HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

**School of Information and Communications Technology**



**PROJECT REPORT**

**A Pixel-Based Approach for Aerial Image Segmentation of Sentinel-2 Imagery**

|  |  |  |
| --- | --- | --- |
| **Supervisor:** | PhD Tran Nguyen Ngoc |  |
| **Student:** | Vu Duc Thang – 20225553  Thang.vd225553@sis.hust.edu.vn | |
| **Specialization:** | Cyber Security K67 | |
| **HANOI, 6/2025** | | |

**ACKNOWLEDGMENT**

I would like to express my deepest gratitude to my supervisor, PhD Tran Nguyen Ngoc, for his valuable guidance, encouragement, and constructive feedback throughout the duration of this project. His insights and expertise have been instrumental in shaping the direction and quality of this work.

I am also thankful to the School of Information and Communications Technology, Hanoi University of Science and Technology, for providing the resources and academic environment necessary for conducting this research.

Special thanks go to my friends and classmates for their support, discussions, and suggestions during the project. Lastly, I would like to acknowledge my family for their unconditional support and motivation throughout my academic journey.

**ABSTRACT**

Rapid urbanization in Hanoi has led to significant challenges in land use management and urban planning. This study explores a pixel-based classification approach for land cover mapping using Sentinel-2 imagery, focusing on central districts of Hanoi. Three classification models—Support Vector Machine (SVM), Random Forest (RF), and a custom-built 1D Convolutional Neural Network (CNN)—were implemented and evaluated.

The methodology involved selecting key spectral bands and incorporating the Normalized Difference Vegetation Index (NDVI) to enhance class separability. Training data were manually digitized and split into training and testing sets. Model performance was evaluated using accuracy metrics and confusion matrices.

Results show that the CNN model achieved the highest overall accuracy (95.26%), outperforming SVM (89.84%) and RF (91.75%), particularly in distinguishing spectrally similar urban land cover classes. The findings demonstrate the effectiveness of deep learning in urban remote sensing and highlight its potential for supporting intelligent land use monitoring in fast-developing cities.

**TABLE OF CONTENTS**

[CHAPTER 1. INTRODUCTION 1](#_Toc202708381)

[1.1 Problem statement 1](#_Toc202708382)

[CHAPTER 2. LITERATURE REVIEW 2](#_Toc202708383)

[2.1 Remote sensing 2](#_Toc202708384)

[2.2 Sentinel – 2 2](#_Toc202708385)

[2.3 QGIS 3](#_Toc202708386)

[2.4 Python and libraries used 4](#_Toc202708387)

[CHAPTER 3. DATA AND METHODOLOGY 6](#_Toc202708388)

[3.1 Data used 6](#_Toc202708389)

[3.2 Study area (Area of Interest) 6](#_Toc202708390)

[3.3 Training Dataset 7](#_Toc202708391)

[3.4 Classification Methodologies 8](#_Toc202708392)

[3.4.1 SVM – Support Vector Machine 8](#_Toc202708393)

[3.4.2 Random Forest 9](#_Toc202708394)

[3.4.3 Convolution Neural Network 9](#_Toc202708395)

[3.5 Results 10](#_Toc202708396)

[3.5.1 SVM – Support Vector Machine 12](#_Toc202708397)

[3.5.2 Random Forest 14](#_Toc202708398)

[3.5.3 Convolution Neural Network 15](#_Toc202708399)

[3.6 Evaluation and Model Comparison 17](#_Toc202708400)

[3.6.1 Overall Accuracy and Class Performance 17](#_Toc202708401)

[3.6.2 Model Performance Comparison 18](#_Toc202708402)

[CHAPTER 4. CONCLUSION 19](#_Toc202708403)

[4.1 Conclusion and Future work sugguest 19](#_Toc202708404)

[4.2 Application of Segmentation Results 19](#_Toc202708405)

[REFERENCES 21](#_Toc202708406)

LIST OF FIGURES

[Figure 2.1. Sensor on-board satellite scans along line AB. As the platform moves forward, an image of the swath region is built up. 2](#_Toc202708410)

