University of Tehran

Earth Observation and Geomatics Engineering

Online ISSN: 2588-4360

Homepage: https://eoge.ut.ac.ir/

Patch-Level Land Use and Land Cover Classification Using a Lightweight Convolutional Neural Network: A Case Study on Optical and SAR Data in Iran

Mohammad Alikhani[1] , Behnam Ebadati[2]✉ , and Reza Attarzadeh[3]

1. Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran. E-mail: m.alikhani1@email.kntu.ac.ir
2. Corresponding author, Department of Geomatics Engineering, ST.C., Islamic Azad University, Tehran, Iran. E-mail: st_b_ebadati@azad.ac.ir
3. Department of Geomatics Engineering, ST.C., Islamic Azad University, Tehran, Iran. E-mail: reza.attarzadeh@iau.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 2025-08-18
Received in revised form 2025-11-16
Accepted 2026-01-15
Available online 2026-05-12

Keywords:
Convolutional Neural Network,
Deep Learning,
LULC,
Earth Observation,
SAR,

ABSTRACT

Open access to Earth observation data, such as Copernicus and Landsat program data, along with remarkable advancements in deep learning, has played a significant role in computer vision tasks. Consequently, there is an increasing demand for well-labeled datasets and the development of optimized models that can be executed on resource-limited hardware. In this study, the Hybrid Multi-Block Convolutional Neural Network (HCNN) was employed for land use and land cover (LULC) classification at the patch level. HCNN was designed to achieve accuracy comparable to or higher than state-of-the-art models while reducing complexity to ensure faster training and optimized performance. To evaluate the model, paired datasets from Sentinel-1 and Sentinel-2 imagery were constructed using Google Earth Engine, covering study areas in Iran. The datasets were co-registered to ensure spatial alignment between the SAR and optical data. The performance of HCNN was benchmarked against three widely used CNN architectures, including DenseNet-121, ResNet-50, and VGG-19. HCNN demonstrated high classification accuracy on both optical and SAR datasets, achieving 99.56% and 99.14%, respectively. These results consistently surpassed the performance of the other CNN architectures, highlighting HCNN's superior accuracy and efficiency across both data types. Moreover, HCNN required substantially less training time, completing in approximately one-eighth of the time needed by the benchmark models, highlighting its efficiency in both accuracy and computational cost.

Cite this article: Alikhani, M., Ebadati, B., & Attarzadeh, R. (2025). Patch-Level Land Use and Land Cover Classification Using a Lightweight Convolutional Neural Network: A Case Study on Optical and SAR Data in Iran, *Earth Observation and Geomatics Engineering*, Volume 9, Issue 1, Pages 47-60. http//doi.org/10.22059/eoge.2026.400906.1187

Publisher: University of Tehran.

DOI: http//doi.org/10.22059/eoge.2026.400906.1187

1. Introduction

The rapid advancement of spaceborne sensors has made remote sensing one of the most important methods for collecting data from the surface (Tshanga et al., 2024). The lack of need for field presence, extensive spatial coverage, and accessibility to hard-to-reach areas has made this technology highly efficient and cost-effective (McClelland et al., 2023). Remote sensing data comes in various forms, including radar, multispectral, and hyperspectral imagery, each offering unique advantages (Lu et al., 2020). Radar sensors, such as Synthetic Aperture Radar (SAR), provide all-weather, day-and-night imaging capabilities by actively emitting signals and measuring their reflections (Sahad, 2025). Multispectral sensors capture data across several broad wavelength bands, effectively identifying land cover types and vegetation health (Zhong et al., 2019). Hyperspectral imaging captures data across a large number of narrow spectral bands, providing rich spectral detail that enhances the detection and classification of subtle surface features (Zhao & Du, 2016). Moreover, the increasing availability of affordable and freely accessible remote sensing datasets, such as those provided by the Copernicus and Landsat programs, has greatly facilitated widespread use and advanced research in this field (Chaves et al., 2020).

Remote sensing is frequently applied in land use and land cover (LULC) classification (Wang et al., 2023). Despite notable advancements, it remains challenging, particularly in heterogeneous environments where diverse surface features coexist (Amin et al., 2024). Furthermore, factors such as spatial scale and geometric distortions can significantly affect the representation of LULC types in satellite imagery. Differences in viewing angles, including acquisitions from ascending and descending orbits, also contribute to this variability (Wang et al., 2023).

Therefore, extracting information using conventional approaches that rely on basic features such as colour, texture, or geometric structure often faces limitations. These limitations become more pronounced when identifying complex surface patterns, particularly in areas with diverse land cover types or varying imaging conditions (Castelluccio et al., 2015). Convolutional Neural Networks (CNNs) are recognized as powerful tools for overcoming these limitations. These networks can capture advanced and conceptual features, enabling the training of models based on this information and achieving better performance in identifying complex surface patterns (Alem & Kumar, 2020). Adaptability to transfer learning techniques can also enhance their performance, especially in scenarios with limited training data (Zhu et al., 2021).

Specifically, models trained based on VGGNet and DenseNet architectures have gained attention in many studies for achieving high classification accuracy (Arathi and Dulhare, 2023; Baral and Aryal, 2023; Gupta et al., 2024; Chowdhury et al., 2024; Li et al., 2020; Wang et al., 2022). These models, initially trained on large and diverse databases such as ImageNet (Deng et al., 2009), have learned to extract deep features that can be transferred to new tasks with limited labeled data. In recent years, these capabilities have become increasingly prominent, particularly in the field of computer vision based on remote sensing data. Several studies have successfully leveraged deep learning to classify land cover at the patch level. Rather than generating large-scale thematic maps, these studies focus on identifying land cover patterns within small patches of satellite imagery. Rengma and Yadav (2024) also used CNN models for LULC classification with Sentinel-2 data, working on 4,000 labeled images spanning 4 classes and obtaining accuracy beyond 90% (Rengma & Yadav, 2024).

In a study by Helber et al. (2018), LULC classification was performed using Sentinel-2 satellite imagery. They introduced the EuroSAT dataset, containing 27,000 labelled images across 10 LULC classes. Evaluations of deep CNN architectures, such as GoogleNet and ResNet50, yielded classification accuracies above 95% (Helber et al., 2018). More recently, Harini et al. (2024) utilized the EuroSAT dataset with a ResNet50-based CNN, achieving classification accuracy exceeding 90% (Harini et al., 2024). Krishnan et al. (2024) tested various CNN architectures on the EuroSAT RGB dataset, including simple CNN, ResNet50 and its variants, as well as VGG16 and VGG-19. Their experiments showed accuracies above 95%, with deeper networks like VGG-16 and VGG-19 reaching more than 97% (Krishnan et al., 2024). Yassine et al. (2021) demonstrated that using all 13 spectral bands of the EuroSAT dataset resulted in accuracy above 95%, which further improved to over 98% when spectral bands were combined with spectral indices (Yassine et al., 2021).

Similarly, other datasets have also been utilized for LULC classification using deep CNNs at the patch level. For instance, Maurya et al. (2024) applied deep CNNs, specifically VGG-19 and ResNet-50, for classification on the UC Merced dataset, which contains 21 classes and 2,100 aerial images. The VGG-19 model achieved a test accuracy of over 95%, while the ResNet-50 model reached over 90% accuracy (Maurya et al., 2024). Zhang et al. (2021) also employed the VGG-19 applied to land use classification using the same dataset, achieving an accuracy over 90% and highlighting the model's capability in classifying high-resolution remote sensing images (Zhang et al., 2021).

In our previous study (Alikhani et al., 2025), a Hybrid Multi-Block Convolutional Neural Network (HCNN) architecture was introduced for patch-level LULC classification. This architecture integrates residual, dense, inception, and squeeze-excitation blocks into a base CNN structure. This design aims to enhance classification accuracy while substantially reducing computational cost. In that study, two models based on the HCNN architecture were trained separately on the multispectral and RGB versions of the EuroSAT dataset. The dataset contains 27,000 images representing ten LULC classes collected from various regions of Europe. The results showed that both HCNN models not only achieved higher classification accuracy compared to baseline CNN architectures but also required significantly less training time, highlighting the

effectiveness of the proposed hybrid design in reducing computational complexity.

