Journal of Geomatics Science and Technology, vol. 14, no. 2, pp. 119-133, December 2024- Research article

Land Cover and Land Use Extraction Based on Deep Learning Methods Using Satellite Images

Pooya Heidari [1]*, Asghar Milan [2], Alireza Gharagozlou [3]

1 Master student in Photogrammetry, Faculty of Civil, Water, and Environmental Engineering, Shahid Beheshti University (SBU), Tehran, Iran, po.heidari@mail.sbu.ac.ir

2 Assistant Professor, Faculty of Civil, Water, and Environmental Engineering, Shahid Beheshti University (SBU), Tehran, Iran, a_milan@sbu.ac.ir

3 Associate Professor, Faculty of Civil, Water, and Environmental Engineering, Shahid Beheshti University (SBU), Tehran, Iran, a_gharagozlo@sbu.ac.ir

(*Received: December 2023, Accepted: July 2024*)

Abstract

Information on land use and cover needs to be gathered due to the growing urban population, city growth, and urbanization. Applications for this data include environmental protection, urban planning, planning for urban infrastructure, and strategic planning to guarantee the sustainable growth of urban areas. The primary source of data on land cover and land use at the moment is remote sensing imagery. Information about land cover and land use can be retrieved from remote sensing images using image classification techniques. In terms of classification accuracy, deep learning techniques recently outperformed other methods for classifying land use and cover. Convolutional neural networks (CNNs), which are quite popular in this field, are one of the significant deep learning classification architectures frequently used in land cover and land use classification. Recently, the convolutional neural network technique known as ResNet has been applied to remote sensing applications, particularly for the classification of land use and cover. ResNet models are an effective choice for classifying land cover and land use because they can handle the vanishing gradient issue. The primary objective of this study is to assess the performance of the Glorot Uniform and Random Uniform weight initializers in the ResNet50, ResNet101, and ResNet152 architectures for extracting the land cover and land use of the EuroSat dataset. The weighted F1 score, IoU indexes, overall accuracy, and kappa coefficient were used to evaluate the accuracy of the results. ResNet101's corresponding values for these indexes were, in turn, 0.8869, 0.7951, 0.8871, and 0.8743. These results indicate that, in terms of classification accuracy, ResNet101 has outperformed the ResNet50 and ResNet152 methods.

Keywords: Land cover and land use, Sustainable development, Deep learning, Convolutional neural network, Kappa coefficien

* Corresponding author

1-Introduction

The population of cities has grown in recent years due to urban growth and development. The proportion of people living in cities worldwide grew from roughly 30% in 1950 to 55% in 2018 and is expected to reach 68% by 2050 [1]. Currently, cities' patterns of land use and cover are changing as a result of urbanization and rapid population growth [2]. Information on land use and cover is essential for many aspects of environmental research and studies of global change, as urban areas grow due to economic development and population growth. Planning for long-term economic development and short-term land management can both benefit from timely and accurate national data on land cover and use. The applications for such information include infrastructure planning, urban planning, monitoring urban growth, and environmental protection. Assessing climate change, soil erosion, urban climate, urban heat islands, ecosystem damage, flood forecasting, and natural disasters can all benefit from using this data. In addition, to ensure sustainable urban development and meet the demands of a growing urban population, an appropriate urban land cover and land use planning method is crucial. The primary data sources for accurate land cover and land use data are remote sensing images, which can also be used to generate land cover and land use data [3].

Since Francis J. Marschner started using aerial photographs to create maps of the whole United States in 1940, remote sensing data has been used to classify land use and cover [4]. Following the launch of a multispectral satellite in July 1972 and the initiation of the Landsat program, research on the classification of land use and cover using remote sensing images has advanced to a certain point. The establishment of the Landsat program and the release of its data brought new challenges in data fusion, temporal change detection, and ecological applications in land cover and land use. Maps of urban land cover and land use are frequently created using data and images from remote sensing [5]. Land cover and land use maps are frequently created using remote sensing data from satellites, aircraft, and drones [6]. The incredible progress in sensor technology has made a vast amount of remote sensing imagery available. These images become the main source for spatial data extraction and can be beneficial in a range of research applications. Large areas can be covered by medium resolution images, that tend to be accessible for free. Information regarding land use and cover can be extracted using them as well. Field surveys and conventional interpretations are the main methods used to extract and retrieve such data, which can be labor-and time-intensive. Furthermore, considering the environment is constantly evolving, data on land use and cover must be updated frequently. As a result, it is important to develop effective and efficient methods for the automated extraction of land cover and land use data, with classification serving as the main method to generate maps of land cover and land use.

Deep learning algorithms have numerous applications in studies regarding remote sensing [7], with image classification being one of its most significant applications [8]. Since image classification is so important in computer vision, deep learning has become one of the most powerful techniques in the field owing to the ongoing advancement of artificial intelligence algorithms [9]. The field of remote sensing has brought more attention to deep learning since 2014, and deep learning algorithms have shown impressive results in a range of image analysis tasks, such as classifying land use and cover [10]. Once compared with other approaches, deep learning-based classification methods have numerous of advantages in regards to classification accuracy. Methods for classifying land cover based on deep learning have a lot of potential for carrying out such tasks [11].

Deep learning can model hierarchical representations of features, and such features can be used to describe urban land cover and land use patterns. Urban land cover and land use can be classified extremely effectively using deep learning models [12]. Deep learning outperforms traditional methods in the extraction of multi-level spatial features from remote sensing images, allowing for high performance in image classification [13]. The most significant of the primary advantages of deep learning approaches is the fact that they can adaptively learn different features from images [14], unlike traditional classification methods that describe land cover and land use based on spectral or spectral-spatial features [15]. High performance and flexibility are achieved by deep learning methods [16], which can also be used to learn high-level semantic features [17] and efficiently organize multiple levels of information to express complex relationships between data [18]. Neural network architecture serves as the foundation for deep learning classifiers. A neural network's input layers are linked to output layers through hidden layers to form its basic structure. These layers convert the input data into output data by using parameters and

activation functions. Each neuron's value is modified by the weights of the connecting nodes to determine how the input values are converted into output values [19]. A neural network with many more layers and parameters than a particular neural network is referred to as "deep learning" [20]. The structure of deep learning is shown in Figure 1.

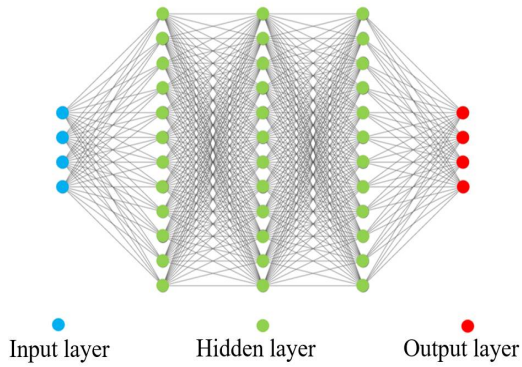

Figure 1: General structure of deep learning

Recurrent neural networks (RNNs) and convolutional neural networks (CNNs) are the two primary architectures used in deep learning [21]. Despite the fact that both architectures are deep learning classifiers for remote sensing data [22], CNNs are a more widely used technique for classifying images [23, 24] and are also used extensively for classifying land use and cover [25, 26]. Compared to other methods, they perform better, are more accurate [27], and tend to be more efficient. The extensive use of CNNs for classifying land cover and land use is mainly owed to such factors [28, 29].

This study's primary goal is examining and evaluating the results of extracting the land cover and land use of the EuroSat dataset using the Glorot Uniform and Random Uniform weight initializers in the ResNet50, ResNet101, and ResNet152 architectures. A brief overview of CNNs, their basic structure, the EuroSat dataset, the ResNet deep learning method, and research materials and methodology will all be covered in the upcoming sections. A brief overview of CNN models and some research on using the ResNet network to classify the EuroSat dataset are presented in the following section.

