Journal of Spatial Information Sciences

...JSIS

APPLICATION OF CONVOLUTIONAL NEURAL NETWORK (CNN) IN IDENTIFICATION AND MAPPING OF URBAN ROAD NETWORK IN PARTS OF BENIN CITY, NIGERIA USING REMOTELY SENSED IMAGERY

Isaac O. Ovu, J.I. Igbokwe, J.O. Ejikeme, A.J. Adeboboye

www.journals.unizik.edu.ng/jsis

APPLICATION OF CONVOLUTIONAL NEURAL NETWORK (CNN) IN IDENTIFICATION AND MAPPING OF URBAN ROAD NETWORK IN PARTS OF BENIN CITY, NIGERIA USING REMOTELY SENSED IMAGERY

Isaac O. Ovu[1], J.I. Igbokwe[2], J.O. Ejikeme[3], A.J. Adeboboye[4]

[1, 2, 3&4]Department of Surveying and Geoinformatics, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria
*Corresponding Author Email: zikbridgeinternational@gmail.com

DOI: https://doi.org/10.5281/zenodo.15709354

Abstract

Accurate mapping of urban road networks is critical for sustainable urban planning, efficient transportation management, and disaster response. This study presents an automated approach for urban road network extraction using Convolutional Neural Networks (CNNs) applied to high-resolution unmanned aerial vehicle (UAV) imagery of Benin City, Nigeria. UAV data were captured using the DJI Matrice 100 platform equipped with a Zen muse X5 camera, providing imagery at a spatial resolution of 10–15 cm. Following comprehensive preprocessing steps—comprising noise reduction, calibration, image enhancement, georeferencing, orthorectification, and mosaicking—training datasets were generated through manual labeling and feature extraction using ArcGIS. The U-Net based CNN model was trained using 80% of the labeled data with the remaining 20% reserved for testing. Data augmentation techniques were employed to enhance model generalization and mitigate over fitting. Model evaluation demonstrated robust performance, achieving a validation accuracy of 91.3%, mean Intersection over Union (IoU) of 0.834, precision of 0.89, recall of 0.85, and F1-score of 0.87. The trained model exhibited computational efficiency, processing images at an average of 127 ms per image using 2.84 GB of GPU memory. Beyond road extraction, the model successfully classified additional urban land cover classes, including buildings, bare ground, vegetation, and water bodies, yielding an overall dataset quality rating of 8.8/10. The study highlights the potential of deep learning models in providing scalable, accurate, and efficient solutions for urban infrastructure mapping.

Keywords: Convolutional Neural Network (CNN), U-Net, UAV imagery, road network extraction, urban planning, Benin City, semantic segmentation, deep learning, remote sensing.

1.0 INTRODUCTION

In recent years, rapid urbanization has led to increasing demands on transportation infrastructure, urban planning, and disaster management systems particularly in developing countries [6]. Accurate and up-to-date mapping of urban road networks is critical for facilitating effective decision-making in these areas. Traditional road network mapping methods, which rely heavily on field surveys and manual digitization of aerial photographs, are often time-consuming, labor-intensive, and susceptible to human error [8]. These limitations underscore the need for more efficient, automated, and scalable solutions for road network extraction, especially in rapidly expanding urban centers such as Benin City, Nigeria.

The integration of high-resolution remotely sensed imagery with advanced machine learning techniques offers a promising solution to these challenges. Among various machine learning models, Convolutional Neural Networks (CNNs) have demonstrated remarkable success in computer vision tasks due to their hierarchical feature extraction capabilities and ability to learn complex spatial patterns[7]. CNNs, particularly architectures like U-Net, are well-suited for semantic segmentation tasks, making them highly effective for pixel-level classification required in road extraction processes [3].

This study investigates the application of CNNs for the automated identification and mapping of urban road networks in parts of Benin City, Nigeria, using high-resolution Unmanned Aerial Vehicle (UAV) imagery. The UAV data, captured using the DJI Matrice 100 equipped with a Zen muse X5 camera, provides detailed spatial information at resolutions between 10–15 cm, enabling the model to accurately delineate road networks and other urban features. Through a structured workflow involving data preprocessing, feature extraction, model training, and evaluation, this research demonstrates the superior accuracy, efficiency, and scalability of CNN-based approaches compared to traditional methods.

By employing state-of-the-art deep learning techniques, this study aims not only to advance the field of automated road network extraction but also to provide valuable insights and tools for urban planners, transportation authorities, and disaster management agencies operating in rapidly evolving urban environments.

2.0 STUDY AREA

The study area covers parts of Oredo and Egor Local Government Areas of Benin City, Edo State, Nigeria. Benin City, located between 06°26′01”N to 06°26′56”N and 05°31′11”E to 05°41′09”E, spans approximately 50 square kilometers. The area is characterized by tropical rainforest vegetation, humid climate, and gently undulating topography. Benin City, located in Edo State, Nigeria, is a vibrant metropolis known for its rich cultural heritage and diverse population. The city's demographic makeup encompasses various ethnic groups such as the Edo (Binis), Esan, and Afemai, contributing to its cultural tapestry. Its population includes people engaged in trade, agriculture, government service, and a growing number involved in the arts and education sectors. The city is a blend of tradition and modernity, with a youthful population contributing to its dynamism and economic activities.

JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 2, ISSUE 2, PP 133–147, 2025 DOI: https://doi.org/10.5281/zenodo.15709354
PUBLISHED 21-05-2025

www.journals.unizik.edu.ng/jsis

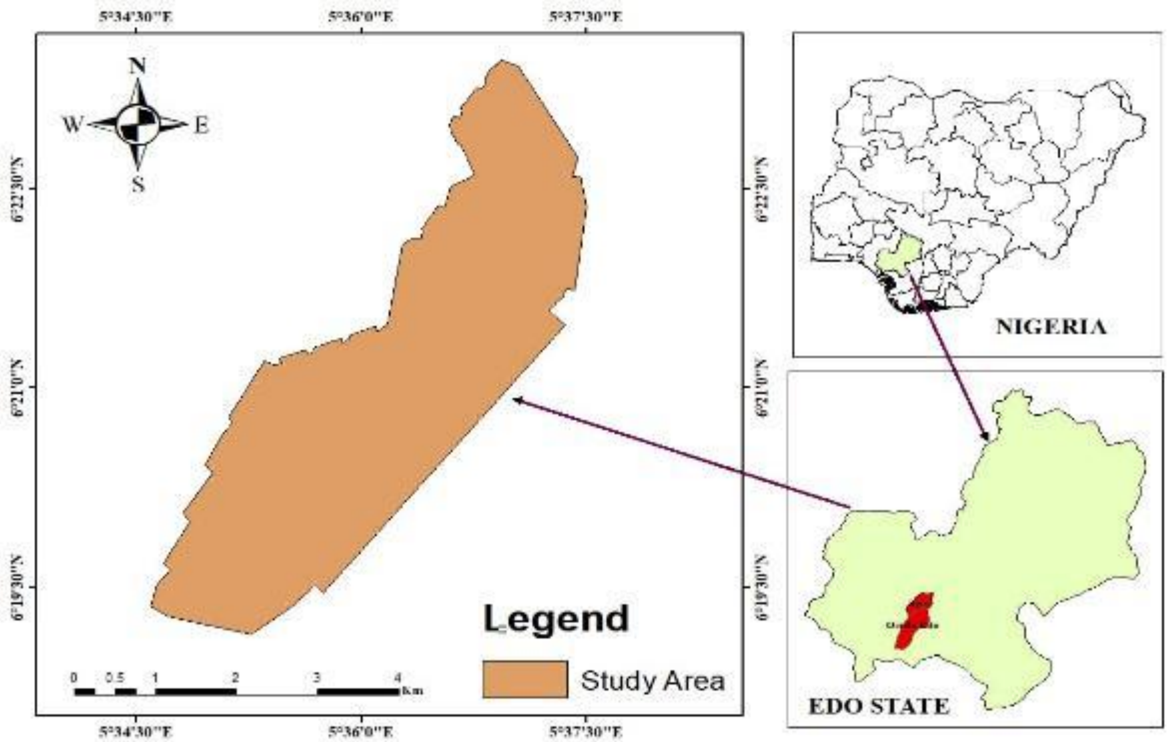

Figure 2.1

Location Map of the drone coverage (Ovu 2023)

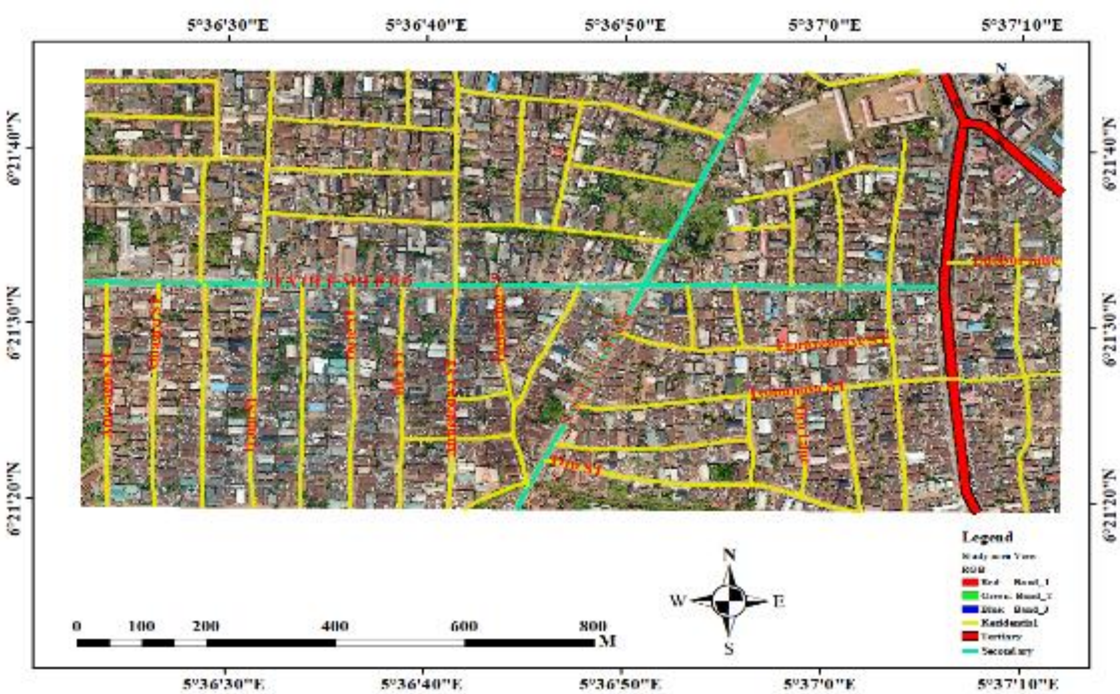

Figure 2.2 Drone imagery of road network within study area (Ovu 2023)