[Figure 2.2. Sentnel-2 SAFE product content 3](#_Toc202708411)

[Figure 3.1. Training shapefile overlay 7](#_Toc202708412)

[Figure 3.2. Reference Model 9](#_Toc202708413)

[Figure 3.3. My Model's layer Architecture 10](#_Toc202708414)

[Figure 3.4. Ground truth 11](#_Toc202708415)

[Figure 3.5. Confusion Matrix of SVM model 12](#_Toc202708416)

[Figure 3.6. SVM Classification result 13](#_Toc202708417)

[Figure 3.7. Confusion Matrix of Random Forest Model 14](#_Toc202708418)

[Figure 3.8. Random Forest Clasification Result 15](#_Toc202708419)

[Figure 3.9. Confusion Matrix of CNN 16](#_Toc202708420)

[Figure 3.10. Training Accuracy 16](#_Toc202708421)

[Figure 3.11. CNN Classification result 17](#_Toc202708422)

LIST OF TABLES

[Table 1. Spectral bands of the Sentinel-2 A satellite imagery 6](#_Toc202708449)

[Table 2. Training and testing sample sizes used in this study. 8](#_Toc202708450)

[Table 3. F1-score comparison 18](#_Toc202708451)

[Table 4. Performance Comparison 18](#_Toc202708452)

# INTRODUCTION

## Problem statement

Hanoi is one of the fastest urbanizing cities in Vietnam, especially in the inner-city areas. Rapid development poses challenges for land management and urban planning. Uncontrolled expansion may lead to land use conflicts, loss of green spaces, and urban structure imbalance.

By applying pixel-based classification techniques on Sentinel-2 imagery using machine learning algorithms such as Random Forest, SVM, and CNN, it is possible to classify and monitor major land cover types in inner-city Hanoi, including:

* High-density residential zones
* Urban green spaces (parks, tree-lined roads)
* Water bodies (e.g., West Lake, To Lich River)
* Bare land or parking lots

Practical Benefits:

* Generate detailed land use maps to support Hanoi's Department of Planning and Architecture.
* Monitor annual land use changes to detect unauthorized land conversions.
* Provide a foundational geospatial database for intelligent land management systems.

# LITERATURE REVIEW

## Remote sensing

Remote sensing is the science of obtaining information about objects or areas from a distance, typically from aircraft or satellites. Remote sensors collect data by detecting the energy that is reflected from Earth. These sensors can be on satellites or mounted on aircraft.

Remote sensing can be divided into two types of methods: Passive remote sensing and Active remote sensing. Passive sensors gather radiation that is emitted or reflected by the object or surrounding areas. Reflected sunlight is the most common source of radiation measured by passive sensors.

Remote sensing of earth’s environment comprises measuring and recording of electromagnetic energy reflected from or emitted by the planet’s surface and atmosphere from a-vantage point above the surface, and relating of such measurements to the nature and distribution of surface materials and atmospheric conditions. Sensors mounted on aircraft or satellite platforms measure the amounts of energy reflected from or emitted by the earth’s surface. These measurements are made at a large number of points distributed either along a one-dimensional profile on the ground below the platform or over a two-dimensional area on either side of the ground track of the platform. The sensors scan the ground below the satellite or aircraft platform and as the platform moves forward, an image of the earth’s surface is built up. Fig. 2.1 shows how a sensor on board satellite scans along line AB. Each scan line of a remotely sensed image is a digital or numerical record of radiance measurements made at regular intervals along the line. A set of consecutive scan lines forms an image (Mather, 1987). Two-dimensional image data can be collected by means of two types of imaging sensors, namely, nadir looking or side looking sensor.A diagram of a sensor

AI-generated content may be incorrect.Sentinel – 2

Figure .. Sensor on-board satellite scans along line AB. As the platform moves forward, an image of the swath region is built up.

Sentinel-2 is a European wide-swath, high-resolution, multi-spectral imaging mission. The full mission specification of the twin satellites flying in the same orbit but phased at 180°, is designed to give a high revisit frequency of 5 days at the Equator.

