov4," *Journal of Artificial Intelligence and Technology*, vol. 2, no. 1, pp. 32–37, 2022.
- [4] A. Shabbir, A. Rasheed, H. Rasheed et al., "Detection of glaucoma using retinal fundus images: a comprehensive review," *Mathematical Biosciences and Engineering*, vol. 18, no. 3, pp. 2033–2076, 2021.
- [5] W. Zhang, P. Du, P. Fu et al., "Attention-aware dynamic self-aggregation network for satellite image time series classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022.
- [6] N. Hamid and J. R. Abdul Hamid, "Multi level image segmentation for urban land cover classifications," *IOP Conference Series: Earth and Environmental Science*, vol. 767, no. 1, Article ID 012024, 2021.

- [7] S. Fatima, N. Aiman Aslam, I. Tariq, and N. Ali, "Home security and automation based on internet of things: a comprehensive review," *IOP Conference Series: Materials Science and Engineering*, vol. 899, no. 1, Article ID 012011, 2020.
- [8] Q. Zou, K. Xiong, Q. Fang, and B. Jiang, "Deep imitation reinforcement learning for self driving by vision," *CAAI Transactions on Intelligence Technology*, vol. 6, no. 4, pp. 493–503, 2021.
- [9] M. Arsalan Aslam, M. Naveed Salik, F. Chughtai, N. Ali, S. Hanif Dar, and T. Khalil, "Image classification based on mid-level feature fusion," in *Proceedings of the 15th International Conference on Emerging Technologies (ICET)*, pp. 1–6, IEEE, Peshawar, Pakistan, February 2019.
- [10] Z. Y. Lv, W. Shi, X. Zhang, and J. A. Benediktsson, "Landslide inventory mapping from bitemporal high-resolution remote sensing images using change detection and multiscale segmentation," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 5, pp. 1520–1532, 2018.
- [11] D. D. Nguyen, L. T. Ngo, L. T. Pham, and W. Pedrycz, "Towards hybrid clustering approach to data classification: multiple kernels based interval-valued fuzzy c-means algorithms," *Fuzzy Sets and Systems*, vol. 279, pp. 17–39, 2015.
- [12] M. Sajid, N. Ali, S. H. Dar et al., "Data augmentation-assisted makeup-invariant face recognition," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–10, Article ID 2850632, 2018.
- [13] M. Saqlain, S. Rubab, M. M. Khan, N. Ali, and S. Ali, "Hybrid approach for shelf monitoring and planogram compliance (hyb-smpc) in retails using deep learning and computer vision," *Mathematical Problems in Engineering*, vol. 2022, pp. 1–18, Article ID 4916818, 2022.
- [14] N. Ali, B. Zafar, F. Riaz et al., "A hybrid geometric spatial image representation for scene classification," *PLoS One*, vol. 13, no. 9, Article ID 0203339, 2018.
- [15] A. Latif, A. Rasheed, U. Sajid et al., "Content-based image retrieval and feature extraction: a comprehensive review," *Mathematical Problems in Engineering*, vol. 2019, pp. 1–21, Article ID 9658350, 2019.
- [16] K. Yadav, M. Yadav, and S. Saini, "Stock values predictions using deep learning based hybrid models," *CAAI Transactions on Intelligence Technology*, vol. 7, no. 1, pp. 107–116, 2022.
- [17] Ke Li, G. Wan, G. Cheng, L. Meng, and J. Han, "Object detection in optical remote sensing images: a survey and a new benchmark," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 296–307, 2020.
- [18] M. Asif, M. Bin Ahmad, S. Mushtaq, K. Masood, T. Mahmood, and A. Ali Nagra, "Long multi-digit number recognition from images empowered by deep convolutional neural networks," *The Computer Journal*, vol. 117, 2021.
- [19] Qi Wang, S. Liu, J. Chanussot, and X. Li, "Scene classification with recurrent attention of vhr remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 1155–1167, 2019.
- [20] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: an overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, 2019.
- [21] Si B. Chen, Q. S. Wei, W.-Z. Wang, J. Tang, B. Luo, and Zu. Y. Wang, "Remote sensing scene classification via multi-branch local attention network," *IEEE Transactions on Image Processing*, vol. 31, pp. 99–109, 2022.
