Employing Transfer Learning in Land-use Land-cover for Risk Management

Marjan Ahangarha [1], Hassan Rezvan [1], Mohammad Javad Valadan Zoej [1], Fahimeh Youssefi [1,2]

[1] Department of Photogrammetry and Remote Sensing, K. N. Toosi University of Technology, Tehran, Iran – (m.ahangarha, h.rezvan, youssefi)@email.kntu.ac.ir; valadanzouj@kntu.ac.ir
[2] Institute of Artificial Intelligence, USX, Shaoxing University, 508 West Huancheng Road, Yuecheng District, Shaoxing, Zhejiang Province, Postal Code 312000, China – youssefi@usx.edu.cn

Keywords: Remote Sensing, Land-use Land-cover classification, EuroSAT, Deep learning, Transfer learning, Disaster management.

Abstract

In disaster management, land-use land-cover (LULC) maps are vital for real-time situational awareness and coordinated responses. These maps aid in coordinating field operations, guiding rescue teams, and identifying high vulnerability areas. They ensure accurate spatial information sharing and management during disaster response efforts. Deep transfer learning models have emerged as powerful tools for LULC classification, addressing challenges like insufficient training data and complex classification tasks. In this research instead of building networks from scratch, pre-trained networks that used EuroSAT benchmark dataset are employed. Several deep learning models including 1-layer CNN, 4-layers CNN, VGG16 and an improved ResNet-50 network as proposed method are considered and compared in this study. The results were analyzed in both quantitative and qualitative ways. In the quantitative mode, the measurement criteria such as Overall Accuracy (OA), F1Score, Precision and Recall were calculated, and in the qualitative mode, the class diagram was drawn in the feature space to check the separability of the classes. Finally, the results show high overall accuracy score of 95.9% indicating the high potential of our proposed network for ResNet-50. The proposed method has resolved insufficient training dataset by implementing data augmentation that it can be solved the problem of lack dataset.

1. Introduction

The number and consequences of natural disasters have shown a dramatic development in past decades (Thenkabail, 2015). Natural disasters are events which are caused by purely natural phenomena and bring damage to human societies (Van Westen, 2000). Disasters – from earthquakes and storms to floods and droughts – kill approximately 40,000 to 50,000 people per year. This is the average over the last few decades. While that's a relatively small fraction of all deaths globally, disasters can have much larger impacts on specific populations. Single extreme events can kill tens to hundreds of thousands of people. In the 20th century, more than a million deaths per year were not uncommon. Disasters have other large impacts, too. Millions of people are displaced – some left homeless – by them each year. And the economic costs of extreme events can be severe, and hard to recover from. This is particularly true in lower-income countries (Ritchie and Rosado, 2022). Therefore, it is necessary to develop models that performs Disaster Risk Management (DRM) elements, including identification, monitoring, management and mitigation strategies (Kumar et al., 2024).

Disaster risk with all its components is a spatial phenomenon. Therefore, it can only be addressed with spatial data. The main source of feeding DRM models is remote sensing information and without such meaningful information, mentioned DRM components would simply not possible (Thenkabail, 2015). During the last decades, remote sensing has become an operational tool in disaster preparedness and warning phases (Van Westen, 2000). Remote sensing images enable providing efficient information about earth's surface over a large area at a low cost (Ali and Johnson, 2022). In this regard detailed LULC map which categorizes and characterizes the earth's terrestrial surface (Rangel et al., 2024), is essential in monitoring and planning.

Machine Learning (ML) algorithms has significantly improved the ability to generate LULC maps. These algorithms identify patterns with exceptional accuracies by means of learning from large amounts of data and building model upon that (Rangel et al., 2024). These methods ranging from conventional methods based on image statistics such as Bayesian and Maximum Likelihood to more complex and advanced methods, such as Support Vector Machine (SVM), Random Forest (RF), K-nearest neighbours (KNN), Decision Trees (DTs) and etc (Ali and Johnson, 2022).

