remote sensing

MDPI

Article

Satellite Image Classification Using a Hierarchical Ensemble Learning and Correlation Coefficient-Based Gravitational Search Algorithm

Kowsalya Thiagarajan [1], Mukunthan Manapakkam Anandan [2], Andrzej Stateczny [3,*], Parameshachari Bidare Divakarachari [4] and Hemalatha Kivudujogappa Lingappa [5]

1 Department of Electronics and Communication Engineering, Muthayammal Engineering College, Rasipuram 637408, India; kowsalya.t.ece@mec.edu.in
2 Department of Computer Science and Engineering, Veltech University, Chennai 600062, India; drmamukunthan@veltech.edu.in
3 Department of Geodesy, Gdansk University of Technology, 80232 Gdansk, Poland
4 Department of Telecommunication Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru 570016, India; paramesh@gsss.edu.in
5 Department of Information Science and Engineering, Sri Krishna Institute of Technology, Bangalore 560090, India; isehod@skit.org.in
* Correspondence: andrzej.stateczny@pg.edu.pl

check for updates

Citation: Thiagarajan, K.; Manapakkam Anandan, M.; Stateczny, A.; Bidare Divakarachari, P.; Kivudujogappa Lingappa, H. Satellite Image Classification Using a Hierarchical Ensemble Learning and Correlation Coefficient-Based Gravitational Search Algorithm. *Remote Sens.* **2021**, *13*, 4351. https://doi.org/10.3390/rs13214351

Academic Editors: Jon Atli Benediktsson, Pedro Melo-Pinto, Riccardo Roncella and Xiaoli Deng

Received: 22 August 2021
Accepted: 27 October 2021
Published: 29 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: Satellite image classification is widely used in various real-time applications, such as the military, geospatial surveys, surveillance and environmental monitoring. Therefore, the effective classification of satellite images is required to improve classification accuracy. In this paper, the combination of Hierarchical Framework and Ensemble Learning (HFEL) and optimal feature selection is proposed for the precise identification of satellite images. The HFEL uses three different types of Convolutional Neural Networks (CNN), namely AlexNet, LeNet-5 and a residual network (ResNet), to extract the appropriate features from images of the hierarchical framework. Additionally, the optimal features from the feature set are extracted using the Correlation Coefficient-Based Gravitational Search Algorithm (CCGSA). Further, the Multi Support Vector Machine (MSVM) is used to classify the satellite images by extracted features from the fully connected layers of the CNN and selected features of the CCGSA. Hence, the combination of HFEL and CCGSA is used to obtain the precise classification over different datasets such as the SAT-4, SAT-6 and Eurosat datasets. The performance of the proposed HFEL–CCGSA is analyzed in terms of accuracy, precision and recall. The experimental results show that the HFEL–CCGSA method provides effective classification over the satellite images. The classification accuracy of the HFEL–CCGSA method is 99.99%, which is high when compared to AlexNet, LeNet-5 and ResNet.

Keywords: accuracy; Convolutional Neural Networks; Correlation Coefficient-Based Gravitational Search Algorithm; ensemble learning; hierarchical framework; satellite image classification

1. Introduction

Remote sensing images are widely considered as an essential source of data related to the Earth's surface. Information about the fundamental land cover from remote sensing images is required for classification applications [1,2]. The advantages of the remote sensing techniques are their low cost and the possibility for huge area coverage [3,4]. This image classification technique is utilized to recognize and detect appropriate information from satellite images [5], since satellite images have sufficient information to perform land cover mapping to deliver data at national, international and local scales [6]. Remotely sensed satellite imaging is used in various applications, such as forestry, regional planning, agriculture and geology, to examine and handle human activities and natural resources [7–9]. The biophysical cover of the Earth's surfaces is considered one of the most important climate

variables. In environmental analysis, adequate information about the land cover is essential to monitor the effects of resource management and climate change [10].

Due to the ease of implementation and accuracy of the statistical parameters, they are broadly utilized in satellite image analysis techniques. However, these statistical parameter-based techniques are time-consuming, and they are only suitable for small areas. The weather and errors related to the photographic equipment create noise in the satellite images [11]. High-resolution satellite images create various issues during scene classification, which are as follows: (1) enhanced images provide more details, but the low-level features present in low-resolution images are inadequate in capturing different image data, and (2) objects existing in similar types of scenes have various orientations and scales [12,13]. In general, the classification method depends on the labeled samples to train the classifiers, where the accuracy is mainly based on the quality and number of the training samples. However, the annotation of the labeled samples generally requires a great deal of time and is difficult to obtain in different real-world contexts [14].

In this research, the automated analysis and classification of remote sensing satellite images is accomplished. This automated classification is beneficial in different real-world applications, such as environmental monitoring, planning, rescuing, searching and so on [15]. The main contributions of this research are as follows:

- The refined and semantic features are extracted from the fully connected layers of AlexNet, LeNet-5 and ResNet. Further, these extracted features are concatenated together to obtain multiple ensembles of features.
- From the extracted features, the optimal set of features is selected using the CCGSA technique. Hence, the combination of multiple ensemble features from the HFEL and optimal features selected from the CCGSA is used to increase the classification accuracy of satellite images.
- Three different datasets, i.e., SAT-4, SAT-6 and Eurosat datasets, are considered to analyze the performance of the HFEL–CCGSA method.

The following existing works are related to the classifications accomplished on the SAT-4 and SAT-6 datasets.

Unnikrishnan, A., Sowmya, V. and Soman, K.P [16] presented a deep learning architecture with hypertuning of the network, reducing the input bands to two (i.e., red and near-infrared (NIR)). Deep learning architectures were designed for three networks, namely VGG, AlexNet and ConvNet. Here, the hypertuning was accomplished over the filters of each convolutional layer. The classification of different classes was accomplished using the modified architecture with a reduced number of filters and two-band information. However, the hypertuned deep learning architecture obtained less accuracy during the classification. Jiang, J et al. [17] developed the Double-Channel Convolutional Neural Network (DCCNN) model for classifying RGB-NIR images using the correlation among the R, G, B and NIR bands. For describing the RGB and NIR image features, the DCCNN had two independent CNN networks. Next, feature fusion was performed at the fully connected layer and the classification was performed at the final layer. This configuration was useful for the effective utilization of various features of RGB-NIR images. Moreover, overfitting was avoided using the net dropout technique, which eliminated 60% of the neurons in the fully connected layer. However, the classification accuracy was lower for the double-channel CNN model with the net dropout technique.

N. Yang et al. [18] presented a training approach, namely greedy DropSample, for increasing the speed of the Convolutional Neural Network (CNN)'s optimization process during image classification. This method was mainly focused on the samples that generated the highest gradients. Moreover, the activations of the network were biased in the training process due to the absence of certain training samples. The samples with less losses were filtered out to increase the speed of the CNN. However, the developed DropSample failed to consider the similarity between the classes. Weng et al. [19] developed the Multi-Dimensional Multi-Grained Scanning Structure (MGSS) method to classify the land cover/land use over remote sensing images. The developed MGSS was used to extract the

MOST WIEDZY Downloaded from mostwiedzy.pl

spatial and spectral information from the images. Next, the prediction was obtained by mapping the probability feature vectors in the residual forest structure. The number of parameters required for the optimization of MGSS was lower when compared to the CNN. However, the gradient passed from the high level affected by a certain event in MGSS. Zhong, Y et al. [20] presented an agile CNN structure, namely SatCNN, for obtaining the classification over High-Spatial-Resolution Remote-Sensing (HSR-RS) images. Here, the intrinsic features from the HSR-RS images were captured using preprocessing with the z-score methods. The developed SatCNN was used to balance the training efficiency and generalization ability of the model. However, the performance of the SatCNN was sensitive to the testing ratio. Specifically, the SatCNN's performance was affected due to the intra- and inter-class complexity of the HSR-RS.

