IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

Patch-Based Land Use Classification Using GoogLeNet, ResNet50, and AlexNet on Satellite Imagery

Ms. Jayashree C. Pasalkar[1], Madiha Sameer Shaikh[2], Sai Pramod Nimbalkar[3], Divya Dattatray Rahane[4]
Professor, Department of Information Technology[1]
Student, Department of Information Technology[2-4]
AISSMS Institute of Information Technology (IOIT), Pune, Maharashtra, India
jayashree.pasalkar@aissmsioit.org[1], 25madiha.shaikh@gmail.com[2],
nimbalkarsai46@gmail.com[3], divyarahane022@gmail.com[4]

Abstract: *The study highlights a patch-based approach for land use classification based on Convolutional Neural Networks (CNNs) on satellite photos. The EuroSAT dataset contains 64 × 64 pixels sized, labeled RGB images representing different land types. This dataset is used to train and evaluate three different models: GoogLeNet, ResNet50, and AlexNet. In order to maintain the uniformity of everything during training and testing, Sentinel-2 images through Google earth engine are divided into patches of equal size without overlap. Images are pre-trained to normalize and add augmentation to make the models more generalized. Accuracy, confusion matrix, performance per class and loss trends are used to measure the level of effectiveness of each model. The trained models are used to analyze actual satellite images and synthesize predictions made by each patch to generate a spatial map of land use. These results can be visualized using geospatial tools to understand which land uses are widespread at various locations. The comparison of the models helps to understand their advantages and disadvantages in detecting spatial features, which in turn helps in selecting the models to be used in land use classification.*

Keywords: Land Use Classification, Remote Sensing, Convolutional Neural Networks (CNN), EuroSAT Dataset, Patch-Based Classification, Sentinel-2 Imagery, Geospatial Analysis

I. INTRODUCTION

A large portion of geospatial analysis is composed of Land Use and Land Cover (LULC) classification. It relies on satellite images to identify and label various types of land surfaces, such as cities, forests, waters, and farms. It is actually useful in planning of cities, monitoring the environment, handling crops, and observing the severity of a disaster. The availability of high-resolution images is now abundant with new remote sensing technologies and satellites, such as Sentinel-2. Nevertheless, these massive and complicated data sets still present a challenge in extracting helpful and ordered data because of land cover, light, and pattern changes.

Previously, the classification of land use was done manually by individuals or through traditional machine learning processes that required manual extraction of features. These techniques had difficulties in their application in various regions and ability to easily record detailed patterns. The use of convolutional neural networks (CNNs) has proven to be a superior alternative since they are able to extract features automatically on raw images. This means they are suitable in remote sensing applications where texture and spatial context are important in detecting various land uses.

This paper proposes a patch-based land use classification system that can deal with the huge satellite images. The input images are split into 64×64 pixel blocks instead of classifying each pixel, which is both difficult and noisy to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

computers. This makes the EuroSAT dataset to be used in training constant. The data consists of labeled RGB images of the various categories of lands. The patch approach allows the model to narrow down on the local spatial aspects such as texture, structure and color. It can also process high-resolution images in large scale.
The model performance has been tested using three popular CNN architectures, GoogLeNet, ResNet50, and AlexNet, and they have been compared in a similar combination. These models are designed differently: GoogLeNet is efficient with inception modules, ResNet50 is capable of learning deeper features with the help of residual connections, and AlexNet is simpler in structure. In order to ensure that the model is more reliable and flexible, data preprocessing methods such as normalization and data augmentation are implemented. Accuracy, class accuracy, confusion matrix and training loss trends were used as metrics to measure the performance. To apply in practice, the satellite image of Sentinel-2 was retrieved through Google Earth Engine and processed these images into a patch to analyze them. The outcomes of the single patches are integrated to produce a complete land use classification map which is subsequently displayed with the help of GIS. The approach is not only efficient in large scale classification but also provides a clear picture of the spatial land patterns. In general, the aim of the work is to construct a scalable framework of deep learning to classify the land use and compare various models with the purpose to identify the optimal model to use in geospatial applications.