Meanwhile, Deep Learning-based (DL) models have also achieved excellent performances in the field. Notably, these DL-based models have often outperformed the ML methods when there is sufficient amount of training data available (Ali and Johnson, 2022). Also, a major advantage of DL approaches is their ability to automatically learn spectral and contextual patterns and hierarchical features from training set to distinguish between different LULC classes (Ali and Johnson, 2022; Pushpalata et al., 2024). Despite the increased complexity that comes with training deep-learning models from the beginning, this approach can sometimes result in overfitting. To avoid this, deep transfer learning has emerged as a novel methodology within ML. Through the utilization of transfer learning, it becomes feasible to substantially lessen both the necessity for extensive training data and the duration required for training within the target domain. This is achieved because the model doesn't need to start its training anew in the target domain, nor do the training and testing datasets need to be entirely separate or identically distributed. The core aim of transfer learning is to enhance the performance of specific learners within particular domains by applying the insights acquired from various yet related source domains. The versatility of transfer learning has positioned it as a notable and intriguing topic within the broader field of machine learning. (Dastour and Hassan, 2023).

Convolution Neural Network (CNN), has been widely used in tasks such as image classification, object detection and segmentation. The architecture consists of convolution layer, pooling layer, activation function and fully connected layer. Various types of architecture can be applied to CNN and among these architectures, VGG-16 and ResNet-50 are two of them. VGG-16 is consisted of 16 convolution layers and fully connected which is usually used to recognize and classify images. The architecture has 13 convolutional layers with 3x3 filters and 3 fully connected layers (Hartatik and Anam, 2023). ResNet is a CNN network architecture that can use hundreds or even thousands convolution layers. Due to the network depth, it is easy to optimize and has better performance in terms of accuracy.

The aim of this paper is to examine different deep learning networks for high-accuracy classification, improving networks through modifying parameters related to deep learning and achieving the best accuracy with less epochs and short training times.

The rest of paper is as follows: Section 2 describes related works in the field, Section 3 highlights the materials and employed methodology, in Section 4 results are illustrated and finally, in Section 5, the conclusion is presented.

2. Related works

In the field of risk management, particularly within the context of LULC mapping, the integration of advanced ML techniques has emerged as a pivotal strategy for enhancing predictive capabilities and decision-making processes. Among these techniques, transfer learning stands out as a promising approach due to its ability to leverage pre-existing knowledge from related domains, thereby significantly reducing the need for extensive datasets and computational resources typically required for training models from scratch. This section aims to explore the existing body of literature surrounding the application of transfer learning in LULC mapping, highlighting its potential benefits and challenges, and setting the stage for a deeper discussion on how these methodologies can be optimized for effective risk management strategies.

In 2021, Naushad et al. investigated the performance of deep transfer learning architectures in LULC classification. Their study was based on two potential architectures, namely, VGG16 and Wide ResNet-50. Their findings illuminate the broader implications of transfer learning within the domain of risk management, suggesting that such methodologies can significantly enhance the accuracy and efficiency of LULC classification efforts. This, in turn, has the potential to inform more precise and effective risk assessment strategies, enabling stakeholders to better anticipate and mitigate environmental risks associated with land use changes.

In 2022, Alem et al. applied deep transfer learning techniques to the UC Merced dataset. Their models achieved OA of 92.46%, 94.38%, and 99.64% using Resnet-50V2, InceptionV3, and VGG-19, respectively.

In 2022, Sobhana et al. marked a significant advancement in the use of ML for disaster management by employing a CNN enhanced with data augmentation techniques to classify the images related to natural disasters such as earthquakes, floods, cyclones, and wildfires. The study's novelty lies in its approach to overcoming the limitations of traditional datasets through data augmentation, which involves manipulating images to create variations that improve the model's learning capability. This method not only addresses the issue of limited labelled data but also enhances the model's ability to recognize disasters under different conditions, thereby improving classification accuracy.

In 2023, Dastour and Hassan evaluated thirty-nine deep transfer learning models were systematically alongside multiple deep transfer learning models for LULC classification using a consistent set of criteria. Among their models, ResNet-50, EfficientNetV2B0, and ResNet-152 were the top performers in terms of kappa and accuracy scores. ResNet-152 required three times longer training time than EfficientNetV2B0 on their test computer, while ResNet-50 took roughly twice as long. ResNet-50 achieved an overall F1Score of 0.967 on the test set.