- [22] N. Ali, B. Zafar, M. K. Iqbal et al., "Modeling global geometric spatial information for rotation invariant classification of satellite images," *PLoS One*, vol. 14, no. 7, Article ID 0219833, 2019.
- [23] X. Wang, S. Wang, C. Ning, and H. Zhou, "Enhanced feature pyramid network with deep semantic embedding for remote sensing scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 9, pp. 7918–7932, 2021.
- [24] M. M. Elkholy, M. S. Mostafa, H. M. Ebeid, and M. Tolba, "Unsupervised hyperspectral band selection with deep autoencoder unmixing," *International Journal of Image and Data Fusion*, vol. 13, no. 3, pp. 244–261, 2021.
- [25] A. Vali, S. Comai, and M. Matteucci, "Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: a review," *Remote Sensing*, vol. 12, no. 15, p. 2495, 2020.
- [26] H. Yang and W. Yang, "Gscctl: a general semi-supervised scene classification method for remote sensing images based on clustering and transfer," *International Journal of Remote Sensing*, pp. 1–25, 2022.
- [27] P. K. Sethy, "Identification of wheat tiller based on alexnet feature fusion," *Multimedia Tools and Applications*, vol. 81, pp. 8309–8316, 2022.
- [28] D. Wang, C. Zhang, and M. Han, "Mlfc net: a multi level feature combination attention model for remote sensing scene classification," *Computers & Geosciences*, vol. 160, Article ID 105042, 2022.
- [29] O. Kechagias-Stamatis and N. Aouf, "Automatic target recognition on synthetic aperture radar imagery: a survey," *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, no. 3, pp. 56–81, 2021.
- [30] L. Gomez-Chova, D. Tuia, G. Moser, and G. Camps-Valls, "Multimodal classification of remote sensing images: a review and future directions," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1560–1584, 2015.
- [31] L. Zhang, L. Zhang, and Bo Du, "Deep learning for remote sensing data: a technical tutorial on the state of the art," *IEEE Geoscience and remote sensing magazine*, vol. 4, no. 2, pp. 22–40, 2016.
- [32] X. X. Zhu, D. Tuia, L. Mou et al., "Deep learning in remote sensing: a comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.
- [33] P. Ghamisi, J. Plaza, Y. Chen, J. Li, and A. J. Plaza, "Advanced spectral classifiers for hyperspectral images: a review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 8–32, 2017.
- [34] U. Chakraborty and D. Chakraborty, "Remote sensing image classification: a survey of support-vector-machine-based advanced techniques," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 33–52, 2017.
- [35] W. Zhou, S. Newsam, C. Li, and Z. Shao, "Patternnet: a benchmark dataset for performance evaluation of remote sensing image retrieval," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 197–209, 2018.
- [36] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: benchmark and state of the art," *Proceedings of the IEEE*, vol. 105, no. 10, pp. 1865–1883, 2017.
- [37] L. He, J. Li, C. Liu, and S. Li, "Recent advances on spectral-spatial hyperspectral image classification: an overview and new guidelines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1579–1597, 2018.

- [38] L. Ma, Yu Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: a meta analysis and review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 152, pp. 166–177, 2019.
- [39] S. Kumar and D. Kumar, "A review of remotely sensed satellite image classification," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 3, p. 1720, 2019.
- [40] G. Cheng, X. Xie, J. Han, L. Guo, and G. S. Xia, "Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3735–3756, 2020.
- [41] M. Govind and S. K. Pandey and Dr, *A Comparative Study on Supervised and Unsupervised Techniques of Land Use and Land Cover Classification*, 2022.
- [42] G. De Luca, "A survey of nisq era hybrid quantum classical machine learning research," *Journal of Artificial Intelligence and Technology*, vol. 2, no. 1, pp. 9–15, 2022.
- [43] M. Alyas Khan, M. Ali, M. Shah et al., "Machine learning-based detection and classification of walnut fungi diseases," *Intelligent Automation & Soft Computing*, vol. 30, no. 3, pp. 771–785, 2021.
- [44] V. Nasiri, A. A. Darvishsefat, H. Arefi, V. C. Griess, S. M. M. Sadeghi, and S. A. Borz, "Modeling forest canopy cover: a synergistic use of sentinel 2, aerial photogrammetry data, and machine learning," *Remote Sensing*, vol. 14, no. 6, p. 1453, 2022.