2. Review on Past Research

2.1. CNN overview and history

When Fukushima first presented this network architecture in 1988, it was not extensively utilized due to the lack of computational hardware for training [30]. In 1993, LeCun et al. [31] solved the handwritten digit classification problem with remarkable results using a gradient-based learning algorithm on CNNs. Then, in 1997, RNN models—which could analyze data like sounds or signals—were developed. Deep learning reached stagnation after the 1990s, essentially as an effect of hardware and data shortages as well as issues with network training [32]. The stagnation was basically ended in 2006 with the introduction of layer-by-layer pretraining, a technique for training deep neural networks [33]. When AlexNet, an advanced and deep CNN, won the ImageNet competition in 2012, CNN models started to become increasingly popular. Other CNN models, like ResNet, have been developed and utilized in numerous applications, such as image classification, following AlexNet's victory in the ImageNet challenge [34]. The timeline for using CNN models is displayed in Figure 2.

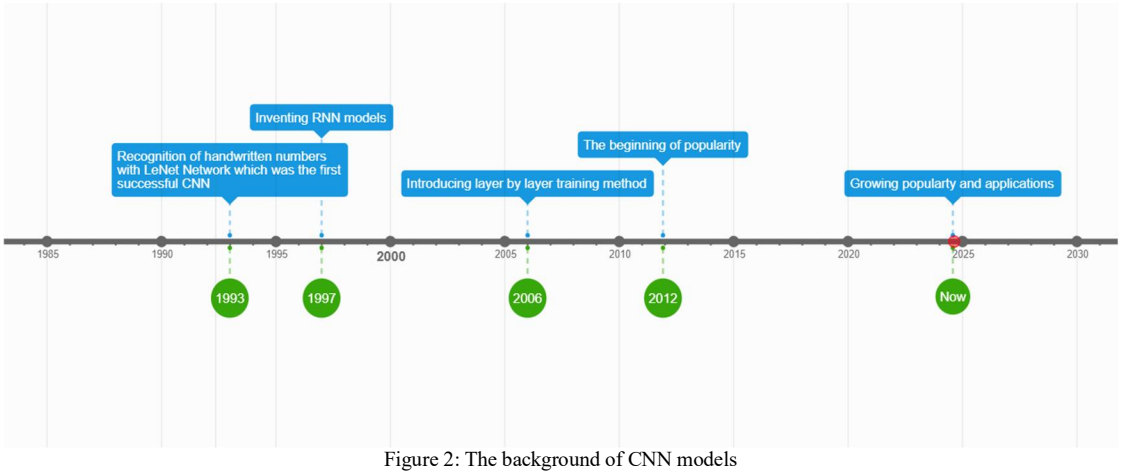

Figure 2: The background of CNN models

2.2. Classifying EuroSat dataset with ResNet models

The vanishing gradient and complexity problems make neural network training more difficult and time- and resource-consuming for deeper neural networks. To address these issues, ResNet was created in 2015, and since then, it has become incredibly popular in the image classification field [35]. Due to these benefits, ResNet has been widely used in the EuroSat dataset for classifying land cover and land use [36].

Kumar and Rohith [37] classified the EuroSat dataset, which included residential areas, river and lake classes, agricultural areas (annual and permanent crops), and residential areas using the ResNet34, ResNet50, ResNet101, and ResNet152 networks. The networks' respective classification accuracy was 0.9855, 0.9835, 0.9773, and 0.9641. Moreover, these four networks had weighted F1 scores of 0.9120, 0.9925, 0.9873, and 0.9813.

Chen et al. [38] classified land cover and land use using the ResNet101 network and the EuroSat dataset. Their classification accuracy was assessed using the following metrics: weighted F1 score, kappa coefficient, and overall accuracy, which were 0.9474, 0.9415, and 0.9471, respectively.

Chen and Tsou [39] classified land cover and land use with an overall accuracy of 0.9463 using the ResNet34 network and the EuroSat dataset.

Dimitrovski et al. [40] used the ResNet50 and ResNet152 networks and the EuroSat dataset. The classification accuracy of these two networks was 0.97 and 0.9740.

Using varying ratios of training to testing data, Yamashkina et al. [41] employed the ResNet50 network and the EuroSat dataset. The study's most accurate training and testing ratios were 90/10 and 80/20, and the network's training and test data ratios had respective accuracy values of 0.9637 and 0.9643.

By utilizing the ResNet50 pre-trained model and the EuroSat dataset, Mahamunkar and Netak [42] were able to attain a classification accuracy of 0.9657.

With a training-to-testing ratio of 90/10, 20/80, Zhang et al. [43] employed the ResNet50 network and the EuroSat dataset; the resulting classification accuracy was 0.7506 and 0.8853, respectively.

These studies indicate that ResNet is a powerful CNN for classifying land use and cover data collected from EuroSat. Therefore, to classify land cover and land use, we used ResNet50, ResNet101, and ResNet152. These networks' identity and convolutional blocks were initialized with Glorot Uniform and Random Uniform weights, respectively. The output of these networks was then assessed using the kappa coefficient, overall accuracy, weighted F1 score, IOU, and confusion matrix—all of which will be covered in the following sections.

3. Methodology and Materials

3.1. EuroSat dataset

Since deep learning techniques rely on data, the development of an appropriate dataset has a big impact on the results. Consequently, to train efficient models, an extensive training set and an extensive number of images are required [44]. Deep learning is capable of extracting extremely complex decision rules because of the vast amount of training data. A number of datasets regarding land use and cover have recently been created [45]. The EuroSat dataset is one of the datasets that is commonly utilized in deep learning to classify land use and cover. This dataset's images are taken from publicly accessible Sentinel-2A satellite imagery. The red, green, and blue bands on the satellite's 13 bands have a spatial resolution of 10 meters.

Figure 3: Sample images of annual crop (A), forest (B), herbaceous vegetation (C), highway (D), industrial areas (E), pasture (F), permanent crop (G), residential areas (H), river (I) and sea and lake (J) classes in the EuroSat dataset

There are 27,000 images in the EuroSat dataset, organized into 10 classes. There are two to three thousand images in each class. Annual crops, forests, herbaceous vegetation, highways, industrial areas, pastures, permanent crops, residential areas, rivers, sea, and lakes comprise the classes included in this dataset. Each image had a size of 64 × 64 pixels [46]. The EuroSat dataset is a medium resolution remote sensing image dataset [47]. A sample image from each class in this dataset is displayed in Figure 3.

3.2. Proposed Method

The research was carried out according to Figure 4.

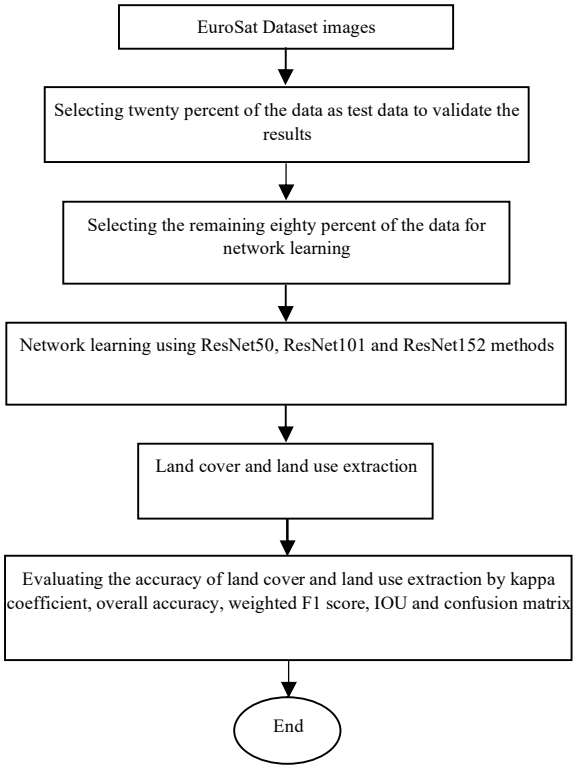

Figure 4: The general workflow in this research

As can be seen in Figure 4, following the acquisition of the images from the EuroSat dataset, 20% of the data were randomly chosen for testing the results, while the remaining 80% of the data was used for network training. For the intention of classifying land use and cover, the ResNet50, ResNet101, and ResNet152 networks were used. Finally, accuracy metrics like the kappa coefficient, overall accuracy, weighted F1 score, Intersection over Union (IOU), and confusion matrix were also used to assess the obtained accuracies in the classification stage. The Convolutional Neural Network, ResNet technique, and classification accuracy evaluation indices are discussed in the following sections.