Each of the satellites in the Sentinel-2 mission carries a single payload: the optical Multi-Spectral Instrument (MSI) that samples 13 spectral bands: four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution. The orbital swath width is 290 km.

**Data Formats**

Sentinel-2 Level-1C and Level-2A User products are made available in Sentinel-SAFE format, including image data in JPEG2000 format, quality indicators (e.g. defective pixels mask), auxiliary data and metadata.

The SAFE format has been designed to act as a common format for archiving and conveying data within ESA Earth Observation archiving facilities.

A diagram of a product

AI-generated content may be incorrect.

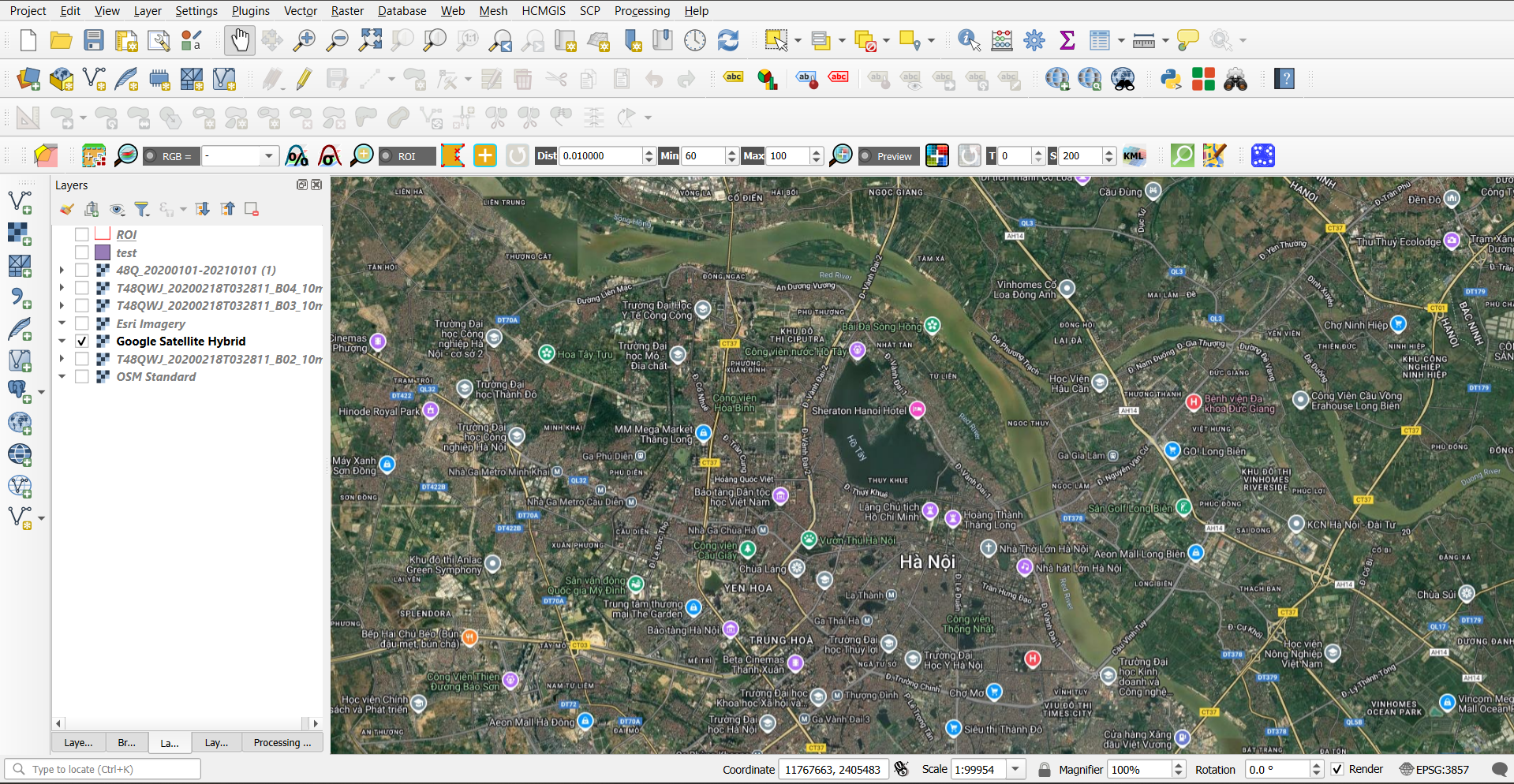
Figure .. Sentnel-2 SAFE product content