In 2023, Kunwar and Ferdush focused on classifying LULC using transfer learning with the ViT model pre-trained on ImageNet-21k. Two datasets were utilized, one with data augmentation and one without. The ViT, VGG-16, and ResNet-50 models were compared in terms of accuracy, F1Score, and training time. The ViT model outperformed the other models in accuracy but required more training time. Data augmentation significantly improved the ViT model's performance, especially in classes like Forest and Sea Lake, although the Pasture class had the lowest accuracy. The study highlighted the importance of hyperparameter tuning and advanced training techniques like learning rate optimization and regularization to achieve optimal results. By fine-tuning the ViT model, the study achieved a remarkable 99.19% accuracy on the EuroSAT RGB dataset, showcasing the effectiveness of transfer learning in LULC classification. The results indicated that the ViT model had higher accuracy and training time compared to VGG-16 and ResNet-50. The ResNet-50 model, on the other hand, exhibited better accuracy than VGG-16 and required less training time for both augmented and non-augmented datasets. Data augmentation techniques were found to be crucial in mitigating overfitting and enhancing dataset diversity, leading to improved model performance. By incorporating model enhancement techniques and careful hyperparameter tuning, the developed model not only achieved exceptional accuracy but also demonstrated its efficiency in mapping class distributions and providing insights into geospatial imagery for environmental conservation and urban development planning.

In 2024, Fayaz et al. focused on land-cover classification using deep learning techniques, specifically transfer learning with Inception-v3 and DenseNet121 architectures, to enhance urban planning, agricultural zoning, and environmental monitoring. The study strategically fine-tuned versions of Inception-v3 and DenseNet121 networks, enabling the models to adapt to the intricacies of land-cover classification by improving pre-existing knowledge from extensive datasets. Through experiments on the UC-Merced_LandUse dataset, the approach achieved impressive results, including 92% accuracy, 93% recall, 92% precision, and a 92% F1Score.

3. Materials and method

3.1 Dataset

The EuroSAT dataset is a comprehensive resource designed for LULC classification, improving advanced satellite imagery. The dataset is based on Sentinel-2 satellite imagery, which provides high-resolution images across multiple spectral bands. This allows for detailed analysis of land cover types. The dataset categorizes land cover into 10 distinct classes. These classes represent various types of land use, such as forests, urban areas, agricultural land, and water bodies, facilitating diverse applications in environmental monitoring and urban planning. There are two versions of the dataset available which are RGB and Multispectral (Helber et al., 2018). In this study the multispectral version of the dataset is used to generate LULC maps.

3.2 Method

Figure 1 illustrates the workflow of the methodology used in this study.

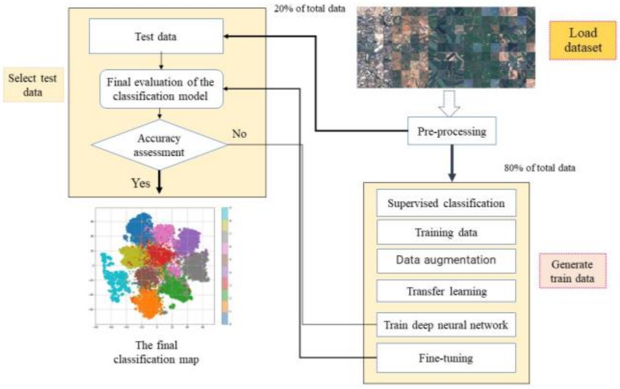

Figure 1. The methodology flowchart.

3.2.1 Train and test data collection: A total of 27,000 images in the database were used to train and test the network. Eighty percent of the data was used for training networks, with 21600 images being used. As mentioned, we dedicated 20 percent of the total data, which represents a total of 5,400 images, for testing the models. These are data that were not present in the training process and are actually new images for the network.

3.2.2 Data augmentation: Data augmentation in data analysis is a technique used to increase the amount of data by adding revised versions of existing data or newly created artificial data from existing information. This method acts as a regularizer and helps to reduce overfitting in training models. In addition, image recognition algorithms are improved when transferring images presented in virtual environments to real-world data. In fact, using data augmentation, it is possible to extract multiple images from one image, which actually increases the performance of the training process of the model. The study used a random rotation of 0.2, random zoom and a random flip horizontally and vertically on the training data.