II. LITERATURE REVIEW

Sr. No.	Title	Author/Year	Key Findings	Limitations
1	Mapping of Land Use and Land Cover (LULC) using EuroSAT and Transfer Learning	Suman Kunwar, Jannatul Ferdush (2023)	Showed that standard statistical methods (MLC) can hit 92% accuracy for broad categories like water and land.	Struggles to tell the difference between similar-looking areas like bare soil and concrete.
2	Building instance classification using street view images	Jian Kang et al. (2018)	Found that "ground-level" street views give much better detail for individual buildings than top-down satellite shots.	Relies heavily on high-quality street images; trees or cars often block the view of the buildings.
3	Classification of Land Cover and Land Use (LULC) Based on ResNet-50 and CNN Models	Yang et al. (2018)	Proved that deep learning (CNNs) can automate database updates with 85% accuracy by looking at both shapes and textures.	Needs a massive amount of perfectly labeled data to work well and is computationally "heavy".
4	Land Cover and Land Use Extraction Based on Deep Learning Methods Using Satellite Images	A. Gharagozlou (2020)	Confirmed that deep learning is far faster and more accurate than manual mapping for tracking city growth.	Struggles with low-resolution images where different land types start to "blur" together.
5	A Comparison of Deep Transfer Learning Methods for Land Use and Land Cover Classification	Dastour & Hassan (2023)	Compared 39 models and found ResNet-50 is one of the best "all-rounders" for accuracy and speed.	Some models are incredibly slow to train, taking 3x longer than others for only minor accuracy gains.
6	Land Use and Land Cover Classification	Luigi Selmi	For LULC classification ResNet50 increased accuracy	RGB-only data limitation, challenge of discriminating

Sr. No.	Title	Author/Year	Key Findings	Limitations
	using a ResNet Deep Learning Architecture		(~98%) via Transfer imaging and data augmentation.	similar land classes.
7	Satellite-Based Land Cover Mapping Through Fine-Tuned ResNet-50 Architecture	Aasish, Abhinav, Adwaith, Daya Naik, Nivin K S	With transfer learning, ResNet-50 achieved ~98.3% accuracy, indicating land cover classification is less complex.	Confusion between similar classes (e.g., vegetation types); limited use of only RGB bands.
8	Google Earth Engine for large-scale land use and land cover mapping: an object-based classification approach	Hossein Shafizadeh-Moghadam et al.	Object-based classification with RF model resulted in high accuracy (~91.7%) classifying large scale datasets.	Requires multiple features, high preprocessing; high computational complexity and large data processing time.
9	Landuse Landcover Change Detection Using Geospatial Semantic Segmentation U-Net-ResNet Architecture	Ranu Sewada, Hemlata Goyal	Multispectral satellite data classify land-use with pixel-level results using U-Net + ResNet34 with high accuracies.	Inferior accuracy for built-up and rock classes; does not generalise due to limited dataset and patch-size.
10	Deep Learning-based Land Use and Land Cover Changes Detection from Satellite Imagery	Mandicou BA, Pape Ibrahima Thiam, Etienne Delay, et al.	U-Net achieved better performance than FCN-8 in detecting LULC change with temporal satellite images.	Performance depends on quality of images and labelled data; class imbalance remains a great challenge.
11	An Analysis of Deep Neural Network Models for Practical Applications	Alfredo Canziani & Eugenio Culurciello, Adam Paszke / 2016	Comparative study highlighting trade-offs between accuracy, computational cost, and scalability of deep networks.	Lacks domain-specific focus on satellite imagery; no experimental validation for land use classification.
12	The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches	Alom et al., 2018	Historical overview of CNN evolution from AlexNet to GoogLeNet and ResNet with key innovations.	Mostly theoretical; lacks experimental validation. Not focused on satellite image classification.
13	Deep AlexNet with Reduced Number of Trainable Parameters for Satellite Image Classification	Anju Unnikrishna, Sowmya V, Soman K P / 2023	Proposes a lightweight AlexNet with reduced parameters maintaining accuracy while lowering computational cost.	May sacrifice performance on highly complex datasets. Limited comparison with more advanced architectures.
14	Image Classification Algorithm Based on	Shaojuan Li, Lizhi Wang, Jia	Introduces enhancements to AlexNet for better feature	Improvements are incremental; still less

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

Sr. No.	Title	Author/Year	Key Findings	Limitations
	Improved AlexNet	Li, Yuan Yao / 2021	extraction and training convergence.	powerful than deeper architectures. May not scale well.
15	SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and <0.5MB Model Size	Forrest N. Iandola et al. / 2016	Achieves AlexNet-level accuracy with significantly fewer parameters using fire modules.	Slight trade-off in accuracy compared to deeper models. Not optimized for satellite imagery tasks.