## QGIS

QGIS is a free and open-source cross-platform desktop geographic information

system (GIS) application that supports viewing, editing, and analysis of geospatial

data.



## Python and libraries used

Python is a high-level, interpreted programming language known for its simplicity, readability, and versatility. It's one of the most popular languages in the world and is widely used in fields such as web development, automation, data analysis, and especially machine learning.

It has become the most popular language for machine learning due to several key advantages. First, its clean and intuitive syntax allows developers and researchers to focus more on problem-solving and algorithm development rather than language complexity. Python offers a rich ecosystem of libraries that are specifically designed for machine learning and data science, such as Scikit-learn for traditional ML algorithms, TensorFlow and Keras for deep learning, NumPy and Pandas for numerical computing and data manipulation, and Matplotlib and Seaborn for data visualization. These libraries make it easier to process large datasets, build complex models, and visualize results efficiently. Additionally, Python has strong community support, meaning there are many resources, tutorials, and open-source contributions available for learners and professionals alike. Its ability to integrate well with other languages and systems also makes it a powerful tool for deploying machine learning solutions in real-world applications. As a result, Python is considered the standard language for machine learning in both academia and industry.

|  |  |
| --- | --- |
| **Library** | **Description** |
| **NumPy** | A library for high-performance numerical computing, especially with multi-dimensional arrays. |
| **Matplotlib.pyplot** | A plotting module for creating 2D visualizations like line charts, histograms, etc. |
| **Scikit-learn** | A popular machine learning library offering tools for classification, regression, and clustering. |
| **GDAL** | A geospatial data library for reading, writing, and transforming raster and vector data formats. |
| **Rasterio** | A user-friendly library for reading, writing, and processing raster (e.g., satellite) imagery. |
| **Pandas** | A data analysis library that provides data structures like DataFrame for handling tabular data. |
| **TensorFlow** | An open-source library by Google for building and training deep learning models. |
| **Keras** | A high-level API for TensorFlow that simplifies building and training neural networks. |

# DATA AND METHODOLOGY

## Data used

Sentinel-2, level 2A (ID = L2A\_T48QWJ\_A024325\_20200218T033333) acquired on 18 Febuarary 2020, was obtained from the Copernicus Open Access Hub. Spectral bands of the Sentinel-2A satellite imagery are shown in Table 1.

Table . Spectral bands of the Sentinel-2 A satellite imagery

|  |  |  |  |
| --- | --- | --- | --- |
| **Spectral Band** | **Resolution** | **Center Wavelength (nm)** | **Band Width** |
| Band 1 | 60 | 443 | 20 |
| **Band 2** | **10** | **490** | **65** |
| **Band 3** | **10** | **560** | **35** |
| **Band 4** | **10** | **665** | **30** |
| Band 5 | 20 | 705 | 15 |
| Band 6 | 20 | 740 | 15 |
| Band 7 | 20 | 783 | 20 |
| **Band 8** | **10** | **842** | **115** |
| Band 8A | 20 | 865 | 20 |
| Band 9 | 60 | 940 | 20 |
| Band 10 | 60 | 1375 | 30 |
| **Band 11** | **20** | **1610** | **90** |
| **Band 12** | **20** | **2190** | **180** |

To ensure high-quality classification while balancing spectral relevance and spatial resolution, a subset of spectral bands from the Sentinel-2 MSI sensor was selected. Specifically, six bands were used in the analysis: Band 2 (Blue), Band 3 (Green), Band 4 (Red), Band 8 (Near Infrared), Band 11 (Short-Wave Infrared 1), and Band 12 (Short-Wave Infrared 2).