The following existing works are related to the classifications accomplished on the EuroSAT dataset.

S. A. Yamashkin et al. [21] solved the issue of classification over HSR-RS images using deep learning methods along with the conditions of labeled data scarcity. The GeoSystemNet model solved the classification issue based on the genetic uniformity of spatially neighboring objects of various scales and hierarchical levels. However, the GeoSystemNet model required a huge amount of freedom degrees to maintain the classification performance. Syrris et al. [22] developed the SatImNet, which is a group of open training, structured and harmonized data with respect to certain rules. Further, the CNN was modeled to obtain the classification of satellite images.

Finally, the satellite image classification carried out on the Office-31 dataset and the NWPU-Merced-Land satellite image dataset was performed as follows: Hu et al. [23] presented the Coordinate Partial Adversarial Domain Adaptation (CPADA) to perform an unsupervised satellite image classification. The CPADA was used to develop a partial transfer learning technique, and the negative transfer was discarded by coordinate loss using the down-weighting outlier satellite images. The classification of CPADA improved because of the domain-invariant features obtained from the CPADA. However, the features were misaligned, due to the deviation that occurred among the predicted and ideal weights.

The better classification of the satellite images resulted in high classification accuracy. The hypertuned deep learning architecture obtained less accuracy during the classification process due to the vanishing gradient problem [16]. Moreover, the net dropout technique was used to avoid the overfitting issues that affect the classification accuracy due to the removal of relevant features [17]. The DropSample developed for the CNN failed to consider the similarity of the classes, which may affect the classification performance [18]. Moreover, a huge amount of freedom degrees is required for the GeoSystemNet model to maintain the classification performance [21].

Solution

In this paper, classification accuracy over the satellite images is increased using the multiple ensemble features and optimal features selected from the CCGSA technique to select the relevant features to avoid the overfitting problem. The correlation coefficient considered in the feature selection process is used to avoid the irrelevant features from the feature set. The classification accuracy of the HFEL–CCGSA method also improved using image data obtained from the hierarchical framework to maintain the gradient in the network.

The overall organization of the paper is as follows: a detailed explanation of the HFEL–CCGSA method is given in Section 2. Section 3 provides the results and discussion of the HFEL–CCGSA method. Finally, the conclusions are made in Section 4.

2. HFEL–CCGSA Method

In the HFEL–CCGSA method, the satellite image classification is performed using HFEL and CCGSA-based feature selection. In 'without CCGSA', the refined and semantic features are extracted from the fully connected layers of AlexNEt, LeNet-5 and ResNet.

MOST WIEDZY Downloaded from mostwiedzy.pl

Next, the CCGSA-based feature selection is used to select the optimal features from the extracted features. Generally, input images have unknown characteristics and an effective model is required to handle the unknown characteristics of images. The developed method randomly selects the number of images in the model for classification. The proposed method analyzes the effect of the pre-processing, such as normalization and augmentation. Then, k-fold cross-validation is applied to analyze the performance of the developed method in remote sensing classification. Both the extracted and selected features are given as input to the MSVM for classifying the satellite images. Therefore, the utilization of hierarchical images, feature extraction from the CNN and optimal feature selection are used to improve the classification performance. Figure 1 shows a block diagram of the HFEL–CCGSA method. The hierarchial framework provides the data in various manners for feature learning, the CCGSA method selects the features based on correlation, and ensemble learning selects the feature set based on the MSVM model.

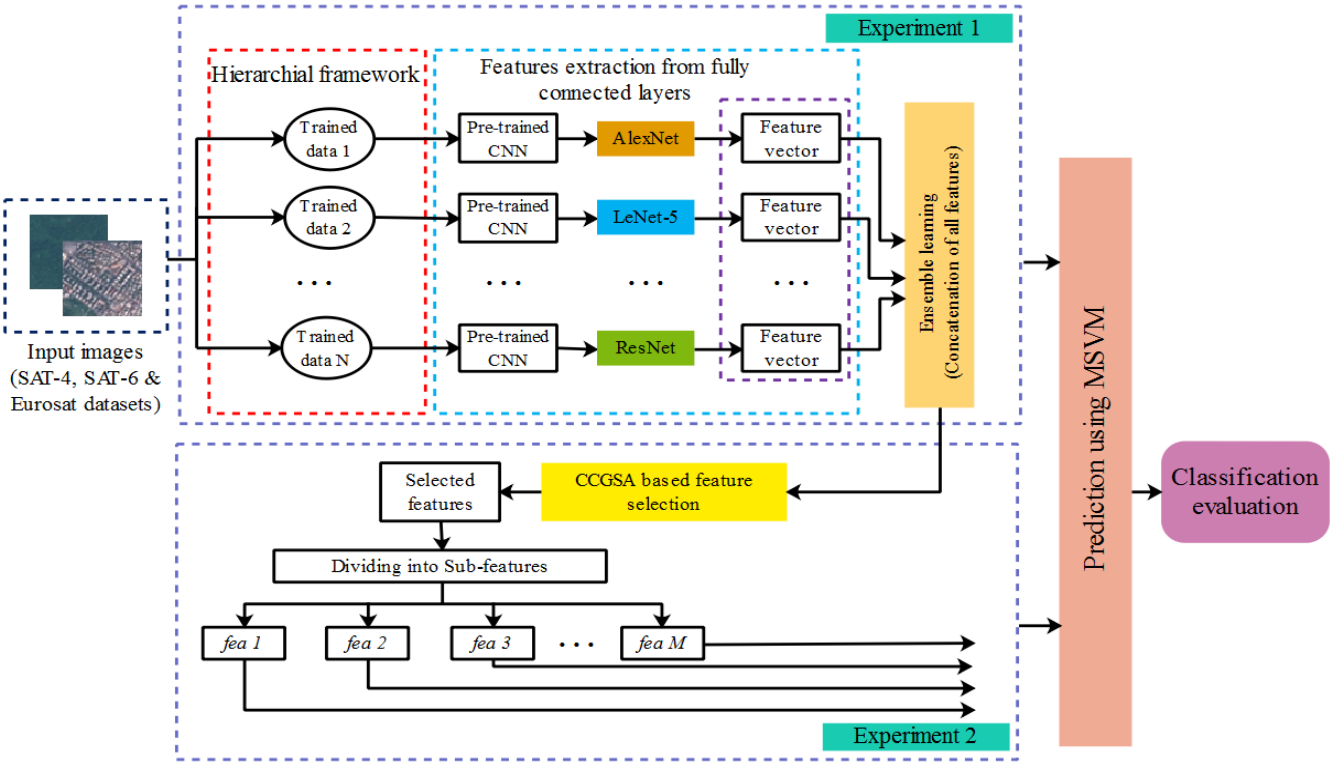

Figure 1. Block diagram of HFEL–CCGSA method.