3.2.1. Convolutional Neural Network

A Convolutional Neural Network consists of convolutional layers, pooling layers, and a fully connected layer [48]. Figure 5 shows the architecture of a convolutional neural network.

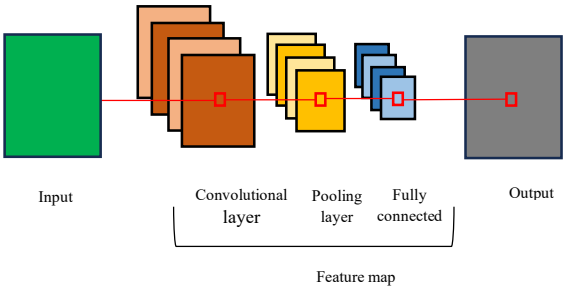

Figure 5: Architecture of Convolutional Neural Network

A convolutional neural network's primary component is the convolutional layer. Therefore, by applying kernels of a specific size to images, image information can be extracted. Typically, the network learns kernels, which focus on fundamental patterns (such as edges, colors, shapes, and textures) that exist in sufficiently deep convolutional neural networks. Basic and comprehensive patterns have been replaced with conceptual representations assigned to certain sample categories. The pooling layer is only used to assign the necessary parameters, like pooling type, kernel size, and stride; it has no parameters to learn. An activation function is added to improve the network's performance. The learned feature representations are mapped to the labeled sample space by the fully connected layers, which serve as classifiers [49].

In each layer, the input image is convolved with a set of K kernels, $W = \{W_1, W_2, \dots, W_k\}$, and a bias term, $B = \{b_1, b_2, \dots, b_k\}$, is added to each feature map, X_k. After that, σ, a non-linear transformation, is applied to these features. This process is repeated for each convolutional layer l according to equation 1 [50].

$$X_k^l = \sigma(W_k^{l-1} * X^{l-1} + b_k^{l-1}) \quad (1)$$

3.2.2. ResNet method

ResNet models are a type of deep convolutional neural networks that use residual learning to make training very deep networks easier [51]. ResNet models come in a variety of versions, the most popular of which are the 5-variant models, each of which has 18, 34, 50, 101, and 152 layers [52].

ResNet's architecture consists of multiple convolutional layers connected by residual connections. The input passes through several convolutional layers in a regular convolutional neural network before being transformed into a prediction. In ResNet, a shortcut connection is added to bypass some of the convolutional layers. This shortcut connection allows the output of the previous layer to be directly added to the output of the next layer, creating a residual block. A residual block is formed using two or more convolutional layers with a shortcut connection. The ResNet structure has proven effective in various computer-vision tasks, including image classification.

The original ResNet152 model took first place in one of the most significant computer vision competitions, the large-scale ImageNet (ILSVRC) visual recognition challenge in 2015. Numerous deep learning models that developed after ResNet were inspired by this popular and significant architecture in the field [53]. ResNet uses residual connections to enhance model performance overall. This method provides it plausible to train much deeper networks and solve optimization problems [54]. ResNet models can also be applied to classify land use and urban land cover [55]. Figure 6 presents the network's structure.

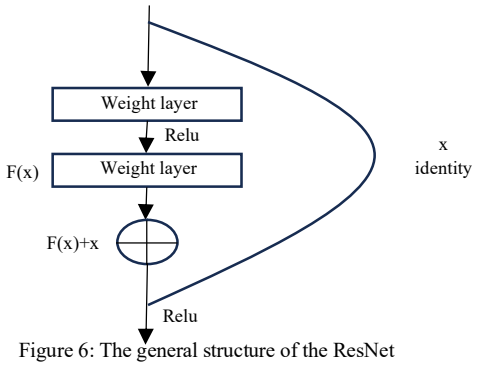

Figure 6: The general structure of the ResNet

The two primary ResNet network components are the identity block and the convolutional block. Equation 2 defines the identity block that is employed in the ResNet network:

$$y = F(x,\{Wi\}) + x \qquad (2)$$

where x and y are the input and output vectors of the considered layers. The function of $F(x,\{Wi\})$ represents the trained residual map. Figure 7 shows the ResNet Identity block.

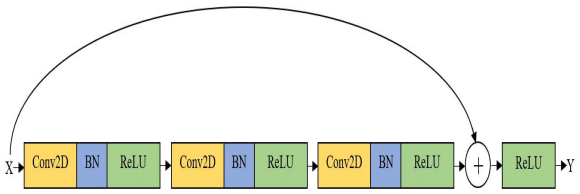

Figure 7: ResNet Identity block

The first part is a two-dimensional convolution layer with a stride of (1,1) and a filter size of (1×1). Batch normalization was used to normalize the channel axis, and the ReLU function was employed to adjust the nonlinear activation function. With a (f×f) filter size, the second component is similar to the first.

In the convolutional block, the input and output dimensions do not match. The shortcut connections W_s a perform linear mapping to change the dimensions between x and F as follows in Equation 3:

$$y = F(x,\{Wi\}) + W_s x \qquad (3)$$

where F is the output of the stacked layer, x and y are the combined input and output vector of the convolutional block. Figure 8 shows the ResNet convolutional block.

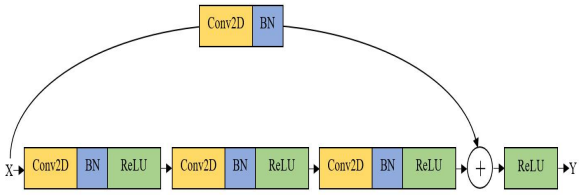

Figure 8: ResNet Convolutional block

In the convolutional block, a 2D convolution layer is added to the shortcut connection, which sets it apart from the identity block. Convolutional blocks are constructed using the same basic principles as identity blocks, with the addition of a 2D convolutional layer to the shortcut. The input x in this shortcut has been resized to match the main path. The output dimensions determine the filter size (1×1) and stride (s, s) of the 2D convolutional layer. Lastly, the output of the main path was combined with the modified shortcut. Its capacity to address the issue of vanishing gradients is the modified shortcut's primary benefit [56].

The general ResNet50, ResNet101, and ResNet152 structures that were employed in this study are explained. These networks are structurally similar, with the primary difference being the number of identity blocks and layers in the network.

The ResNet networks used in this study are displayed in Figure 9.

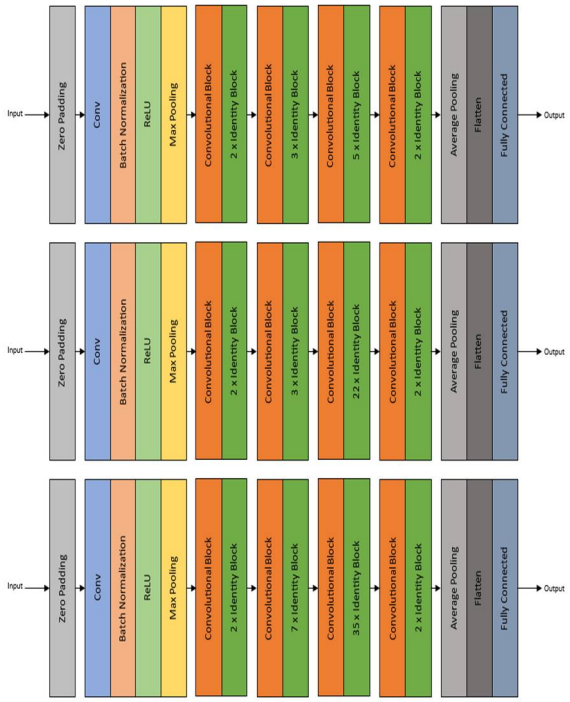

Figure 9: Overall structure of ResNet50 network (top), ResNet101 network (middle), and ResNet152 network (bottom)

Figure 9 shows how similar the ResNet50, ResNet101, and ResNet152 networks' structures are to each other. First, each of the three networks receives the image as input, which is then passed through a (3×3) zero padding layer. In the first stage, all three networks share a stride of (2×2) and use a 2D convolutional layer with 64 filters of size (7×7). Channel standardization is achieved through batch normalization, following which the ReLU activation function is employed. Finally, a max pooling layer featuring a (2×2) stride is implemented.