- [45] D. Cabrera, L. Cabrera, and E. Cabrera, "Perspectives organize information in mind and nature: empirical findings of point-view perspective (p) in cognitive and material complexity," *Systems*, vol. 10, no. 3, p. 52, 2022.
- [46] N. Ali, K. B. Bajwa, R. Sablatnig et al., "A novel image retrieval based on visual words integration of sift and surf," *PLoS One*, vol. 11, no. 6, Article ID 0157428, 2016.
- [47] M. Wang, M. Wander, S. Mueller, N. Martin, and J. B. Dunn, "Evaluation of survey and remote sensing data products used to estimate land use change in the United States: evolving issues and emerging opportunities," *Environmental Science & Policy*, vol. 129, pp. 68–78, 2022.
- [48] R. Jacobsen, C. A. Bernabel, M. Hobbs, N. Oishi, M. Puig-Hall, and Z. Shannon, "Machine learning: paving the way for more efficient disaster relief," *AIAA SCITECH 2022 Forum*, p. 0397, 2022.
- [49] Z. Zheng, S. Du, H. Taubenböck, and X. Zhang, "Remote sensing techniques in the investigation of aeolian sand dunes: a review of recent advances," *Remote Sensing of Environment*, vol. 271, Article ID 112913, 2022.
- [50] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geoscience Frontiers*, vol. 7, no. 1, pp. 3–10, 2016.
- [51] N. Ali, K. B. Bajwa, R. Sablatnig, and Z. Mehmood, "Image retrieval by addition of spatial information based on histograms of triangular regions," *Computers & Electrical Engineering*, vol. 54, pp. 539–550, 2016.
- [52] A. Ben Hamida, A. Benoit, P. Lambert, and C. Ben Amar, "3 d deep learning approach for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4420–4434, 2018.
- [53] B. Zafar, R. Ashraf, N. Ali et al., "Intelligent image classification-based on spatial weighted histograms of concentric circles," *Computer Science and Information Systems*, vol. 15, no. 3, pp. 615–633, 2018.
- [54] B. Zafar, R. Ashraf, N. Ali, M. Ahmed, S. Jabbar, and S. A. Chatzichristofis, "Image classification by addition of spatial information based on histograms of orthogonal vectors," *PLoS One*, vol. 13, no. 6, Article ID 0198175, 2018.
- [55] M. Jain, M. Bajwa, H. Kumar et al., "Agriculture assistant for crop prediction and farming selection using machine learning model with real-time data using imaging through uav drone," in *Emergent Converging Technologies and Bio-medical Systems*, pp. 311–330, Springer, Singapore, 2022.
- [56] C. Pati, A. K. Panda, A. K. Tripathy, S. K. Pradhan, and S. Patnaik, "A novel hybrid machine learning approach for change detection in remote sensing images," *Engineering Science and Technology, an International Journal*, vol. 23, no. 5, pp. 973–981, 2020.
- [57] Y. Li, J. Ma, and Y. Zhang, "Image retrieval from remote sensing big data: a survey," *Information Fusion*, vol. 67, pp. 94–115, 2021.
- [58] H. Shirmard, E. Farahbakhsh, R. D. Müller, and R. Chandra, "A review of machine learning in processing remote sensing data for mineral exploration," *Remote Sensing of Environment*, vol. 268, Article ID 112750, 2022.
- [59] A. J. Moshayedi, A. S. Roy, A. Kolahdooz, and Y. Shuxin, "Deep learning application pros and cons over algorithm," *EAI Endorsed Transactions on AI and Robotics*, vol. 1, pp. 1–13, 2022.
- [60] K. D. P. Pijush, S. Pal, M. Mukhopadhyay, and S. P. Singh, "Big data classification: techniques and tools," in *Applications of Big Data in Healthcare*, pp. 1–43, Elsevier, Landran, India, 2021.
- [61] M. Sajid, N. Ali, N. I. Ratyal et al., "Deep learning in age-invariant face recognition: a comparative study," *The Computer Journal*, vol. 65, no. 4, pp. 940–972, 2022.
- [62] K. Schulz, R. Hänsch, and U. Sörgel, "Machine learning methods for remote sensing applications: an overview," *Earth resources and environmental remote sensing/GIS applications*, Article ID 1079002, 2018.
- [63] M. Awais, W. Li, S. Hussain et al., "Comparative evaluation of land surface temperature images from unmanned aerial vehicle and satellite observation for agricultural areas using in situ data," *Agriculture*, vol. 12, no. 2, p. 184, 2022.