2.1. Image Acquisition

To evaluate the HFEL–CCGSA method, the tests were carried out on three different datasets, namely the SAT-4, SAT-6 and Eurosat datasets. Both the SAT-4 and SAT-6 datasets were acquired from the NAIP dataset [24]. The SAT-4 multispectral dataset has 500,000 images, including four different classes of images, such as trees, barren land, grasslands and all other land covers. Next, the SAT-6 dataset has 405,000 images and contains different classes of images, such as water bodies, grasslands, barren land, buildings, trees and roads. The images that exist in both the SAT-4 and SAT-6 datasets have the size of 28 × 28 with 1 m spatial resolution, and each image contains red, blue, green and near-infrared (NIR). Figure 2 shows sample images from the SAT-4 and SAT-6 datasets.

MOST WIEDZY Downloaded from mostwiedzy.pl

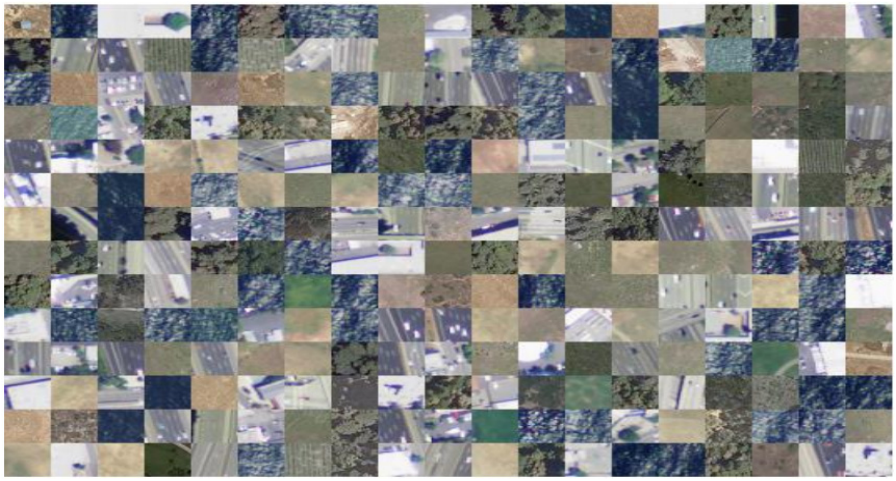

Figure 2. Sample images from SAT-4 and SAT-6 datasets.

In the Eurosat dataset [25], the satellite images are acquired from European cities that are dispersed over 34 countries. The Eurosat dataset has 27,000 labeled and geo-referenced images, where each image has a size of 64 × 64. The Eurosat dataset contains 10 different classes of images, such as residential buildings, industrial buildings, seas and lakes, herbaceous vegetation, highways, pastures, rivers, annual crops, permanent crops and forests, where each class has 2000–3000 images. Additionally, the Eurosat images have 13 bands, namely red, blue, green, aerosols, red edge 1, red edge 2, red edge 3, red edge 4, shortwave infrared 1, shortwave infrared 2, NIR, water vapor and cirrus. Figure 3 shows sample images from the Eurosat dataset.

Figure 3. Sample images from Eurosat dataset.

2.2. Hierarchical Framework

In a hierarchical framework, an ordered set of tasks is accomplished to generate the coarse data of the images. An illustration of the hierarchical framework used in the HFEL–CCGSA method is shown in Figure 4. For example, the input image can be directly given to the CNN for feature extraction or it can be generated after performing the pre-processing. The hierarchy contains the arrangement of items to represent the data at various levels or at the same level. Here, images are arranged as raw images, pre-processed and augumented images to extract features for better representation.

MOST WIEDZY Downloaded from mostwiedzy.pl

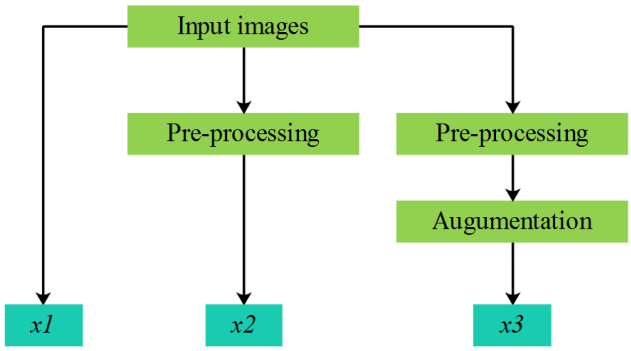

Figure 4. Hierarchical framework.

2.2.1. Image Pre-Processing

After accomplishing the image acquisition, histogram equalization and normalization are performed in order to enhance the image quality. The visual quality of satellite images is improved by varying the pixel value range using normalization, also referred to as contrast stretching. Equation (1) shows the common formulae of the normalization technique.

$$x2 = (x1 - min)\frac{newmax - newmin}{max - min} + newmin \quad (1)$$

where $x1$ specifies the input image subjected to the preprocessing; min and max represent the minimum and maximum intensity values, i.e., from 0 to 255; moreover, $x2$ indicates the image after normalization, and the new minimum and maximum values are represented as $newmin$ and $newmax$, respectively. After normalization, the new image ranges from 0 to 1 and it is subjected to histogram equalization. Next, histogram equalization is applied to adjust the image contrast using the histogram values. Histogram equalization is considered an effective technique to provide a better image without losing information such as points, image patches and edges [26].

2.2.2. Data Augmentation

After performing the pre-processing, the image $(x2)$ is subjected to data augmentation. During data augmentation, the pre-processed image is rotated according to the angles 0°, 90°, 180° and 270°, and it is flipped from left to right. In this way, this documentation generates eight augmented images for a single pre-processed image.

From this hierarchical framework, there are three different image outputs, such as the $x1$-raw input image, $x2$-pre-processed image and $x3$-pre-processed and augmented image. Similarly, the same hierarchical framework is carried out for N amount of input sample images. Further, these outputs from the hierarchical framework are subjected to the feature extraction process, which is explained in the following section.

2.3. Feature Extraction from the CNN

In the HFEL–CCGSA method, ensemble learning is performed by concatenating the features from the CNN. There are three different CNNs, namely AlexNet, LeNet-5 and Resnet, used to extract the optimal features from the input images $x1$, $x2$ and $x3$, respectively. Here, the feature vectors from the fully connected layers are obtained and these feature vectors are concatenated together as follows.