III. PROBLEM STATEMENT

Land use classification using satellite imagery is a complex task due to variability and big data. For accurate classification of land, this project applies Convolutional Neural Network (CNN) on tiled EuroSAT data.

IV. OBJECTIVES

- Develop a land use classification architecture that incorporates CNNs to improve its capability to identify areas using patches.
- Test the EuroSAT dataset by training and testing the model on different types of land.
- Apply preprocessing algorithms such as data normalization and data augmentation to ensure that the models are able to process new data.
- Compare GoogLeNet, ResNet50 and AlexNet to identify the most successful design.
- Check model performance using metrics such as test accuracy, class accuracy and confusion matrices.
- Prepare land use maps by applying the trained models on real-world Sentinel-2 satellite imagery.
- Visualize the results using GIS software such as QGIS for evaluation of land use predictions of different tiles to check spatial accuracy and comparison.
- Not only recommend precise technical corrective actions on future work in the area but also spot current restrictions.

V. METHODOLOGY

The research highlights a patch-based land use classification technique using Convolutional Neural Networks (CNNs). The choice of this method is that it is effective in capturing the spatial and textural details in the satellite images. The models can learn localized features, unlike classifying each pixel individually, which is delicate and sensitive to noise. These characteristics incorporate items such as color, texture, and structures that distinguish unmistakably the land use types such as farming, forests, or houses.

The patches are of 64×64 pixels to remain compatible with the EuroSAT dataset. It is advisable to match the size of the training with the prediction to make proper predictions. Smaller patches can easily lose context, whereas larger patches can overlap land use types, reducing accuracy. Therefore, 64×64 size is a good compromise on maintaining the clarity of classes and space.

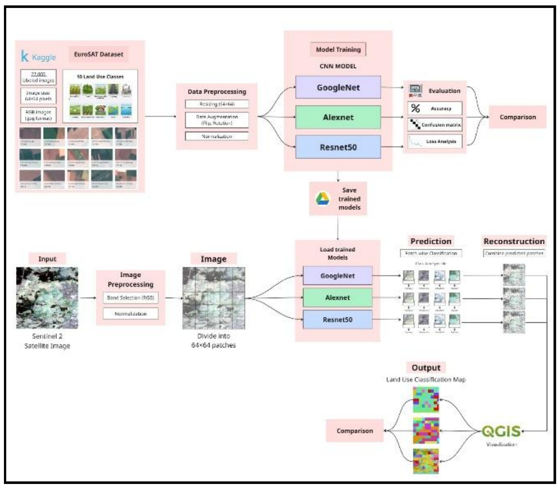
The annual crop, forest, herbaceous vegetation, highway, industrial, pasture, permanent crop, residential, river, sea lake are the land use categories included in the EuroSAT dataset that is trained and is supervised by learning. In order to ensure that the model can be generalized, preprocessing strategies such as normalization and data augmentation have been used. Normalization ensures that the input scale remains constant in various configurations and augmentation, such as flipping and rotating, make models allow other orientations of satellite images. Three CNN models were employed to test the effectiveness of the various architectures: GoogLeNet, ResNet50, and AlexNet. This helps to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

determine which model performance is the most optimal to handle spatial hierarchies and features. Accuracy and confusion matrices were used for measuring the performance to identify patterns. A patch-wise identification approach was used to process real-world satellite imagery in the inference stage to ensure consistency, with complete reconstruction of the interpretable classified patch-based land use output.

A. Proposed System Architecture

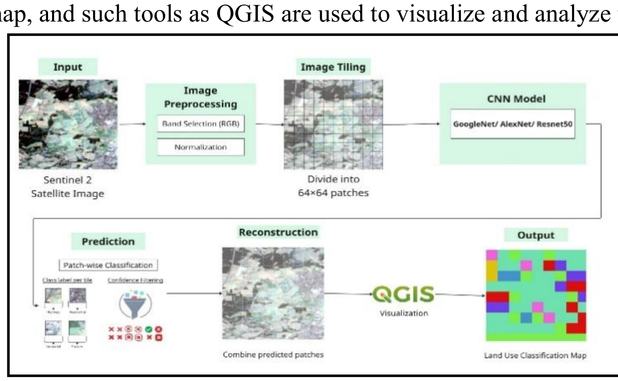
This system proposes a procedure for applying deep learning to classify land use. It begins with the EuroSAT dataset, in which CNN models, such as GoogLeNet, ResNet50, and AlexNet, are trained on 64×64 RGB patches. These patches are prepared with normalization and augmentation procedures. Accuracy, confusion matrix, and loss trends are used to evaluate the models. To apply in the real-world, Sentinel-2 images are pre-processed and divided into patches and each patch is classified with these trained models and filtered with confidence levels. These outputs are reconstructed into a complete land use map and displayed using GIS tools and each model can be analysed and compared.