These bands were chosen due to their proven effectiveness in distinguishing key land cover types, especially in urban environments. Bands 2, 3, and 4 provide information from the visible spectrum and are essential for true color composites. Band 8 captures near-infrared reflectance, which is sensitive to vegetation health and biomass. Bands 11 and 12, in the shortwave infrared range, are particularly useful for detecting built-up areas, moisture content, and distinguishing impervious surfaces. All selected bands offer either 10 m or 20 m resolution, enabling detailed classification over urban districts while maintaining computational efficiency.

## Study area (Area of Interest)

In this study, a central urban region of Hanoi, Vietnam was selected to evaluate the performance of different classification algorithms using Sentinel-2 imagery. The study area includes six densely populated and diverse districts: Tay Ho, Ba Dinh, Dong Da, Cau Giay, Hoan Kiem, and Hai Ba Trung. These districts represent the urban core of Hanoi, characterized by a complex mixture of land cover types such as residential zones, commercial buildings, transportation infrastructure, parks, and small water bodies. The area also includes historical sites, government buildings, and high-rise developments, reflecting rapid urbanization. The heterogeneous landscape and mixed land use make this region a suitable testbed for evaluating classification algorithms under challenging urban conditions

## Training Dataset

In this study, the training dataset was created using manual digitization in QGIS over atmospherically corrected Sentinel-2 imagery. Based on visual interpretation and high-resolution basemaps, a shapefile was constructed to delineate homogeneous regions representing distinct land cover classes. The labeled classes include residential, impervious surface, vegetation, water, bare land, and others.

Polygons were carefully digitized to ensure spatial distribution across the six study districts: Tay Ho, Ba Dinh, Dong Da, Cau Giay, Hoan Kiem, and Hai Ba Trung, covering a variety of urban forms and land use patterns.

After digitization, the shapefile was rasterized to align with the spatial resolution of the Sentinel-2 imagery. A total of 18,681 labeled pixels were extracted from the shapefile. The dataset was then randomly split into 80% training (14,944 pixels) and 20% testing (3,737 pixels) sets. This ensured that the training and validation data were spatially independent while maintaining class diversity.

A black background with different colored labels

AI-generated content may be incorrect.

Figure .. Training shapefile overlay

The spatial distribution of the labeled training areas is shown in Figure 3.1, and a summary of the number of pixels per class is provided in Table 2. These training and testing datasets were subsequently used to evaluate the performance of different classification algorithms.

Table . Training and testing sample sizes used in this study.

|  |  |  |
| --- | --- | --- |
| **Land Cover** | **Training (pixels)** | **Testing (pixels)** |
| Water | 4109 | 1028 |
| Built Area | 7661 | 1915 |
| Vegetation | 1691 | 423 |
| Others | 1483 | 371 |

During analysis, it became evident that the vegetation class was frequently confused with others class, where are bare land or construction sites, due to similar spectral reflectance in certain bands. To address this, the Normalized Difference Vegetation Index (NDVI) was added as an additional predictor.

The value range of the NDVI is -1 to 1. Negative values of NDVI (values approaching -1) correspond to water. Values close to zero (-0.1 to 0.1) generally correspond to barren areas of rock, sand, or snow. Low, positive values represent shrub and grassland (approximately 0.2 to 0.4), while high values indicate temperate and tropical rainforests.

The normalized difference vegetation index, abbreviated NDVI, is defined as

With Sentinel-2, the index look like this:

By incorporating NDVI alongside selected spectral bands (red and NIR from Sentinel-2), the models achieved improved distinction between ambiguous classes during the feature extraction step in QGIS. The inclusion of NDVI contributed to stronger separability, particularly between vegetation and bare/construction areas, enhancing overall classification accuracy.