Building on the effectiveness of the HCNN architecture demonstrated in the case study on the EuroSAT dataset, the present study evaluates its performance on new datasets derived from a different geographical region. Specifically, the HCNN performance is assessed using SAR and optical datasets generated from Sentinel-1 and Sentinel-2 imagery, covering six LULC classes in Iran. Furthermore, its performance is compared with DenseNet-121, ResNet-50, and VGG19 models trained on generated datasets. The following sections provide a detailed description of the procedures for dataset creation and the characteristics of the data. Additionally, the architecture of the proposed model is described, which was designed to achieve high classification accuracy while minimizing computational complexity.

2. Methodology

2.1. Earth Observation Data

In this study, Earth observation data from the Sentinel-1 and Sentinel-2 missions, which are freely provided by the European Space Agency (ESA), were used. These data were acquired exclusively over various regions of Iran from the beginning to the end of 2024.

The Sentinel-1A satellite was launched in April 2014, followed by Sentinel-1B in April 2016. Both satellites are equipped with C-band SAR sensors, which allow image acquisition under various weather situations, in daylight and at night. SAR data are collected in the Interferometric Wide Swath (IW) mode, which is the most commonly used mode for land monitoring. This mode provides a spatial resolution of approximately 10 meters and a swath width of around 250 kilometers. The data are acquired using dual polarization modes, specifically VV and VH (Vanderhoof et al., 2023). Until 2022, the combined operation of Sentinel-1A and Sentinel-1B enabled a revisit cycle of 6 days for any given location. However, in December 2021, Sentinel-1B went out of service due to a power system failure, and its operation was officially terminated in 2022 (Nhangumbe et al., 2023).

Sentinel-2, analogous to Sentinel-1, operates with two complementary satellites named Sentinel-2A and Sentinel-2B. These satellites were designed to enhance both temporal and spatial coverage. Sentinel-2A was launched in June 2015, and Sentinel-2B followed in March 2017. This mission was developed to monitor land surfaces and natural resources, particularly for accurate observation of vegetation cover, agricultural lands, forests, urban areas, water resources, and environmental changes. Sensors carried by both satellites record imagery in 13 spectral bands, ranging from visible to near and shortwave infrared wavelengths. These bands are provided at three spatial resolutions: 10, 20, and 60 meters. With the simultaneous operation of both satellites, the revisit frequency of the Sentinel-2 satellite mission at the equator is roughly every 5 days, allowing for high-resolution temporal and spatial monitoring of surface phenomena (Phiri et al., 2020).

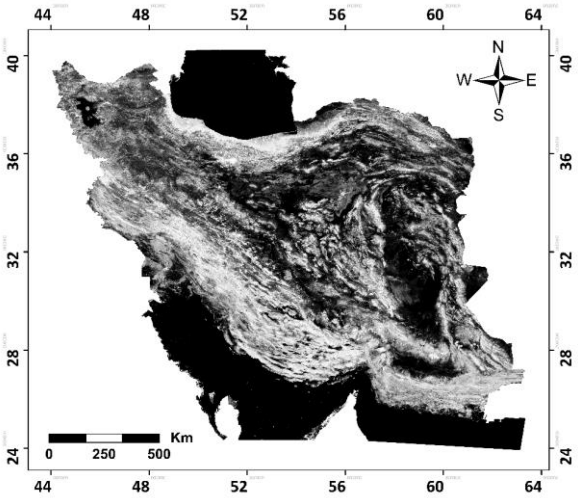

Figure 1. Sentinel-1 mosaic of Iran

2.2. Data Preprocessing

Data acquisition and preprocessing were conducted using the Google Earth Engine (GEE) platform. Most datasets available in the GEE Data Catalog have already undergone basic preprocessing. In other words, essential steps such as orbit file application, thermal noise removal, and radiometric calibration are automatically applied to the data, particularly for Sentinel-1 products. For example, SAR image backscatter values are pre-converted to decibels (dB) units as part of the radiometric calibration process.

However, to improve data quality, some custom preprocessing steps were necessary. For Sentinel-1 data, a circular median filter with a radius of 30 meters was applied to reduce speckle noise. For Sentinel-2 optical data, pixels affected by clouds were identified and masked utilizing the Cloud Mask component. These masked pixels were then replaced using a temporal compositing method. Additionally, all spectral bands with spatial resolutions other than 10 meters were resampled to 10 meters to ensure uniform spatial resolution across the dataset. In the next step, pairs of 64×64-pixel image patches were extracted from various regions across Iran. These patches were generated after performing spatial co-registration between the SAR images and their corresponding optical images, ensuring that each image pair represented a common LULC class. This process was carried out semi-automatically using the Google Earth Engine API within the Google Colab programming environment. The simultaneous use of both platforms significantly accelerated the extraction and generation of image patch pairs. All image patches were visually verified against very high-resolution Google Earth imagery to ensure label correctness. As a result, two separate datasets, one based on optical and SAR imagery, were created, sharing identical class labels and spatially aligned image patches.

Figure 2. Spatially matched patch extraction from optical and SAR images

2.3. Paired Datasets

Each dataset contains 600 geo-referenced image patches for each of the six land cover classes, including Barrenland, Farmland, Forest, River, Sea/Lake, and Urban, resulting in a total of 3,600 patches per dataset. All samples were collected at a 10-meter spatial resolution and extracted as fixed-size 64 × 64 pixel patches to ensure consistency across regions and sensor types.

The optical dataset consists of 12 Sentinel-2 spectral bands, while the SAR dataset comprises four channels generated from VV and VH polarizations acquired in both ascending and descending passes of Sentinel-1 imagery. This dual-polarization and dual-orbit configuration provides enhanced sensitivity to surface roughness, vegetation structure, and moisture conditions. A detailed overview of the datasets, including the number of samples per class and their geographic distribution across Iran, is provided in Table 1. In addition, Figure 3 shows the overall proportional distribution of samples, demonstrating the spatial diversity and balance of the collected data.

Table 1. Class-wise sample counts and geographic coverage

Class	Count	Geographic Coverage
Barrenland	600	Isfahan, Tehran, Semnan, Qom, Kerman, Yazd, Markazi, S. Khorasan, Alborz
Farmland	600	Khuzestan, E. Azarbaijan, Ardabil, Qazvin, Tehran, Lorestan, Qom, Isfahan, Alborz
Forest	600	Gilan, Mazandaran, Golestan
River	600	Khuzestan, E. Azarbaijan, Ilam, Kurdestan, Ardabil, Kermanshah, Lorestan
Sea/Lake	600	W. Azarbaijan, Bushehr, Hormozgan, Khuzestan, Mazandaran, Gilan, Golestan
Urban	600	Tehran, Khuzestan, E. Azarbaijan, Ardabil, Isfahan, Bushehr, Hamedan, Mazandaran
Total	3600	-

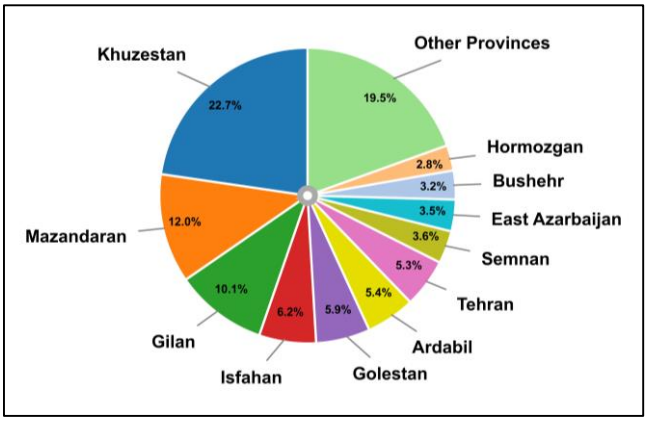

Figure 3. Proportional distribution of samples across Iran

2.4. Data Augmentation and Data Splitting

A data augmentation method was employed to increase the number of available samples, addressing the limited dataset size caused by the relatively small number of discernible classes in SAR imagery compared to optical data. To achieve this, each input image was rotated by 90° and 270° and mirrored across horizontal and vertical axes. After augmentation, the dataset size increased to 10,800 samples, which were split into training and testing sets in an 80:20 ratio, yielding 8,640 samples for training and 2,160 samples for evaluation in each of the optical and SAR datasets.