A convolutional block and two identity blocks consist of the second stage in each of the three networks. In this stage, three sets of [256, 64, 64] filters with a size of (3×3) and a stride of (1×1) are used by both the convolutional and identity blocks.

The third stage of the networks ResNet50, ResNet101, and ResNet152 comprises one convolutional block and three identity blocks, and one convolutional block and seven identity blocks, respectively. In this stage, three sets of [512, 128, 128] filters with (3×3) size and (2×2) stride are used by both the convolutional and identity blocks.

In the fourth stage, one convolutional block and five identity blocks form the structure of the ResNet50 network, one convolutional block and twenty-two identity blocks makes up the ResNet101 network, and one convolutional block provides through the ResNet152 network. In this stage, three sets of [1024, 256, 256] filters with a size of (3×3) and a stride of (2×2) are used by both the convolutional and identity blocks.

All three networks have two identity blocks and one convolutional block in the fifth stage. In this stage, three sets of [1024, 256, 256] filters with a size of (3×3) and a stride of (2×2) are used by both the convolutional and identity blocks.

In the final stage, an average pooling layer of size (2×2) is applied. Subsequently, the output passes through a flattening layer, and eventually, the input is reduced to the number of classes using a fully connected layer. For this, the softmax activation function applies.

3.2.3. Loss function and optimizer

One of the most important steps in a deep learning-based classification is choosing the loss function. In classification problems, the categorical cross-cntropy loss function is usually applied. Equation 4 contains the mathematical formula for this loss function:

$$J(\Theta) = \sum_{i=1}^{p} \sum_{k=1}^{K} y_k^i \log(\hat{y}_k^i) \quad (4)$$

As indicated in the above formula, y_k^i and $\hat{y}_k^i$ show the correct and predicted probabilities that the i-th sample belongs to the k-th class. K represents the total number of target classes, p represents the total number of samples, and (Θ) represents the hypothesized set as a function of weights and biases [57].

Selecting an optimizer function is an essential step in deep learning-based classification. In this study, Adam optimizer was used. This optimizer works effectively with large data sets or parameters, needs less memory, and is computationally productive [58].

3.2.4. Accuracy evaluation

As generating maps using remote sensing data, accuracy evaluation is vital because it allows one to assess the performance of different classifiers as well as the sampling impact. Furthermore, land cover classification products must undergo validation and accuracy assessment when trying to demonstrate the quality of the remote sensing products [59].

3.2.4.1. Kappa coefficient and Overall accuracy

The most widely used accuracy assessment method for displaying an image's classification accuracy is the kappa index. However, image classification is validated using overall accuracy [60]. The overall accuracy ranges from zero to one hundred, where one hundred represents the highest accuracy and zero represents the lowest. From zero to one, the kappa coefficient indicates the accuracy, with zero denoting the lowest accuracy and one referring to the highest accuracy [61]. Equation 5 defines the formula for calculating the kappa coefficient [62]:

$$Kappa = \frac{N\sum_{i,j=1}^{n} X_{i,j} - \sum_{i,j=1}^{n}(Y_i \times Z_j)}{N^2 - \sum_{i,j=1}^{n}(Y_i \times Z_j)} \quad (5)$$

In the above formula, N represents the total number of ground truth points, n represents the total number of land cover and land use classes, $X_{i,j}$ is the sum of correctly classified points in row i and column j, Y_i represents the total number of points in the rows and Z_j represents the total number of points in the columns.

A classified image's overall accuracy is calculated by comparing each pixel's classification to the known land cover conditions determined from relevant ground truth data [63]. Equation 6 defines the general formula for overall accuracy (OA) as follows [64]:

$$OA = \frac{Number\ of\ correct\ pixels}{Number\ of\ pixels} \times 100 \quad (6)$$

3.2.4.2. F1 score

The F1 score, which measures the balance between recall and precision, is a helpful quantitative measure of learning data. The precision measurements, also known as the positive predictive value, quantifies the proportion of correctly identified pixels in each class. The recall, also referred to as sensitivity, shows the number of real pixels found in each class [65]. Equations 7, 8, and 9 define the precision, recall, and F1 formula score [66]:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (7)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (8)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

The weighted F1 score determines a weighted average with accounting for the dataset's class balance. In Equation 10, the weighted F1 score formula is mentioned [67]:

$$weighted\ F1\ score = \frac{\sum_{i=1}^{k} F1\ score_i}{\sum_{i=1}^{k} N_i} \quad (10)$$

3.2.4.3. Intersection over Union (IoU)

Intersection over Union (IoU) represents the ratio of correctly classified pixels to the total number of pixels between the reference and the obtained classification. The formula for IoU is shown in Equation 11 [68]:

$$IoU = \frac{TP}{TP + FP + FN}, \quad (11)$$

where TP, FP, and FN indicate True Positive, False Positive, and False Negative, respectively.

3.2.4.4. Confusion matrix

The conventional technique to assess uncertainty in land cover and land use data is the confusion matrix. This matrix has been applied to numerous indices to assess the degree to which data estimates and the ground truth match up. It provides a basis for assessing classification accuracy and identifying errors [69]. In remote sensing, the confusion matrix has been used for quite a while. The land cover classifications produced by the classification method are represented by rows in this matrix, while the categories identified by the ground truth, or land cover reality, are represented by columns. The number of observations associated to each class combination is indicated by the cell values. Typically, in order to test the classification method, these observations are a sample of points for which ground truth data has been collected [70]. A confusion matrix needs to be used to evaluate accuracy in order to ensure the level of classification accuracy [71].

3.2.5. Weight initialization

Weight initialization is a crucial step in training a neural network and involves adjusting weights throughout the training process until the loss

converges to a minimum. Consequently, weight initialization plays a direct role in driving the convergence of a network, so choosing an appropriate weight initialization method is essential in the training process. A well-selected weight initialization enables accelerated network training and improved performance [72]. In our research, we used Random Uniform weight initializer for the identity block of ResNet networks and Glorot Uniform weight initializer for the convolutional block.

3.2.5.1. Glorot Uniform weight initializer

This approach to weight initialization sets weights based on the input and output units in the layers of the network, with the goal of maintaining consistent variance in activations and gradients. This approach significantly improved the training of deep neural networks [73].

3.2.5.2. Random Uniform weight initializer

In this approach, the values of all weights are assigned random numbers, typically selected from a normal or uniform distribution. The influence of exploding/vanishing gradients in deep neural networks is significantly reduced by this method [74].

4. Implementing the proposed method and presenting the results

Each network used in this study was trained following the instructions in the previous section. The number of epochs used to train the networks in this study was 30 epochs. In each network, 80% of the EuroSat dataset was used for network training and 20% of the EuroSat dataset was used as test data to evaluate the results and obtain the classification accuracy for each network. The categorical cross-entropy loss function and Adam activation function were used in the ResNet deep learning networks. Figure 10 shows the classification accuracy plots over 30 epochs for the training and test datasets of these three deep learning networks.

On the other hand, the loss plots show a decreasing trend over each epoch for both the training and test datasets. In addition, the classification accuracy of the test data is lower than the accuracy of the training data in the classification accuracy plot, while the cost of the test data is higher than the cost of the training data in the loss plot. This indicates that our models are converging.