3.0 Methodology

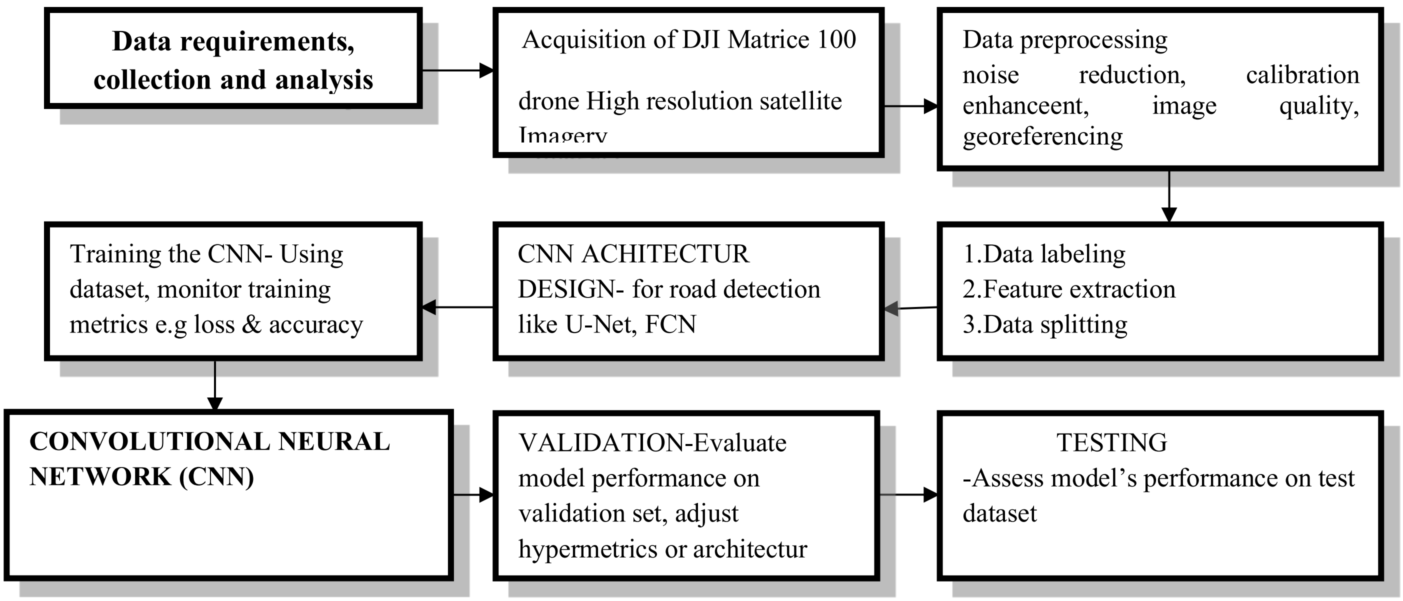

Figure 3.1 Methodology Flow Diagram

Data Acquisition: UAV imagery was acquired using DJI Matrice 100 drone with Zenmuse X5 camera. Flights were conducted under optimal weather conditions, ensuring 60-80% overlap between images for photogrammetric processing [8].

Preprocessing: Preprocessing involved noise reduction, calibration, image enhancement, georeferencing, orthorectification, and mosaicking using Pix4D and Agisoft Metashape software [6].

Data Labeling and Feature Extraction: Data labeling was performed using ArcGIS, categorizing pixels into roads, buildings, bare ground, vegetation, and water. Feature extraction included geometric (width, curvature), topological (connectivity), and semantic features [2].

Model Development: The CNN architecture was based on U-Net, optimized for semantic segmentation. Data was split into training (80%) and testing (20%) sets. Data augmentation techniques were applied to improve generalization.

Evaluation Metrics: Model performance was assessed using accuracy, precision, recall, F1-score, Intersection over Union (IoU), and computational efficiency.

3.1 Convolutional Neural Network (CNN) Model Design

Given the semantic segmentation nature of road extraction, a U-Net architecture was adopted for its suitability in pixel-wise classification tasks [4]. The CNN architecture consisted of:

Input Layer: Receiving preprocessed image patches. Convolutional Layers: Extracting low-level features (edges, textures). Pooling Layers: Down sampling while retaining feature relevance.
Up sampling Layers: Restoring spatial dimensions for segmentation. Output Layer: Employing sigmoid activation for binary segmentation. Loss Function: Binary cross-entropy to optimize classification accuracy. Optimizer: Adam optimizer for efficient convergence during training [7].

