Extended Classification of Geographical Features: Desert, Mountain, and Lake Detection from Satellite Images Using Deep Feature Fusion

Zahraa Msheik, Lynn Hamieh, Michael Geha

Extended Work based on: Automatic Desert/Mountain Detection from Satellite Image Using Deep Transfer Learning [16]

***Abstract*—Today, satellite images are used widely to track the Earth and study changes in the environment. In recent years, deep learning has made it possible to analyze these images more accurately and with far less manual effort. In this work, we present an enhanced deep-learning approach that automatically identifies three types of land regions desert, mountain, and lake. This expands on earlier studies that focused only on desert and mountain separation and introduces a more complex three-class classification task.**

Our workflow includes image collection and resizing, deep feature extraction using DenseNet models, feature reduction, feature fusion, classification, cross-validation of three-fold to evaluate consistency. Two DenseNet variants are used to extract features, which are then reduced by half and combined in sequence to create a richer and more informative feature representation. This fused feature vector is then used to classify the input satellite images into the three defined categories.

The results show that the proposed multi-model fusion strategy reaches a maximum accuracy of 99.25% with the FFV-SM classifier, outperforming each individual DenseNet model and previously reported two-class results. The Random Forest method, which was used in earlier desert–mountain studies, also performs well in the new three-class setting, achieving 98.70% accuracy. These outcomes demonstrate that combining deep feature extraction with classical machine-learning classifiers is an effective and reliable direction for improving land-cover recognition in satellite imagery.

Keywords: Deep Learning, Satellite Image, DenseNet, Feature Fusion, Multi-Class Classification, Lake Detection, Environmental Monitoring

I. Introduction

Satellite images are now one of the main tools used to study the Earth and understand how different environments change over time. They cover large areas, provide continuous monitoring, and store information that helps researchers compare old landscapes with new ones. Because of these advantages, satellite imagery has become very useful for tracking long-term environmental conditions in many places around the world. Using remote sensing data, the capabilities have been widely exploited in environmental monitoring and desertification assessment studies [1], [4].

Working directly in desert, mountain, or lake areas is not always easy. Many of these regions are far away, difficult to reach, and expensive to study using field methods. For that reason, using satellite images offers a more practical and reliable solution. Automatic image analysis also helps researchers learn more about the condition of these regions, follow changes in land cover, study climate effects, and support decisions that protect natural environments. Still, getting accurate results from satellite images requires advanced techniques and well-designed systems. The effectiveness of satellite imagery for large-scale land-cover monitoring and environmental analysis has been enhanced using recent advances in multispectral data analysis and artificial intelligence [2], [5], [6].

In this work, we present a deep learning approach for identifying desert, mountain, and lake regions from satellite images. Most earlier studies focused on only two classes, mainly desert and mountain. However, the goal is extended to a three-class classification problem, which is more realistic and more challenging. The method is to use pre-trained DenseNet models to extract the features of the image, and test different classifiers, including SoftMax and other machine-learning techniques.

We analyzed three DenseNet versions: DenseNet121, DenseNet169, and DenseNet201. After testing with SoftMax, features from DenseNet121 and DenseNet201 gave the strongest results. The features were then merged into a single feature set and tested with SoftMax, Decision Tree, and Random Forest. The SoftMax classifier combined with the fused features achieved the best accuracy, reaching 99.2%, which shows that images of deserts, mountains, and lakes can be clearly separated using the method.

The work introduces a fused-feature framework based on DenseNet121 and DenseNet201, showing strong performance with the SoftMax classifier. It also expands land-cover classification by adding lake detection, making the approach more complete and applicable to real environmental studies.

The paper is organized as follows: Section II reviews related work, Section III explains the method, and Sections IV and V present the results and conclusion. Accuracy, precision, sensitivity, and specificity are used to measure performance. The outcomes show that the proposed method is reliable and effective for classifying satellite images into desert, mountain, and lake regions.

II. Literature Review

In recent studies, automated and algorithmically processed image examination has been widely used in multiple fields, particularly in environmental assessment and land-cover analysis. In this context, satellite images (SI) are commonly used to assess the performance of developed deep learning techniques (DLTs), as they provide rich spatial and visual information about geographical regions. Several studies have examined satellite images using different DLTs and classifiers. For example, the works reported in [7] and [8] applied convolutional neural network–based models for satellite image classification and demonstrated that deep learning approaches can achieve higher accuracy than traditional image-processing methods. However, these studies mainly focused on relatively simple or binary classification tasks and did not address more complex multi-class landscape classification problems. Additional related research on environmental and land-cover analysis using satellite imagery can be found in [9] and [10], which further confirm the effectiveness of deep learning for satellite image examination, but similarly rely on limited classification settings or single-feature representations.