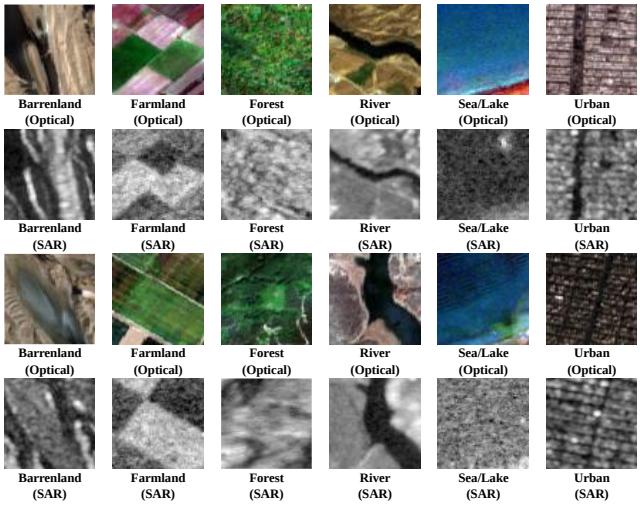

Figure 4. Samples of optical and SAR datasets

2.6. Applied Convolutional Neural Network Models

Convolutional Neural Networks (CNNs) form the foundation of many deep learning tasks across image processing. By applying convolution and pooling operations to multidimensional matrices, they are capable of extracting meaningful features, reducing computational complexity, and enhancing processing speed (Alzubaidi et al., 2021).

This process is depicted in Figure 5, demonstrating how a basic CNN operates in image classification tasks. In contrast to the previous study, in which the proposed HCNN architecture was not directly compared with established models, this work employs three well-known CNN architectures, namely DenseNet-121, VGG-19, and ResNet-50, to provide a direct performance comparison alongside the HCNN model. These architectures were originally designed and trained by their developers for three-channel image data. However, in this case, their input channels were modified to match the characteristics of the optical and SAR data used.

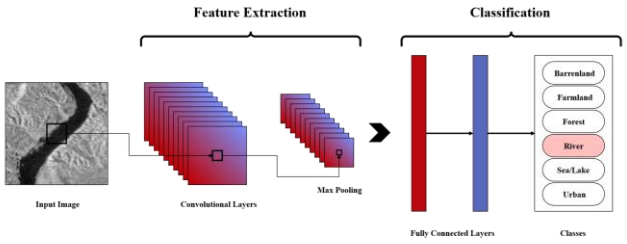

Figure 5. Structure of a basic CNN for image classification

2.6.1. Hybrid Multi-Block Convolutional Neural Network

The Hybrid Multi-Block Convolutional Neural Network (HCNN) is a deep convolutional network architecture designed to improve patch-level LULC classification accuracy. It achieves this by integrating multiple convolutional blocks. The network begins with an initial convolutional layer with a 3×3 kernel to extract low-level features, followed by batch normalization to stabilize learning and max-pooling to reduce spatial dimensions. A residual block with two convolutional layers and a skip connection ensures efficient gradient flow and facilitates deeper learning. Dense blocks follow, concatenating the outputs of several convolutional layers to enhance feature propagation and reuse. An Inception block captures multi-scale spatial patterns through parallel convolutions with varying kernel sizes. A squeeze-excitation block adaptively recalibrates the importance of feature channels, emphasizing informative spatial features while reducing computational cost. Global average pooling condenses spatial information before passing the features to fully connected layers for final classification. Ablation studies confirmed that each of these blocks contributes to the overall performance, with the residual, dense, and inception blocks having the greatest impact on accuracy improvements, while the squeeze-excitation block mainly enhances computational efficiency (Alikhani et al., 2025). The combination of these blocks balances accuracy and computational cost, allowing the model to learn complex hierarchical features from both optical and SAR datasets. The details of the HCNN architecture are presented in Table 2.

Table 2. The HCNN architecture summary

ID	Layer Type	Filters	Shape / Unit
1	Conv2D	64 (3×3)	64 × 64 × 64
2	Batch Normalization	-	64 × 64 × 64
3	MaxPooling2D	2×2	32 × 32 × 64
4	Conv2D	64 (3×3)	32 × 32 × 64
5	Conv2D	64 (3×3)	32 × 32 × 64
6	Add (Skip Connection)	-	32 × 32 × 64
7	MaxPooling2D	2×2	16 × 16 × 64
8	Conv2D	64 (3×3)	16 × 16 × 64
9	Concatenate	-	16 × 16 × 64
10	Conv2D	128 (3×3)	16 × 16 × 128
11	Concatenate	-	16 × 16 × 128
12	Conv2D	192 (3×3)	16 × 16 × 192
13	Concatenate	-	16 × 16 × 192
14	Conv2D	256 (1×1)	16 × 16 × 256
15	Conv2D	256 (3×3)	16 × 16 × 256
16	Conv2D	256 (5×5)	16 × 16 × 256
17	Concatenate	-	16 × 16 × 192
18	Global Average Pooling	-	192
19	Dense	-	12
20	Dense	-	192
21	Reshape	-	1 × 1 × 192
22	Multiply Feature Maps	-	16 × 16 × 192
23	Global Average Pooling	-	192
24	Dense (ReLU)	-	512
25	Dropout	-	512
26	Dense (Softmax)	-	10

2.6.2. DenseNet-121

Comprising 121 layers, DenseNet-121 is a deep CNN that employs dense connections, allowing each layer to access the outputs of all earlier layers (Weng et al., 2017). It consists of multiple dense blocks, separated by transition layers that include convolution and pooling operations. The architecture promotes feature reuse, reducing the number of parameters while maintaining representational power (Chollet, 2016; Howard et al., 2017). Each layer contributes a fixed number of feature maps, controlled by the growth rate. DenseNet-121 was also trained on ImageNet. Within each block, batch normalization, ReLU activation, and 1×1 and 3×3 convolutions are applied. This structure improves gradient flow and mitigates the vanishing gradient problem. DenseNet-121's compact design allows efficient feature propagation across deep layers (Thepade & Chauhan, 2023).

2.6.3. ResNet-50

ResNet-50, a deep CNN with 50 layers, is built upon a residual learning structure. It is composed of residual blocks, which introduce shortcut connections to help gradients flow more easily through the network. This design reduces the impact of gradient vanishing in multi-layer neural networks,

improving the learning process (Harini et al., 2024). The network is primarily used for image processing tasks, such as feature extraction, by automatically learning spatial hierarchies (He et al., 2015). ResNet-50 is applied in various computer vision applications, including classification, object detection, and segmentation. It has been employed in numerous studies across different fields, showcasing its versatility in image-related tasks (Srivani Bobba, 2024; Wan et al., 2024; Zhang et al., 2024).

2.6.4. VGG-19

VGG-19 is a deep CNN introduced by the Visual Geometry Group, University of Oxford (Simonyan & Zisserman, 2014). It has 19 layers, including 16 convolutional layers, 3 fully connected layers, and multiple max-pooling layers. The network uses small 3×3 filters in all convolutional layers, allowing it to capture fine details in images while maintaining a consistent architecture (Du et al., 2023). Every convolutional layer is accompanied by a ReLU activation function to introduce non-linearity, and subsequently, a max-pooling layer is applied for spatial dimensionality reduction. The VGG-19 was trained on the ImageNet dataset, which contains over a million images across 1,000 categories. Compared to VGG-16, VGG-19 has three additional convolutional layers, increasing the number of parameters and computational cost. Both models end with fully connected layers and a SoftMax classifier for final predictions (Mascarenhas & Agarwal, 2021)

2.7. CNN Model Training

The Adam optimizer with an initial learning rate of 0.001 was selected to optimize the models. Training was conducted over 50 epochs with a batch size of 64, using categorical cross-entropy as the loss function. A dynamic learning rate schedule was applied, where the learning rate decreased by 20% if validation accuracy showed no improvement over two successive epochs. The models were trained separately on optical and SAR datasets, applying the same optimization settings and learning rate adjustments to maintain consistency.