[Figure 1: Proposed System Architecture]

B. Data Flow

The data flow chart explains the processing of satellite imagery during their entire process. Preparation of a Sentinel-2 image is initially done by choosing RGB bands, normalizing and breaking it down into 64×64 patches. Then CNN models such as GoogLeNet, ResNet50, or AlexNet process each patch to obtain class prediction and confidence scores. Any results with low confidence are set as unknown. Lastly, the patches are then classified and combined together to create a complete land use map, and such tools as QGIS are used to visualize and analyze the data.



[Figure 2: Data Flow Diagram]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

VI. IMPLEMENTATION

The PyTorch deep learning system was used to construct the system. The primary training data is the EuroSAT dataset of approximately 27,000 labeled RGB images divided into ten classes, such as Annual Crop, Forest, Industrial, Residential, etc. To test the effectiveness of this set, it was split into the training part and the testing one. The data preprocessing was done using torchvision transforms. It includes downsizing pictures to 64×64 pixels and normalizing them with ImageNet statistics. To reinforce the model, Random horizontal and vertical flips and rotations were used in the training to provide the models with different spatial views.

CNN has three architectures, namely GoogLeNet, ResNet50, and AlexNet. Their final connected layers were adjusted to the ten classes. The training involves Adam and cross-entropy loss. To maintain things constant, learning rate scheduling and gradient clipping was added to monitor progress using loss curves.

To be practical, Sentinel-2 images are retrieved in Google Earth Engine and stored as GeoTIFFs. The raster data was processed with the Rasterio library, getting RGB bands and breaking the images into non-overlapping 64×64 tiles. The tiles are passed through the trained models in order to make predictions and confidence scores. A threshold is used to eliminate low-confidence results, which was identified as unknown.

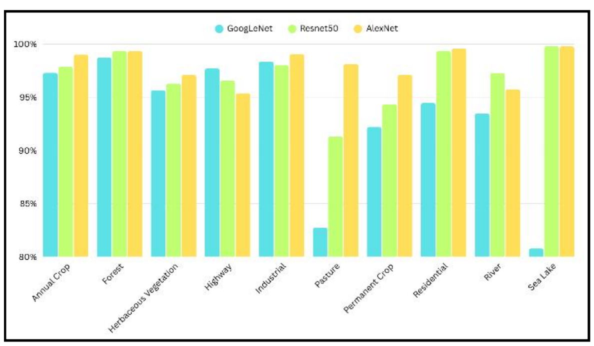
The last stage was to assemble the predicted tiles into a complete resolution classification map with important spatial metadata such as coordinate reference system (CRS). The completed map is stored as a GeoTIFF, and it is loaded into QGIS to be visualized, whereby each land use category is represented using a different color. The models were then compared to determine which one is the best when doing big geospatial tasks in terms of accuracy and spatial consistency.

VII. RESULT AND ANALYSIS

A. Evaluation (Test Accuracy for Each Model)

ResNet50 outperformed the other models by getting the highest accuracy of 97.28 percent, surpassing both GoogLeNet with 93.48 percent and AlexNet with 94.05 percent. It shows that more powerful models with residual relationships are better at capturing the detailed characteristics of satellite images. AlexNet, despite being simpler, was also performing well, with GoogLeNet being a little lower.

B. Class wise Accuracy

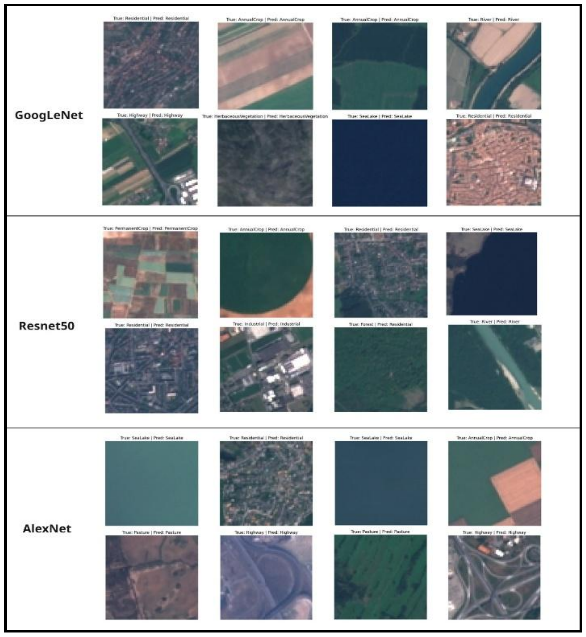


[Figure 3: Class Wise Accuracy]