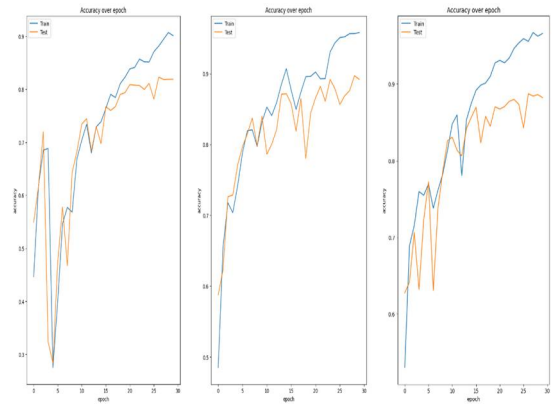

Figure 10: Plots of classification accuracy in the ResNet50 network (right), ResNet101 network (center), and ResNet152 network (left) for the training and test datasets at each epoch.

Furthermore, Figure 11 shows the loss plots over 30 epochs for the training and test datasets of all three networks.

As can be seen from the classification accuracy graphs, the accuracy of both training and test datasets increases as the number of epochs increases during network training.

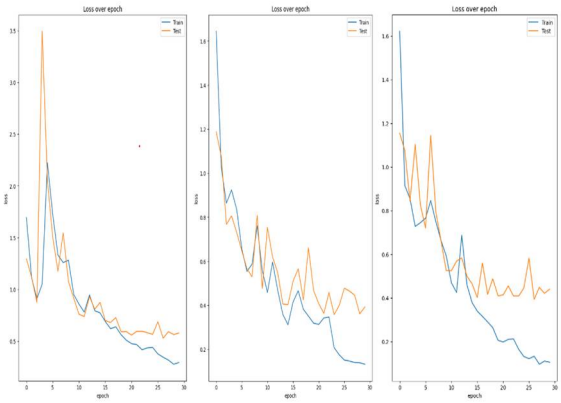

Figure 11: Loss plots in the ResNet50 network (right), ResNet101 network (center), and ResNet152 network (left) for the training and test datasets at each epoch.

Figure 12 shows the confusion matrices of these three networks for land cover and land use classification of EuroSat dataset.

The main diagonal values of the confusion matrix indicate the true positive values. In the confusion matrices provided, the columns represent the predicted values while the rows represent the true values. These matrices can be used to calculate, various classification accuracy indices such as overall accuracy, kappa coefficient, F1 score, and IoU (Intersection over Union).

Table 1 shows precision, recall, and F1 score indices for the ten classification classes of the EuroSat dataset for the ResNet50, ResNet101, and ResNet152 networks.

The precision, recall, and F1 score metrics of the ResNet50 model demonstrate strong performance over a range of land cover and land use classes in the EuroSat dataset. With precision values of 0.9088 for rivers, 0.9640 for forests, and 0.9757 for lakes and seas, it effectively reduces false positives. Recall values for industrial areas are 0.9204, lakes and seas are 0.9748, and forests are notable at 0.9135. However, challenges are evident in the highway 0.8270 and herbaceous vegetation 0.8046 classes.

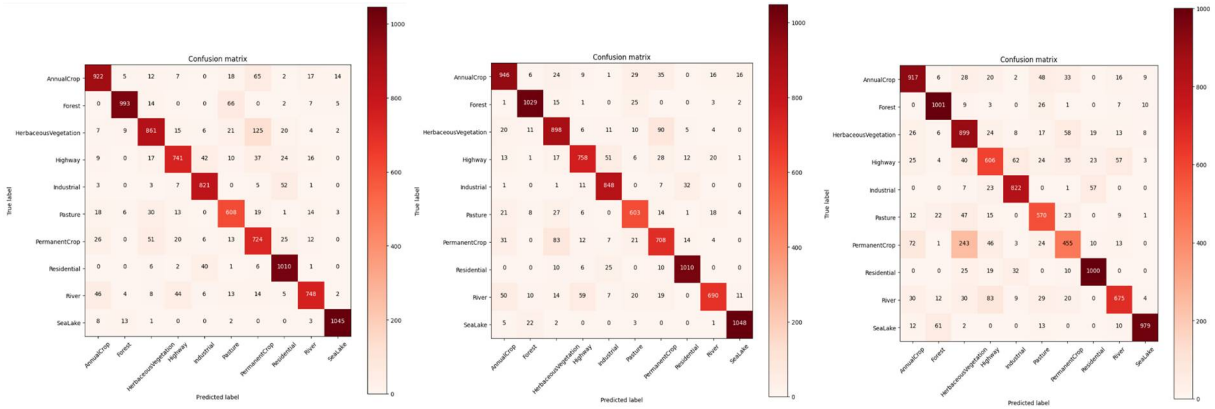

Figure 12: Confusion matrix of ResNet50 network (left), ResNet101 network (middle), and ResNet152 network (right) for land cover and land use classification in the EuroSat dataset.

Table 1: Precision, Recall, and F1 score metrics for the ResNet50, ResNet101, and ResNet152 networks for the ten classification classes of the EuroSat dataset.

Index	Network	Class									
		Annual crop	Forest	Herbaceous vegetation	Highway	Industrial area	Pasture	Permanent crop	Residential area	River	Sea and Lake
Precision	ResNet50	0.8873	0.9640	0.8584	0.8727	0.8914	0.8085	0.7276	0.8851	0.9088	0.9757
	ResNet101	0.8694	0.9466	0.8230	0.8732	0.8926	0.8410	0.7771	0.9404	0.9126	0.9685
	ResNet152	0.8382	0.8993	0.6759	0.7222	0.8763	0.7589	0.7154	0.9017	0.8437	0.9654
Recall	ResNet50	0.8681	0.9135	0.8046	0.8270	0.9204	0.8539	0.8255	0.9474	0.8404	0.9748
	ResNet101	0.8743	0.9563	0.8511	0.8357	0.9422	0.8589	0.8045	0.9519	0.7840	0.9694
	ResNet152	0.8498	0.9470	0.8339	0.6894	0.9032	0.8154	0.5247	0.9208	0.7567	0.9090
F1 Score	ResNet50	0.8776	0.9381	0.8306	0.8492	0.9056	0.8306	0.7735	0.9152	0.8733	0.9752
	ResNet101	0.8718	0.9514	0.8369	0.8540	0.9167	0.8498	0.7906	0.9461	0.8435	0.9690
	ResNet152	0.8439	0.9225	0.7466	0.7054	0.8896	0.7862	0.6054	0.9111	0.7978	0.9363

The model's overall efficacy is highlighted by F1 scores such as 0.9752 for lakes and seas, 0.9381 for forests, and 0.8733 for rivers.

For the ResNet101 model, precision values of 0.9685 for lakes and seas, 0.9466 for forests, and 0.9404 for residential areas were achieved. Strong recall values are observed in lakes and seas 0.9694, forests 0.9563, and residential areas 0.9519, reflecting the model's accuracy. F1 scores of 0.9514 for forests, 0.9690 for lakes and seas, and 0.9461 for residential areas further attest to its effectiveness.

The ResNet152 model's evaluation on the EuroSat dataset results in significant precision values of 0.8993 for forests, 0.9017 for residential areas, and 0.9654 for lakes and seas. Herbaceous vegetation, pasture, and permanent crops with lower precision scores present challenges. In contrast to difficulties in permanent crops and highways, strong recall values are found in forests, residential areas, and industrial areas. The model's effectiveness is demonstrated by F1 scores of 0.9225 for forests, 0.9363 for lakes and seas, and 0.9111 for residential areas.

Table 2 displays kappa coefficient, overall accuracy, weighted F1 score and IoU indexes for ResNet50, ResNet101, and ResNet152 networks.

Table 2: Kappa coefficient, Overall Accuracy, weighted F1 score, and weighted IoU indexes for the ResNet50, ResNet101, and ResNet152 networks.