2.8. Performance Assessment

2.8.1 Standard Validation

The performance of the all models was assessed using four primary statistical measures. Overall Accuracy (OA) quantifies the proportion of correct predictions out of the total samples, indicating how accurately the model classifies the entire dataset (Congalton & Green, 2019). The F1-score, which evaluates the similarity between predicted and true values, was also used to assess predictive performance; higher scores reflect better accuracy and agreement (Lipton et al., 2014). The Kappa Coefficient (KC) measures the agreement between predicted and actual labels while accounting for chance agreement. It indicates how much better the model performs compared to classification by chance (Cohen, 1960). Finally, the Root Mean Square Error (RMSE) was employed to measure the average deviation between predicted and actual values, providing an overall indication of the model's predictive accuracy (Endris et al., 2021).

2.8.2 Cross Validation

To ensure that the high accuracies of the HCNN model are not dependent on a single train and test division, a five-fold cross validation strategy was applied to both optical and SAR datasets. The entire dataset was divided into five equal subsets. In each iteration, four subsets were used for training and one subset was used for validation. This procedure was repeated five times so that each subset served once as the validation fold. All training settings used in the main experiments were kept unchanged during cross validation to maintain consistency. After each iteration, OA, F1-Score, KC and RMSE were calculated. When the five iterations were completed, the results were summarized as the mean value and the standard deviation across the folds. This procedure provides a more reliable estimate of model stability and generalization, avoiding dependence on a single random split.

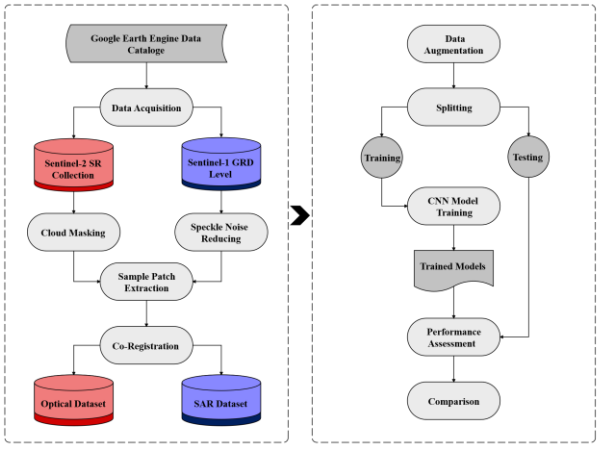

Figure 6. Methodology diagram

3. Results

3.1. Performance Assessment Results

The HCNN, DenseNet-121, ResNet-50, and VGG-19 models were independently trained on optical and SAR datasets, resulting in a total, 8 models that were analyzed for performance assessment. Figure 7 shows the validation accuracy of these models during training on optical and SAR datasets over 50 epochs. The models trained on the SAR dataset achieved overall accuracies exceeding 93%, while those trained on the optical dataset attained accuracies above 99%. Table 3 summarizes the performance assessment results of these models using statistical metrics. Among the evaluated models, HCNN consistently achieved the highest overall accuracy across both datasets, reaching 99.56% on the optical dataset and 99.14% on the SAR dataset. DenseNet-121, ResNet-50, and VGG-19 achieved slightly

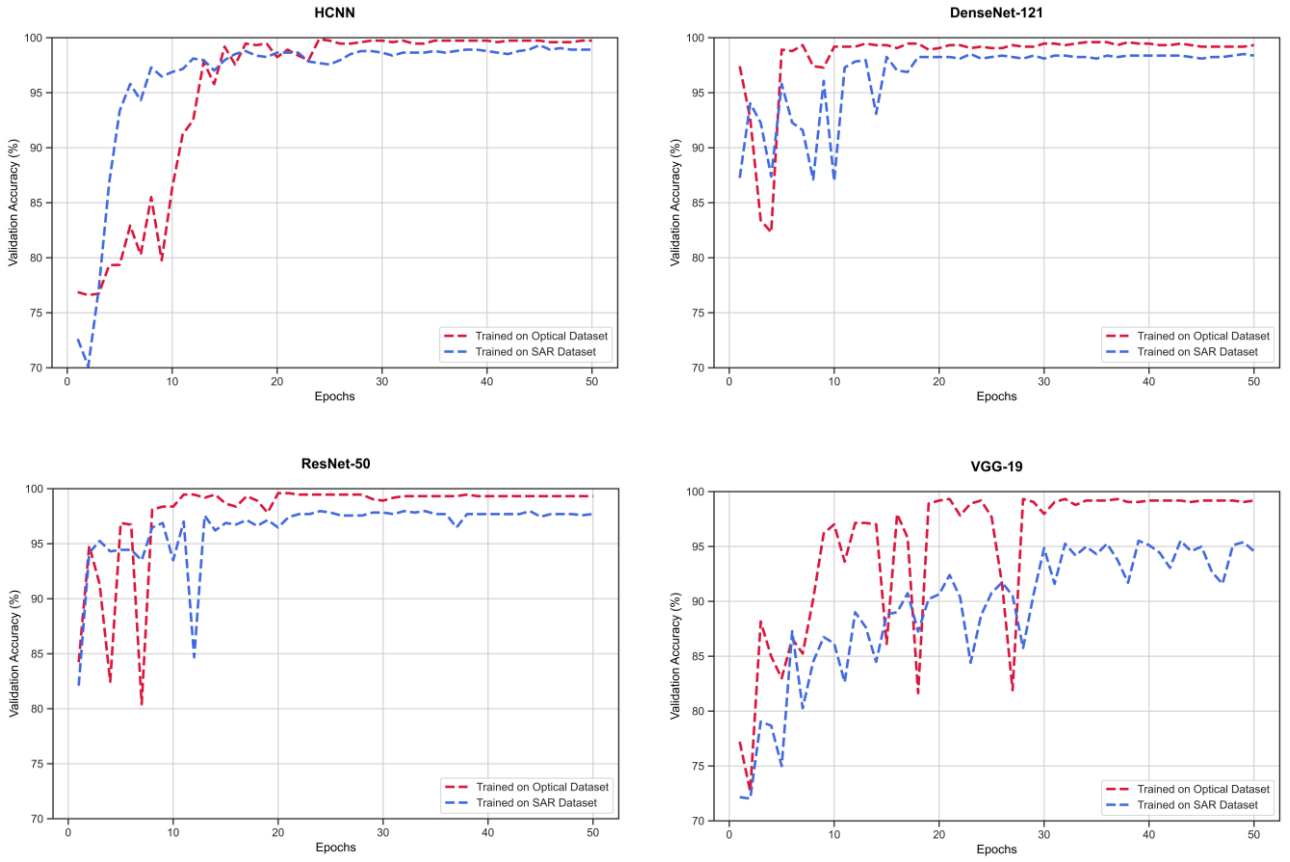

Figure 7. Validation accuracy trends during the training of models

lower accuracies, with 99.22%, 99.25%, and 98.89% on optical data, and 98.36%, 97.65%, and 94.56% on SAR data, respectively. More specifically, the accuracy of models in classifying samples of each class was independently assessed.

Table 3. Statistical metrics values of models

Model	Dataset	OA	F1- Score	KC	RMSE
HCNN	Optical	99.56	99.46	99.30	0.07
DenseNet-121		99.22	99.08	98.85	0.13
ResNet-50		99.25	99.01	98.87	0.09
VGG-19		98.89	98.59	98.48	0.11
HCNN	SAR	99.14	98.33	98.73	0.24
DenseNet-121		98.36	97.89	97.80	0.37
ResNet-50		97.65	97.30	97.10	0.49
VGG-19		94.56	94.12	93.49	0.62

The class-wise accuracy rates are reported in Table 4. The class-wise performance comparison of the models demonstrates that HCNN generally outperformed the other models across most classes and datasets. In the Barrenland class, HCNN achieved the highest accuracy on the SAR dataset with 98.35%, which is notably higher than DenseNet-121 with 95.04%, ResNet-50 with 93.38%, and VGG-19 with 92.56%. In the City class, HCNN again delivered a strong performance, achieving 98.36% on SAR data, slightly surpassing DenseNet-121 at 98.36%, ResNet-50 at 96.72%, and VGG-19 at 92.62%. For the Farmland class, HCNN reached 99.39% on optical data and 98.27% on SAR data, outperforming DenseNet-121, ResNet-50, and VGG-19, particularly on optical data, where the other models fell slightly behind.