Both AlexNet and ResNet50 performed quite decent on the majority of the classes in terms of class accuracy. Both had gone above 99 percent in Residential, Sea Lake, and Forest classes. GoogLeNet was, conversely, weak on similar looking land cover with poorer results in Pasture (82.75) and Sea Lake (80.79). Interestingly, AlexNet actually performed well in Pasture and Permanent Crop classes even though it is designed in a simpler manner.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, April 2026 Impact Factor: 8.2

C. Prediction on Test Images

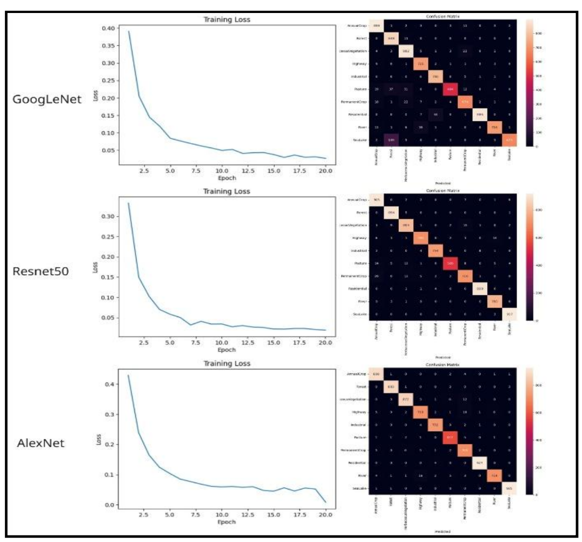


[Figure 4: Prediction on Test Images]

D. Training Loss

The learning curve in ResNet50 was smooth and stabilized, which indicates excellent learning and generalization. GoogLeNet was more prone to fluctuation and AlexNet reached the goal quicker but was a little unstable, which is likely due to its simplicity. These trends coincide with the accuracy outcomes, with ResNet50 as the most consistent and the most successful.

E. Confusion Matrix

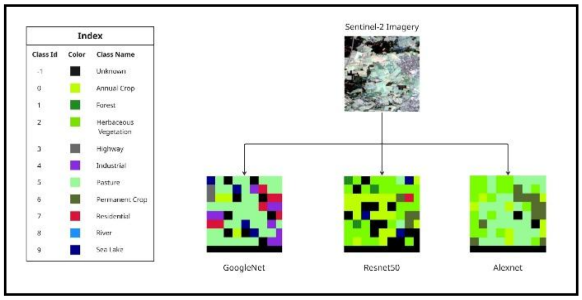


[Figure 5: Training Loss Graph and Confusion Matrix]

DOI: 10.48175/IJARSCT-33281

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

The confusion matrix indicates that GoogLeNet was more problematic in distinguishing similar classes such as SeaLake and Forest, which implies that it is not easy to distinguish similar spectral patterns. However, ResNet50 was able to better differentiate between classes with fewer errors, even in challenging classes. AlexNet was also stable, although it confused classes at times with similar textures such as Pasture and Herbaceous Vegetation.



[Figure 6: Patch wise Land Use Classified Output]

The outcome of patch-classification reveals the degree to which the three models perceive satellite information of the real world differently. All the models could identify key features, such as the constructed area on the right of the image. Residential was the name of this area in both GoogLeNet and ResNet50. Nevertheless, ResNet50 labels fewer residential patches than GoogLeNet, which means that it is more cautious in its predictions.

In farming areas, GoogLeNet and AlexNet mainly saw patches as Pasture. On the other hand, ResNet50 predicted them as Herbaceous Vegetation and PermanentCrop. This refers to the various perspectives on vegetation of these models. Also, ResNet50 correctly identified darker regions as Forest, whereas GoogLeNet occasionally detected these areas as SeaLake, indicating that it is more sensitive to spectral variations.

AlexNet usually grouped numerous patches into a few dominant groups such as PermanentCrop and Herbaceous Vegetation. It also had the smallest number of unknown patches, and this could indicate that it is overconfident in its predictions. ResNet50, on the other hand, had more unknowns, implying that it has more stringent confidence criteria.