## Classification Methodologies

### SVM – Support Vector Machine

In land cover classification studies, the Support Vector Machine (SVM) classifier was employed using the Radial Basis Function (RBF) kernel, which is known for its strong performance in remote sensing classification tasks – according to Knorn et al. and Shi and Yang. Therefore, I used the RBF kernel to implement the SVM algorithm. There are two key parameters were tuned to optimize the model:

* **C (Cost parameter):** Controls the trade-off between maximizing the margin and minimizing classification error. Higher values of C may lead to overfitting, while lower values allow more classification flexibility.
* **γ (Gamma parameter):** Influences the shape of the decision boundary. Lower γ values lead to smoother boundaries, while higher values make the boundary more complex and sensitive to training samples.

Following the study of Li et al. and pretested to our dataset, in this study, to find the optimal parameters for SVM:

* **C** values tested: 2⁻² to 2⁷
* γ values tested: 2⁻⁵ to 2⁴

### Random Forest

The Random Forest (RF) classifier was applied as one of the three machine learning methods for land cover classification using Sentinel-2 imagery. RF is a robust, ensemble-based method that constructs multiple decision trees and aggregates their predictions to improve classification accuracy and reduce overfitting.

Two key parameters were tuned in the RF model:

* **ntree:** The number of trees in the forest. Higher values generally improve stability and performance, up to a point where the accuracy plateaus.
* **mtry:** The number of randomly selected features considered at each split. This influences both model accuracy and computational efficiency.

In this study, Sentinel-2 imagery with 6 input bands was used. A range of values was tested to identify the best configuration:

* **ntree** values tested: 100, 200 and 500.
* **mtry** values tested: from 1 to 10.

### Convolution Neural Network

The design of the Convolutional Neural Network (CNN) model in this study was inspired by the work of Ali and Johnson (2022), who proposed a deep learning approach for land-use and land-cover (LULC) classification in semi-arid regions using Sentinel-2 imagery. Their research demonstrated the effectiveness of a patch-based 2D CNN model trained on medium-resolution remote sensing data, achieving high classification accuracy in complex, spectrally similar landscapes. In particular, they tested both 4-band and 10-band Sentinel-2 composites and found that the 4-band CNN model (using blue, green, red, and near-infrared bands) yielded better performance in terms of accuracy and computational efficiency.

A diagram of layers of layers

AI-generated content may be incorrect.

Figure .. Reference Model

In this project, I implemented a 1D Convolutional Neural Network (CNN) for land cover classification based on Sentinel-2 imagery. The design was inspired by the 2D CNN architecture proposed by Ali & Johnson (2022), which used a patch-based approach (e.g., 5×5×B) to extract both spatial and spectral features via 2D convolutions and max-pooling layers. However, my model takes a different approach by focusing on the spectral dimension only, using Conv1D layers, which better matched the structure of my training data.

Specifically, my input data was prepared as 1D spectral sequences, where each sample represents the spectral values of a single pixel across selected bands and indices (B2, B3, B4, B8, B11, B12, NDVI). This format is naturally suited for Conv1D, as it emphasizes spectral feature learning without relying on local spatial context. As a result, I reshaped the data into sequences and used Conv1D layers with kernel size 2 to capture transitions across neighboring spectral bands.

A colorful object with a pointy

AI-generated content may be incorrect.

Figure .. My Model's layer Architecture

The model architecture includes three convolutional blocks with decreasing filters (64, 32, 16), each followed by Batch Normalization and Dropout for regularization. Unlike the reference model, I did not use max-pooling, which is often optional in Conv1D settings depending on the structure and size of the input. The final classification is performed by two fully connected dense layers (16 and 128 units) and an output layer with softmax activation for multi-class prediction.

## Results

**Ground Truth Data:**

**A map of a city

AI-generated content may be incorrect.**

Figure .. Ground truth

To generate reliable ground truth data for supervised classification, the Semi-Automatic Classification Plugin (SCP) in QGIS was utilized. SCP is a widely used open-source tool that streamlines the process of collecting training and validation samples directly from satellite imagery. Using this plugin, Regions of Interest (ROIs) were manually digitized by visually interpreting Sentinel-2 imagery and high-resolution basemaps. Each ROI polygon was assigned to a specific land cover class, including residential, impervious surfaces, vegetation, water, bare land, and others.