Examining the dominant information contained in a satellite image database remains one of the key tasks in satellite image analysis, as the quality of extracted features directly influences detection accuracy. Earlier works confirm that DLT-based satellite image examination is widely adopted by researchers due to its ability to automatically learn discriminative features from complex image data. In our previous work [16], a deep learning framework based on DenseNet models was proposed for classifying desert and mountain regions and achieved high detection accuracy within a binary classification framework. While effective, that approach was limited to two landscape classes and did not explore feature fusion or additional geographical categories. The present study builds on this prior work by extending the classification task to include lake regions and by introducing a deep feature fusion strategy that combines features extracted from multiple DenseNet models, enabling improved discrimination in a more challenging multi-class setting.

III. Methodology

The performance of the proposed algorithm driven data analysis framework depends on the approach used to examine the selected satellite image (SI) database. In this study, a DenseNet (DN)-based deep learning technique (DLT) is developed to effectively analyze desert, mountain, and lake images, with the various stages illustrated in Fig. 1. The chosen images are extracted from the SI repository described in [11]. All images are resized to 224×224×3 pixels, followed by appropriate data augmentation to enhance data diversity. During training, SI images are transformed into deep feature representations of size 1×1×1000 using pre-trained DenseNet models and are used for three-class classification with a SoftMax (SM) classifier. Among the evaluated models, DenseNet-121 and DenseNet-201 produced superior results and were therefore selected for feature fusion. The fused feature vector (FFV) is generated using a 50% feature reduction step followed by serial concatenation. In this work, the 50% feature reduction is performed using a random dropout strategy, where half of the extracted feature elements are randomly retained from each DenseNet output vector. Specifically, 500 features are selected from the 1000-dimensional feature vector of each model, and the reduced feature vectors are concatenated serially to form a final FFV of size 1×1×1000. Performance evaluation is conducted using a three-class classification framework involving desert, mountain, and lake categories. The performance of the model is evaluated using performance metrics that include true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), which are evaluated using the confusion matrix (CM). Final evaluation metrics include accuracy (AC), precision (PR), sensitivity (SE), and specificity (SP). The experimental results confirm that the FFV-based DLT achieves a classification accuracy of 99.2% with the FFV-SM classifier. These results show the robustness and effectiveness of the proposed approach in accurately identifying desert, mountain, and lake regions in satellite imagery.

A. SI Database

A representative image dataset was selected to demonstrate the utility of the proposed DLT. In this study, the SI dataset presented in [11] was employed. From this dataset, 1,200 images were selected for each of the three classes—namely desert, mountain, and lake.

Various data augmentation techniques were implemented, including image rotation, horizontal and vertical flipping, slight zooming, and adjustments to brightness and contrast. These augmentations helped increase the diversity of the dataset and effectively reduced overfitting. This exposed the model to more diverse and en-

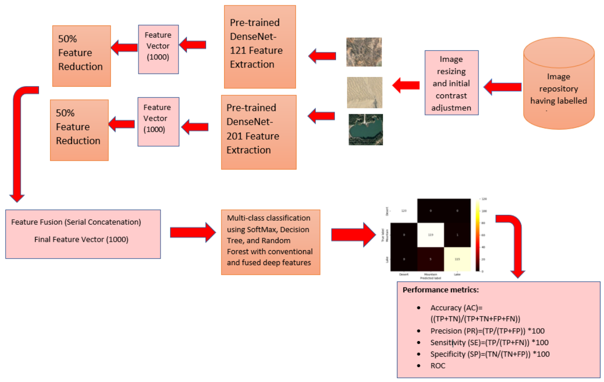


Fig. 1: Developed a deep learning framework for classifying of the chosen SI database.