Network	Index			
	Overall accuracy	Kappa coefficient	Weighted F1 score	IoU
ResNet50	0.8804	0.8669	0.8808	0.7853
ResNet101	0.8871	0.8743	0.8869	0.7951
ResNet152	0.8233	0.8032	0.8213	0.6991

The results show that when using the EuroSat dataset for land cover and land use classification accuracy indices, the ResNet101 network performs better than both the ResNet50 and ResNet152 networks. However, the ResNet101 network and the

ResNet50 network have very similar numerical values for all indexes. The ResNet101 network's Overall Accuracy, Kappa coefficient, weighted F1 score, and weighted IoU were, in that order, 0.8871, 0.8743, 0.8869, and 0.7951. These index values for the ResNet50 network were 0.8804, 0.8669, 0.8808, and 0.7853, respectively. These values are very close to the ResNet101 index values. In addition, these index values were 0.8233, 0.8032, 0.8213, and 0.6991 for the ResNet152 network, in that order. The classification accuracy index results obtained for the ResNet152 network are lower when compared to the ResNet101 and ResNet50 networks.

Figure 13 demonstrates some of the predictions using ResNet50 model which is from test data.

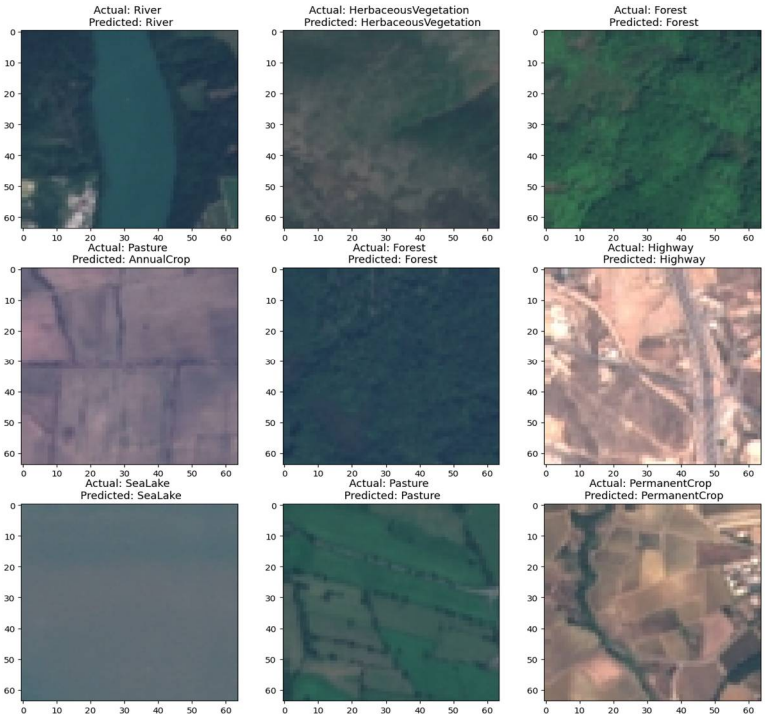

Figure 13: The ResNet50 sample results.

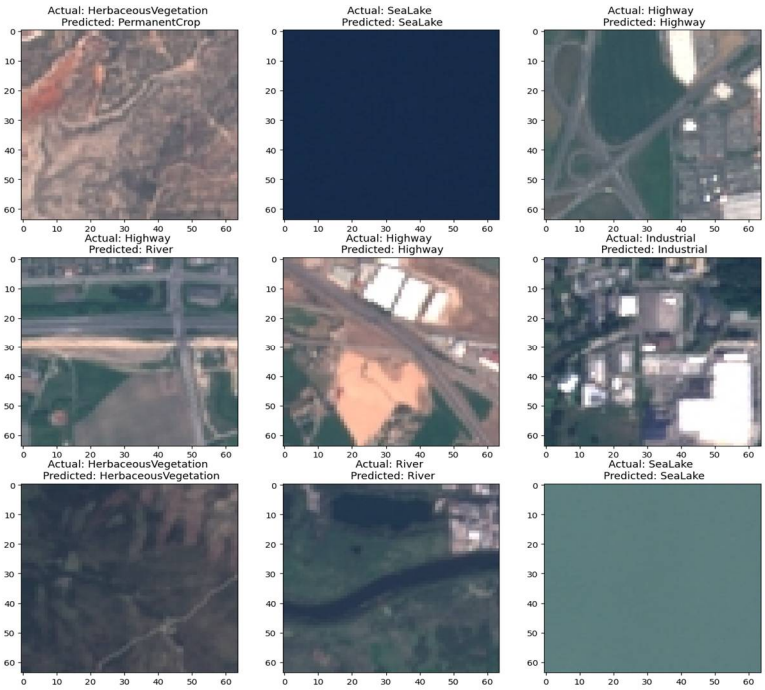

Figure 14: The ResNet101 sample results.

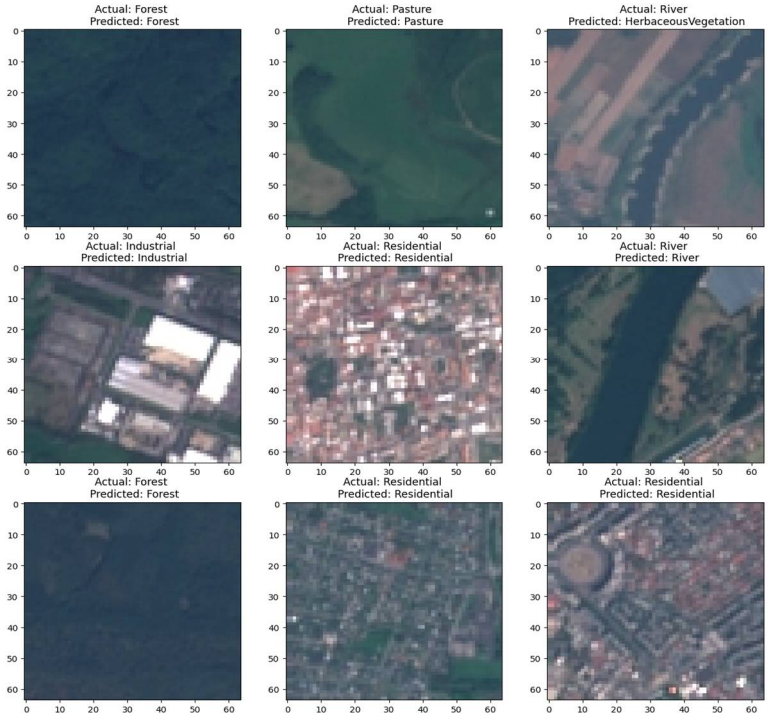

Figure 15: The ResNet152 sample results.

River, herbaceous vegetation, forest, highway, sea, lake and permanent crop classes were all correctly predicted by the ResNet50 model; however, the pasture class was predicted incorrectly.

Some of the predictions made with the ResNet101 model based on test data are displayed in Figure 14.

Correct predictions were made for industrial areas, lakes, rivers, and sea, but incorrect predictions were made for highways and herbaceous vegetation in the ResNet101 model.

Figure 15 shows some of the predictions using the ResNet152 model based on test data.

The model accurately predicted the classes of residential, industrial, pasture, and forest areas, but the river class was predicted incorrectly.