- [64] M. Sarafanov, E. Kazakov, N. O. Nikitin, and A. V. Kalyuzhnaya, "A machine learning approach for remote sensing data gap-filling with open bsource implementation: an example regarding land surface temperature, surface albedo and ndvi," *Remote Sensing*, vol. 12, no. 23, p. 3865, 2020.
- [65] A. Shabbir, N. Ali, J. Ahmed et al., "Satellite and scene image classification based on transfer learning and fine tuning of resnet50," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–18, Article ID 5843816, 2021.
- [66] M. Kadhim and M. Abed, M. Huk, M. Maleszka, and E. Szczerbicki, "Convolutional neural network for satellite image classification," in *Asian Conference on Intelligent Information and Database Systems*, pp. 165–178, Springer, Cham, 2019.
- [67] Q. Liu, S. Basu, S. Ganguly et al., "Deepsat v2: feature augmented convolutional neural nets for satellite image classification," *Remote Sensing Letters*, vol. 11, no. 2, pp. 156–165, 2020.
- [68] M. A. Shafaey, H. Ebied, M. N. Al-Berry et al., "Deep learning for satellite image classification," in *International Conference on Advanced Intelligent Systems and Informatics*, pp. 383–391, Springer, 2018.
- [69] Q. Zhu, Y. Zhong, L. Zhang, and D. Li, "Adaptive deep sparse semantic modeling framework for high spatial resolution

- image scene classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. 1–16, 2018.
- [70] J. Xie, N. He, L. Fang, and A. Plaza, “Scale-free convolutional neural network for remote sensing scene classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6916–6928, 2019.
- [71] J. Fang, Y. Yuan, X. Lu, and Y. Feng, “Robust space-frequency joint representation for remote sensing image scene classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7492–7502, 2019.
- [72] Y. Liu, Y. Zhong, and Q. Qin, “Scene classification based on multiscale convolutional neural network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 12, pp. 7109–7121, 2018.
- [73] X. Lu, H. Sun, and X. Zheng, “A feature aggregation convolutional neural network for remote sensing scene classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7894–7906, 2019.
- [74] R. Zhu, Li Yan, Mo Mo, and Yi Liu, “AttentionBased deep feature fusion for the scene classification of HighResolution remote sensing images,” *Remote Sensing*, vol. 11, no. 17, p. 1996, 2019.
- [75] W. Zhang, P. Tang, and L. Zhao, “Remote sensing image scene classification using cnn-capsnet,” *Remote Sensing*, vol. 11, no. 5, p. 494, 2019.
- [76] G. S. Xia, J. Hu, F. Hu et al., “Aid: a benchmark data set for performance evaluation of aerial scene classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3965–3981, 2017.
- [77] F. Zhang, Bo Du, and L. Zhang, “Scene classification via a gradient boosting random convolutional network framework,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1793–1802, 2016.
- [78] R. S. A. Kareem, A. G. Ramanjineyulu, R. Rajan et al., “Multilabel land cover aerial image classification using convolutional neural networks,” *Arabian Journal of Geosciences*, vol. 14, no. 17, p. 1681, 2021.
- [79] Y. Zhong, F. Fei, and L. Zhang, “Large patch convolutional neural networks for the scene classification of high spatial resolution imagery,” *Journal of Applied Remote Sensing*, vol. 10, no. 2, Article ID 025006, 2016.
- [80] M. Verma, N. Gupta, B. Tolani, and R. Kaushal, “Explainable custom cnn architecture for land use classification using satellite images,” vol. 6, pp. 304–309, in *Proceedings of the Sixth International Conference on Image Information Processing (ICIIP)*, vol. 6, IEEE, Shimla, India, February 2021.
- [81] M. Wang, X. Zhang, X. Niu, F. Wang, and X. Zhang, “Scene classification of high-resolution remotely sensed image based on resnet,” *Journal of Geovisualization and Spatial Analysis*, vol. 3, no. 2, p. 16, 2019.
- [82] B. Zafar, R. Ashraf, N. Ali et al., “A novel discriminating and relative global spatial image representation with applications in cbir,” *Applied Sciences*, vol. 8, no. 11, p. 2242, 2018.
- [83] Ma Zhong, Z. Wang, C. Liu, and X. Liu, “Satellite imagery classification based on deep convolution network,” *International Journal of Computer and Information Engineering*, vol. 10, no. 6, pp. 1155–1159, 2016.