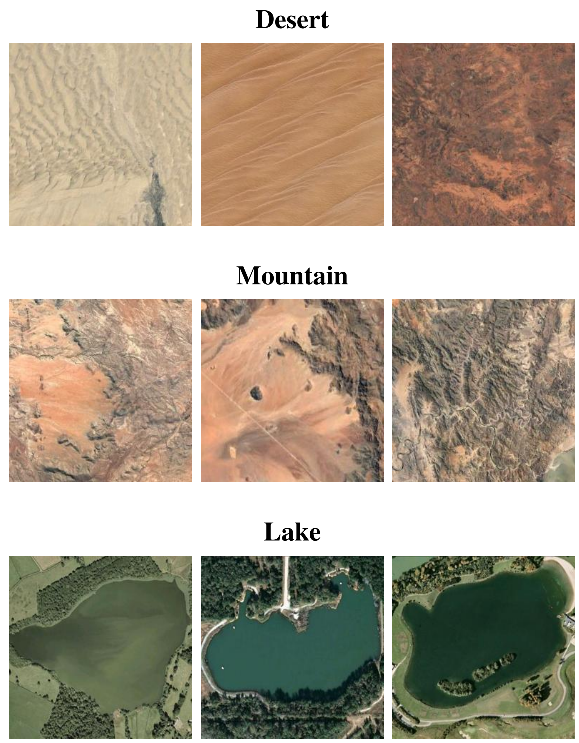


Fig. 2: Representative images from the expanded SI dataset showing desert, mountain, and lake landscapes.

riched training samples that improved its generalization capability.

For experimental evaluation, 800 images from each class were used for training, while the remaining 400 images per class were reserved for testing. In addition, threefold cross-validation was employed to ensure a reliable and robust assessment of the model performance. The dataset was divided into three folds, and in each iteration, a different subset of 800 images per class was used for training, while the remaining samples were used for testing. This procedure was repeated three times, and the highest classification accuracy achieved across the three runs was reported.

Fig. 2 shows representative images from the desert, mountain, and lake classes, highlighting the diversity and quality of the dataset.

B. DL Model

The study focused on examining pre-trained DenseNet Models . This investigation utilizes DN-models DN-121, DN-169, and DN-201 [12], [13]. Compared to other extensively studied deep learning architectures like VGG and ResNet variants, DN-models have significantly fewer configurable parameters (less than 20.2 million). The initial configurations applied to the chosen DN-models include: number of epochs set to 40, monitored metrics being accuracy and loss, pooling method as Max, optimizer as SGD, learning rate at 1e-4, activation function as ReLU, and a batch size of 8 [14]. This research first implements the selected models on the SI database and assesses the accuracy achieved using SM and individual features.Based on the acquired data, the FFV is produced, and relevant information about these values is accessible in [15], [16].

C. Model Interpretability using Grad-CAM

To ensure that the deep learning model makes decisions based on meaningful and relevant image regions, we applied Gradient-weighted Class Activation Mapping (Grad-CAM). Grad-CAM generates heatmaps that highlight the areas of an image that most strongly influence the model's predictions. While the underlying mechanism involves backpropagating gradients from the target class, here we provide only a brief overview, as the audience may already be familiar with this technique.

By applying Grad-CAM to sample images from each class (desert, mountain, and lake), we can visualize the attention of the network and confirm that it focuses on sensible and distinctive features relevant to each landscape. This not only validates the interpretability of the model but also supports the reliability of the classification results reported in the study

D. Performance Verification

The effectiveness of the proposed deep learning tool (DLT) was initially validated using a simpler two-class classification scenario (desert vs. mountain) as a preliminary evaluation, before extending the analysis to a more complex three-class problem that included lake landscapes. All experiments were conducted on a system equipped with an Intel Core i7-11800H processor operating at 2.30 GHz, 16 GB of RAM, and an NVIDIA GeForce RTX 3070 Laptop GPU with 8 GB of dedicated VRAM. The model and the experiment were carried out using Python 3.10 and the TensorFlow 2.8.

Standard performance metrics for model evaluation included accuracy (AC), precision (PR), sensitivity (SE), and specificity (SP), all of which were derived from the confusion matrix (CM). In addition, receiver operating characteristic (ROC) analysis was used to further assess the reliability of the classification results.

A DenseNet-121 model trained on a three-class dataset without feature fusion was used as the baseline configuration. The proposed enhancements, such as the introduction of the lake class and the feature-fusion vector (FFV-SM), consistently outperformed the baseline model and yielded more robust and accurate classification performance.

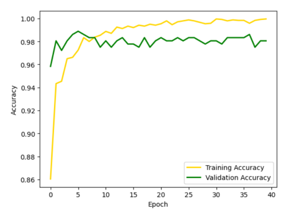
IV. Results and Discussion

At first, the work began by checking the performance of the DenseNet models on different satellite images. The model extracted features from the images , and the SoftMax layer handled the final outcome. Lake images was added to the dataset, which made things more diverse and not repetitive. This extra class, along with FFV-SM fusion approach, helped raise the results higher. Figures 3-5 show how DenseNet-121 and DenseNet-201 performed while training and validating the desert, mountain, and lake classes. the accuracy and loss lines remained stable, so the models did not drift or show signs of memorizing the data.