5. Conclusion

Accurate information on land cover and land use plays a crucial role in environmental applications, infrastructure planning, and ensuring sustainable urban development. Nowadays, a lot of remote sensing imagery is available for obtaining information about land use and cover due to advancements in satellite technologies. Classification techniques must be applied in order to extract this data from remote sensing images. Deep learning techniques are widely used in remote sensing studies for land cover and land use classification because they lead to accurate classification results. These techniques are capable of modeling hierarchical characteristics, which are crucial for land cover and land use classification. Convolutional neural networks (CNNs) are common deep learning architectures for classifying land cover and land use due to their performance, efficiency, and accuracy. ResNet is one of the CNNs used to categorize land cover and land use. It uses residual learning techniques to improve network training for the determination of land cover and land use information. The vanishing gradient problem can be dealt with by ResNet models, making them an effective choice for classifying land use and cover. Since deep learning relies on data, choosing the right dataset for network training affects the output and outcomes significantly. Deep learning methods can be applied to the EuroSat dataset to extract information about land use and cover. The primary objective of this study is to assess the performance of the Glorot Uniform and Random Uniform weight initializers in the ResNet50, ResNet101, and ResNet152 architectures for extracting the land cover and land use of the EuroSat dataset. The performance of these networks was then evaluated using the kappa coefficient, overall accuracy, weighted F1 score, and intersection over union (IoU). The corresponding outcomes of these indexes for the ResNet50 network were 0.8804, 0.8669, 0.8808, and 0.7853. The ResNet152 network's index values were 0.8233, 0.8032, 0.8213, and 0.6991, while the ResNet101 network's corresponding indexes were 0.8871, 0.8743,

0.8869, and 0.7951. These results demonstrate that numerical values for the Kappa coefficient, overall accuracy, weighted F1 score, and IoU do not significantly differ between the ResNet50 and ResNet101 networks. However, these two networks outperformed the ResNet152 network in every classification accuracy index. Overall, the ResNet101 network outperformed the ResNet50 and ResNet152 techniques in each accuracy assessment index.

References

[1] Alexander, C. (2020). Normalised difference spectral indices and urban land cover as indicators of land surface temperature (LST). International Journal of Applied Earth Observation and Geoinformation, 86, 102013.

[2] Reis, S. (2008). Analyzing land use/land cover changes using remote sensing and GIS in Rize, North-East Turkey. Sensors, 8(10), 6188-6202.

[3] Lv, Q., Dou, Y., Niu, X., Xu, J., Xu, J., & Xia, F. (2015). Urban land use and land cover classification using remotely sensed SAR data through deep belief networks. Journal of Sensors, 2015.

[4] Marschner, F. J., & Anderson, J. R. (1967). Major land uses in the United States. US Geological Survey.

[5] Jozdani, S. E., Johnson, B. A., & Chen, D. (2019). Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. Remote Sensing, 11(14), 1713.

[6] Zhang, X., Han, L., Han, L., & Zhu, L. (2020). How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery?. Remote Sensing, 12(3), 417.

[7] Lee, S. H., Han, K. J., Lee, K., Lee, K. J., Oh, K. Y., & Lee, M. J. (2020). Classification of landscape affected by deforestation using high-resolution remote sensing data and deep-learning techniques. Remote Sensing, 12(20), 3372.

[8] Li, W., Fu, H., Yu, L., Gong, P., Feng, D., Li, C., & Clinton, N. (2016). Stacked Autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping. International journal of remote sensing, 37(23), 5632-5646.

[9] Bhatti, U. A., Yu, Z., Yuan, L., Zeeshan, Z., Nawaz, S. A., Bhatti, M., ... & Wen, L. (2020). Geometric algebra applications in geospatial artificial intelligence and remote sensing image processing. IEEE Access, 8, 155783-155796.

[10] Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. ISPRS journal of photogrammetry and remote sensing, 152, 166-177.

[11] Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., ... & Zhang, L. (2020). Deep learning in environmental remote sensing: Achievements and challenges. Remote Sensing of Environment, 241, 111716.

[12] Huang, B., Zhao, B., & Song, Y. (2018). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. Remote Sensing of Environment, 214, 73-86.

[13] Digra, M., Dhir, R., & Sharma, N. (2022). Land use land cover classification of remote sensing images based on the deep learning approaches: a statistical analysis and review. Arabian Journal of Geosciences, 15(10), 1003.

[14] Tong, X. Y., Xia, G. S., Lu, Q., Shen, H., Li, S., You, S., & Zhang, L. (2020). Land-cover classification with high-resolution remote sensing images using transferable deep models. Remote Sensing of Environment, 237, 111322.

[15] Zhang, P., Ke, Y., Zhang, Z., Wang, M., Li, P., & Zhang, S. (2018). Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery. Sensors, 18(11), 3717.

[16] Zhao, J., Wang, L., Yang, H., Wu, P., Wang, B., Pan, C., & Wu, Y. (2022). A land cover classification method for high-resolution remote sensing images based on NDVI deep learning fusion network. Remote Sensing, 14(21), 5455.

[17] Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. IEEE Geoscience and remote sensing magazine, 4(2), 22-40.

[18] Cheng, X., He, X., Qiao, M., Li, P., Hu, S., Chang, P., & Tian, Z. (2022). Enhanced contextual representation with deep neural networks for land cover classification based on remote sensing images. International Journal of Applied Earth Observation and Geoinformation, 107, 102706.

[19] Kavhu, B., Mashimbye, Z. E., & Luvuno, L. (2021). Climate-based regionalization and inclusion of spectral indices for enhancing transboundary land-use/cover classification using deep learning and machine learning. Remote Sensing, 13(24), 5054.

[20] Storie, C. D., & Henry, C. J. (2018, July). Deep learning neural networks for land use land cover mapping. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium (pp. 3445-3448). IEEE.

[21] Luus, F. P., Salmon, B. P., Van den Bergh, F., & Maharaj, B. T. J. (2015). Multiview deep learning for land-use classification. IEEE Geoscience and Remote Sensing Letters, 12(12), 2448-2452.

[22] Campos-Taberner, M., García-Haro, F. J., Martínez, B., Izquierdo-Verdiguier, E., Atzberger, C., Camps-Valls, G., & Gilabert, M. A. (2020). Understanding deep learning in land use classification based on Sentinel-2 time series. Scientific reports, 10(1), 17188.

[23] Zhu, M., He, Y., & He, Q. (2019). A review of researches on deep learning in remote sensing application. International Journal of Geosciences, 10(1), 1-11.

[24] Naushad, R., Kaur, T., & Ghaderpour, E. (2021). Deep transfer learning for land use and land cover classification: A comparative study. Sensors, 21(23), 8083.

[25] Alem, A., & Kumar, S. (2020, June). Deep learning methods for land cover and land use classification in remote sensing: A review. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) (pp. 903-908). IEEE.

[26] Aspri, M., Tsagkatakis, G., & Tsakalides, P. (2020). Distributed training and inference of deep learning models for multi-modal land cover classification. Remote Sensing, 12(17), 2670.

[27] Jin, B., Ye, P., Zhang, X., Song, W., & Li, S. (2019). Object-oriented method combined with deep convolutional neural networks for land-use-type classification of remote sensing images. Journal of the Indian Society of Remote Sensing, 47, 951-965.

[28] Zang, N., Cao, Y., Wang, Y., Huang, B., Zhang, L., & Mathiopoulos, P. T. (2021). Land-use mapping for high-spatial resolution remote sensing image via deep learning: A review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 5372-5391.

[29] Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanesh, F., & Zhang, Y. (2018). Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. Remote Sensing, 10(7), 1119.

[30] Fukushima, K. (1988). Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural networks*, *1*(2), 119-130.

[31] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324.

[32] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*.

[33] Erhan, D., Courville, A., Bengio, Y., & Vincent, P. (2010, March). Why does unsupervised pre-training help deep learning?. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (pp. 201-208). JMLR Workshop and Conference Proceedings.

[34] Vasavi, S., Somagani, H. S., & Sai, Y. (2023). Classification of buildings from VHR satellite images using ensemble of U-Net and ResNet. *The Egyptian Journal of Remote Sensing and Space Sciences*, *26*(4), 937-953.

[35] Borawar, L., & Kaur, R. (2023, March). ResNet: Solving vanishing gradient in deep networks. In *Proceedings of International Conference on Recent Trends in Computing: ICRTC 2022* (pp. 235-247). Singapore: Springer Nature Singapore.