- [84] N. Laban, B. Abdellatif, M. Ebied, H. A. Shedeed, and M. F. Tolba, “Multiscale satellite image classification using deep learning approach,” *Studies in Computational Intelligence: Machine Learning and Data Mining in Aerospace Technology*, Springer, Cham, pp. 165–186, 2020.
- [85] X. Chang, C. Li, J. Wu, Li Lei, Q. Li, and A. Benjamin, “Land cover extraction of remote sensing images with parallel convolutional network,” in *Proceedings of the International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, pp. 1–6, IEEE, Shenzhen, China, September 2021.
- [86] H. Ebied, M. Al Berry, and M. Tolba, “Deep learning for satellite image classification,” in *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics*, pp. 383–391, Springer, Cham, Switzerland, March 2018.
- [87] N. He, L. Fang, S. Li, A. Plaza, and J. Plaza, “Remote sensing scene classification using multilayer stacked covariance pooling,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 12, pp. 6899–6910, 2018.
- [88] N. B. Crews and K. A. Crews, “Mapping vegetation morphology types in a dry savanna ecosystem: integrating hierarchical object-based image analysis with random forest,” *International Journal of Remote Sensing*, vol. 35, no. 3, pp. 1175–1198, 2014.
- [89] N. He, L. Fang, S. Li, J. Plaza, and A. Plaza, “Skip-connected covariance network for remote sensing scene classification,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 5, pp. 1461–1474, 2020.
- [90] R. Shofiyati, “The role of remote sensing technology (latest geospatial) for agriculture in Indonesia,” in *2nd International Conference on Smart and Innovative Agriculture (ICoSIA 2021)*, pp. 243–248, Atlantis Press, Holland, Netherlands, 2022.
- [91] H. Jiang, M. Peng, Y. Zhong et al., “A survey on deep learning-based change detection from high-resolution remote sensing images,” *Remote Sensing*, vol. 14, no. 7, p. 1552, 2022.
- [92] Y. Tao, M. Xu, F. Zhang, Bo Du, and L. Zhang, “Unsupervised-restricted deconvolutional neural network for very high resolution remote-sensing image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 6805–6823, 2017.
- [93] A. Romero, C. Gatta, and G. Camps Valls, “Unsupervised deep feature extraction for remote sensing image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1349–1362, 2016.
- [94] Z. H. Choudhury, “Encryption and encoding of facial images into quick response and high capacity color 2d code for biometric passport security system,” 2022, <https://arxiv.org/abs/2203.15738>.
- [95] Z. Xiong, F. Mo, X. Zhao, F. Xu, X. Zhang, and Y. Wu, “Dynamic texture classification based on 3d ica learned filters and Fisher vector encoding in big data environment,” *Journal of Signal Processing Systems*, vol. 10, pp. 1–15, 2022.
- [96] Y. D. Kwadiki and K. Kwadiki, “Content-based image retrieval using integrated features and multi-subspace randomization and collaboration,” *International Journal of System Assurance Engineering and Management*, pp. 1–11, 2022.
- [97] V. Babic and Z. Babić, “Unsupervised quaternion feature learning for remote sensing image classification,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 4, pp. 1521–1531, 2016.
- [98] S. Bujwid, A. Pieropan, H. Azizpour, A. Maki, and M. Miquel, “An Analysis of Over-sampling Labeled Data in Semi-supervised Learning with Fixmatch,” 2022, <https://arxiv.org/pdf/2201.00604.pdf>.
- [99] K. Abhas, R. Bloch, and J. Lamond, *Cities and Flooding: A Guide to Integrated Urban Flood Risk Management for the 21st century*, World Bank Publications, 2012.

- [100] D. Lin, K. Fu, Y. Wang, G. Xu, and X. Sun, "Marta gans: unsupervised representation learning for remote sensing image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 2092–2096, 2017.
- [101] P. Ladosz, L. Weng, M. Kim, and H. Oh, *Exploration in Deep Reinforcement Learning: A Survey*, Information Fusion, 2022.
- [102] S. Milani, N. Topin, M. Veloso, and F. Fang, "A Survey of Explainable Reinforcement Learning," 2022, <https://arxiv.org/abs/2005.06247>.