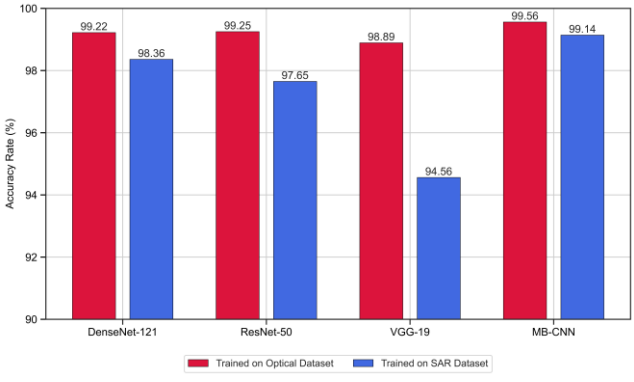

Figure 8. Overall accuracy comparison of the models

Remarkably, HCNN achieved perfect accuracy for the River class on both optical and SAR datasets, whereas the other models showed lower performance, with ResNet-50

Table 4. Class-wise accuracy rates of the models.

Class Name	HCNN		DenseNet-121		Resnet-50		VGG19	
	Optical	SAR	SAR	Optical	Optical	SAR	Optical	SAR
Barrenland	99.68	98.35	95.04	99.59	99.73	95.04	99.59	98.35
Farmland	99.18	98.36	98.36	99.21	99.52	98.36	99.21	98.36
Forest	99.39	98.27	99.19	99.16	99.21	99.19	99.16	98.27
River	99.61	99.84	99.15	98.36	99.28	99.15	98.36	99.84
Sea/Lake	100	99.91	98.41	99.74	99.19	98.41	99.74	100
Urban	99.49	99.87	100	99.43	98.36	100	99.43	100

and VGG-19 scoring below 100%. In the Forest class, HCNN also achieved the highest accuracy on SAR data, at 99.84%, and maintained a competitive performance on optical data. Similarly, for the Sea\Lake class, HCNN recorded 99.49% on optical data and 100% on SAR data, closely matching or exceeding other models. Overall, HCNN consistently provided the highest or comparable accuracy across all classes compared to DenseNet-121, ResNet-50, and VGG-19.

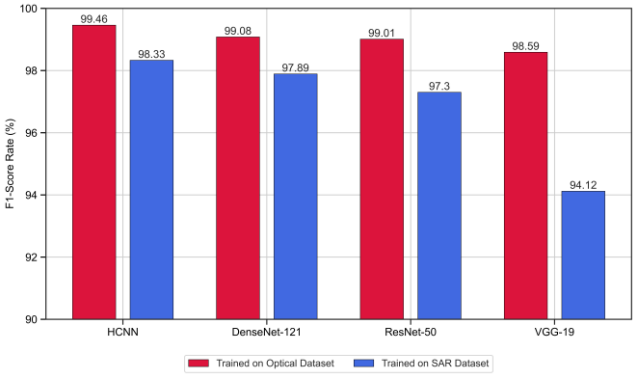

Figure 9. F1-Score comparison of the models

Similarly, the performance of models was evaluated using F1-scores, Kappa coefficients, and RMSE to compare their classification accuracy. As shown in Figure 9, HCNN achieved the highest F1-scores with 99.46% on the optical dataset and 98.33% on the SAR dataset, clearly outperforming DenseNet-121, ResNet-50, and VGG-19, which achieved lower scores ranging between 94.12% and 99.08%. The Kappa coefficients, presented in Figure 10, further confirmed the consistent and reliable performance of HCNN, reaching 99.30% on optical data and 98.73% on SAR data. In comparison, DenseNet-121 achieved kappa values of 98.85% on optical data and 97.80% on SAR data. ResNet-50 yielded similar results with 98.87% and 97.10%, whereas VGG-19 performed slightly lower with 98.48% and 93.49%, respectively. Finally, the RMSE values, depicted in Figure 11, highlighted the accuracy of HCNN, with the lowest errors of 0.08 on optical and 0.24 on SAR datasets, compared to higher RMSE values for the other models.

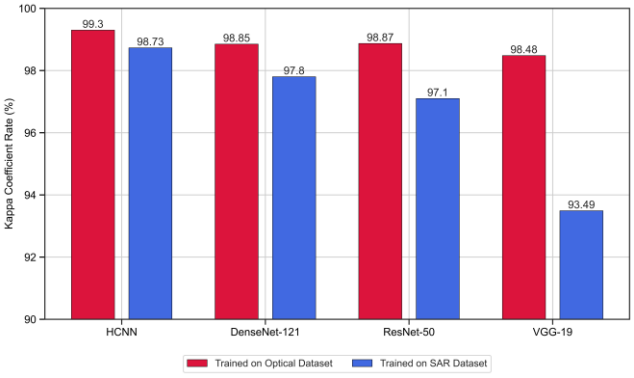

Figure 10. Kappa coefficient comparison of the models

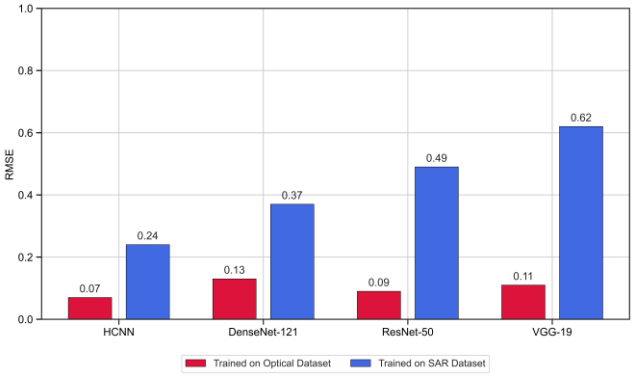

Figure 11. RMSE comparison of the models

The cross-validation results of the HCNN model trained on optical and SAR datasets are presented in Table 5, showing nearly consistent performance across all folds and demonstrating the model's reliability and stability. As shown in Table 5, the optical data maintains OA values between 99.49% and 99.58%, with F1-Score and KC following the same stable pattern. The RMSE values for the optical dataset remain tightly concentrated around 0.24, indicating highly consistent error behavior across folds. The SAR data also demonstrates strong stability, with OA values ranging from 98.87% to 99.03% and similarly consistent F1-Score and KC values. Although the RMSE values for SAR are naturally higher than those of the optical dataset due to the radiometric characteristics of SAR imagery, their variation across folds remains minimal.

Table 5. Five-fold cross validation results for HCNN model

Dataset	Fold	OA	F1- Score	KC	RMSE
Optical	Fold 1	99.52	99.47	99.01	0.245
	Fold 2	99.58	99.53	99.15	0.240
	Fold 3	99.49	99.44	98.92	0.248
	Fold 4	99.55	99.50	99.07	0.243
	Fold 5	99.51	99.46	98.96	0.246
SAR	Fold 1	98.92	98.65	98.10	0.276
	Fold 2	99.03	98.74	98.29	0.270
	Fold 3	98.87	98.60	98.01	0.280
	Fold 4	98.95	98.68	98.16	0.275
	Fold 5	99.01	98.72	98.23	0.272

3.2. Network Complexity Analysis

Table 6 presents the number of layers, parameters in millions, and Floating-Point Operations (FLOPs) of the models, which reflect the structural complexity of each network. FLOPs quantify the total number of operations a model performs, serving as a measure of its computational complexity. As shown, DenseNet-121, despite having the highest number of layers, has fewer parameters compared to VGG-19 and ResNet-50. In contrast, VGG-19 has the fewest layers but the highest number of parameters, indicating a heavier architecture. The HCNN, with only 26 layers, has the lowest number of parameters and the smallest FLOPs. This demonstrates its simpler architecture and lower computational demand compared to the other models.

Table 6. Network complexity comparison of models

Models	Layers	Parameters (M)	FLOPs (M)
HCNN	26	0.59	242.10
DenseNet-121	121	6.96	479.71
ResNet-50	50	25.62	684.99
VGG-19	19	127	3450.76

To help understand the complexity of the networks, Table 7 reports the training time for each one. The results show that HCNN achieved the fastest training times, with 392 seconds on the optical dataset and 279 seconds on the SAR dataset. In contrast, VGG-19 was the slowest model, requiring nearly 8 times longer than HCNN to train. DenseNet-121 and ResNet-50 had similar training durations, but were still considerably longer than HCNN.