[36] Harini, M., Selvavarshini, S., Narmatha, P., Anitha, V., Selvi, S. K., & Manimaran, V. (2024, January). Resnet-50 Integrated with Attention Mechanism for Remote Sensing Classification. In *International Conference on Advances in Distributed Computing and Machine Learning* (pp. 255-265). Singapore: Springer Nature Singapore.

[37] Rohith, G., & Kumar, L. S. (2020, June). Remote sensing signature classification of agriculture detection using deep convolution network models. In *International Conference on Machine Learning, Image Processing, Network Security and Data Sciences* (pp. 343-355). Singapore: Springer Singapore.

[38] Chen, G., Zhang, X., Tan, X., Cheng, Y., Dai, F., Zhu, K., ... & Wang, Q. (2018). Training small networks for scene classification of remote sensing images via knowledge distillation. *Remote Sensing*, *10*(5), 719.

[39] Chen, F., & Tsou, J. Y. (2022). Assessing the effects of convolutional neural network architectural factors on model performance for remote sensing image classification: An in-depth investigation. *International Journal of Applied Earth Observation and Geoinformation*, *112*, 102865.

[40] Dimitrovski, I., Kitanovski, I., Kocev, D., & Simidjievski, N. (2023). Current trends in deep learning for Earth Observation: An open-source benchmark arena for image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *197*, 18-35.

[41] Yamashkina, E. O., Yamashkin, S. A., Platonova, O. V., & Kovalenko, S. M. (2022). Development of a neural network model for spatial data analysis. *Rossiiskii Tekhnologicheskii Zhurnal [Russian Technological Journal]*, (10), 28-37.

[42] Mahamunkar, G. S., & Netak, L. D. (2021, December). Comparison of various deep CNN models for land use and land cover classification. In *International Conference on Intelligent Human Computer Interaction* (pp. 499-510). Cham: Springer International Publishing.

[43] Zhang, Z., Mi, X., Yang, J., Wei, X., Liu, Y., Yan, J., ... & Yu, T. (2023). Remote sensing image scene classification in hybrid classical–quantum transferring CNN with small samples. *Sensors*, *23*(18), 8010.

[44] Cheng, G., Xie, X., Han, J., Guo, L., & Xia, G. S. (2020). 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, 13, 3735-3756.

[45] Rousset, G., Despinoy, M., Schindler, K., & Mangeas, M. (2021). Assessment of deep learning techniques for land use land cover classification in southern new Caledonia. Remote Sensing, 13(12), 2257.

[46] Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(7), 2217-2226.

[47] Neumann, M., Pinto, A. S., Zhai, X., & Houlsby, N. (2019). In-domain representation learning for remote sensing. arXiv preprint arXiv:1911.06721

[48] Bhosle, K., & Musande, V. (2019). Evaluation of deep learning CNN model for land use land cover classification and crop identification using hyperspectral remote sensing images. Journal of the Indian Society of Remote Sensing, 47(11), 1949-1958.

[49] Li, W., Liu, H., Wang, Y., Li, Z., Jia, Y., & Gui, G. (2019). Deep learning-based classification methods for remote sensing images in urban built-up areas. Ieee Access, 7, 36274-36284.

[50] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60-88.

[51] Stoimchev, M., Kocev, D., & Džeroski, S. (2023). Deep Network Architectures as Feature Extractors for Multi-Label Classification of Remote Sensing Images. Remote Sensing, 15(2), 538.

[52] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[53] Fahmi, H. (2022). Patch based Classification using ResNet for Land Cover changes detection of Batu City. MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi (Journal of Computer Science and Information Technology), 14(2), 64-69.

[54] Wang, L., Wang, J., Liu, Z., Zhu, J., & Qin, F. (2022). Evaluation of a deep-learning model for multispectral remote sensing of land use and crop classification. The Crop Journal, 10(5), 1435-1451.

[55] Hosseiny, B., Abdi, A. M., & Jamali, S. (2022). Urban land use and land cover classification with interpretable machine learning–A case study using Sentinel-2 and auxiliary data. Remote Sensing Applications: Society and Environment, 28, 100843.

[56] Shehab, L. H., Fahmy, O. M., Gasser, S. M., & El-Mahallawy, M. S. (2021). An efficient brain tumor image segmentation based on deep residual networks (ResNets). *Journal of King Saud University-Engineering Sciences*, *33*(6), 404-412.

[57] Dastour, H., & Hassan, Q. K. (2023). A Comparison of Deep Transfer Learning Methods for Land Use and Land Cover Classification. *Sustainability*, *15*(10), 7854.

[58] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[59] Macarringue, L. S., Bolfe, É. L., & Pereira, P. R. M. (2022). Developments in land use and land cover classification techniques in remote sensing: A review. Journal of Geographic Information System, 14(1), 1-28.

[60] Fan, F., Wang, Y., & Wang, Z. (2008). Temporal and spatial change detecting (1998–2003) and predicting of land use and land cover in Core corridor of Pearl River Delta (China) by using TM and ETM+ images. Environmental monitoring and assessment, 137, 127-147.

[61] Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote sensing of environment, 37(1), 35-46.

[62] Foody, G. M. (2002). Status of land cover classification accuracy assessment. Remote sensing of environment, 80(1), 185-201.

[63] Rwanga, S. S., & Ndambuki, J. M. (2017). Accuracy assessment of land use/land cover classification using remote sensing and GIS. International Journal of Geosciences, 8(04), 611.

[64] Tilahun, A., & Teferie, B. (2015). Accuracy assessment of land use land cover classification using Google Earth. American Journal of Environmental Protection, 4(4), 193-198.

[65] Liu, Y., Fan, B., Wang, L., Bai, J., Xiang, S., & Pan, C. (2018). Semantic labeling in very high resolution images via a self-cascaded convolutional neural network. ISPRS journal of photogrammetry and remote sensing, 145, 78-95.

[66]. Temenos, A., Temenos, N., Kaselimi, M., Doulamis, A., & Doulamis, N. (2023). Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP. IEEE Geoscience and Remote Sensing Letters, 20, 1-5.

[67] Huang, X., Yang, D., He, Y., Nelson, P., Low, R., McBride, S., ... & Guarraia, M. (2023). Land cover mapping via crowdsourced multi-directional views: The more directional views, the better. *International Journal of Applied Earth Observation and Geoinformation*, *122*, 103382.

[68] Cherif, E., Hell, M., & Brandmeier, M. (2022). DeepForest: Novel deep learning models for land use and land cover classification using multi-temporal and-modal sentinel data of the amazon basin. remote sensing, 14(19), 5000.

[69] Abdu, H. A. (2019). Classification accuracy and trend assessments of land cover-land use changes from principal components of land satellite images. International Journal of Remote Sensing, 40(4), 1275-1300.

[70] Hay, A. M. (1988). The derivation of global estimates from a confusion matrix. International Journal of Remote Sensing, 9(8), 1395-1398.

[71] Abdelkareem, O. E. A., Elamin, H. M. A., Eltahir, M. E. S., Adam, H. E., Elhaja, M. E., Rahamtalla, A. M., ... & Elmar, C. (2018). Accuracy assessment of land use land cover in Umabdalla natural reserved forest, South Kordofan, Sudan. International journal of agricultural and environmental sciences, 3(1), 5-9.

[72] Narkhede, M. V., Bartakke, P. P., & Sutaone, M. S. (2022). A review on weight initialization strategies for neural networks. *Artificial intelligence review*, *55*(1), 291-322.

[73] Desai, C. (2024). Impact of Weight Initialization Techniques on Neural Network Efficiency and Performance: A Case Study with MNIST Dataset. *International Journal Of Engineering And Computer Science*, *13*(04).

[74] Li, H., Krček, M., & Perin, G. (2020). A comparison of weight initializers in deep learning-based side-channel analysis. In *Applied Cryptography and Network Security Workshops: ACNS 2020 Satellite Workshops, AIBlock, AIHWS, AIoTS, Cloud S&P, SCI, SecMT, and SiMLA, Rome, Italy, October 19–22, 2020, Proceedings 18* (pp. 126-143). Springer International Publishing.