International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Deep Learning Based Land Cover Classification Using Satellite Imaginary for Sustainable Urban Planning

[1]M. Durga Sridevi
[1]M.Tech, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
durgasridevi12@gmail.com

[2]Mr.K.T.V. Subbarao
[2]Assistant Professor, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
kanurisubbarao@gamail.com

Abstract*— Land cover classification plays a critical role in understanding urban development, environmental changes, and spatial patterns required for sustainable urban planning. The rapid availability of high-resolution satellite imagery has created opportunities for automated and large-scale analysis of urban and natural land-cover conditions. However, manual interpretation and conventional image-processing techniques are time-consuming and may have difficulty distinguishing complex land-cover patterns. This literature survey examines deep learning-based approaches for automated land cover classification from satellite imagery, with particular emphasis on Convolutional Neural Networks (CNNs), spatial feature extraction, image preprocessing, and multi-class classification. Existing machine-learning and deep-learning methods are analyzed according to their classification capabilities, advantages, limitations, computational requirements, and applicability to urban environments. The survey further examines the use of satellite-derived information for identifying buildings, roads, vegetation, water bodies, and other land-cover categories. Based on the identified limitations, an integrated artificial-intelligence framework is formulated that combines satellite image acquisition, preprocessing, feature extraction, deep learning-based classification, accuracy assessment, and urban planning decision support. The framework is intended to support automated land-cover mapping, urban expansion monitoring, environmental-change analysis, and detection of urban sprawl. The study also identifies challenges related to dataset diversity, spatial resolution, class imbalance, computational complexity, seasonal variations, model generalization, and explainability. Overall, deep learning-based satellite image classification provides a promising foundation for intelligent, efficient, and sustainable urban land management.

Keywords*— Land Cover Classification, Satellite Imagery, Deep Learning, Convolutional Neural Networks, Remote Sensing, Image Classification, Urban Planning, Urban Growth, Urban Sprawl, Vegetation Mapping, Water Body Detection, Sustainable Development, Environmental Monitoring, Smart Cities, Deep Learning-Based Remote Sensing.

I. INTRODUCTION

Land cover classification is an important component of remote sensing and sustainable urban planning, providing valuable information about the distribution of buildings, roads, vegetation, water bodies, and other surface characteristics. Rapid urbanization continuously changes land-cover patterns and creates challenges related to infrastructure development, environmental protection, transportation planning, and natural-resource management. Satellite imagery provides a practical source of large-scale spatial information for observing these changes over time. Recent advances in artificial intelligence and deep learning have created opportunities for automated land-cover classification and large-scale satellite image analysis [1]–[4].

Conventional land-cover mapping generally depends on manual interpretation, field surveys, traditional image-processing techniques, or conventional machine-learning algorithms. Although these approaches can provide useful classification results, manual interpretation becomes time-consuming when large geographical regions and high-resolution satellite images are considered. Deep-learning techniques can automatically learn representative spatial features from satellite imagery and reduce dependence on manually designed features [5]–[8]. Convolutional Neural Networks (CNNs) have therefore become widely investigated for extracting complex visual patterns and classifying different land-cover categories.

Satellite images contain substantial spatial variability caused by differences in illumination, seasonal conditions, geographical characteristics, image resolution, and land-cover appearance. Buildings, roads, vegetation, and water bodies may exhibit similar spectral or spatial characteristics in certain conditions, making accurate classification challenging [9]–[11]. Deep-learning models can learn hierarchical representations from image data and provide improved capabilities for distinguishing complex land-cover patterns.

Accurate land-cover classification can support urban expansion monitoring, environmental-change assessment, urban-sprawl detection, infrastructure planning, and sustainable resource management [12]–[16]. However, challenges including limited labeled datasets, class imbalance, computational requirements, spatial-resolution differences, seasonal variations, and model generalization remain important research issues.

This survey therefore aims to:

1. Review deep-learning and machine-learning approaches for satellite-image-based land-cover classification.
2. Examine CNN-based techniques for extracting spatial features from remote-sensing imagery.
3. Analyze classification approaches for buildings, roads, vegetation, water bodies, and related land-cover categories.
4. Compare the strengths, limitations, accuracy, and computational requirements of existing approaches.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

5. Examine the influence of image resolution, preprocessing, dataset diversity, and environmental variations on classification performance.
6. Formulate an integrated deep-learning framework for automated land-cover classification and sustainable urban planning.
7. Identify evaluation requirements, research challenges, and future directions for reliable and intelligent satellite-image analysis.