2.3.1. AlexNet

In the HFEL–CCGSA method, AlexNet [27,28] is used to extract the feature vectors from the given input, where the structure of the network has eight different layers, i.e., five convolution layers and three fully connected layers. AlexNet uses an effective activation function, namely the rectified linear unit (RELU), which prevents the gradient vanishing issue. The RELU's gradient is always 1, even when the input is not less than 0. The

MOST WIEDZY Downloaded from mostwiedzy.pl

RELU is also used to improve the training speed, and Equation (2) defines the RELU activation function.

$$y = max(0, x1) \tag{2}$$

where y is the neuron output and $x1$ is the input. Next, AlexNet contains various small sub-networks, which may encounter the overfitting issue. Thus, some layers are dropped out to avoid this. On the other hand, some of the neurons are trained in each iteration during the dropout period. The generalization is improved in AlexNet by minimizing joint adaption among the neurons. The output of AlexNet is an average of all the sub-networks. Therefore, the dropout is used to maximize the robustness. Next, the convolutional layers are used for the automatic extraction of features, and the extracted features are reduced by the pooling layer. The convolution of AlexNet is represented in the following Equation (3).

$$C(h, w) = (x1 \times m), (h, w) = \sum_b \sum_c x1(h - b, w - c)m(b, c) \tag{3}$$

where the input image is represented as $x1$; the width and height of the input image are h and w, respectively; the convolutional kernel is represented as m; the width and height of the convolutional kernel are b and c, respectively. The convolution is used to acquire the features from the image and these parameters are exchanged in order to minimize the model's complexity. Moreover, the feature map is reduced using max pooling in AlexNet. Then, cross-channel normalization (i.e., a local normalization technique) is used to accomplish feature generalization.

$$Softmax\ (x1)_i = \frac{\exp(x1_i)}{\sum_{j=1}^{n} \exp(x1_j)} for\ i = 0,\ 1,\ 2, \ldots k \tag{4}$$

Equation (4) is a Softmax function that is used as an activation function in the fully connected layers for mapping the output between the range of 0 and 1. Next, the feature vectors from the fully connected layers of AlexNet, i.e., fa, are extracted from the output of the fully connected layers and used for ensemble learning.

2.3.2. LeNet-5

LeNet-5 [29] is generally a modified version of the CNN that is applied to the pre-processed image to extract the feature vectors. LeNet-5 has two convolutional layers and two max pooling layers followed by a fully connected layer. The process of LeNet-5 is described as follows.

First, feature extraction is performed using the convolution layer, where each layer has a number of convolutional kernels. In this layer, the input matrix is convolved with the convolution kernel. Consider the input $x2 = \{x2_{ij} | i = 1, 2, \ldots I, j = 1, 2, \ldots J\}$, where I and J are the input image and amount of data in the input image, respectively. The convolution kernel of LeNet-5 is $CK = \{ck_{p,q} | p = 0, 1, \ldots SCK - 1,\ q = 0, 1, \ldots SCK - 1\}$, where SCK defines the size of the convolution kernel. Equation (5) shows the results from the convolutional layer.

$$oc_{i,j} = \left\{ f\left(\sum_{p=0}^{SCK-1} \sum_{q=0}^{SCK-1} ck_{p,q} x2_{i+m,j+q} + ot \right) \right\}_{i=1,\ 2,\ldots I; j=1,2,\ldots,J} \tag{5}$$

where the output obtained from the convolution is represented as $oc_{i,j}$, the offset term is represented as ot and the activation function is represented as $f(.)$. Similar to AlexNet, this LeNet-5 uses the RELU as an activation function.

The dimension of the data is minimized by accomplishing feature selection using the pooling layer. This pooling layer uses the maximum pooling, which used to obtain

MOST WIEDZY Downloaded from mostwiedzy.pl

the points with high values. The operation of maximum pooling $(pool(.))$ is expressed in Equation (6).

$$mp_n^l = pool\left(mp_n^{l-1}\right) \tag{6}$$

where the lth layer and its former layer are represented as mp_n^l and mp_n^{l-1}, respectively, and n represents the nth sample

In general, the fully connected layer is the last layer of the CNN. Here, the RELU activation function is used by each neuron and this RELU links the neurons with those from previous layers. This fully connected layer combines the local information that can differentiate the types of classes one from another. Equation (7) shows the output of the fully connected layer (l).

$$fc_n^l = f\left(ck^l \cdot fc_n^{l-1} + ot^l\right) \tag{7}$$

Hence, the integration of the convolution layer, pooling layer, RELU activation function and fully connected layer is used to accomplish the feature extraction from the input image.

2.3.3. ResNet

ResNet [30] is used to obtain the semantic features from the input image based on its deeper network structure using a residual block. ResNet has five convolutional layer stages, which include 101 layers that are deeper with less redundancy. Due to the deeper network architecture, ResNet effectively extracts the features from the pre-processed and augmented image $(x3)$. The output from the layers of ResNet is expressed in Equation (8).

$$H(x3) = F(x3) + x3 \tag{8}$$

where $H(x3)$ is the required output features from ResNet and $F(x3)$ is the network map. Hence, the ResNet features are referred to as the features extracted from the residual network. Moreover, it is worth stating that the features from AlexNet and LeNet-5 are concatenated along with the features extracted using Resnet (i.e., ensemble learning). The concatenation of all features $(FV1)$ from AlexNet, LeNet-5 and Resnet is expressed in Equation (9).

$$FV1 = \{fa,\ fc,\ H\} \tag{9}$$

The AlexNet features f_a, LeNet-5 features f_c, and ResNet feature H are combined for the feature selection process. The formulated $FV1$ is a feature vector obtained from the combination of hierarchical framework and ensemble learning, which involves appropriate features obtained from the input images. $FV1$ is given as one input for the MSVM, and the selected features from the feature vector are given as an additional input to the MSVM. The feature selection process over the concatenated features is explained in the following section.

2.4. Feature Selection Using CCGSA

In the HFEL–CCGSA method, CCGSA-based feature selection is used to select the optimal features in each stage to minimize the redundancy and interferences caused by the identical features. In some existing feature selection strategies [31,32], classification accuracy is taken as a primary factor for selecting the features. However, the feature selection using the GSA considers the correlation coefficient [33] as the primary factor. The reason for not selecting the classification accuracy is that it requires an additional classifier in the feature selection process. The process of the HFEL–CCGSA method already contains two levels, i.e., it includes HFEL and a feature selection process for classification purposes. Instead of increasing the complexity of the HFEL–CCGSA method, the GSA considers the correlation coefficient of the features; hence, the CCGSA is developed for selecting the features. The solution to the feature selection issue is represented using a binary representation. The selection of the respective feature is represented by each element

MOST WIEDZY Downloaded from mostwiedzy.pl

of the binary solution. Hence, a solution of the CCGSA considers 1 (selected) or 0 (not selected) in each dimension.

Generally, the GSA [34] is designed for solving the optimization issue in the continuous search space, but in order to use the CCGSA in the process of searching for features, solutions are required to convert the results into binary form. Therefore, the continuous solutions are converted into binary form using the transfer function that is expressed in Equation (10).

At this time, each solution present in the velocity vector defines the probability of flipping the respective feature from selected to not selected and vice versa. Specifically, flipping occurs when the dimension has high velocity values; otherwise, when the dimension has lower velocity values, there is no flipping in the feature selection.

$$TF\left(v_d^i(t)\right) = \left|\tanh\left(v_d^i(t)\right)\right| \tag{10}$$

where the transfer function is represented as TF; the ith element velocity in dimension d and iteration t is represented as $v_d^i(t)$.

The output from Equation (10) is utilized to change the solution location to 1 or 0 as shown in Equation (11).

$$FV2(t+1) = \begin{cases} -FV1 & r < TF\left(v_k^i(t)\right) \\ FV1 & r \geq TF\left(v_k^i(t)\right) \end{cases} \tag{11}$$

where the random number generated between the range of 0 and 1 is represented as r.