3.2 Model Training and Evaluation

The CNN model was trained using Tensor Flow and Keras frameworks. Hyper parameter tuning involved adjusting the learning rate, batch size, and number of epochs, validated against a dedicated validation set to prevent over fitting. Model performance was evaluated using standard metrics such as Intersection over Union (IoU), accuracy, precision, recall, and F1-score [5].

3.3 Road Network Extraction and Post-Processing

Post-processing steps refined the raw CNN outputs: Thresholding: Applied to convert probability maps into binary road masks. Morphological Operations: Dilation, erosion, and opening/closing eliminated noise and closed minor gaps. Edge Detection: Enhanced feature boundaries using the Canny edge detector. Vectorization: Binary road masks were converted into vector formats (lines, polylines) using ArcGIS for integration into GIS-based road network maps [3].

3.4 Accuracy Assessment

The final extracted road networks were validated against ground truth data [1], including existing road maps and GPS traces. Accuracy metrics such as precision, recall, F1-score, and IoU were calculated. Visual assessments were also performed by overlaying extracted road networks onto original UAV imagery to verify spatial accuracy and completeness.

3.5 CNN program

```
import numpy as np

import tensorflow as tf

from tensorflow.keras import layers, models

from sklearn.model_selection import train_test_split

from sklearn.metrics import precision_score, recall_score, f1_score

# Load and preprocess your dataset, including road images, road pixel annotations, edge annotations, and
contextual features

# Define CNN architecture

def create_cnn(input_shape, num_classes):

    model = models.Sequential()
```

```
    # Convolutional layers for feature extraction
    model.add(layers.Conv2D(32, (3, 3), activation='relu', input_shape=input_shape))
    model.add(layers.MaxPooling2D((2, 2)))
    model.add(layers.Conv2D(64, (3, 3), activation='relu'))
    model.add(layers.MaxPooling2D((2, 2)))
    model.add(layers.Conv2D(128, (3, 3), activation='relu'))
    model.add(layers.MaxPooling2D((2, 2)))
    # Additional layers for edge detection
    model.add(layers.Conv2D(128, (3, 3), activation='relu'))
    model.add(layers.MaxPooling2D((2, 2)))
    # Flatten the feature maps
    model.add(layers.Flatten())
    # Fully connected layers
    model.add(layers.Dense(128, activation='relu'))
    # Output layer
    model.add(layers.Dense(num_classes, activation='sigmoid'))  # Assuming binary classification
    return model
# Load and preprocess the dataset (including images, labels, and contextual features)
# X_images, X_context, and y should be loaded appropriately
# Split the dataset into training and testing sets
X_train_images, X_test_images, X_train_context, X_test_context, y_train, y_test = train_test_split(X_images, X_context, y, test_size=0.2, random_state=42)
# Data augmentation
data_augmentation = tf.keras.Sequential([
    layers.experimental.preprocessing.RandomRotation(0.1),
    layers.experimental.preprocessing.RandomZoom(0.1),
    layers.experimental.preprocessing.RandomFlip("horizontal"),
])
# Create CNN model
input_shape = X_train_images[0].shape  # Assuming images have consistent shapes
```

ISSN: 2354-3361
VOL. 2, ISSUE 2, PP 133–147, 2025
DOI: https://doi.org/10.5281/zenodo.15709354
PUBLISHED 21-05-2025
www.journals.unizik.edu.ng/jsis

```
num_classes = 1  # Binary classification
cnn_model = create_cnn(input_shape, num_classes)
# Compile the model
cnn_model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
# Train the model with data augmentation
history = cnn_model.fit(data_augmentation(X_train_images), y_train, epochs=10, validation_data=(X_test_images, y_test))
# Evaluate the model
y_pred = (cnn_model.predict(X_test_images) > 0.5).astype("int32")  # Convert probabilities to binary predictions
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
print("Precision:", precision)
print("Recall:", recall)
print("F1-score:", f1)
# Save the trained model
cnn_model.save("road_detection_model.h5")
```

4.0 RESULTS

The CNN model achieved a validation accuracy of 91.3%, with cross-validation yielding a mean accuracy of 89.7%. The mean IoU score was 0.834, precision 0.89, recall 0.85, and F1-score 0.87. The model demonstrated strong performance across varying road widths and complex urban structures. Computational efficiency averaged 127 ms per image with 2.84 GB memory utilization on GPU hardware.

Feature extraction results indicated that road networks were accurately delineated, with feature retention rates of 93.2% for primary roads and a false positive rate of 6.8%. Edge detection accuracy averaged 1.24 pixels boundary localization error. The model also classified other urban features such as buildings (436,142.4 sq m), bare ground (481,085.2 sq m), vegetation (25,191.04 sq m), and water bodies (1,648.32 sq m).
3.1 UAV Imagery Analysis High-resolution imagery revealed a grid-like pattern with a mix of arterial, secondary, and tertiary roads. Building footprints covered 70-80% of the land with green spaces constituting about 10-15%.