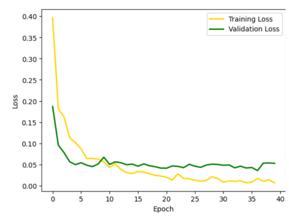
The overall accuracy reached 99.2%. This value is considerably high , though it needs to be read carefully. These three landscapes look quite different from each other, so differentiating between them is not the hardest task. To avoid the model from copying patterns, the dataset was rearranged and expanded, pre-trained weights were used, and the training and test images stayed seperate. This approach made the model work for the results instead of memorizing data.

Grad-CAM was implemented to see what parts of the images guided the predictions (Fig.6). The heatmaps highlighted the key areas: mountain edges, water surfaces, and smooth sand areas. This shows that the model was responding to the landscape features rather than random marks in the background. The deeper convolution layer (CL1-CL4) also showed how the images transformed from raw pixels to well-defined shapes and textures. The first results (Fig.5 and Table 1) illustrated steady performance across all three types, with minor errors. This was validated by the ROC curves that showed clear distinction between classes. Adding the fusion improved the outcome again, ending with 99.2% accuracy . Compared to the other methods, such as RF-based fusion at 98.7%, the difference looks small but did not decrease across the folds. Robustness matters, especially when images fall between categories.

Overall, the model managed all three image types and was able to differentiate between satellite images with strong accuracy. Further work may look at wider Earth-observation datasets and additional environmental groups. The key advancement in results come from adding lake images to the dataset and using the FFV-SM fusion method, which form the main novelty of this approach.

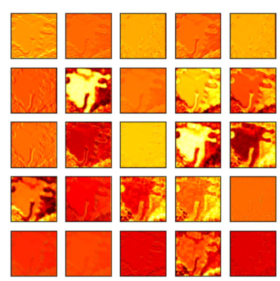


(a) Accuracy

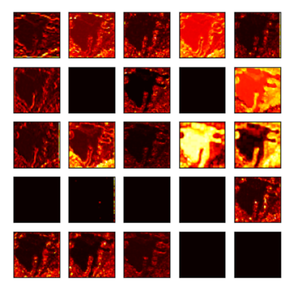


(b) Loss

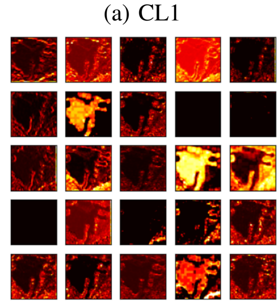
Fig. 3: Training phase experimental results of DN-201 model with SoftMax classification .



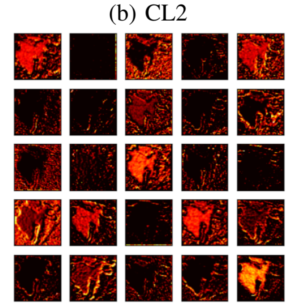
(a) CL1



(b) CL2



(c) CL3

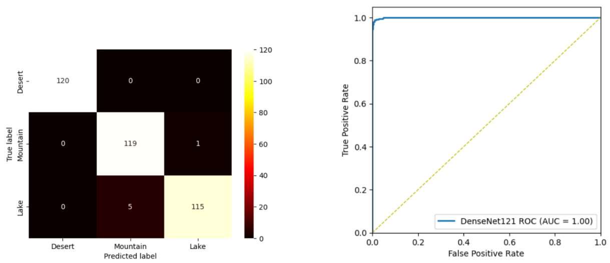


(d) CL4

Fig. 4: Several convolutional-layer (CL) outcomes of DN-121.

V. Conclusion

This study presented a deep learning architecture for the effective classification of desert, mountain, and lake landscapes from satellite images. The proposed



(a) CM (b) ROC_curve

Fig. 5: Preliminary results obtained using DN-121 with SoftMax-based classification.

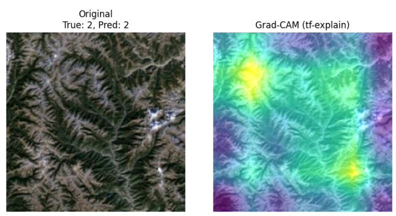


Fig. 6: Grad-CAM visualization obtained using the DenseNet-121 model for a representative satellite image. The highlighted regions indicate the discriminative areas that guided the classification decision.

framework employed pre-trained DenseNet models and incorporated a feature-fusion vector (FFV-SM), demonstrating strong potential in capturing discriminative patterns across multiple landscape types. The inclusion of the lake class further complicated and diversified the dataset, thereby confirming the robustness of the proposed model. The best-performing approach achieved a classification accuracy of 99.2%, surpassing earlier two-class configurations. In addition, threefold cross-validation validated the reliability and robustness of the framework across all classes.