3.2.3 Models: Four models have been considered for this study. 1-layer CNN, 4-layers CNN, VGG-16 and finally an improved ResNet-50 model. 1-layer CNN model consists of one Conv2D layer. In 4-layers CNN, unlike simple CNN, four Conv2D layers are used. Dropout is used in both CNN models. Two other methods, VGG-16 and ResNet-50 are pre-trained models whose weights are updated using imagenet and finetuning in the model. In improved ResNet-50 model, dropout, learning rate schedule, and kernel regularizer are employed in the architecture. Also, unfreezing last 10 layers is done with the aim of improving performance of the model training and fine-step.

To achieve this goal, Tensorflow and Keras libraries, which are the most popular deep learning frameworks, have been used. Also, the process of coding and running the code in the Google Colab environment has been done with GPU, which accelerates the implementation process. Also, the optimizer, Adam, with a learning rate of 0.000001, and the sparse categorical cross entropy loss function were used. All models are trained with 50 epochs.

3.2.4 Accuracy validation: The evaluation of the models has been done both quantitatively and qualitatively or visually. Quantitative evaluation includes calculating OA, Precision, Recall and F1Score and visual evaluation is conducted by comparing the output of classes through intra-class and inter-class criteria.

4. Results

As mentioned earlier, the results and evaluations have been examined in both quantitative and qualitative ways. The results of each model are illustrated and discussed below.
It should be noted that for convenience, the names of the classes are coded as shown in Table 1.

Code	Class name
1	Annual Crop
2	Forest
3	Herbaceous Vegetation
4	Highway
5	Industrial
6	Pasture
7	Permanent Crop
8	Residential
9	River
10	Sea Lake

Table 1. Encoding of class names

4.1 1-layer CNN

The separability diagram of the classes in the feature space is shown in Figure 2.

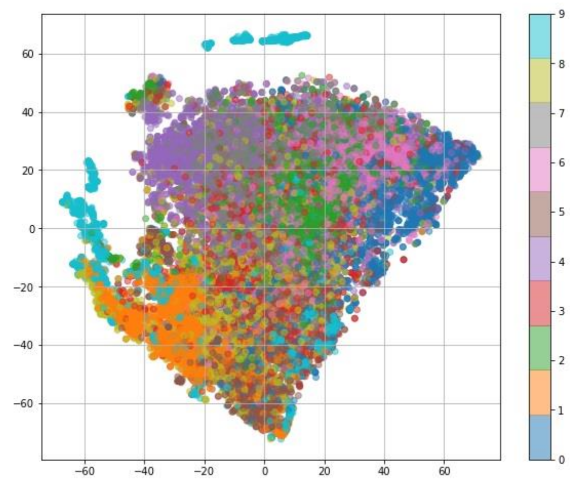

Figure 2. Classification diagram in feature space for 1-layer CNN.

As can be seen in Figure 2, 1-layer CNN was not a suitable classifier. Because, the model could not distinguish different classes well and the class diagram in feature space is mixed. However, it can be said that classes such as Forest, Industrial, Sea Lake and Annual Crop have been distinguished better. But in general, the network has not performed well in classification.

The confusion matrix for 1-layer CNN model is shown in Figure 3.

According to Table 2, this network has an OA of 56.64%, which is not a good overall accuracy, although we cannot expect more from a simple CNN network with one layer. In general, as is evident in the Figure 2, The accuracy of residential class is lower than other classes and after that the highway class has the lowest value. Also, the highway class has the lowest precision. The highest accuracy is for the sea Lake class, which is a more distinct class, and this network has classified the distinct classes better.

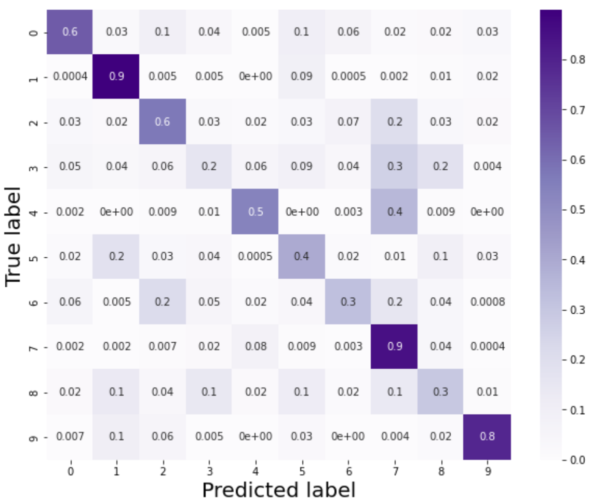

Figure 3. Confusion matrix for 1-layer CNN.