The remainder of this paper discusses existing land-cover classification methods, the proposed comprehensive deep-learning framework, satellite-image and land-cover representation, classification analysis, urban planning decision support, evaluation requirements, research challenges, future research directions, and concluding observations.

II. EXISTING METHODS FOR DEEP LEARNING-BASED LAND COVER CLASSIFICATION

Existing satellite-image-based land-cover classification approaches can be broadly classified into manual interpretation, spectral and pixel-based methods, conventional machine learning, CNN-based classification, deep feature extraction, semantic segmentation, and advanced attention-based approaches.

A. Manual and Traditional Satellite Image Interpretation

Traditional land-cover mapping relies on manual image interpretation, field observations, expert knowledge, and conventional image-processing techniques. Although useful for small areas, these methods are time-consuming, subjective, and difficult to scale for large satellite-image collections. Variations in illumination, season, atmospheric conditions, and image resolution further affect interpretation reliability.

B. Spectral and Pixel-Based Classification

Pixel-based methods classify land-cover categories using spectral characteristics of individual pixels. They can identify vegetation, water bodies, buildings, and other surfaces efficiently. However, they have limited capability to represent spatial relationships between neighboring pixels, which can result in fragmented classification, particularly in complex urban environments.

C. Conventional Machine-Learning Classification

Methods such as Support Vector Machines, Random Forest, and Decision Trees provide automated classification using extracted spectral, spatial, and textural features. Their performance depends strongly on feature engineering and training-data quality, limiting their ability to represent highly complex spatial patterns.

D. CNN-Based Classification

CNNs automatically learn hierarchical spatial features such as edges, textures, boundaries, and object patterns directly from satellite imagery. They reduce dependence on handcrafted features and provide effective land-cover classification. However, performance can be affected by dataset diversity, class similarity, and computational requirements.

E. Deep Feature and Semantic Segmentation Approaches

Pretrained deep networks can extract meaningful representations through transfer learning, while semantic segmentation provides pixel-level land-cover maps. These approaches are useful for detailed building, road, vegetation, and water-body mapping but generally require substantial annotated data and computational resources.

TABLE I: COMPARISON OF EXISTING LAND COVER CLASSIFICATION METHODS

Method	Technique	Strength	Limitation
Manual	Visual analysis	Simple	Time-consuming
Pixel-based	Spectral features	Fast	Limited spatial information
Machine Learning	RF, SVM	Automated	Feature dependent
CNN	Deep feature learning	High-level features	Data required
Segmentation	Pixel-level learning	Detailed mapping	Computationally complex
Proposed	**CNN-based classification**	**Automated land-cover mapping**	**Requires validation**

F. Advanced and Sustainable Urban-Planning Approaches

Attention, multi-scale, hybrid, and encoder-decoder architectures can improve representation of land-cover objects with different sizes and spatial characteristics. Classified imagery can support urban expansion, vegetation monitoring, infrastructure planning, and urban-sprawl analysis. However, dataset limitations, class imbalance, seasonal variation, resolution differences, and model generalization remain challenges. An integrated deep-learning framework should therefore connect satellite acquisition, preprocessing, feature extraction, classification, change analysis, and sustainable urban-planning decision support.

III. PROPOSED DEEP LEARNING-BASED LAND COVER CLASSIFICATION AND SUSTAINABLE URBAN PLANNING SYSTEM

The proposed system considers satellite-image-based land-cover classification as an integrated sequence of interconnected processing stages:

Satellite Imagery + Geographic Data
↓
Image Acquisition
↓
Data Preprocessing
↓
Image Enhancement and Normalization
↓
Spatial Feature Extraction
↓
CNN-Based Deep Feature Learning
↓
Land Cover Classification
↓
Building / Road / Vegetation / Water Body Identification
↓
Land Cover Change and Urban-Sprawl Analysis
↓
Urban Planning Decision Support
↓
Planner Monitoring and Intervention

The architecture consists of five major functional layers: satellite-image acquisition and preprocessing, spatial feature extraction, deep-learning classification, land-cover analysis, and sustainable urban-planning decision support. The first layer receives satellite imagery covering the target geographical region. The input imagery may contain different land-cover characteristics and spatial patterns that must be prepared before classification.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

The preprocessing layer performs image resizing, normalization, noise reduction, and other required operations to obtain consistent image representations. Appropriate preprocessing is important because variations in image quality, resolution, illumination, atmospheric conditions, and acquisition periods can influence classification performance.