The derivation of the fitness function is important to select the appropriate features from the feature set $(FV1)$. In this CCGSA, the efficiency of the solution is represented by the solution's physical mass, i.e., the optimal feature set has a high mass when compared to the worst feature subset. The optimal selection of the features is obtained by considering the number of features in the subset and the correlation coefficient. Here, the selected feature subset should have a lower number of features and less inter-correlation of features. The fitness function used in the CCGSA is expressed in Equation (12).

$$Fitness = min\left[\delta\frac{R\overline{r_{cf}}}{\sqrt{R + R(R-1)\overline{r_{ff}}}} + \beta\frac{|R|}{|A|}\right] \tag{12}$$

where the first term represents the correlation coefficient; the amount of selected features is represented as R; the amount of total features in the $FV1$ is represented as A; the average feature to feature inter-correlation and feature to class correlation are represented as $\overline{r_{ff}}$ and $\overline{r_{cf}}$, respectively. Therefore, this feature selection is used to eliminate the redundant and irrelevant features from the feature vectors. The selected features $(FV2)$ from the feature vector are divided into M sets and are given as the second input to the voting-based prediction. The features from CNN models have a size of $n \times 1000$ and the feature selection model of CCGSA selects 0.8 correlated features from extracted features.

2.5. Classification Using MSVM

Finally, the features from 'without CCGSA' (i.e., concatenated features) and HFEL-CCGSA (i.e., selected features) are given as input to the MSVM in order to obtain the precise classification of satellite images. To create the MSVM, a multiple-binary Support Vector Machine (SVM) classifier is combined during the classification process [35]. Here, the One Versus All (OVA) method is used to solve the issues related to the multi-class SVM. The C-SVM models are established in the OVA for C amount of classes.

3. Results and Discussion

The results and discussion of the satellite image classification using the HFEL–CCGSA method are given in this section. The HFEL–CCGSA method was implemented and

MOST WIEDZY Downloaded from mostwiedzy.pl

simulated using MATLAB R2019a, where the PC was operated with an i9 Intel core processor, 128 GB RAM, 3 TB hard disk and Windows 10 operating system (64-bit). The HFEL-CCGSA method was analyzed with three different datasets, namely the SAT-4, SAT-6 and Eurosat datasets. Here, the performance is analyzed using different metrics such as accuracy, precision and recall, which are defined as follows.

(i) Accuracy

Accuracy is defined as the total number of precise predictions obtained using the HFEL–CCGSA method, and Equation (13) expresses the accuracy.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{13}$$

where TP and TN are the true positive and true negative; FP and FN are the false positive and false negative.

(ii) Precision

Precision is the ratio among the true positives and the sum of true positives and false positives and is expressed in Equation (14).

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{14}$$

(iii) Recall

Recall is the ratio among the true positives and true positives and false negatives, which is shown in Equation (15).

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{15}$$

3.1. Quantitative Analysis on SAT-4 Dataset

This section provides the quantitative analysis of the SAT-4 dataset, which contains 500,000 satellite images. In these images, 70% of the data were used for training and 30% of the data were used for testing purposes. Here, four different classes were considered for classification, namely grassland, trees, barren land and a class that covered all land cover classes excluding grassland, trees and barren land. The quantitative analysis of the HFEL–CCGSA method for the SAT-4 dataset was carried out in three different ways, which are given as follows.

Table 1 shows the performance comparison between the case 'without CCGSA' and the HFEL–CCGSA method on the SAT-4 dataset. Here, the case 'without CCGSA' was individually processed using the MSVM classifier. From Table 1, it can be concluded that the combination in the case 'without CCGSA' performed well when compared to the individual experiments. The reason for the integration of 'without CCGSA' is that it uses multiple ensembles of features at the same time for the classification performance. Moreover, the features from the images (i.e., hierarchical framework) lead to improved classification.

Tables 2 and 3 show the performance analysis of the individual CNN and HFEL–CCGSA (AlexNet + LeNet-5 + ResNet) and different feature selection methods and HFEL–CCGSA, respectively. The combination of CNNs, i.e., AlexNet + LeNet-5 + ResNet, provides better classification because it extracts the feature vectors from the fully connected layers of all CNNs. As the satellite images from the datasets have different angles and perceptions, different types of CNN are used in the HFEL–CCGSA to obtain the refined and semantic features from the images. Hence, the combination of different CNNs, namely AlexNet, LeNet-5 and ResNet, can be used to obtain better classification. Moreover, the PSO and BDA were used for the comparison of the different feature selection methods, because all the optimization algorithms consider velocity to update the location of the population. From Table 3, it can be seen that the classification performance was improved using the CCGSA,

MOST WIEDZY Downloaded from mostwiedzy.pl

because it selects the features mainly based on the correlation coefficient of features. Hence, the CCGSA can be used to avoid the irrelevant features from the image features, which helps to improve the accuracy up to 99.99% for satellite image classification.

Table 1. Performance analysis for 'without CCGSA' and HFEL–CCGSA method on SAT-4 dataset.

Experiments	Performance	Classes				
		Grassland	Tree	Barren Land	Others	Overall
without CCGSA	Precision (%)	98.87	97.12	98.15	97.12	97.81
	Recall (%)	100	98.84	97.68	98.49	98.75
	Accuracy (%)	98.40	97.68	98.09	98.76	98.23
HFEL–CCGSA (Experiment 1 + Experiment 2)	Precision (%)	99.98	99.94	99.96	99.97	99.96
	Recall (%)	99.79	99.96	99.97	99.93	99.91
	Accuracy (%)	100	99.98	100	99.98	99.99

Table 2. Performance analysis for individual CNN and HFEL–CCGSA (AlexNet + LeNet-5 + ResNet) on SAT-4 dataset.

CNN	Performance	Classes				
		Grassland	Tree	Barren Land	Others	Overall
EL with AlexNet	Precision (%)	99.45	98.97	97.18	99.17	98.69
	Recall (%)	99.08	97.17	99.38	98.45	98.52
	Accuracy (%)	99.97	98.99	99.76	99.19	99.47
EL with LeNet-5	Precision (%)	98.47	99.08	98.07	99.03	98.66
	Recall (%)	97.08	97.67	98.79	98.97	98.12
	Accuracy (%)	98.76	99.08	98.68	97.04	98.39
EL with ResNet	Precision (%)	99.01	97.56	98.07	97.43	98.01
	Recall (%)	98.43	98.89	98.69	99.18	98.79
	Accuracy (%)	97.78	97.04	99.02	97.99	97.95
HFEL–CCGSA (AlexNet + LeNet-5 + ResNet)	Precision (%)	99.98	99.94	99.96	99.97	99.96
	Recall (%)	99.79	99.96	99.97	99.93	99.91
	Accuracy (%)	100	99.98	100	99.98	99.99

Table 3. Performance analysis for different feature selection methods and HFEL–CCGSA on SAT-4 dataset.