The accuracy evaluation table for this network is below:

Class	Accuracy	Precision	Recall	F1Score
1	93.87%	0.60	0.76	0.67
2	93.53%	0.87	0.64	0.74
3	90.12%	0.57	0.54	0.56
4	88.24%	0.19	0.40	0.26
5	93.44%	0.54	0.71	0.61
6	90.35%	0.47	0.45	0.46
7	91.46%	0.33	0.58	0.42
8	85.62%	0.85	0.42	0.56
9	89.94%	0.37	0.39	0.38
10	96.5%	0.78	0.88	0.82

Table 2. Quantitative analysis of 1-layer CNN network

4.2 4-layers CNN

The separation of classes in the 4-layers CNN model is shown in Figure 4.

This network has made a better distinction than the 1-layer CNN network. So that we have less mixed classes, but there are still some false classifications. For example, the highway class is one of the mixed classes whose pixels can be seen in other classes. Another mixed class is Permanent Crop, which is also present in its surrounding classes, such as Herbaceous Vegetation class. In the two classes Sea Lake and Pasture, we see the scattering and separation of pixels in groups in different places of the map, which represents the very inappropriate classification of these two classes.

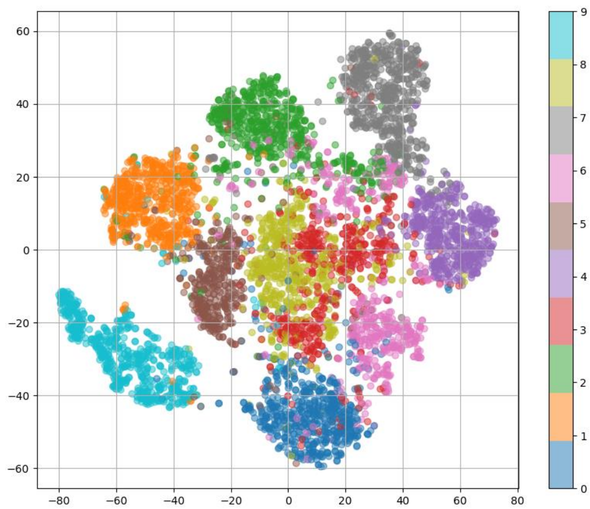

Figure 4. Classification diagram in feature space for 4-layers CNN.

The confusion matrix and quantitative measures are shown in Figure 5 and Table 3.

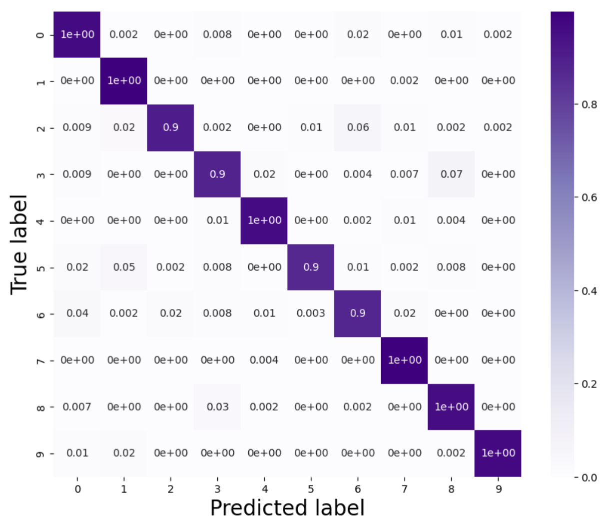

Figure 5. Confusion matrix for 4-layers CNN.