SCP provided efficient tools for sample labeling, spectral signature analysis, and conversion of vector ROIs into raster training datasets. This workflow ensured consistency in sample quality and spatial distribution across the study area. The collected samples were subsequently split into 80% for training and 20% for testing, aligning with best practices for supervised learning. By leveraging SCP's semi-automated capabilities, the ground truth dataset was created with enhanced accuracy and reduced manual effort, supporting the effective training and evaluation of machine learning and deep learning classifiers.

### SVM – Support Vector Machine

A graph of numbers and a number

AI-generated content may be incorrect.The performance of the Support Vector Machine (SVM) classifier was evaluated using a confusion matrix generated from predictions within the ROI based on ground truth labels. As shown in Figure 3.2, the classifier achieved high accuracy across most land cover classes, with particularly strong results in the Built Area and Water classes.

Figure .. Confusion Matrix of SVM model

The Built Area class showed the highest correct classification, with 447,214 pixels correctly identified. However, there was some confusion between Built Area and Vegetation, with 31,789 pixels misclassified as vegetation, and around 10,000 pixels misclassified as Others. Similarly, the Vegetation class had 39,281 correctly classified pixels, but also exhibited notable confusion with Others (8,507 pixels) and Built Area (5,501 pixels).

The Water class was classified with very high accuracy, achieving over 100,000 correctly classified pixels, and only a small number (1,639 pixels) were incorrectly predicted as Built Area. The Others class was also reasonably well classified, with 52,871 correct predictions, although some misclassification occurred into Vegetation (9,043 pixels) and Built Area (5,198 pixels).

Overall, the SVM classifier demonstrated strong performance with high classification accuracy, particularly for well-separated spectral classes like Water and Built Area. Most errors were associated with confusion between Vegetation and Others, which is understandable due to the spectral similarity between vegetated areas, bare soil, and construction zones. This suggests that while SVM is robust, additional features such as vegetation indices or spatial context may further improve classification accuracy for spectrally mixed classes.

A map of a city

AI-generated content may be incorrect.**Result:**

Figure .. SVM Classification result

### Random Forest

A graph with numbers and symbols

AI-generated content may be incorrect.Similar to the SVM classifier, the Random Forest (RF) model achieved strong overall classification performance, particularly for the Built Area and Water classes. As shown in Figure 3.6, RF successfully classified 476,669 Built Area pixels and over 100,000 Water pixels correctly—comparable to SVM’s accuracy in these well-defined classes.

Figure .. Confusion Matrix of Random Forest Model

However, compared to SVM, Random Forest exhibited slightly more confusion in some class boundaries, especially between Vegetation and Built Area, and between Others and Built Area. For instance, RF misclassified 15,494 Vegetation pixels as Built Area, which is notably higher than SVM’s 5,501 misclassified in the same category. Similarly, 14,396 Others pixels were mislabeled as Built Area by RF, while SVM had a lower misclassification count in this case.

On the other hand, RF showed slightly better performance in minimizing misclassifications from Water to Others—with only 735 pixels confused, compared to 0 in SVM, but SVM showed a slightly better ability to isolate Water from Built Area (SVM: 1,639 pixels vs. RF: 2,719 pixels).

Despite these differences, both classifiers consistently demonstrated excellent performance for major land cover classes. Yet, SVM showed a more stable ability to separate spectrally similar categories, such as Vegetation and Others, which are often confused due to overlapping spectral characteristics in urban environments. Meanwhile, RF provided competitive results with faster training and better interpretability, making it a strong alternative where computational efficiency or model explainability is prioritized.

**Result:**

A map of a city

AI-generated content may be incorrect.