Feature Selection Methods	Performance	Classes				
		Grassland	Tree	Barren Land	Others	Overall
Particle Swarm Optimization (PSO)	Precision (%)	98.49	97.06	97.37	98.16	97.77
	Recall (%)	99.08	99.28	98.78	97.33	98.61
	Accuracy (%)	98.19	98.15	99.14	98.02	98.37
Binary Dragonfly Algorithm (BDA)	Precision (%)	99.07	99.12	98.06	98.79	98.76
	Recall (%)	98.46	99.88	97.67	99.02	98.75
	Accuracy (%)	98.97	97.76	98.44	97.55	98.18
HFEL–CCGSA (CCGSA)	Precision (%)	99.98	99.94	99.96	99.97	99.96
	Recall (%)	99.79	99.96	99.97	99.93	99.91
	Accuracy (%)	100	99.98	100	99.98	99.99

MOST WIEDZY Downloaded from mostwiedzy.pl

3.2. Quantitative Analysis on SAT-6 Dataset

This section provides the quantitative analysis of the SAT-6 dataset, which has 40,500 satellite images. These 40,500 satellite images contain six different land cover classes, such as grassland, trees, barren land, roads, buildings and water bodies. From the 40,500 satellite images, 70% of the images were used for training and 30% of the images were used for testing purposes. The performance analysis of the HFEL–CCGSA method for the SAT-6 dataset was accomplished in three different ways, which are as follows.

The comparison among the case 'without CCGSA' and the HFEL–CCGSA method for the SAT-6 dataset is presented in Table 4. Table 4 shows that the HFEL–CCGSA achieves higher classification accuracy when compared to the case 'without CCGSA'. The ensemble of multiple semantic and refined features can be used to obtain a higher classification accuracy of 99.99% than in the case 'without CCGSA'.

Table 4. Performance analysis for 'without CCGSA' and HFEL–CCGSA method on SAT-6 dataset.

Experiments	Performance	Classes						
		Grassland	Trees	Barren Land	Roads	Buildings	Water Bodies	Overall
without CCGSA	Precision (%)	98.46	99.78	99.67	98.42	98.06	98.46	98.8
	Recall (%)	99.08	97.67	97.99	99.05	99.45	98.05	98.54
	Accuracy (%)	97.37	98.06	99.07	97.09	98.67	99.05	98.21
HFEL–CCGSA (Experiment 1 + Experiment 2)	Precision (%)	99.88	99.95	99.96	99.93	99.98	99.94	99.94
	Recall (%)	99.97	99.98	100	99.99	99.97	99.89	99.96
	Accuracy (%)	99.99	99.98	99.99	100	99.98	100	99.99

The performance comparison of the individual CNNs and HFEL-CCGSA (AlexNet + LeNet-5 + ResNet) and the different feature selection methods and CCGSA for the SAT-6 dataset are shown in Tables 5 and 6, respectively. From Tables 5 and 6, it can be concluded that the performance of the HFEL–CCGSA method is better than that of the individual CNNs and PSO and BDA methods. The reduction in outliers in the ensemble learning (i.e., AlexNet + LeNet-5 + ResNet) can be used to improve the classification performance of the HFEL–CCGSA method. On the other hand, the correlation coefficient-based feature selection method can be used to extract the refined features from the extracted feature vectors, which can be used to increase the accuracy up to 99.99% compared to the other feature selection methods.

Table 5. Performance analysis for individual CNNs and HFEL–CCGSA (AlexNet + LeNet-5 + ResNet) on SAT-6 dataset.

CNN	Performance	Classes						
		Grassland	Trees	Barren Land	Roads	Buildings	Water Bodies	Overall
EL with AlexNet	Precision (%)	98.45	99.94	98.16	97.08	99.47	98.47	98.59
	Recall (%)	99.98	99.76	99.47	98.02	98.89	98.73	99.14
	Accuracy (%)	99.12	98.74	99.08	98.57	99.37	99.43	99.05
EL with LeNet-5	Precision (%)	98.06	99.52	97.69	98.77	98.84	98.64	98.58
	Recall (%)	98.15	97.17	98.87	99.34	98.58	97.35	98.24
	Accuracy (%)	99.01	98.05	98.64	99.33	99.22	99.03	98.88
EL with ResNet	Precision (%)	97.24	98.99	97.08	98.42	99.33	98.11	98.19
	Recall (%)	99.05	99.15	98.49	99.67	98.42	99.08	98.97
	Accuracy (%)	98.46	98.66	97.88	98.94	98.64	99.57	98.69
HFEL–CCGSA (AlexNet + LeNet-5 + ResNet)	Precision (%)	99.88	99.95	99.96	99.93	99.98	99.94	99.94
	Recall (%)	99.97	99.98	100	99.99	99.97	99.89	99.96
	Accuracy (%)	99.99	99.98	99.99	100	99.98	100	99.99

MOST WIEDZY Downloaded from mostwiedzy.pl

Table 6. Performance analysis for different feature selection methods and HFEL–CCGSA on SAT-6 dataset.

Feature Selection Methods	Performance	Classes						
		Grassland	Trees	Barren Land	Roads	Buildings	Water Bodies	Overall
PSO	Precision (%)	98.89	98.77	97.08	98.46	99.52	99.08	98.63
	Recall (%)	99.33	97.45	98.69	99.67	98.68	98.88	98.78
	Accuracy (%)	98.45	98.42	99.07	97.37	99.47	99.77	98.75
BDA	Precision (%)	99.08	98.47	98.55	97.66	99.06	98.94	98.62
	Recall (%)	98.62	98.66	99.11	98.88	98.79	98.63	98.78
	Accuracy (%)	98.33	99.08	99.03	98.67	98.79	98.99	98.81
HFEL–CCGSA (CCGSA)	Precision (%)	99.88	99.95	99.96	99.93	99.98	99.94	99.94
	Recall (%)	99.97	99.98	100	99.99	99.97	99.89	99.96
	Accuracy (%)	99.99	99.98	99.99	100	99.98	100	99.99

3.3. Quantitative Analysis on Eurosat Dataset

This section provides the quantitative analysis of the Eurosat dataset, which contains 27,000 satellite images. In these images, 70% of the data were used for training and 30% of the data were used for testing purposes. The Eurosat dataset has 12 different land use and land cover classes, such as residential buildings, industrial buildings, highways, permanent crops, pastures, forests, herbaceous vegetation, annual crops, rivers, seas and lakes. The quantitative analysis of the HFEL–CCGSA method for the Eurosat dataset was performed in three different ways, which are given as follows.

Tables 7–9 show the performance comparison for the different experiments, CNNs and feature selection methods, respectively. From the analysis, it can be concluded that the HFEL–CCGSA method provides better performance than the individual experiments, the CNNs and the PSO and BDA feature selection methods. For example, the classification accuracy of the HFEL–CCGSA method is 99.49%, which is high when compared to the case 'without CCGSA'. The HFEL–CCGSA method achieves better performance due to its ensemble of multiple features and the appropriate feature selection from the images.

3.4. Comparative Analysis

The comparative analysis of the HFEL–CCGSA method with existing satellite image classification methods is described in this section. The existing methods considered for the comparison were the two-band AlexNet [16], hyperparameter-tuned AlexNet [16], two-band ConvNet [16], hyperparameter-tuned ConvNet [16], two-band VGG [16], hyperparameter-tuned VGG [16], DCCNN [17], MGSS [19] and GeoSystemNet [21]. Here, the comparison was made between three different datasets, namely the SAT-4, SAT-6 and Eurosat datasets.

Table 7. Performance analysis for 'without CCGSA' and HFEL–CCGSA method on Eurosat dataset.

Experiments	Performance	Overall
without CCGSA	Precision (%)	98.15
	Recall (%)	99.67
	Accuracy (%)	98.56
HFEL–CCGSA	Precision (%)	98.93
	Recall (%)	99.15
	Accuracy (%)	99.49

MOST WIEDZY Downloaded from mostwiedzy.pl

Table 8. Performance analysis for individual CNNs and HFEL–CCGSA (AlexNet + LeNet-5 + ResNet) on Eurosat dataset.