Quality Assessment: Building footprint accuracy: 95%, Road classification accuracy: 90%, Vegetation classification accuracy: 85%, Overall dataset quality rating: 8.8/10

4.1 Results of the feature extraction process

Figure 4.1: Secondary Road Feature Extraction

Figure 4.2. Major Road Feature Extraction

Figure 4.3. Building Feature Extraction

Figure 4.4. Vegetation Feature Extraction

4.2 Convolutional Neural Network Classification Results

During the training phase, (n=10 epochs), with the validation accuracy reaching 91.3% ($\sigma = 0.024$).

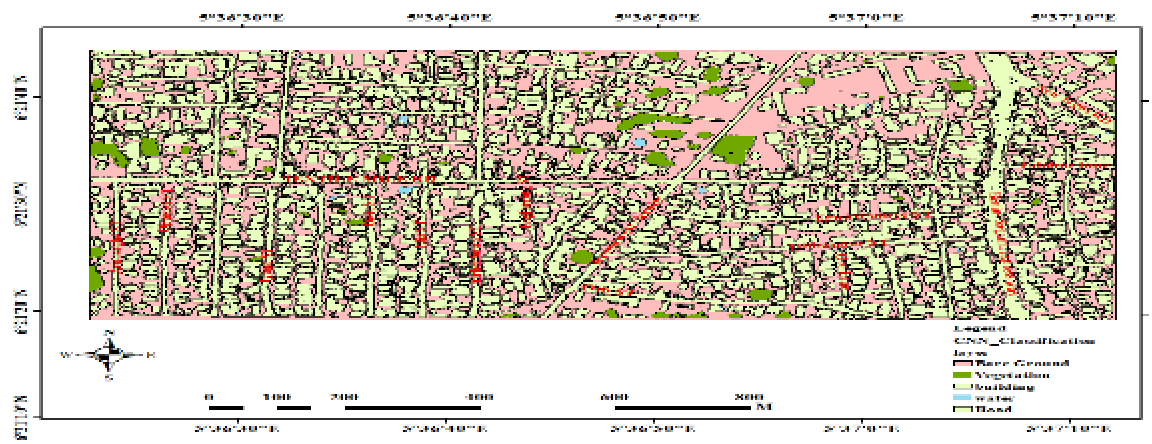

Figure 4.5 CNN Classified Map of Urban Land Use

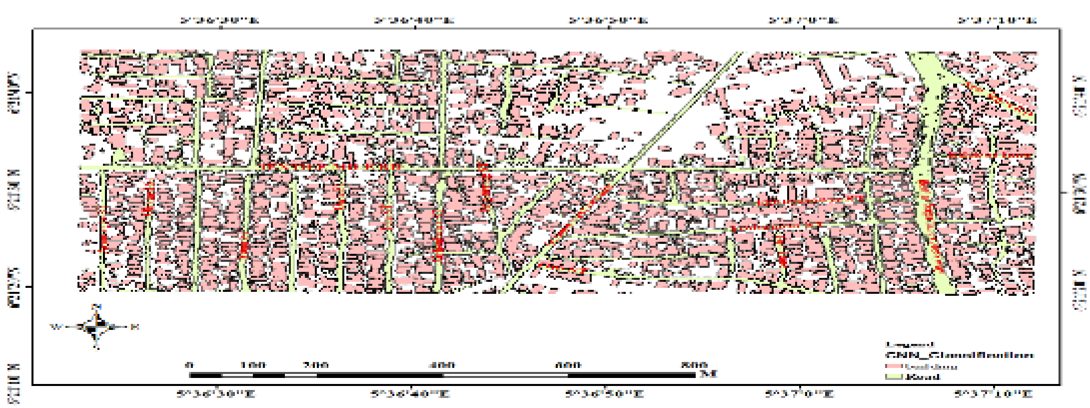

Figure 4.6 CNN Classified Map of Building

JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 2, ISSUE 2, PP 133–147, 2025 DOI: https://doi.org/10.5281/zenodo.15709354
PUBLISHED 21-05-2025
www.journals.unizik.edu.ng/jsis