Given its strong performance, the framework underscored the novelty of combining multi-model feature fusion with an expanded three-class dataset, enabling precise multi-class land-cover classification. In future work, this framework could be extended to a broader range of satellite imagery and additional types of land-cover, such as forests or urban areas, incorporating advanced fusion strategies or transformer-based architectures to further enhance classification performance.

TABLE I: Performance evaluation metrics achieved using differnt DN-variants

Model	TP	FN	TN	FP	AC	PR	SE	SP
DenseNet121	118.0	2.0	238.0	2.0	98.8889	98.3333	98.3333	99.1667
DenseNet169	116.0	4.0	236.0	4.0	97.7778	96.6667	96.6667	98.3333
DenseNet201	117.6667	2.3333	237.6667	2.3333	98.7037	98.0556	98.0556	99.0278
FFV-SM	118.6667	1.3333	238.6667	1.3333	99.2593	98.8889	98.8889	99.4444
FFV-DT	108.6667	11.3333	228.6667	11.3333	93.7037	90.5556	90.5556	95.2778
FFV-RF	117.6667	2.3333	237.6667	2.3333	98.7037	98.0556	98.0556	99.0278

VI. References

References

[1] Liu, Z., Si, J., Deng, Y., Jia, B., Li, X., He, X., Zhou, D., Wang, C., Zhu, X., Qin, J., Ndayambaza, B., Wang, B. (2023). Assessment of land desertification and its drivers in semi-arid alpine mountains: A case study of the Qilian Mountains Region, Northwest China. Remote Sensing, 15(15), 3836.

[2] Lim, S. L., Sreevalsan-Nair, J., Daya Sagar, B. S. (2024). Multispectral data mining: A focus on remote sensing satellite images. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 14(2), e1522.

[3] Davis, D. S. (2019). Object-based image analysis: A review of developments and future directions of automated feature detection in landscape archaeology. Archaeological Prospection, 26(2), 155–163.

[4] Fichera, C. R., Modica, G., Pollino, M. (2012). Land cover classification and change-detection analysis using multi-temporal remote sensed imagery and landscape metrics. European Journal of Remote Sensing, 45(1), 1–18.

[5] alukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y. A., Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations: A review. Remote Sensing, 12(7), 1135.

[6] Lemenkova, P. (2024). Artificial intelligence for computational remote sensing: Quantifying patterns of land cover types around Cheetham Wetlands, Port Phillip Bay, Australia. Journal of Marine Science and Engineering, 12(8), 3390.

[7] Pritt, M., Chern, G. (2017). Satellite image classification with deep learning. In Proceedings of the IEEE Applied Imagery Pattern Recognition Workshop (AIPR) (pp. 1–7). IEEE.

[8] Tumpa, P. P., Islam, M. S. (2024). Lightweight parallel convolutional neural network with SVM classifier for satellite imagery classification. IEEE Transactions on Artificial Intelligence.

[9] Wang, Z., Shi, Y., Zhang, Y. (2023). Review of desert mobility assessment and desertification monitoring based on remote sensing. Remote Sensing, 15(18), 4412.

[10] Ramadasan, S., Vijayakumar, K., Raju, V., Prabha, S. (2024). Automatic lake and river detection from satellite images using MobileNet scheme. In Proceedings of the International Conference on Science, Technology, Engineering and Management (ICSTEM) (pp. 1–5). IEEE.

[11] Kaggle SkyView Dataset.

[12] Shan, L., Wang, W. (2021). DenseNet-based land cover classification network with deep fusion. IEEE Geoscience and Remote Sensing Letters, 19, 1–5.

[13] Keras. (2024). Keras applications documentation. https://keras.io/api/applications/

[14] Geetha, R., Ramadasan, S., Vijayakumar, K., Prabha, S. (2024). Automatic classification of plant leaf images into healthy and disease class with EfficientNet: A study. In Proceedings of the International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI) (pp. 1–5). IEEE.

[15] Ramadasan, S., Augasthega, R., Vijayakumar, K., Prabha, S. (2024). Detection of foot-ulcer from digital photographs using MobileNet variants with feature fusion. In Proceedings of the Ninth International Conference on Science, Technology, Engineering and Mathematics (ICONSTEM) (pp. 1–6). IEEE.

[16] Kadry, S., Prudhvi, M. N., Narayanan, M., Rajinikanth, V. (2025). Automatic desert–mountain detection from satellite images using deep transfer learning. In Lecture Notes in Networks and Systems. Springer.

[17] Msheik Z., Multiclass Geographical Classification Dataset, GitHub, 2025. https://github.com/zahraamsheik313/multiclass-geographical-classification