CNN	Performance	Overall
EL with AlexNet	Precision (%)	98.42
	Recall (%)	99.11
	Accuracy (%)	98.99
EL with LeNet-5	Precision (%)	98.22
	Recall (%)	97.54
	Accuracy (%)	98.42
EL with ResNet	Precision (%)	97.38
	Recall (%)	96.44
	Accuracy (%)	97.45
HFEL–CCGSA (AlexNet + LeNet-5 + ResNet)	Precision (%)	98.93
	Recall (%)	99.15
	Accuracy (%)	99.49

Table 9. Performance analysis for different feature selection methods and HFEL–CCGSA on Eurosat dataset.

Feature Selection Methods	Performance	Overall
PSO	Precision (%)	98.04
	Recall (%)	98.74
	Accuracy (%)	96.48
BDA	Precision (%)	97.11
	Recall (%)	98.57
	Accuracy (%)	97.57
HFEL–CCGSA (CCGSA)	Precision (%)	98.93
	Recall (%)	99.15
	Accuracy (%)	99.49

Table 10 shows the comparative analysis of the HFEL–CCGSA method for three different datasets, namely the SAT-4, SAT-6 and Eurosat datasets. From Table 10 and Figure 5, it can be concluded that the HFEL–CCGSA method performs well when compared to the existing satellite image classification techniques. The classification accuracy of the HFEL–CCGSA method is 99.99% for both the SAT-4 and SAT-6 datasets, which is high when compared to the existing classification techniques. Additionally, the classification accuracy of the HFEL–CCGSA method is 99.49% for the Eurosat dataset, which is high when compared to the GeoSystemNet model [21]. The DCCNN [17] achieves lower classification accuracy, because of the net dropout technique used to avoid the overfitting issue. Moreover, the classification performance of the GeoSystemNet model [21] is maintained only when using a huge amount of freedom degrees. However, the classification accuracy of the HFEL–CCGSA method is improved when using both low- and high-level image data and an ensemble of multiple features. Furthermore, the accuracy of the HFEL–CCGSA method is improved by selecting the optimal features during the classification.

MOST WIEDZY Downloaded from mostwiedzy.pl

Table 10. Comparative analysis of HFEL–CCGSA method on 5-fold cross-validation.

Dataset	Method	Classification Accuracy (%)
SAT-4 dataset	2-Band AlexNet [16]	99.66
	Hyperparameter-Tuned AlexNet [16]	98.45
	2-Band ConvNet [16]	99.03
	Hyperparameter-Tuned ConvNet [16]	98.45
	2-Band VGG [16]	99.03
	Hyperparameter-Tuned VGG [16]	98.59
	DCCNN [17]	98.00
	MGSS [19]	99.97
	HFEL–CCGSA	99.99
SAT-6 dataset	2-Band AlexNet [16]	99.08
	Hyperparameter-Tuned AlexNet [16]	97.43
	2-Band ConvNet [16]	99.10
	Hyperparameter-Tuned ConvNet [16]	97.48
	2-Band VGG [16]	99.15
	Hyperparameter-Tuned VGG [16]	97.95
	DCCNN [17]	97.00
	MGSS [19]	99.95
	HFEL–CCGSA	99.99
Eurosat dataset	GeoSystemNet [21]	95.30
	HFEL–CCGSA	99.49

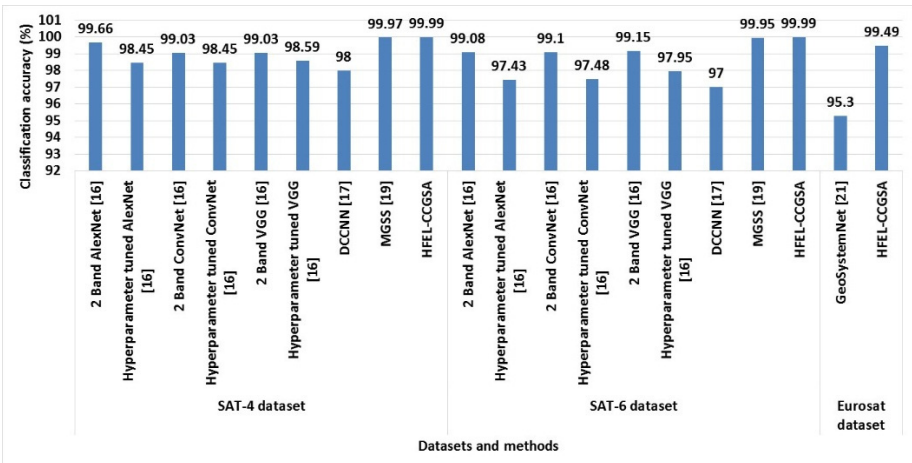

Figure 5. Comparison graph for classification accuracy.

4. Conclusions

In this paper, the combination of HFEL and CCGSA is used for the precise classification of satellite images from the SAT-4, SAT-6 and Eurosat datasets. The hierarchical framework includes data pre-processing and augmentation, which are used to generate the images. Subsequently, three different CNNs, namely AlexNet, LeNet-5 and ResNet, are used to extract appropriate features from hierarchical images. On the other hand, the CCGSA is used to eliminate the redundant features from the selected features based on the correlation coefficient. Further, the MSVM effectively classifies satellite images using both multiple ensemble features and selected features. From the performance analysis, it is concluded

MOST WIEDZY Downloaded from mostwiedzy.pl

that the HFEL–CCGSA method provides better performance than the existing methods. The classification accuracy of the HFEL–CCGSA method analyzed on the Eurosat dataset is 99.49%, which is lower when compared to the GeoSystemNet model.