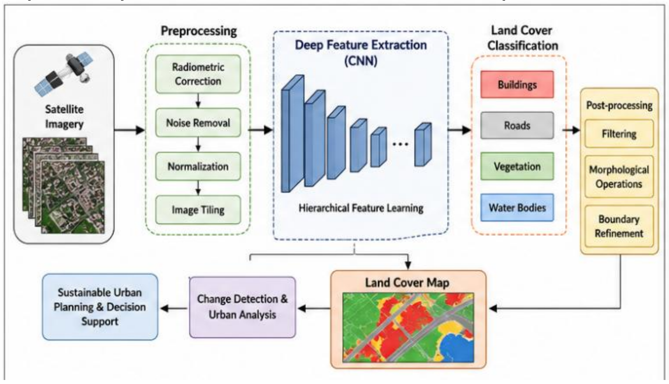

Fig. 1. Overall Architecture of the Proposed Deep Learning-Based Land Cover Classification System

The spatial feature extraction layer analyzes the processed satellite images to identify meaningful visual characteristics. The proposed framework uses a CNN-based deep-learning model to automatically learn hierarchical representations from satellite imagery. The learned features can capture edges, textures, shapes, spatial patterns, and object characteristics associated with different land-cover categories.

The classification layer uses the extracted deep features to classify satellite imagery into predefined categories such as buildings, roads, vegetation, and water bodies. The model produces a land-cover representation that can be used for subsequent spatial analysis. The classification process reduces dependence on manual interpretation and enables automated processing of large volumes of satellite imagery.

The land-cover analysis layer examines the classified output to identify spatial distributions and changes in urban and environmental regions. Building expansion, road development, vegetation reduction, and changes in water-body coverage can be analyzed from the generated land-cover maps. Such information can provide useful evidence for understanding urban growth and land-use transformation.

Finally, the sustainable urban-planning decision-support layer presents classified land-cover information and detected changes to planners through an accessible interface. The system can support urban expansion monitoring, infrastructure planning, environmental assessment, vegetation management, and urban-sprawl analysis. The framework is intended to assist planning decisions rather than completely replace human judgement.

IV. SATELLITE IMAGE DATA AND LAND-COVER REPRESENTATION

A. Satellite Image Acquisition and Processing

The proposed system considers satellite imagery as the primary input for automated land-cover analysis. Images can cover urban, semi-urban, and surrounding natural regions containing buildings, roads, vegetation, water bodies, and other surface categories.

The preprocessing stage prepares satellite images for deep-learning analysis. It can handle differences in image dimensions, resolution, brightness, noise, and other image characteristics. Images are resized into a consistent representation and normalized before being provided to the deep-learning model.

Image preprocessing is important because inconsistent input characteristics can negatively influence feature extraction and classification. Satellite images acquired at different times may also exhibit variations caused by seasonal conditions, atmospheric effects, illumination, and environmental changes. Appropriate preprocessing can therefore improve the consistency of the classification process.

B. Land-Cover Category Representation

The proposed framework represents the study region using multiple land-cover categories. The primary categories considered in the project are:

- **Buildings**
- **Roads**
- **Vegetation**
- **Water Bodies**

Buildings represent constructed urban structures and provide information about the spatial expansion and density of developed areas. Roads represent transportation-related land cover and can provide information about urban connectivity and infrastructure development.

Vegetation represents green areas, agricultural regions, parks, forests, and other plant-covered surfaces. Vegetation information is important for environmental assessment and sustainable urban development. Water bodies represent rivers, lakes, reservoirs, ponds, and other water-covered regions and can support water-resource monitoring.