Table 3 shows to accuracy measures for 4-layers CNN model:

Class	Accuracy	Precision	Recall	F1Score
1	97.1%	0.84	0.89	0.86
2	98.57%	0.95	0.91	0.93
3	95.13%	0.88	0.71	0.79
4	96.15%	0.78	0.83	0.81
5	97.32%	0.74	0.97	0.84
6	98.34%	0.87	0.92	0.90
7	95.96%	0.65	0.88	0.75
8	96.63%	0.97	0.77	0.86
9	97.52%	0.91	0.86	0.88
10	99.03%	0.96	0.94	0.95

Table 3. Quantitative analysis of 4-layers CNN network

According to Table 3, the OA of the 4-layers CNN method is 85.88%, which is much better than the simple CNN network. The lowest accuracy criterion of this network is related to the Herbaceous Vegetation class, although the difference with the Pasture class is also small. The highest accuracy criterion here is also for the sea Lake class. Generally, the accuracy values for this network are close to each other.

4.3 VGG-16

As Figure 6 shows, this network has a better visual view than the 1-layer and 4-layers CNN. For example, the dispersion of the Pasture class in this network has completely disappeared and for the Sea Lake class it has also decreased so much that it can be considered a continuous class. In this network, the highway class has less mixed pixels that can be ignored. This network has actually provided a better output than the two first networks. There are fewer mixed pixels and false classification. Also, classes have better cohesion.

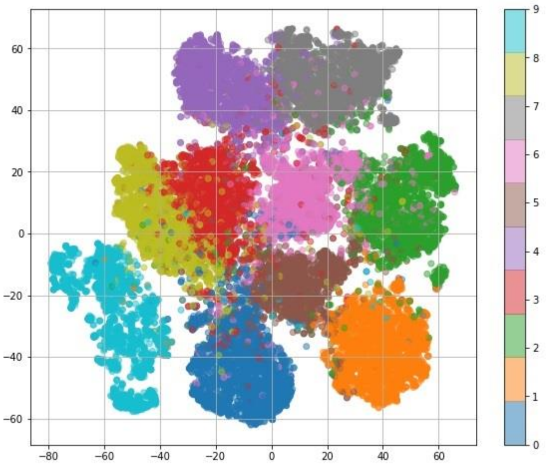

Figure 6. Classification diagram in feature space for VGG-16.

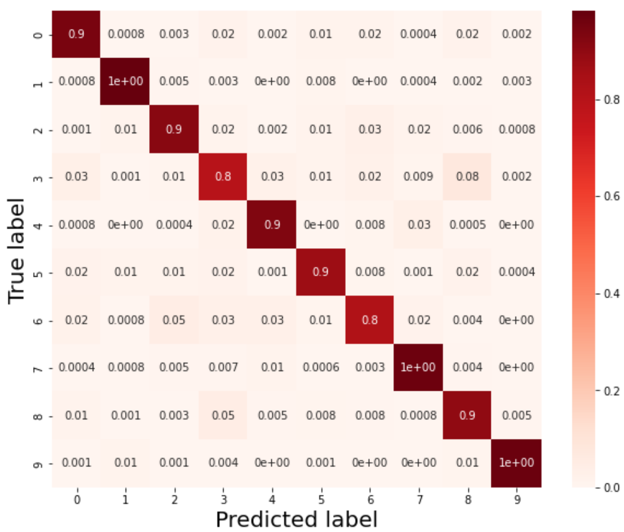

Figure 7. Confusion matrix for simple VGG-16.

According to Table 4, the values of Recall, F1Score and precision are in the range of 0.8 to 0.9, which indicates the good classification of this network. Also, the OA for this model is 91.22%. In fact, our main goal is that the classifier can distinguish similar classes well. The VGG-16 network has been able to separate these similar classes to an acceptable extent. For example, it was able to classify Annual Crop and Permanent Crops, both of which have agricultural coverage, with the least mixed pixels. Also, two classes, Forest and Pasture, which have vegetation, are well separated. Also, the Highway and Residential classes, which may be similar, are well separated. In addition, this network has less false classification compared to previous networks.