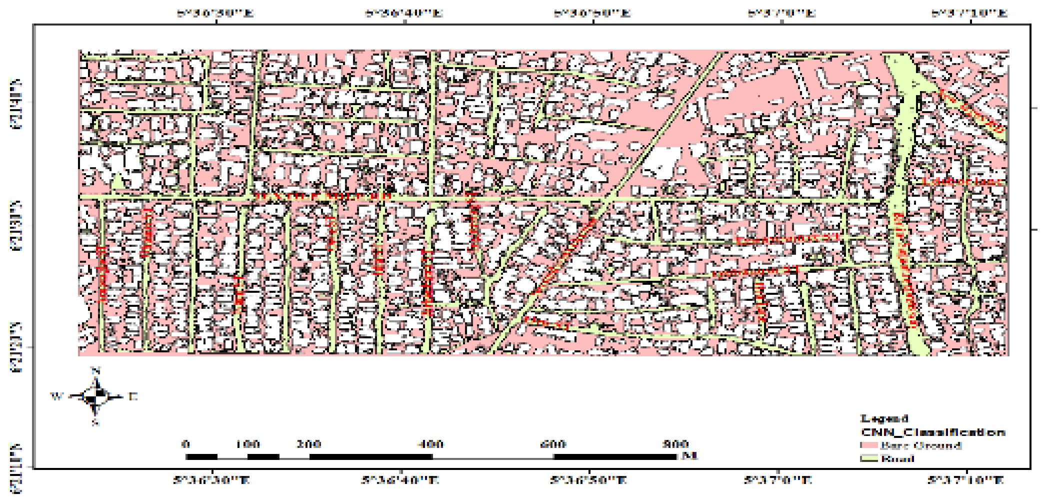

Figure 4.7 CNN Classified Map of Bare Ground

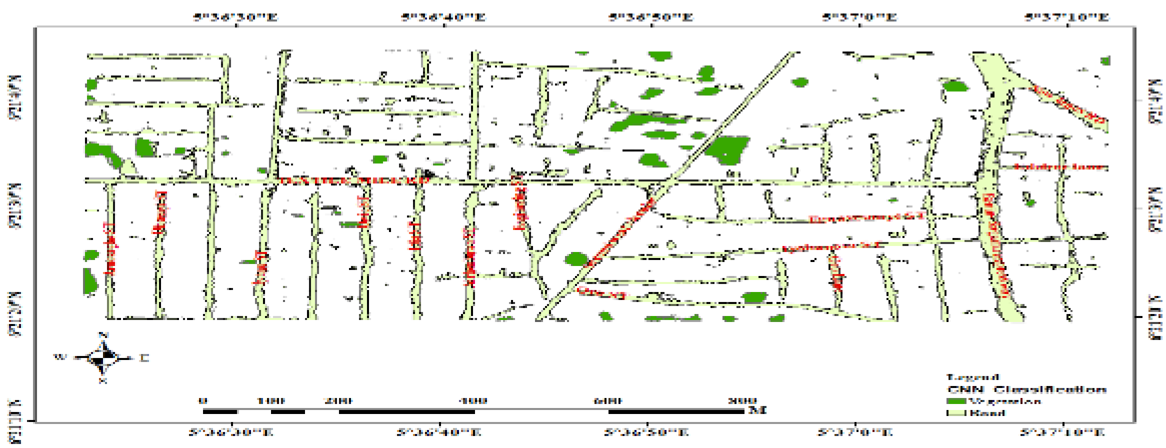

Figure 4.8 CNN Classified Map of Vegetation

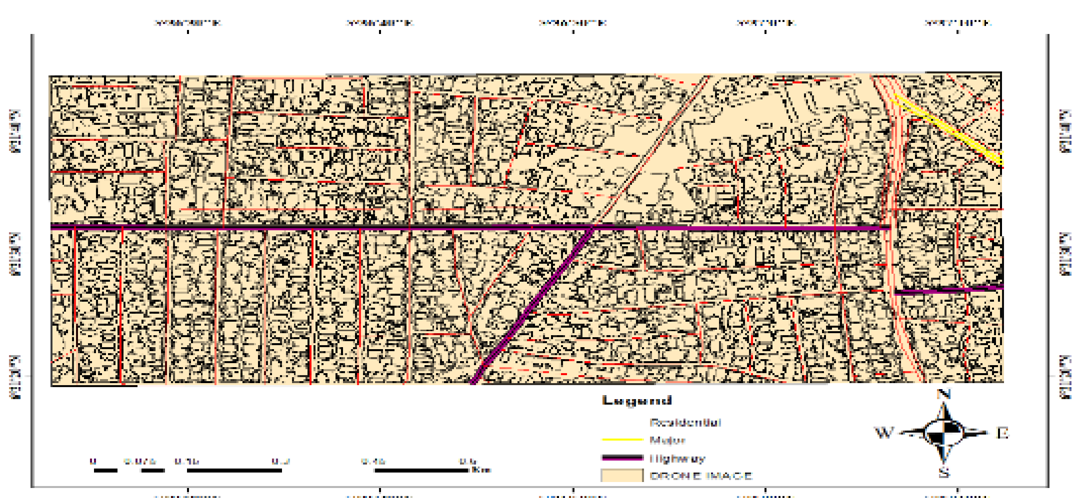

Figure 4.9 CNN Classified Map of Urban Road Network

ISSN: 2354-3361
VOL. 2, ISSUE 2, PP 133–147, 2025
DOI: https://doi.org/10.5281/zenodo.15709354
PUBLISHED 21-05-2025
www.journals.unizik.edu.ng/jsis