- [103] M. Maroto-Gómez, Á. Castro-González, J. C. Castillo, M. Malfaz, and M. Á. Salichs, "An adaptive decision-making system supported on user preference predictions for human-robot interactive communication," *User Modeling and User-Adapted Interaction*, vol. 554, pp. 1–45, 2022.
- [104] X. Shen, B. Liu, Y. Zhou, J. Zhao, and M. Liu, "Remote sensing image captioning via variational autoencoder and reinforcement learning," *Knowledge-Based Systems*, vol. 203, Article ID 105920, 2020.
- [105] K. Huang, W. Nie, and N. Luo, "Fully polarized sar imagery classification based on deep reinforcement learning method using multiple polarimetric features," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 10, pp. 3719–3730, 2019.
- [106] T. Hermosilla, M. A. Wulder, J. C. White, and N. C. Coops, "Land cover classification in an era of big and open data: optimizing localized implementation and training data selection to improve mapping outcomes," *Remote Sensing of Environment*, vol. 268, Article ID 112780, 2022.
- [107] S. A. Medjahed and M. Ouali, "Band selection based on optimization approach for hyperspectral image classification," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 21, no. 3, pp. 413–418, 2018.
- [108] P. Boggavarapu and P. K. L. N. Boggavarapu, "Improved whale optimization based band selection for hyperspectral remote sensing image classification," *Infrared Physics & Technology*, vol. 119, Article ID 103948, 2021.
- [109] Z. Zhang and G. Zhang, "Genetic algorithm-based parameter optimization for eo 1 hyperion remote sensing image classification," *European Journal of Remote Sensing*, vol. 53, no. 1, pp. 124–131, 2020.
- [110] F. Huang, J. Lu, J. Tao, and L. Li, "Research on optimization methods of elm classification algorithm for hyperspectral remote sensing images," *IEEE Access*, vol. 7, pp. 201970–108089, Article ID 108070, 2019.
- [111] A. Rezaeipannah and M. Mojarad, "Modeling the scheduling problem in cellular manufacturing systems using genetic algorithm as an efficient meta-heuristic approach," *Journal of Artificial Intelligence and Technology*, vol. 1, no. 4, pp. 228–234, 2021.
- [112] A. Rida, A. Hicham, and N. Abderrahim, "Optimization of object-based image analysis with genetic programming to generate explicit knowledge from worldview-2 data for urban mapping," in *Geospatial Intelligence*, pp. 157–169, Springer, Cham, Switzerland, 2022.
- [113] A. Jamshed, B. Mallick, and R. Kumar Bharti, "An Efficient Pattern Mining Convolution Neural Network (Cnn) Algorithm with Grey Wolf Optimization (Gwo)," 2022, <https://arxiv.org/abs/2204.04704>.
- [114] F. Xie, C. Lei, F. Li, D. Huang, and J. Yang, "Unsupervised hyperspectral feature selection based on fuzzy c-means and grey wolf optimizer," *International Journal of Remote Sensing*, vol. 40, no. 9, pp. 3344–3367, 2019.
- [115] A. Kumar and S. Lekhraj, "Grey wolf optimizer and other metaheuristic optimization techniques with image processing as their applications: a review," *IOP Conference Series: Materials Science and Engineering*, vol. 1136, no. 1, Article ID 012053, 2021.
- [116] H. Faris, I. Aljarah, M. A. Al-Betar, and S. Mirjalili, "Grey wolf optimizer: a review of recent variants and applications," *Neural Computing & Applications*, vol. 30, no. 2, pp. 413–435, 2018.
- [117] W. Song, S. Li, L. Fang, and T. Lu, "Hyperspectral image classification with deep feature fusion network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3173–3184, 2018.
- [118] J. Zhu, L. Fang, and P. Ghamisi, "Deformable convolutional neural networks for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 8, pp. 1254–1258, 2018.
- [119] H. Kwon and H. Kwon, "Going deeper with contextual cnn for hyperspectral image classification," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4843–4855, 2017.
- [120] G. Cheng, Z. Li, J. Han, X. Yao, and L. Guo, "Exploring hierarchical convolutional features for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6712–6722, 2018.
- [121] R. Cao, L. Fang, T. Lu, and N. He, "Self-attention-based deep feature fusion for remote sensing scene classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 1, pp. 43–47, 2021.
- [122] T. Deselaers, D. Keysers, and H. Ney, "Features for image retrieval: an experimental comparison," *Information Retrieval*, vol. 11, no. 2, pp. 77–107, 2008.