Below, in Table 4 the quantitative measures are shown:

Class	Accuracy	Precision	Recall	F1Score
1	98.61%	0.93	0.93	0.93
2	99.24%	0.98	0.95	0.96
3	98.24%	0.90	0.92	0.91
4	96.77%	0.83	0.84	0.83
5	98.39%	0.94	0.90	0.92
6	98.4%	0.91	0.93	0.92
7	97.34%	0.83	0.89	0.86
8	98.73%	0.96	0.92	0.94
9	97.31%	0.89	0.85	0.87
10	99.42%	0.95	0.99	0.97

Table 4. Quantitative analysis of VGG-16

4.4 Improved ResNet-50

According to Figure 8, it is clear that the classes are well separated and there are few mixed pixels. Some of these mixed pixels belong to Sea Lake and Forest classes. In other classes, this classification has been done well.

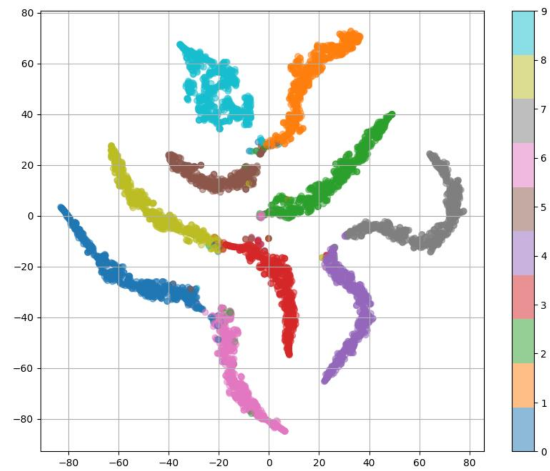

Figure 8. Classification diagram in feature space for improved ResNet-50.

The confusion matrix for this model is shown in Figure 9:

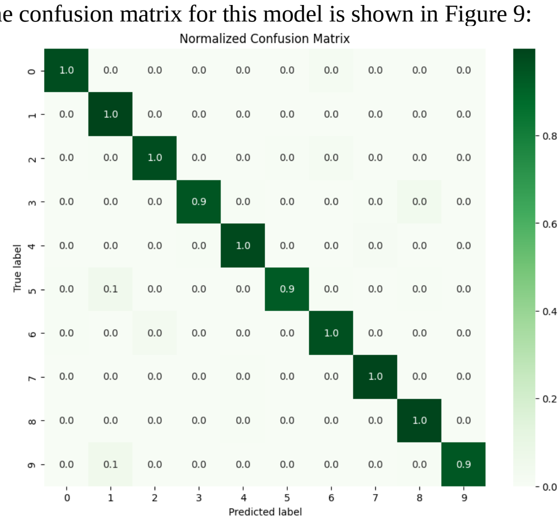

Figure 9. Confusion matrix for improved ResNet-50.

Below, in Table 5 the quantitative measures are shown:

Class	Accuracy	Precision	Recall	F1Score
1	100%	1	1	1
2	98%	0.83	1	0.91
3	100%	1	1	1
4	100%	1	1	1
5	100%	1	1	1
6	99%	1	0.90	0.95
7	100%	1	1	1
8	100%	1	1	1
9	100%	1	1	1
10	99%	1	0.90	0.95

Table 5. Quantitative analysis of improved ResNet-50

As it is clear from Table 5, a very good classification has been done quantitatively and the high values of accuracy and F1Score are proof of this claim. Also, the accuracy of the improved ResNet-50, is 95.9% which is the highest among all models considered in this study. Unlike the previous models, the accuracy of the Sea Lake class is lower compared to other classes. As showed in Figure 8 and discussed earlier, Sea Lake and Forest classes have mixed pixels and this has made them less accurate than other classes. On the other hand, it is clear that similar classes related to vegetation are well separated from each other.