The classification of these categories provides a structured representation of the geographical environment. The resulting land-cover map can subsequently be used for urban-growth monitoring and planning analysis.

C. Spatial Feature Representation

Satellite images contain important spatial information related to object shape, texture, boundaries, patterns, and neighboring regions. The proposed CNN-based architecture automatically extracts these characteristics from the input imagery.

Early network layers can learn basic visual characteristics such as edges and textures, while deeper layers can capture more complex spatial patterns. These representations help distinguish different land-cover categories even when their visual characteristics are complex.

For example, roads may be represented through elongated spatial patterns, buildings through geometric structures, vegetation through texture and distribution patterns, and water bodies through relatively homogeneous regions. Learning these characteristics automatically reduces dependence on handcrafted feature engineering.

D. Deep Feature Representation

The extracted CNN features provide a high-level representation of satellite-image content. These features are subsequently provided to the classification component for identifying the appropriate land-cover category.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Deep feature learning can capture nonlinear relationships within satellite imagery and provide stronger representations than conventional pixel-based approaches. It can also support automated classification across large collections of images.

However, the quality of learned representations depends on the diversity and quality of training data. Images containing different geographical regions, seasons, resolutions, and environmental conditions should therefore be considered during model development.

TABLE II: LAND-COVER CLASSES AND URBAN-PLANNING APPLICATIONS

Land-Cover Class	Main Purpose
Buildings	Urban expansion monitoring
Roads	Infrastructure planning
Vegetation	Environmental monitoring
Water Bodies	Water-resource management

E. Urban Land-Cover Representation

The classified satellite imagery provides a spatial representation of urban and environmental land-cover conditions. The system can use this representation to determine the distribution of buildings, roads, vegetation, and water bodies within a geographical region.

Comparing land-cover maps generated from different observation periods can provide information about changes in urban development. Increasing building coverage may indicate urban expansion, while changes in vegetation or water-body coverage may indicate environmental transformation.

This representation provides a foundation for connecting deep-learning classification with sustainable urban-planning applications.

V. DEEP LEARNING-BASED LAND COVER CLASSIFICATION AND URBAN ANALYSIS

A central component of the proposed framework is the combination of deep spatial feature learning and multi-class land-cover classification. Conventional pixel-based methods may consider individual spectral characteristics, whereas the proposed CNN-based approach learns spatial relationships from image regions.

The CNN processes satellite-image data and extracts increasingly complex features through multiple learning stages. These features are used to distinguish buildings, roads, vegetation, and water bodies. The classification output provides a structured land-cover map representing the spatial distribution of different categories.

The classification process can also support analysis of urban development. Building and road classifications can provide information about the growth of built-up areas and transportation infrastructure. Vegetation classification can help identify green-space distribution and potential vegetation loss. Water-body classification can support monitoring of aquatic resources and changes in surface-water coverage.

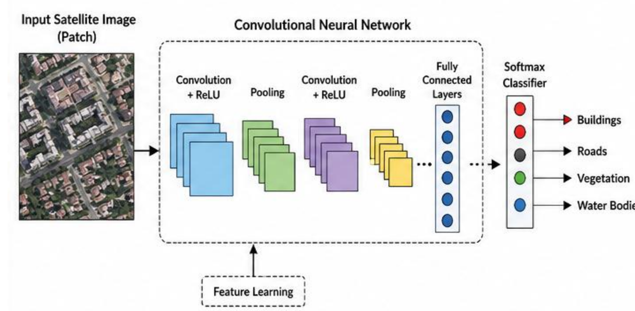

Fig. 2. Deep Learning-Based Satellite Image Classification and Sustainable Urban Planning Analysis

The proposed system can further compare classified land-cover information across different observation periods. Such comparison can assist in identifying urban expansion, changes in vegetation, modifications in water bodies, and development of new infrastructure.

The framework therefore extends beyond simple image classification toward intelligent spatial analysis. The resulting information can be presented to urban planners as supporting evidence for infrastructure development, environmental management, land-use planning, and sustainable urban growth.

The proposed approach also provides a foundation for incorporating additional satellite-image categories in future implementations. Depending on dataset availability, categories such as bare soil, agricultural land, forest, industrial areas, and recreational spaces can be incorporated into the classification system.