Table 7. Training time (Seconds) comparison of models

Dataset	HCNN	DenseNet-121	ResNet-50	VGG-19
Optical	392	3273	3135	3369
SAR	279	2211	2157	2361

Overall, models trained on the optical dataset required more training time than their counterparts trained on the SAR dataset, due to the higher number of input bands. However, in the case of HCNN, the time required to train the model on each dataset was relatively similar.

3.3. Ablation Study

For the HCNN model, an ablation study was conducted to evaluate the contribution of each architectural block, including Residual, Dense, Inception, and SE, on both optical and SAR datasets. For the optical dataset, removing the residual block reduced accuracy from 99.56% to 94.17%, with training time slightly decreasing from 392 to 351 seconds. Excluding the dense or inception blocks decreased accuracy to 96.78% and 96.69%, with training times of 331 and 326 seconds, respectively. Removing the SE block had minimal impact on accuracy, slightly increasing it to 98.47%, while notably increasing training time to 520 seconds. Eliminating all blocks resulted in the lowest accuracy of 89.17% and the shortest training time of 280 seconds. Similarly, for the SAR dataset, removing the residual block decreased accuracy from 99.14% to 92.85%, with training time slightly decreasing from 279 to 245 seconds. Excluding the dense or inception blocks lowered accuracy to 95.60% and 95.42%, with training times of 230 and 225 seconds, respectively. Removing the SE block slightly increased accuracy to 98.00%, but increased training time to 395 seconds. The model without any blocks achieved an accuracy of 87.90% and the shortest training time of 200 seconds.

Table 8. Ablation Study of HCNN Blocks

Dataset	Residual	Dense	Inception	SE	OA	Time
Optical	✗	✓	✓	✓	94.17	351
	✓	✗	✓	✓	96.78	331
	✓	✓	✗	✓	96.69	326
	✓	✓	✓	✗	98.47	520
	✗	✗	✗	✗	89.17	280
	✓	✓	✓	✓	99.56	392
SAR	✗	✓	✓	✓	92.85	245
	✓	✗	✓	✓	95.60	230
	✓	✓	✗	✓	95.42	225
	✓	✓	✓	✗	98.00	395
	✗	✗	✗	✗	87.90	200
	✓	✓	✓	✓	99.14	279

As shown in Table 8, each block contributes to model accuracy, especially the residual block, which has the most significant impact. The SE block, while having minimal effect on accuracy, plays a critical role in improving computational efficiency. By compressing channel-wise information through global pooling and lightweight layers, the SE block reduces computational load and accelerates training, explaining the increase in training time when it is removed. Overall, these results highlight the importance of each architectural block in balancing classification accuracy and computational efficiency for both optical and SAR datasets.

4. Discussion

In the initial phase of this study, two independent datasets were constructed based on optical data from Sentinel-2 and SAR data from Sentinel-1. Both datasets included the same six LULC classes and consisted of paired image patches. The HCNN architecture was previously trained on the EuroSAT dataset, demonstrating its effectiveness, particularly in reducing training time. The primary objective of creating the new datasets was to evaluate HCNN's performance using data from a previously unstudied region. Additionally, the construction of paired datasets enabled a consistent and controlled comparison of the model's performance when trained on optical versus SAR imagery.

In the next step, the HCNN-based model, alongside DenseNet-121, ResNet-50, and VGG-19, was trained on each dataset independently.

Performance evaluation demonstrated that HCNN consistently outperformed the other CNN architectures across both optical and SAR datasets. This superiority was particularly evident in computational efficiency, with HCNN-based models requiring up to eight times less training time than the slowest model, VGG-19. Overall, HCNN achieved the highest classification accuracy, reaching 99.56% on SAR data and 99.14% on optical data, while maintaining a substantially lower computational cost.

The five-fold cross-validation results further confirm the reliability of the HCNN model. The performance metrics, including OA, F1-Score, KC, and RMSE, show minimal variation across all folds for both optical and SAR datasets. This indicates that the network behaves consistently when trained and evaluated on different subsets of the data, demonstrating strong generalization ability and ruling out the possibility that the high accuracies are due to a favorable data split. These results highlight the effectiveness of HCNN in delivering both high accuracy and efficiency, making it a powerful alternative for patch-level LULC classification.

All trained models achieved high and closely matched performance levels. Nevertheless, among them, the HCNN architecture showed the most consistent and reliable results across both data types. Despite its lighter structure and fewer parameters, HCNN achieved equal or higher accuracy than deeper architectures in most classes. Notably, the HCNN trained on SAR data achieved 100% accuracy in classifying Urban and Sea/Lake classes.

In contrast, the VGG-19 model, which had the highest number of parameters, exhibited relatively lower performance on SAR data and required the longest training time. On the other hand, HCNN achieved the shortest training time on both datasets without compromising accuracy. These characteristics make it a practical choice for environments with limited computational resources or time constraints. Overall, models trained on optical data generally achieved higher accuracy compared to their SAR-trained counterparts, reflecting the richer spectral information and higher quality of Sentinel-2 imagery. Nevertheless, the performance differences were not substantial, and SAR-based models still demonstrated notable accuracy in certain classes, such as Urban. While most prior studies on patch-level LULC classification have focused primarily on optical data, the results of this study indicate that SAR imagery remains a reliable and effective alternative, particularly when optical data are unavailable. However, creating a patch-level dataset similar to EuroSAT with a larger number of LULC classes using SAR data presents additional challenges. This difficulty also imposed certain limitations in the present study. It largely stems from challenges faced by the dataset creator in correctly distinguishing classes due to similar backscatter patterns in radar imagery, especially in regions like Iran, which are predominantly covered by arid and mountainous landscapes.

Table 9. Comparative accuracy with prior works

Model	Data Type	OA	Reference
HCNN*	Multispectral	99.14	This Study
	SAR	99.56	This Study
DenseNet-121*	Multispectral	99.22	This Study
	SAR	98.36	This Study
ResNet-50*	Multispectral	99.25	This Study
	SAR	97.65	This Study
VGG-19*	Multispectral	98.89	This Study
	SAR	94.56	This Study
CNN (Early Fusion)	Multispectral + SAR	88.12	Irfan et al., 2025
CNN (Late Fusion)	Multispectral + SAR	83.96	Irfan et al., 2025
ResNet	Multispectral + SAR	72.15	Irfan et al., 2025
GoogleNet	Multispectral + SAR	66.95	Irfan et al., 2025
VGG19	Multispectral	95.00	Rengma & Yadav, 2024
ResNet-50	Multispectral	90.00	Rengma & Yadav, 2024
Random Forest	Multispectral	94.00	Rengma & Yadav, 2024
Vision Transformers	RGB	99.19	Kunwar & Ferdush, 2024
VGG19	RGB	98.06	Kunwar & Ferdush, 2024
ResNet-50	RGB	98.52	Kunwar & Ferdush, 2024
ResNet-101	Multispectral	99.66	Celik & Çelik, 2025
GoogleNet	Multispectral	96.69	Celik & Çelik, 2025
VGG16	Multispectral	98.55	Celik & Çelik, 2025
CNN	Multispectral	97.23	Vir, 2025
EfficientNet	Multispectral	98.10	Sariçayir & Ozcan, 2025
InceptionV3	Multispectral	98.50	Marzougui et al., 2025
CNN	Multispectral	87.12	Marzougui et al., 2025
MobileNet	Multispectral	94.50	Marzougui et al., 2025
ResNet	Multispectral	92.88	Le et al., 2025

The performance comparison across recent studies demonstrates the effectiveness of deep learning and traditional machine learning approaches on patch-based or EuroSAT-like datasets. As shown in Table 9, the evaluated

models, including HCNN, DenseNet-121, ResNet-50, and VGG-19, which are marked with an asterisk (*) in the table, consistently achieve high overall accuracy. The HCNN reaches 99.56% on optical data and 99.14% on SAR data, outperforming the other deep architectures included in the comparison. Other deep learning models reported in similar works generally achieve accuracies ranging from 87% to 99%, depending on the data modality and network depth, and they are collectively represented as CNN in Table 9. Traditional machine learning methods such as Random Forest also show competitive performance but remain slightly below modern deep architectures. This comparison indicates that although many recent models perform well on patch-level multispectral or SAR imagery, the HCNN achieves higher accuracy and efficiency. These results confirm the robustness of the proposed architecture for LULC classification on optical and SAR datasets.