- [123] C. Zhang, Y. Chen, X. Yang et al., "Improved remote sensing image classification based on multi-scale feature fusion," *Remote Sensing*, vol. 12, no. 2, p. 213, 2020.
- [124] Y. Ren, X. Zhang, Y. Ma et al., "Full convolutional neural network based on multi-scale feature fusion for the class imbalance remote sensing image classification," *Remote Sensing*, vol. 12, no. 21, p. 3547, 2020.
- [125] C. Mu, J. Liu, Yi Liu, and Y. Liu, "Hyperspectral image classification based on active learning and spectral-spatial feature fusion using spatial coordinates," *IEEE Access*, vol. 8, pp. 6768–6781, 2020.
- [126] Q. Zhu, Y. Zhong, Y. Liu, L. Zhang, and D. Li, "A deep-local-global feature fusion framework for high spatial resolution imagery scene classification," *Remote Sensing*, vol. 10, no. 4, p. 568, 2018.
- [127] E. Li, J. Xia, P. Du, and C. Lin, "Integrating multilayer features of convolutional neural networks for remote sensing scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, pp. 5653–5665, 2017.
- [128] Y. Yuan, J. Fang, X. Lu, and Y. Feng, "Remote sensing image scene classification using rearranged local features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 3, pp. 1779–1792, 2019.
- [129] L. Kumar Tyagi, R. Kant, and A. Gupta, "A comparative analysis of various local feature descriptors in content-based image retrieval system," *Journal of Physics: Conference Series*, vol. 1854, Article ID 012043, 2021.
- [130] M. J. Alvarez, E. Gonzalez, F. Bianconi, J. Armesto, and A. Fernandez, "Colour and texture features for image retrieval in granite industry," *Dyna*, vol. 77, no. 161, pp. 121–130, 2010.
- [131] M. Ahmed and M. Taha, "A brief survey on modern iris feature extraction methods," *Engineering and Technology Journal*, vol. 39, no. 1, pp. 123–129, 2021.

- [132] Yu-xing Li, S.-bin Jiao, and X. Gao, "A novel signal feature extraction technology based on empirical wavelet transform and reverse dispersion entropy," *Defence Technology*, vol. 17, no. 5, pp. 1625–1635, 2021.
- [133] H. S. Munawar, R. Aggarwal, Z. Qadir, S. I. Khan, A. Z. Kouzani, and M. A. P. Mahmud, "A gabor filter-based protocol for automated image-based building detection," *Buildings*, vol. 11, no. 7, p. 302, 2021.
- [134] A. S. Ismail, X. Gao, and C. Deng, "Sar image classification based on texture feature fusion," in *Proceedings of the IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP)*, pp. 153–156, IEEE, Xi'an, China, September 2014.
- [135] H. Bi, L. Xu, X. Cao, and Z. Xu, "Polsar image classification based on three-dimensional wavelet texture features and Markov random field," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3921–3928, IEEE, Fort Worth, TX, USA, December 2017.
- [136] R. Seifi Majdar and H. Ghassemian, "A probabilistic svm approach for hyperspectral image classification using spectral and texture features," *International Journal of Remote Sensing*, vol. 38, no. 15, pp. 4265–4284, 2017.
- [137] X. G. Kai, Q. L. Zhi, J. J. Wang et al., "The effect of cryogenic treatment on the microstructure and properties of ti-6al-4v titanium alloy," *Materials Science Forum*, vol. 747, pp. 899–903, 2013.
- [138] P. Khare and A. Khare, "Integration of wavelet transform, local binary patterns and moments for content-based image retrieval," *Journal of Visual Communication and Image Representation*, vol. 42, pp. 78–103, 2017.
- [139] M. Sajjad, A. Ullah, J. Ahmad, N. Abbas, S. Rho, and S. W. Baik, "Integrating salient colors with rotational invariant texture features for image representation in retrieval systems," *Multimedia Tools and Applications*, vol. 77, no. 4, pp. 4769–4789, 2018.
- [140] X.-yang Wang, Yu-nan Liu, H. Xu, P. Wang, and H.-ying Yang, "Robust copy-move forgery detection using quaternion exponent moments," *Pattern Analysis & Applications*, vol. 21, no. 2, pp. 451–467, 2018.