Overall, the integration of satellite-image preprocessing, CNN-based spatial feature extraction, multi-class classification, land-cover change analysis, and planning-oriented interpretation provides a comprehensive framework for automated land-cover mapping and sustainable urban planning.

VI. LAND-COVER CHANGE ANALYSIS AND SUSTAINABLE URBAN PLANNING

The classified land-cover information generated by the proposed framework can provide valuable spatial evidence for sustainable urban planning. Rather than limiting the system to assigning a class to individual satellite-image regions, the classification results can be further analyzed to understand the distribution and transformation of urban and environmental areas. Buildings, roads, vegetation, and water bodies provide complementary information about the physical development and environmental condition of a region.

Building classification can be used to identify developed areas and examine patterns of urban expansion. Increasing built-up regions across different observation periods may indicate urban growth and the conversion of previously undeveloped land. Such information can assist planners in understanding the direction and intensity of urban development.

Road classification provides information about transportation infrastructure and its spatial relationship with developed areas. Identifying newly developed roads can help indicate emerging urban regions and provide supporting information for infrastructure planning. The combined analysis of buildings and

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

roads can therefore provide a more meaningful representation of urban development than analyzing either category independently.

Vegetation classification is important for evaluating the environmental characteristics of urban regions. Green areas can contribute to ecological balance, recreational planning, and environmental sustainability. Monitoring changes in vegetation coverage through satellite imagery can help identify regions experiencing vegetation reduction or development pressure.

Water-body classification can support the monitoring of rivers, lakes, reservoirs, ponds, and other surface-water regions. Changes in the spatial extent of water bodies can provide useful information for environmental monitoring and land-use management.

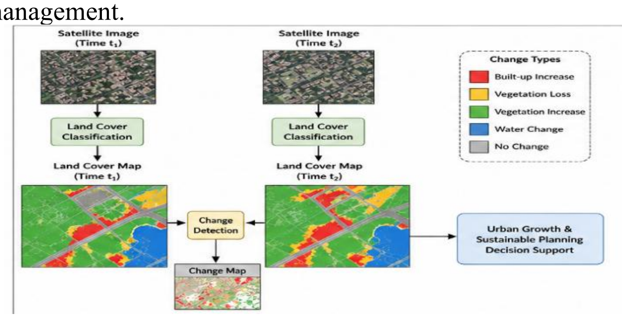

Fig. 3. Land-Cover Change Analysis for Sustainable Urban Planning

The proposed framework can compare classified satellite images acquired at different time periods to identify changes in land-cover distribution. This temporal analysis can support the detection of urban expansion, vegetation changes, infrastructure development, and modifications in water-body coverage.

The resulting information can be provided to urban planners through a decision-support interface. Planners can use the classified maps and detected changes as supporting information for land-use planning, infrastructure development, environmental management, and urban-growth assessment.

The framework does not replace expert planning decisions. Instead, it provides automated and consistent spatial information that can reduce the effort required for large-scale land-cover assessment. Human experts can review the generated classifications and incorporate additional geographical, demographic, economic, and regulatory information before making final planning decisions.

VII. EXPLAINABLE DEEP LEARNING AND DECISION SUPPORT

Deep-learning models can provide powerful classification capabilities, but their decision-making processes may be difficult to interpret. For applications related to urban planning, understanding why a particular region has been classified as a building, road, vegetation area, or water body can improve confidence in the generated results.

The proposed framework can therefore incorporate an explainable analysis component to identify important image regions contributing to the classification decision. Visual explanations can help users understand whether the model is focusing on meaningful spatial patterns or irrelevant image characteristics.

Explainability is particularly important when satellite images contain visually complex regions. For example, a building may be surrounded by roads and vegetation, while water bodies may contain reflections, shadows, or surrounding structures. An explanation mechanism can assist planners and analysts in reviewing uncertain or potentially incorrect classifications.

The system can provide classification results together with confidence information and representative image regions. Low-confidence predictions can be highlighted for additional expert examination. This creates a human-in-the-loop process in which automated deep learning performs large-scale analysis while experts validate important or uncertain results.