4.1. Limitations and Applicability

Despite the superior performance of the Hybrid Multi-Block Convolutional Neural Network, HCNN, in patch-level LULC classification on Sentinel-2 optical data and Sentinel-1 SAR data from Iran, several limitations must be considered for a comprehensive evaluation. The HCNN is designed for patch-level classification and assigns a single dominant label to each patch with a size of 64×64 pixels. As a result, it cannot perform pixel-level semantic segmentation or object-based analysis, and it does not generate fine-grained thematic maps that preserve detailed boundaries between land cover categories. This limitation restricts its direct use in applications that require spatial delineation, including urban planning, precision agriculture, or environmental monitoring at scales smaller than the patch level.

The architecture, however, shows strong flexibility in handling different types of input data. It processes multispectral imagery with 12 bands from Sentinel-2 and SAR imagery with four polarization channels, VV and VH, from both ascending and descending orbits. This demonstrates its robustness in both optical and radar domains without needing a redesign tailored to each modality. The lightweight structure of the model also reduces the training time to nearly one-eighth of well-known benchmark models such as DenseNet-121, ResNet-50, and VGG-19, which makes it suitable for deployment on systems with limited computational resources, including edge devices and cloud-based platforms.

The current evaluation is limited to Iran and relies on datasets created from Sentinel imagery that include six LULC classes: Barrenland, Farmland, Forest, River, Sea/Lake, and Urban. Although these datasets provide regionally relevant benchmarks, the performance of the model on diverse global environments, such as tropical rainforests, arid deserts, or polar landscapes, is still unknown. Validation on larger and more diverse datasets, for example multi-continental collections like BigEarthNet or global Sentinel mosaics, would be essential for assessing generalization and reducing the impact of regional spectral or textural similarities. While the HCNN is optimized for LULC classification, its hybrid combination of residual blocks, dense blocks, inception modules, and squeeze excitation components is suitable for many other patch-based remote sensing tasks.

To address the limitation related to the absence of pixel-wise mapping, which is critical for many operational remote sensing tasks, an effective future extension would be to redesign the HCNN within an encoder–decoder structure similar to U-Net architectures. This adaptation would enable dense prediction and support the creation of high-resolution LULC maps that are required for policy-oriented applications such as sustainable land management and climate impact assessments.

5. Conclusion

In this study, a hybrid convolutional neural network (HCNN) was introduced for LULC classification. Unlike most previous works that primarily relied on optical data, this research incorporated both optical and SAR imagery to better evaluate the true capability of the proposed model. The results demonstrated that the HCNN effectively captures complementary information from these data sources and achieves competitive performance in complex classification scenarios. In particular, the HCNN consistently outperformed widely used CNN architectures, including DenseNet-121, ResNet-50, and VGG-19, while also reducing computational complexity, demonstrating higher classification accuracy.

It should be noted, however, that the current architecture has been designed and tested only for patch-level image classification, in contrast to pixel-wise approaches such as U-Net, which enable the generation of thematic maps. While the HCNN shows strong potential for the LULC classification task, its structure could be further extended to support pixel-level segmentation tasks. Moreover, given the generality of the architecture, it may also be applicable to domains beyond remote sensing, such as medical imaging. Although these applications have not yet been tested, future research is recommended to explore these possibilities.

Data Availability Declaration

The optical and SAR datasets generated in this study (without data augmentation) are publicly available and can be accessed via the following link: https://github.com/moalikhani77/optical-and-sar-datasets-for-land-use-and-land-cover-classification/

References

Abdulraheem, M. I., Zhang, W., Li, S., Moshayedi, A. J., Farooque, A. A., & Hu, J. (2023). Advancement of remote sensing for soil measurements and applications: A comprehensive review. Sustainability, 15(21), 15444. https://doi.org/10.3390/su152115444

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. https://doi.org/10.1109/ICRITO48877.2020.9197824

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 8, 1-74. https://doi.org/10.1186/s40537-021-00444-8

Amin, G., Imtiaz, I., Haroon, E., Saqib, N. U., Shahzad, M. I., & Nazeer, M. (2024). Assessment of machine learning algorithms for land cover classification in a complex mountainous landscape. Journal of Geovisualization and Spatial Analysis, 8(2), 34. https://doi.org/10.1007/s41651-024-00195-z

Arathi, B., & Dulhare, U. N. (2023, March). Classification of cotton leaf diseases using transfer learning-densenet-121. In Proceedings of third international conference on advances in computer engineering and communication systems: ICACECS 2022 (pp. 393-405). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-9228-5_33

Baral, S., Aryal, J., & Sharma, N. (2023). Remote Sensing Image Classification Using Transfer Learning Based Convolutional Neural Networks: An Experimental Overview. Authorea Preprints. https://doi.org/10.36227/techrxiv.22581457.v1

Bobba, S. (2024). Leveraging Pre-trained Deep Learning Models for Remote Sensing Image Classification: A Case Study with ResNet50 and EfficientNet. American Journal of Science, Engineering and Technology, 9(3), 150-162. https://doi.org/10.11648/j.ajset.20240903.11

Celik, F., & Çelik, K. (2025). Prediction of Land Image Classification using Squeeze Synchronization layer and Convolve Craft Focus Module in ResNet 101 Model. International Journal of Engineering and Geosciences, 10(2), 197–206. https://doi.org/10.26833/ijeg.1538708

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258). https://doi.org/10.48550/arXiv.1610.02357

Chowdhury, M. T., Rahman, H., Sumon, M. I., & Talha, A. (2024, May). Classification of satellite images with VGG19 and Convolutional Neural Network (CNN). In 2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT) (pp. 397-402). IEEE. https://doi.org/10.1109/InCACCT61598.2024.10551185

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), 37-46. https://doi.org/10.1177/001316446002000104

Congalton, R. G., & Green, K. (2019). Assessing the accuracy of remotely sensed data: principles and practices. CRC press. https://doi.org/10.1201/9780429052729

Dahhani, S., Raji, M., Hakdaoui, M., & Lhissou, R. (2022). Land cover mapping using sentinel-1 time-series data and machine-learning classifiers in agricultural sub-saharan landscape. Remote Sensing, 15(1), 65. https://doi.org/10.3390/rs15010065

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee. https://doi.org/10.1109/CVPR.2009.5206848

Du, X., Sun, Y., Song, Y., Sun, H., & Yang, L. (2023). A comparative study of different CNN models and transfer learning effect for underwater object classification in side-scan sonar images. Remote Sensing, 15(3), 593. https://doi.org/10.3390/rs15030593

ED Chaves, M., CA Picoli, M., & D. Sanches, I. (2020). Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. Remote Sensing, 12(18), 3062. https://doi.org/10.3390/rs12183062

Endris, H. S., Hirons, L., Segele, Z. T., Gudoshava, M., Woolnough, S., & Artan, G. A. (2021). Evaluation of the skill of monthly precipitation forecasts from global prediction systems over the Greater Horn of Africa. Weather and Forecasting, 36(4), 1275-1298. https://doi.org/10.1175/waf-d-20-0177.1

Gupta, N., Singh, J., Singh, S., Joshi, G., & Mittal, A. (2023, August). Investigations on deep learning pre-trained model VGG-19 using transfer learning for remote sensing image classification on benchmark datasets. In International Conference on Artificial Intelligence on Textile and Apparel (pp. 127-139). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-8476-3_11

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. https://doi.org/10.1007/978-981-97-1841-2_19

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). https://doi.org/10.48550/arXiv.1512.03385

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. https://doi.org/10.48550/arXiv.1709.00029

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861. https://doi.org/10.48550/arXiv.1704.04861

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708). https://doi.org/10.1109/cvpr.2017.243

Irfan, A., Li, Y., E, X., & Sun, G. (2025). Land Use and Land Cover Classification with Deep Learning-Based Fusion of SAR and Optical Data. Remote Sensing, 17(7). https://doi.org/10.3390/rs17071298

Karaköse, E. (2024). An Efficient Satellite Images Classification Approach Based on Fuzzy Cognitive Map Integration with Deep Learning Models Using Improved Loss Function. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3461871

Kunwar, S., & Ferdush, J. (2024). Mapping of Land Use and Land Cover (LULC) Using EuroSAT and Transfer Learning. Revue Internationale de Géomatique, 2024, 1–13. https://doi.org/10.32604/rig.2023.047627