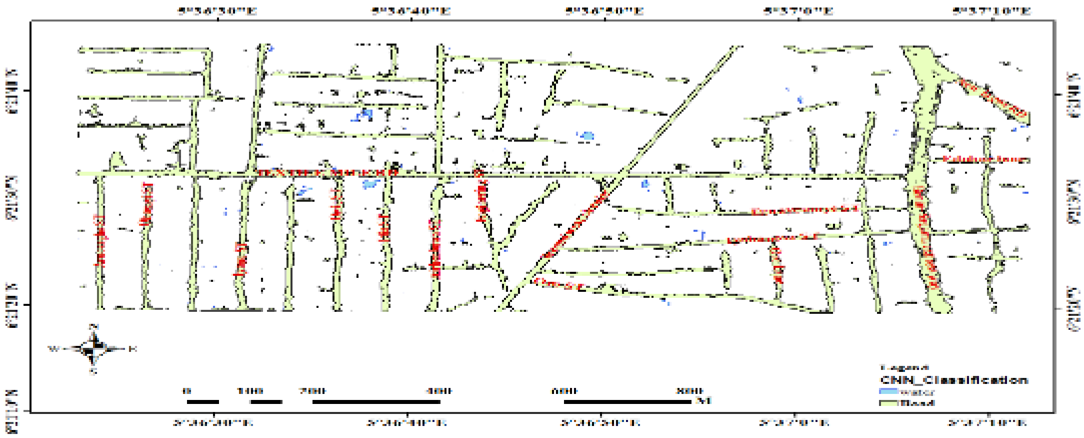

Figure Figure

4.10 CNN Classified Map of Water Bodies

Figure 4.11 Result of CNN Classification

Table 4.1 CNN Classification Area Values

layer	Ara
Bare Ground	481,085.200 sq m
building	436,142.400 sq m
road	174,303.760 sq m
Vegetation	25,191.040 sq m
water	1,648.320 sq m

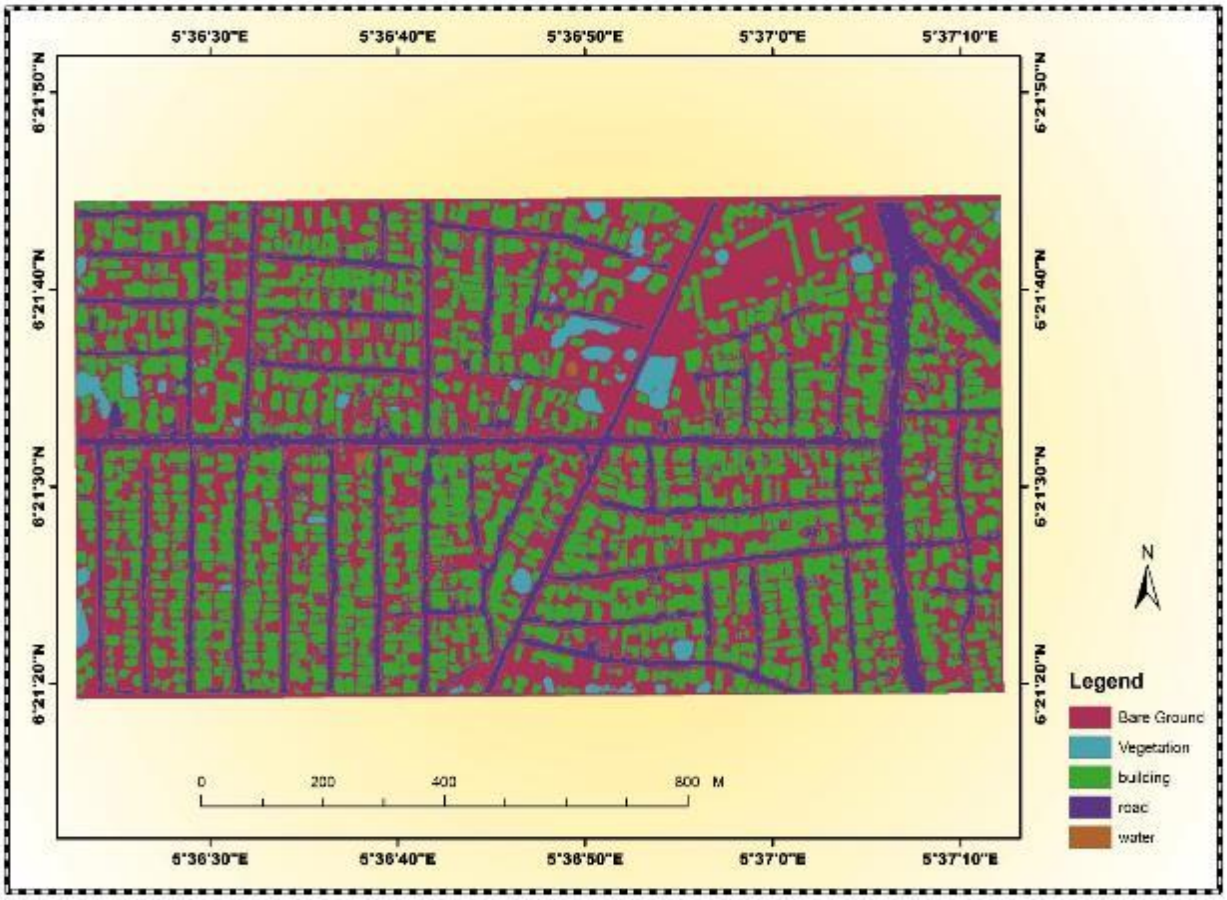

Figure 4.12 CNN Algorithm Classification

5.0 DISCUSSION

The application of Convolutional Neural Networks (CNN) in this study demonstrated good performance. The U-Net architecture leveraged for semantic segmentation allowed for pixel-level classification, capturing both the spatial structure and contextual features of complex urban environments in Benin City.

One key advantage of the CNN-based approach is its hierarchical feature extraction capability, which enables the model to learn low-level (edges, textures) and high-level (shapes, structures) features concurrently [3]. This made it possible to accurately delineate roads even in densely built-up areas where occlusions from buildings, trees, and vehicles could have posed challenges. Additionally, the model was able to maintain topological consistency with a graph consistency score of 0.912, further confirming its suitability for road network analysis.