Author Contributions: The paper investigation, resources, data curation, writing—original draft preparation, writing—review and editing, and visualization were performed by K.T. and M.M.A. The paper conceptualization, software, validation, and formal analysis, and methodology were performed by A.S. Supervision, project administration, and final approval of the version to be published were conducted by P.B.D. and H.K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in NAIP dataset at doi:10.1145/2820783.2820816, reference number [24] and Eurosat dataset at doi: 10.1109/JSTARS.2019.2918242, reference number [25].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Banerjee, B.; Bovolo, F.; Bhattacharya, A.; Bruzzone, L.; Chaudhuri, S.; Mohan, B.K. A new self-training-based unsupervised satellite image classification technique using cluster ensemble strategy. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 741–745. [CrossRef]
2. Zhang, C.; Chen, Y.; Yang, X.; Gao, S.; Li, F.; Kong, A.; Zu, D.; Sun, L. Improved remote sensing image classification based on multi-scale feature fusion. *Remote Sens.* **2020**, *12*, 213. [CrossRef]
3. Pan, Z.; Xu, J.; Guo, Y.; Hu, Y.; Wang, G. Deep learning segmentation and classification for urban village using a worldview satellite image based on U-Net. *Remote Sens.* **2020**, *12*, 1574. [CrossRef]
4. Xia, M.; Tian, N.; Zhang, Y.; Xu, Y.; Zhang, X. Dilated multi-scale cascade forest for satellite image classification. *Int. J. Remote Sens.* **2020**, *41*, 7779–7800. [CrossRef]
5. Bekaddour, A.; Bessaid, A.; Bendimerad, F.T. Multi spectral satellite image ensembles classification combining k-means, LVQ and SVM classification techniques. *J. Indian Soc. Remote Sens.* **2015**, *43*, 671–686. [CrossRef]
6. Pelletier, C.; Valero, S.; Inglada, J.; Champion, N.; Marais Sicre, C.; Dedieu, G. Effect of training class label noise on classification performances for land cover mapping with satellite image time series. *Remote Sens.* **2017**, *9*, 173. [CrossRef]
7. Bhatt, M.S.; Patalia, T.P. Content-based high-resolution satellite image classification. *Int. J. Inf. Technol.* **2019**, *11*, 127–140. [CrossRef]
8. Senthilnath, J.; Kulkarni, S.; Benediktsson, J.A.; Yang, X.S. A novel approach for multispectral satellite image classification based on the ba1t algorithm. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 599–603. [CrossRef]
9. do Nascimento Bendini, H.; Fonseca, L.M.G.; Schwieder, M.; Körting, T.S.; Rufin, P.; Sanches, I.D.A.; Leitao, P.J.; Hostert, P. Detailed agricultural land classification in the Brazilian cerrado based on phenological information from dense satellite image time series. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101872. [CrossRef]
10. Pelletier, C.; Webb, G.I.; Petitjean, F. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sens.* **2019**, *11*, 523. [CrossRef]
11. Ngo, L.T.; Mai, D.S.; Pedrycz, W. Semi-supervising Interval Type-2 Fuzzy C-Means clustering with spatial information for multi-spectral satellite image classification and change detection. *Comput. Geosci.* **2015**, *83*, 1–16. [CrossRef]
12. Liu, Q.; Hang, R.; Song, H.; Li, Z. Learning multiscale deep features for high-resolution satellite image scene classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 117–126. [CrossRef]
13. Yu, H.; Yang, W.; Xia, G.S.; Liu, G. A color-texture-structure descriptor for high-resolution satellite image classification. *Remote Sens.* **2016**, *8*, 259. [CrossRef]
14. Yang, W.; Yin, X.; Xia, G.S. Learning high-level features for satellite image classification with limited labeled samples. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4472–4482. [CrossRef]
15. Zhang, G.; Lu, S.; Zhang, W. CAD-Net: A context-aware detection network for objects in remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 10015–10024. [CrossRef]
16. Unnikrishnan, A.; Sowmya, V.; Soman, K.P. Deep learning architectures for land cover classification using red and near-infrared satellite images. *Multimed. Tools Appl.* **2019**, *78*, 18379–18394. [CrossRef]
17. Jiang, J.; Liu, F.; Xu, Y.; Huang, H. Multi-spectral RGB-NIR image classification using double-channel CNN. *IEEE Access* **2019**, *7*, 20607–20613. [CrossRef]

MOST WIEDZY Downloaded from mostwiedzy.pl

18. Yang, N.; Tang, H.; Yue, J.; Yang, X.; Xu, Z. Accelerating the Training Process of Convolutional Neural Networks for Image Classification by Dropping Training Samples Out. *IEEE Access* **2020**, *8*, 142393–142403. [CrossRef]
19. Weng, L.; Qian, M.; Xia, M.; Xu, Y.; Li, C. Land Use/Land Cover Recognition in Arid Zone Using A Multi-dimensional Multi-grained Residual Forest. *Comput. Geosci.* **2020**, *144*, 104557. [CrossRef]
20. Zhong, Y.; Fei, F.; Liu, Y.; Zhao, B.; Jiao, H.; Zhang, L. SatCNN: Satellite image dataset classification using agile convolutional neural networks. *Remote Sens. Lett.* **2017**, *8*, 136–145. [CrossRef]
21. Yamashkin, S.A.; Yamashkin, A.A.; Zanozin, V.V.; Radovanovic, M.M.; Barmin, A.N. Improving the Efficiency of Deep Learning Methods in Remote Sensing Data Analysis: Geosystem Approach. *IEEE Access* **2020**, *8*, 179516–179529. [CrossRef]
22. Syrris, V.; Pesek, O.; Soille, P. SatImNet: Structured and Harmonised Training Data for Enhanced Satellite Imagery Classification. *Remote Sens.* **2020**, *12*, 3358. [CrossRef]
23. Hu, J.; Tuo, H.; Wang, C.; Zhong, H.; Pan, H.; Jing, Z. Unsupervised satellite image classification based on partial transfer learning. *Aerosp. Syst.* **2019**, *3*, 21–28. [CrossRef]
24. Basu, S.; Ganguly, S.; Mukhopadhyay, S.; DiBiano, R.; Karki, M.; Nemani, R. DeepSat-A Learning framework for Satellite Imagery. In Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 3–6 November 2015; ACM SIGSPATIAL: New York, NY, USA, 2015; pp. 1–10.
25. Helber, P.; Bischke, B.; Dengel, A.; Borth, D. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 2217–2226. [CrossRef]
26. Abdullah-Al-Wadud, M.; Kabir, M.H.; Dewan, M.A.A.; Chae, O. A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **2007**, *53*, 593–600. [CrossRef]
27. Han, X.; Zhong, Y.; Cao, L.; Zhang, L. Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. *Remote Sens.* **2017**, *9*, 848. [CrossRef]
28. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]
29. Wei, G.; Li, G.; Zhao, J.; He, A. Development of a LeNet-5 gas identification CNN structure for electronic noses. *Sensors* **2019**, *19*, 217. [CrossRef] [PubMed]
30. Zhang, J.; Sun, J.; Wang, J.; Yue, X.G. Visual object tracking based on residual network and cascaded correlation filters. *J. Ambient. Intell. Humaniz. Comput.* **2020**, *12*, 8427–8440. [CrossRef]
31. Taradeh, M.; Mafarja, M.; Heidari, A.; Faris, H.; Aljarah, I.; Mirjalili, S.; Fujita, H. An Evolutionary Gravitational Search-based Feature Selection. *Inf. Sci.* **2019**, *497*, 219–239. [CrossRef]
32. Mafarja, M.; Aljarah, I.; Heidari, A.A.; Faris, H.; Fournier-Viger, P.; Li, X.; Mirjalili, S. Binary dragonfly optimization for feature selection using time-varying transfer functions. *Knowl.-Based Syst.* **2018**, *161*, 185–204. [CrossRef]
33. Akhtar, S.; Hussain, F.; Raja, F.; Ehatisham-ul-Haq, M.; Baloch, N.; Ishmanov, F.; Zikria, Y. Improving Mispronunciation Detection of Arabic Words for Non-Native Learners Using Deep Convolutional Neural Network Features. *Electronics* **2020**, *9*, 963. [CrossRef]
34. Rashedi, E.; Nezamabadi-Pour, H.; Saryazdi, S. GSA: A gravitational search algorithm. *Inf. Sci.* **2009**, *179*, 2232–2248. [CrossRef]
35. Gangsar, P.; Tiwari, R. Taxonomy of induction-motor mechanical-fault based on time-domain vibration signals by multiclass SVM classifiers. *Intell. Ind. Syst.* **2016**, *2*, 269–281. [CrossRef]

MOST WIEDZY Downloaded from mostwiedzy.pl