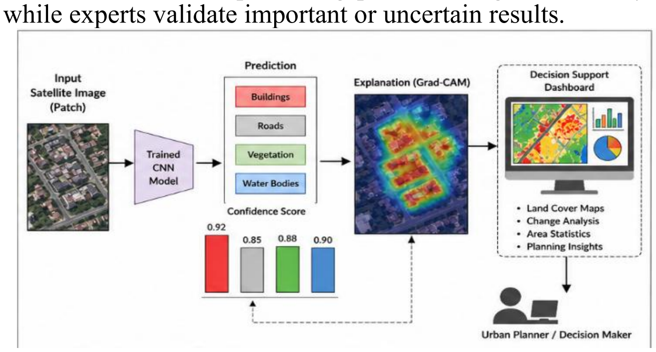

Fig. 4. Explainable AI-Based Land-Cover Classification and Urban Planning Decision Support

The decision-support layer can present classified land-cover maps, category distributions, detected changes, and potentially suspicious or uncertain regions. Such information can support urban expansion assessment, vegetation monitoring, water-resource observation, and infrastructure planning.

The explainable component can also improve the practical acceptance of deep-learning systems by making classification outputs easier to inspect. This is especially relevant for planning applications where decisions may have long-term environmental and socioeconomic consequences.

VIII. EVALUATION FRAMEWORK

Evaluation is an essential component for determining whether the proposed deep-learning framework can reliably classify satellite imagery. The evaluation process should consider both classification performance and practical applicability to urban-planning scenarios.

The primary evaluation measures can include accuracy, precision, recall, and F1-score. Accuracy provides an overall measure of correctly classified samples, while precision evaluates the correctness of predictions for individual land-cover categories. Recall measures the ability of the system to identify relevant samples, and the F1-score provides a combined measure of precision and recall.

A confusion matrix can additionally be used to identify which land-cover categories are most frequently confused. This is particularly important because buildings and roads, or vegetation and other land surfaces, may exhibit similar characteristics in certain satellite images.

For spatial classification and segmentation applications, additional measures such as Intersection over Union (IoU) can be considered to evaluate the spatial overlap between predicted and reference regions. Mean Intersection over Union can provide an overall assessment across multiple land-cover categories.

The evaluation should also consider the diversity of satellite imagery. Testing should include images from different geographical regions and, where possible, different acquisition periods. Such evaluation can provide information about the ability of the model to generalize beyond the training data.

The proposed framework can be evaluated against conventional machine-learning methods and alternative deep-learning architectures. This comparison can determine whether CNN-based feature learning provides meaningful improvements over traditional approaches.

Practical evaluation should also examine computational requirements, processing time, memory usage, and scalability. A model that provides high classification accuracy but requires excessive computational resources may be difficult to deploy for large-scale urban monitoring.

IX. RESEARCH CHALLENGES AND FUTURE ROADMAP

Despite the progress of deep learning in satellite-image classification, several challenges remain. One major challenge is the availability of sufficiently large and diverse labeled datasets. Creating accurate annotations for satellite imagery can require substantial time and expert effort.

Class imbalance is another important issue. Some land-cover categories may occupy much larger areas than others, causing the model to learn dominant categories more effectively while producing weaker predictions for less-represented classes.

Variations in satellite-image resolution also affect classification performance. High-resolution imagery can provide detailed information about buildings and roads, while lower-resolution imagery may provide broader geographical coverage. Developing models capable of handling multiple resolutions remains an important research direction.

Seasonal and environmental variations can further influence classification. Vegetation appearance can change substantially across seasons, while atmospheric conditions and illumination can modify satellite-image characteristics. Models should therefore be evaluated across diverse environmental conditions.

Another challenge is cross-region generalization. A model trained using images from one geographical region may not perform equally well in another region because of differences in architecture, road structures, vegetation, geography, and urban development patterns.

Future research can investigate transfer learning, domain adaptation, attention mechanisms, transformer-based architectures, multi-source satellite data, and multimodal remote-sensing analysis. Integration of optical imagery with additional spatial or environmental information may improve classification robustness.

Future systems can also incorporate continuous change monitoring and predictive urban-growth analysis. Instead of only identifying existing land cover, AI systems could estimate potential future development patterns and provide early information for sustainable planning.