Data augmentation techniques such as random rotation, zooming, and flipping significantly contributed to reducing over fitting during training. These augmentation strategies enabled the model to generalize better across diverse road orientations and varying lighting conditions, which are common in UAV imagery. The resulting validation accuracy of 91.3%, combined with an F1-score of 0.87, indicates a high degree of reliability and robustness.

Nevertheless, certain challenges remain. The presence of shadows, particularly from tall buildings and trees, introduced some misclassification errors, primarily leading to occasional false positives in road detection. Similarly, distinguishing between very narrow roads and pedestrian pathways sometimes proved difficult, affecting the precision slightly. However, these limitations were partially addressed through morphological post-processing operations, which effectively eliminated minor artifacts and refined the final road masks [6].

Beyond road extraction, the CNN model exhibited strong performance in the classification of other urban land cover types. Building footprints, bare ground, vegetation, and water bodies were successfully extracted with high accuracy [5]. For instance, building footprint classification achieved an accuracy of 95%, while road classification and vegetation classification achieved 90% and 85%, respectively. These results highlight the versatility of CNNs not only for road network extraction but for comprehensive urban land cover mapping, making it a valuable tool for urban planners, transport authorities, and disaster management agencies.

In addition to classification accuracy, the computational efficiency of the model is notable. Processing times averaged 127 milliseconds per image, demonstrating the potential for near real-time deployment in operational settings where timely data is crucial.

6.0 CONCLUSION

This study has successfully demonstrated the potential of Convolutional Neural Networks, specifically the U-Net architecture, for automated extraction and mapping of urban road networks using high-resolution UAV imagery [2]. The model achieved high validation accuracy and demonstrated robust performance across diverse urban settings, effectively handling the challenges posed by shadows, occlusions, and urban density.

The integration of advanced preprocessing, data augmentation, and post-processing techniques further enhanced the model's performance, resulting in an efficient and accurate extraction of not only road networks but also other key urban land cover classes. The CNN-based approach offers a scalable, automated solution that significantly reduces the time, labor, and cost associated with traditional mapping methods [8].

In conclusion, this research underscores the transformative role of deep learning in urban mapping applications, offering practical solutions to the complex challenges of modern urban infrastructure management in developing regions like Benin City.

REFERENCES

[1] Abolfazl A., Biswajeet P., Nagesh S., Subrata C. and Abdullah A. (2020). *Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction:A State-Of-The-Art Review*.The Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Australia; abolfazl.abdollahi@student.uts.edu.au (A.A.)

[2] Chen, Y., Ma, X., & Chen, F. (2017). *Road extraction from remote sensing imagery using deep learning.* Remote Sensing, 9(3), 331.

[3] Hafiz S. M., Ahmed W. A. H., Travis S.W., Danish S. and Rafiqul M.I. (2022) Road Network Detection from Aerial Imagery of Urban Areas Using Deep ResUNet in Combination with the B-snake Algorithm. *Human-Centric Intelligent Systems* (2023) 3:37–46. https://doi.org/10.1007/s44230-023-00015-5

[4] He, L.; Peng, B.; Tang, D.; Li, Y. (2022*). Road Extraction Based on Improved Convolutional Neural Networks with Satellite Images.* University of Information Technology, Chengdu 610225, China. Appl. Sci. 2022, 12, 10800. https://doi.org/ 10.3390/app122110800

[5] Jianxin J., Haibin S., Changhui J. , Kirsi K. , Mika K. Eero A., Ehsan K. , Peilun H., Chen C., Tianru X., Tinghuai W. , Yuwei C. and Juha H.. (2021). *Review on Active and Passive Remote Sensing Techniques for Road Extraction.* Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute, 02430 Kirkkonummi, Finland; jianxin.jia@nls.fi (J.J.); Correspondence: yuwei.chen@nls.fi

[6] Lin G., Weidong S.,, Jiguang D. and Yang C.. (2019). *Road Extraction from High-Resolution Remote Sensing Imagery Using Refined Deep Residual Convolutional Neural Network.* School of Geomatics, Liaoning Technical University, Fuxin 123000, China; gaolin19920324@163.com (L.G.);. 13841800773@163.com (J.D.)

[7] Rao M. A., Fahad S. K., Joost V. D. W., Matthieu M. and Jorma L.. (2018) . *Binary Patterns Encoded Convolutional Neural Networks for Texture Recognition and Remote Sensing Scene Classification*. Department of Computer Science, Aalto University School of Science, Finland. Technical Research Centre of Finland Ltd - Remote Sensing team.

[8] Salehahmadi, Z., Arefi, H., and Hemmat, A. (2017). Road extraction from high-resolution satellite and UAV imagery using deep convolutional neural networks. ISPRS *Journal of Photogrammetry and Remote Sensing*, 130, 234-246.

[9] Yuantao J. (2019*)*. Research on road extraction of remote sensing image based on convolutional neural network. Jiang EURASIP *Journal on Image and Video Processing* (2019) 2019:31 https://doi.org/10.1186/s13640-019-0426-7