- [141] A. Nazir, R. Ashraf, T. Hamdani, and N. Ali, "Content based image retrieval system by using hsv color histogram, discrete wavelet transform and edge histogram descriptor," in *Proceedings of the international conference on computing, mathematics and engineering technologies (iCoMET)*, pp. 1–6, IEEE, Sukkur, Pakistan, April 2018.
- [142] M. I. Thusnavis Bella and A. Vasuki, "An efficient image retrieval framework using fused information feature," *Computers & Electrical Engineering*, vol. 75, pp. 46–60, 2019.
- [143] R. Ashraf, M. Ahmed, U. Ahmad, M. A. Habib, S. Jabbar, and K. Naseer, "Mdcbir-mf: multimedia data for content-based image retrieval by using multiple features," *Multimedia Tools and Applications*, vol. 79, no. 13-14, pp. 8553–8579, 2020.
- [144] M. K. Alsmadi, "Content based image retrieval using color, shape and texture descriptors and features," *Arabian Journal for Science and Engineering*, vol. 45, no. 4, pp. 3317–3330, 2020.
- [145] P. Alexey, S. Hitesh, I. Babakov, S. Parkhi, and G. Buddhawar, "Content-based image retrieval using color, texture and shape features," *Key Engineering Materials*, vol. 685, pp. 872–876, 2016.
- [146] B. S. Phadikar, A. Phadikar, and G. K. Maity, "Content-based image retrieval in dct compressed domain with mpeg-7 edge descriptor and genetic algorithm," *Pattern Analysis & Applications*, vol. 21, no. 2, pp. 469–489, 2018.
- [147] L. K. Sharmila and T. S. Sharmila, "An efficient framework for image retrieval using color, texture and edge features," *Computers & Electrical Engineering*, vol. 70, pp. 580–593, 2018.
- [148] G. Alimjan, T. Sun, Yi Liang, H. Jumahun, and Yu Guan, "A new technique for remote sensing image classification based on combinatorial algorithm of svm and knn," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 07, Article ID 1859012, 2018.
- [149] L. N. Neelamsetti and P. Neelamsetti, "Multi temporal land use classification using hybrid approach," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 18, no. 2, pp. 289–295, 2015.
- [150] A. Qayyum, A. Saeed Malik, N. M. Saad et al., "Image classification based on sparse-coded features using sparse coding technique for aerial imagery: a hybrid dictionary approach," *Neural Computing & Applications*, vol. 31, no. 8, pp. 3587–3607, 2019.
- [151] Ce Zhang, X. Pan, H. Li et al., "A hybrid mlp-cnn classifier for very fine resolution remotely sensed image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp. 133–144, 2018.
- [152] R. Nijhawan, H. Sharma, H. Sahni, and A. Batra, "A deep learning hybrid cnn framework approach for vegetation cover mapping using deep features," in *Proceedings of the 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pp. 192–196, IEEE, Jaipur, India, April 2017.
- [153] S. V. S. Prasad, T. S. Savithri, and I. Murali Krishna, "Comparison of accuracy measures for rs image classification using svm and ann classifiers," *International Journal of Electrical and Computer Engineering*, vol. 7, no. 3, p. 1180, 2017.
- [154] M. D. Yang, K. S. Huang, Yi H. Kuo, H. P. Tsai, and L.-M. Lin, "Spatial and spectral hybrid image classification for rice lodging assessment through uav imagery," *Remote Sensing*, vol. 9, no. 6, p. 583, 2017.
- [155] J. Akshya and P. L. K. Priyadarsini, "A hybrid machine learning approach for classifying aerial images of flood-hit areas," in *Proceedings of the International Conference on Computational Intelligence in Data Science (ICCIDS)*, pp. 1–5, IEEE, Chennai, India, October 2019.
- [156] Y. Hua, L. Mou, and X. Zhu, "Recurrently exploring class-wise attention in a hybrid convolutional and bidirectional lstm network for multi-label aerial image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 149, pp. 188–199, 2019.
- [157] K. Bitner and K. Bittner, "Automatic building footprint extraction from multi-resolution remote sensing images using a hybrid fcn," *ISPRS International Journal of Geo-Information*, vol. 8, no. 4, p. 191, 2019.
- [158] P. Deepan and L. R. Sudha, *Scene classification of remotely sensed images using ensembled machine learning models*, pp. 535–550, Environmental Science, Springer, Singapore, 2021.