Lang, F., Zhu, Y., Zhao, J., Hu, X., Shi, H., Zheng, N., & Zha, J. (2024). Flood Mapping of Synthetic Aperture Radar (SAR) Imagery Based on Semi-Automatic Thresholding and Change Detection. Remote Sensing, 16(15), 2763. https://doi.org/10.3390/rs16152763

Le, T.-D., Ha, V., Nguyen, T., Eappen, G., Thiruvasagam, P., Chou, H., Tran, D.-D., Nguyen-Kha, H., Garcés-Socarrás, L., González Rios, J. L., Merlano Duncan, J. C., & Chatzinotas, S. (2025). GLUSE: Enhanced Channel-Wise Adaptive Gated Linear Units SE for Onboard Satellite Earth Observation Image Classification. https://doi.org/10.48550/arXiv.2504.12484

Li, W., Wang, Z., Wang, Y., Wu, J., Wang, J., Jia, Y., & Gui, G. (2020). Classification of high-spatial-resolution remote sensing scenes method using transfer learning and deep convolutional neural network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 1986-1995. https://doi.org/10.1109/JSTARS.2020.2988477

Lipton, Z. C., Elkan, C., & Narayanaswamy, B. (2014). Thresholding classifiers to maximize F1 score. arXiv preprint arXiv:1402.1892. https://doi.org/10.48550/arXiv.1402.1892

Lu, B., Dao, P. D., Liu, J., He, Y., & Shang, J. (2020). Recent advances of hyperspectral imaging technology and applications in agriculture. Remote Sensing, 12(16), 2659. https://doi.org/10.3390/rs12162659

Marzougui, M., Sampedro, G. A., Almadhor, A., Alsubai, S., Al Hejaili, A., & Abbas, S. (2025). Deep Learning-Based Spatial Pattern Modeling for Land Use and Land Cover Classification Using Satellite Imagery. Meteorological Applications, 32(5), e70064. https://doi.org/10.1002/met.70064

Mascarenhas, S., & Agarwal, M. (2021, November). A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification. In 2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON) (Vol. 1, pp. 96-99). IEEE. https://doi.org/10.1109/CENTCON52345.2021.9687944

McClelland, J., Riedel, T., Beyer, F., Gerighausen, H., & Golla, B. (2023). Concepts for Open Access Interdisciplinary Remote Sensing with ESA Sentinel-1 SAR Data. In Environmental Informatics (pp. 57-71). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-46902-2_4

Mienye, I. D., Swart, T. G., Obaido, G., Jordan, M., & Ilono, P. (2024). Deep Convolutional Neural Networks: A Comprehensive Reviews. https://doi.org/10.20944/preprints202408.1288.v1

Nhangumbe, M., Nascetti, A., & Ban, Y. (2023). Multi-temporal Sentinel-1 SAR and Sentinel-2 MSI data for flood mapping and damage assessment in Mozambique. ISPRS International Journal of Geo-Information, 12(2), 53. https://doi.org/10.3390/ijgi12020053

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: A review. Remote sensing, 12(14), 2291. https://doi.org/10.3390/rs12142291

Rengma, N. S., & Yadav, M. (2024). Generation and classification of patch-based land use and land cover dataset in diverse Indian landscapes: A comparative study of machine learning and deep learning models. Environmental Monitoring and Assessment, 196(6), 568. https://doi.org/10.1007/s10661-024-12719-7

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. https://doi.org/10.1007/978-981-15-6315-7_28

Sahad, A. J. Z. (2025). SARrawSim: Synthetic Aperture Radar Raw Data Simulator. SoftwareX, 29, 102019. https://doi.org/10.1016/j.softx.2024.102019

Sariçayir, B., & Ozcan, C. (2025). EfficientNet Deep Learning Model for Satellite Image Classification Using the EuroSAT Dataset.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

Tshanga M, M., Ncube, L., & van Niekerk, E. (2024). Remote sensing insights into subsurface-surface relationships: Land Cover Analysis and Copper Deposits Exploration. *Earth Science Informatics*, *17*(5), 3979-4000. https://doi.org/10.1007/s12145-024-01423-2

Thepade, S. D., & Chauhan, S. (2023). Enhancing Land Use Identification Through Fusion of Densenet121 DCNN and Thepade SBTC Features Using Machine Learning Algorithms and Ensembles. SN Computer Science, 4(6), 772. https://doi.org/10.1007/s42979-023-02325-8

Thepade, S. D., Adrakatti, S. S., Chauhan, S., Chitnis, V., & Bhalerao, A. (2023, August). Comparing the Deep Convolutional Neural Network Models for Identification of Aerial Image Land Use. In 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA) (pp. 1-7). IEEE. https://doi.org/10.1109/ICCUBEA58933.2023.10392161

Vanderhoof, M. K., Alexander, L., Christensen, J., Solvik, K., Nieuwlandt, P., & Sagehorn, M. (2023). High-frequency time series comparison of Sentinel-1 and Sentinel-2 satellites for mapping open and vegetated water across the United States (2017–2021). Remote Sensing of Environment, 288, 113498. https://doi.org/10.1016/j.rse.2023.113498

Vir, A. (2025). Balanced Multi-Task Attention for Satellite Image Classification: A Systematic Approach to Achieving 97.23% Accuracy on EuroSAT Without Pre-Training. https://doi.org/10.48550/arXiv.2510.15527

Wan, J., Li, B., Wang, K., Teng, X., Wang, T., & Mao, B. (2024). An Improved ResNet50 for Environment Image Classification. Procedia Computer Science, 242, 1000-1007. https://doi.org/10.1016/j.procs.2024.08.246

Wang, H., Yang, H., Huang, Y., Wu, L., Guo, Z., & Li, N. (2023). Classification of land cover in complex terrain using gaofen-3 SAR ascending and descending orbit data. Remote Sensing, 15(8), 2177. https://doi.org/10.3390/rs15082177

Wang, Q., Wang, C., Wu, H., Zhao, C., Teng, G., Yu, Y., & Zhu, H. (2022). A two-stage low-altitude remote sensing papaver somniferum image detection system based on YOLOv5s+ DenseNet121. Remote Sensing, 14(8), 1834. https://doi.org/10.3390/rs14081834

Wang, Y., Sun, Y., Cao, X., Wang, Y., Zhang, W., & Cheng, X. (2023). A review of regional and Global scale Land Use/Land Cover (LULC) mapping products generated from satellite remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 206, 311-334. https://doi.org/10.1016/j.isprsjprs.2023.11.014

Weng, Q., Mao, Z., Lin, J., & Guo, W. (2017). Land-use classification via extreme learning classifier based on deep convolutional features. IEEE Geoscience and Remote Sensing Letters, 14(5), 704-708. https://doi.org/10.1109/LGRS.2017.2672643

Xu, X., Li, D., Liu, H., Zhao, G., Cui, B., Yi, Y., ... & Du, J. (2024). Comparative Validation and Misclassification Diagnosis of 30-Meter Land Cover Datasets in China. Remote Sensing, 16(22), 4330. https://doi.org/10.3390/rs16224330

Zhang, C., Iqbal, I., Bhatti, U. A., Liu, J., Awwad, E. M., & Sarhan, N. (2024). ResNet50 in remote sensing and agriculture: evaluating image captioning performance for high spectral data. Environmental Earth Sciences, 83(23), 1-16. https://doi.org/10.1007/s12665-024-11950-2

Zhang, Z., Cui, X., Zheng, Q., & Cao, J. (2021). Land use classification of remote sensing images based on convolution neural network. Arabian Journal of Geosciences, 14, 1-6. https://doi.org/10.1007/s12517-021-06587-5

Zhao, W., & Du, S. (2016). Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. IEEE Transactions on Geoscience and Remote Sensing, 54(8), 4544-4554. https://doi.org/10.1109/TGRS.2016.2543748

Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. Remote sensing of environment, 221, 430-443. https://doi.org/10.1016/j.rse.2018.11.032

Zhu, X. X., Montazeri, S., Ali, M., Hua, Y., Wang, Y., Mou, L., ... & Bamler, R. (2021). Deep learning meets SAR: Concepts, models, pitfalls, and perspectives. IEEE Geoscience and Remote Sensing Magazine, 9(4), 143-172. https://doi.org/10.48550/arXiv.2006.10027