In conclusion, ResNet50 provided the most balanced and reliable results compared to all other models, which captured the basic land use patterns. GoogLeNet was more variable and AlexNet was simpler although it tended to over simplify classifications.

VIII. CONCLUSION

This research shows how land use identified using features in imagery and classified patch wise with the help of neural networks, which were trained on the EuroSAT dataset. The system is capable of identifying significant patterns in the land use mapping projects by splitting satellite images into smaller, manageable tiles/patches and experimenting with models such as GoogLeNet, ResNet50 and AlexNet. This is done to ensure that the inference process remains the same as training to maintain efficiency even with detailed data.

A comparison of various models reveals that each of them is unique in its own right, yet such issues as ambiguity of boundaries and suchlike spectral characteristics still exist. Nonetheless, it is possible to use our framework as a scalable and interpretable approach for land use classification despite these problems. It serves as a base to further learning applications in geospatial technology.

In order to make the system work more efficient, it was important to address some challenges that were observed during testing. One of the important objectives is to reduce edge predictions that are of unknown quality. It can be accomplished through trial and error using overlapping tiles or by resizing them according to the requirements. This would be more extensive and ensure the final outcome is more uniform.

The addition of other different kinds of bands such as Near-Infrared (NIR) and Short-Wave Infrared (SWIR) can help in distinguishing various types of important features like greenery and water. Shifting to patch-based approaches to more intricate approaches such as U-Net or DeepLab may provide us with precision down to the pixel and provide clarity in mixed land use regions.

REFERENCES

[1] S. Kunwar and J. Ferdush, "Mapping of land use and land cover (LULC) using EuroSAT and transfer learning," arXiv preprint arXiv:2401.02424, 2023.

[2] J. Kang, M. Korner, Y. Wang, H. Tausenböck, and X. X. Zhu, "Building instance classification using street view images," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 145, pp. 44-59, 2018.

[3] C. Yang, F. Rottensteiner, and C. Heipke, "Classification of land cover and land use based on convolutional neural networks," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 4, pp. 251-258, 2018.

[4] A. Gharagozlou, "Land Cover and Land Use Extraction Based on Deep Learning Methods Using Satellite Images."

[5] H. Dastour and Q. K. Hassan, "A comparison of deep transfer learning methods for land use and land cover classification," Sustainability, vol. 15, no. 10, p. 7854, 2023.

[6] L. Selmi, "Land Use and Land Cover Classification using a ResNet Deep Learning Architecture," Living Planet Symposium (ESA), 2022.

[7] Aasish, Abhinav, Adwaith, D. Naik, and Nivin K. S., "Satellite-Based Land Cover Mapping Through Fine-Tuned ResNet-50 Architecture," Journal of Emerging Technologies and Innovative Research (JETIR), vol. 12, no. 7, 2025.

[8] H. Shafizadeh-Moghadam, M. Khazaei, S. K. Alavipanah, and Q. Weng, "Google Earth Engine for large-scale land use and land cover mapping: An object-based classification approach using spectral, textural and topographical factors," GIScience & Remote Sensing, vol. 58, no. 6, pp. 914-928, 2021.

[9] R. Sewada and H. Goyal, "Landuse Landcover Change Detection Using Geospatial Semantic Segmentation U-Net-ResNet Architecture," 2023 IEEE 11th Region 10 Humanitarian Technology Conference (R10-HTC), Rajkot, India, Oct. 2023.

[10] Mandicou BA, P. I. Thiam, E. Delay, C. A. Ngom, I. Diop, and A. Bah, "Deep Learning-based Land Use and Land Cover Changes Detection from Satellite Imagery," Proc. 7th International Conference on Machine Vision and Applications (ICMVA), Singapore, 2024.

[11] A. Canziani, E. Culurciello, and A. Paszke, "An analysis of deep neural network models for practical applications," 2016.

[12] M. Z. Alom et al., "The history began from AlexNet: A comprehensive survey on deep learning approaches," 2018.

[13] A. Unnikrishna, V. Sowmya, and K. P. Soman, "Deep AlexNet with reduced number of trainable parameters for satellite image classification," 2023.

[14] S. Li, L. Wang, J. Li, and Y. Yao, "Image classification algorithm based on improved AlexNet," 2021.

[15] F. N. Iandola et al., "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size," 2016.