Explainable AI, uncertainty estimation, and human-in-the-loop validation should also receive greater attention. These capabilities can improve transparency and enable planners to distinguish highly reliable predictions from classifications requiring expert verification.

X. CONCLUSION

Deep-learning-based satellite image classification provides an effective foundation for automated land-cover mapping and sustainable urban planning. CNN-based feature learning can identify complex spatial patterns associated with buildings, roads, vegetation, and water bodies while reducing dependence on manual feature engineering. Integrating classification with land-cover change analysis, explainable AI, evaluation, and decision support can further improve its practical value. The proposed framework can assist urban planners in monitoring urban expansion, infrastructure development, vegetation changes, water resources, and urban sprawl. Future research should focus on robust generalization, diverse datasets, explainability, computational efficiency, and predictive urban-growth intelligence.

REFERENCES

[1] P. Zhang, Y. Ke, Z. Zhang, M. Wang, P. Li, and S. Zhang, "Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery," *Sensors*, vol. 18, no. 11, Art. no. 3717, 2018, doi: 10.3390/s18113717.

[2] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 7, pp. 2217–2226, 2019, doi: 10.1109/JSTARS.2019.2918242.

[3] M. Zhu, B. Wu, Y. N. He, and Y. Q. He, "Land cover classification using high resolution satellite image based on deep learning," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLII-3-W10, pp. 685–690, 2020, doi: 10.5194/isprs-archives-XLII-3-W10-685-2020.

[4] J. Zhao *et al.*, "A land cover classification method for high-resolution remote sensing images based on NDVI deep learning fusion network," *Remote Sens.*, vol. 14, no. 21, Art. no. 5455, 2022, doi: 10.3390/rs14215455.

[5] A. Boguszewski, D. Batorski, N. Ziemba-Jankowska, T. Dziedzic, and A. Zambrzycka, "LandCover.ai: Dataset for automatic mapping of buildings, woodlands, water and roads from aerial imagery," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, 2021, pp. 1102–1110, doi: 10.1109/CVPRW53098.2021.00121.

[6] G. Sumbul and B. Demir, "A deep multi-attention driven approach for multi-label remote sensing image classification," *IEEE Access*, vol. 8, pp. 95934–95946, 2020, doi: 10.1109/ACCESS.2020.2995805.

[7] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, "Classification of remote sensing images using EfficientNet-B3 CNN model with attention," *IEEE Access*, vol. 9, pp. 14078–14094, 2021, doi: 10.1109/ACCESS.2021.3051085.

[8] H. Li, B. Hu, Q. Li, and L. Jing, "CNN-based tree species classification using airborne lidar data and high-resolution

satellite image," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2020, pp. 2679–2682, doi: 10.1109/IGARSS39084.2020.9324011.

[9] A. Rakhlin, A. Davydow, and S. Nikolenko, "Land cover classification from satellite imagery with U-Net and Lovász-Softmax loss," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, 2018, pp. 262–266, doi: 10.1109/CVPRW.2018.00048.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[11] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 3431–3440, doi: 10.1109/CVPR.2015.7298965.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[13] R. Fan, R. Feng, L. Wang, J. Yan, and X. Zhang, "Semi-MCNN: A semisupervised multi-CNN ensemble learning method for urban land cover classification using submeter HRRS images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4973–4987, 2020, doi: 10.1109/JSTARS.2020.3018202.

[14] C. Zhang, X. Pan, H. Li, A. Gardiner, I. Sargent, J. Hare, and P. M. Atkinson, "A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 140, pp. 133–144, 2018, doi: 10.1016/j.isprsjprs.2017.07.014.

[15] S. Chaib, H. Liu, Y. Gu, and H. Yao, "Deep feature fusion for VHR remote sensing scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4775–4784, 2017, doi: 10.1109/TGRS.2017.2700322.

[16] Y. Sun and W. Zheng, "HRNet- and PSPNet-based multiband semantic segmentation of remote sensing images," *Neural Comput. Appl.*, vol. 35, no. 12, pp. 8667–8675, 2023, doi: 10.1007/s00521-022-08111-2.