5. Conclusion

In this research, we addressed the challenges related to disaster and risk management. One of the most important tools that help us in this field is the generating of LULC maps of the study area. For this, we used EuroSAT based on Sentinel-2 images. The proposed dataset consists of 10 classes covering 13 different spectral bands with a total of 27,000 labeled images and reference ground. Deep CNNs have been used for this dataset. For the EuroSAT dataset, we analyzed the performance of RGB spectral bands. The purpose of this research was investigating how to implement learning transfer architectures for LULC classification. The results showed that transfer learning is a very reliable method that can produce the best overall results. Also, the use of data augmentation techniques increased the diversity of the dataset, because these techniques only enhanced the visual changes of each training image without creating new spectral or topological information. Model improvement techniques such as changing the optimizer, and learning rate, modifying loss functions, for more efficiency of the models, improves the performance and finally reduces the required computing time. Our improved ResNet-50 and VGG-16 architectures produced better results than CNN architecture, since they used the pre-trained models. Generating datasets with inter and intra-class variability supported by strong deep learning architectures with data augmentation techniques can effectively increase the representation strength of deep learning network. Therefore, the proposed architecture is effective exploitation of existing satellite datasets and deep learning methods to achieve the best performance. These applications can be extended to several real-world Earth observation applications for remote sensing scene analysis. Finally, we reached an overall accuracy of about 95.9% for the proposed ResNet-50 network, which actually shows that the pre-trained networks are very suitable for classification.

References

Naushad, R., Kaur, T., Ghaderpour, E. Deep Transfer Learning for Land Use and Land Cover Classification: A Comparative Study. *Sensors* 2021, *21*, 8083. https://doi.org/10.3390/s21238083.

Sobhana, M., Chaparala, S.C., Indira, D.N., Kumar, K.K. 2022. A disaster classification application using convolutional neural network by performing data augmentation. *Indonesian Journal of Electrical Engineering and Computer Science*, DOI: http://doi.org/10.11591/ijeecs.v27.i3.pp1712-1720.

Dastour, H., Hassan, Q.K. A Comparison of Deep Transfer Learning Methods for Land Use and Land Cover Classification. *Sustainability* 2023, 15, 7854. https://doi.org/10.3390/su15107854.

Kunwar, S., Ferdush, J. 2023. Mapping of Land Use and Land Cover (LULC) using EuroSAT and Transfer Learning. *arXiv preprint arXiv:2401.02424*.

Fayaz, M., Nam, J., Dang, L.M., Song, H.-K., Moon, H. Land-Cover Classification Using Deep Learning with High-Resolution Remote-Sensing Imagery. *Appl. Sci.* 2024, *14*, 1844. https://doi.org/10.3390/app14051844.

Ali, K., Johnson, B.A. Land-Use and Land-Cover Classification in Semi-Arid Areas from Medium-Resolution Remote-Sensing Imagery: A Deep Learning Approach. *Sensors*, 2022, 22, 8750. https://doi.org/10.3390/ s22228750.

P. Helber, B. Bischke, A. Dengel, D. Borth, "Introducing Eurosat: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification," IGARSS 2018 - 2018 *IEEE International Geoscience and Remote Sensing Symposium*, Valencia, Spain, 2018, pp. 204-207, doi: 10.1109/IGARSS.2018.8519248.

Rangel, A., Terven, J.R., Esparza, D.M., Chavez-Urbiola, E.A. 2024. Land Cover Image Classification. *ArXiv, abs/2401.09607*.

V Pushpalatha, P B Mallikarjuna, H N Mahendra et al. Deep Learning-Based Land Use and Land Cover Classification for Change Detection Studies, 18 July 2024, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-4606544/v1]

Thenkabail, P. S. 2015. *Remote Sensing of Water Resources, Disasters, and Urban Studies*. https://doi.org/10.1201/b19321.

Hannah Ritchie, Pablo Rosado 2022, "Natural Disasters" Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/natural-disasters' [Online Resource]

Kumar, D. , Bassill, N., Ghosh, S., 2024. Analyzing recent trends in deep-learning approaches: a review on urban environmental hazards and disaster studies for monitoring, management, and mitigation toward sustainability. *International Journal on Smart Sensing and Intelligent Systems*,17(1) -. https://doi.org/10.2478/ijssis-2024-0014.

Alem, A., Kumar, S. Transfer learning models for land cover and land use classification in remote sensing image. *Appl. Artif. Intell.* 2022, *36*, 2014192.

Van Westen, C. J., 2000. Remote sensing for natural disaster management. *International archives of photogrammetry and remote sensing*, 33(B7/4; PART 7), 1609-1617.

H. Hartatik, M. K. Anam. 2023. Comparison of Convolutional Neural Network Architecture on Detection of Helmet Use by Humans. *Elinvo (Electronics, Informatics, and Vocational Education)*, *8*(1), 44–54. https://doi.org/10.21831/elinvo.v8i1.52104