Figure .. Random Forest Clasification Result

### Convolution Neural Network

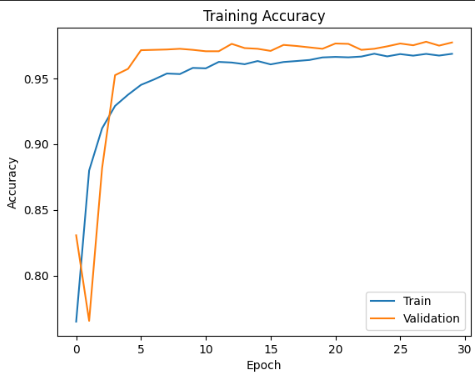
The performance of the proposed 1D CNN model was evaluated using an external ground truth dataset, and the results are summarized in the classification report and confusion matrix shown in Figure 3.8. The CNN achieved an overall accuracy of 95%, with particularly high performance in the Water (F1-score: 0.99) and Built Area (F1-score: 0.98) classes. The Vegetation class had a slightly lower F1-score of 0.76, and the Others class followed with 0.84. The model achieved a macro average F1-score of 0.89, indicating strong performance across all classes despite class imbalance and spectral confusion.

**A screenshot of a graph

AI-generated content may be incorrect.**

Figure .. Confusion Matrix of CNN

The confusion matrix confirms that most misclassifications occurred between Vegetation and Built Area, and between Others and Vegetation, which is consistent with what was observed in the SVM and RF models. However, compared to the previous models, the CNN showed a slight improvement in generalization for the Others class, with 56,773 correctly classified pixels, compared to 52,871 in SVM and 47,168 in RF. Meanwhile, Vegetation classification remained challenging, as expected, though CNN still outperformed RF in this category (F1-score: 0.76 vs. 0.72).

The training accuracy plot (Figure 3.9) demonstrates that the model converged quickly, reaching over 95% validation accuracy within the first 100 epochs and continuing to improve slightly through 30 epochs. The minimal gap between training and validation curves indicates good generalization and low risk of overfitting.

At the very beginning of training, the model weights are randomly initialized. As the model starts learning from the training data in the first few epochs, may not yet generalize well to unseen validation data.

Figure .. Training Accuracy

During epoch 1 or 2, the model might start overfitting very slightly or making unstable predictions as it adjusts its weights for the first time.

In summary, the CNN model demonstrated comparable or better performance than SVM and RF in most categories, with superior generalization capability. Its ability to learn spectral patterns from the input sequences enabled it to reduce confusion in spectrally overlapping classes, supporting its effectiveness in pixel-based land cover classification using Sentinel-2 imagery.

**Result:**

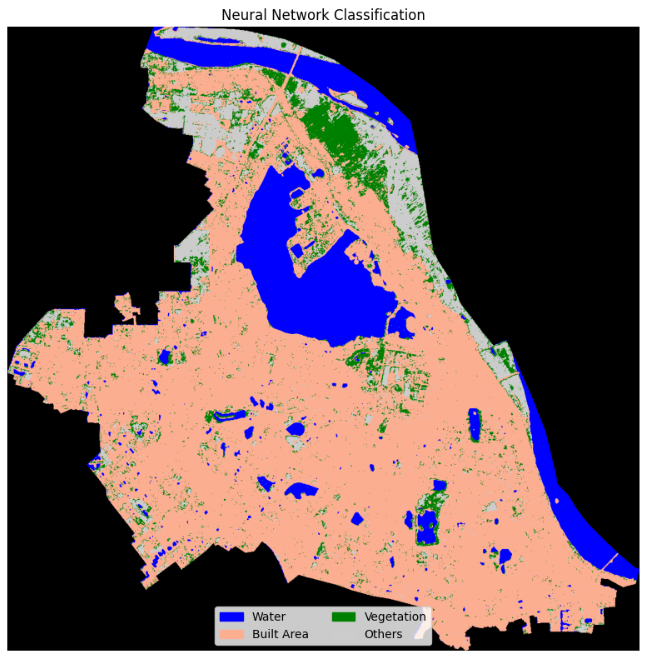
****

Figure .. CNN Classification result

## Evaluation and Model Comparison

To assess the effectiveness of different classification approaches for Sentinel-2-based land cover mapping in urban Hanoi, three models—Support Vector Machine (SVM), Random Forest (RF), and a custom-built 1D Convolutional Neural Network (CNN)—were trained and tested using the same dataset and evaluated with the same ground truth samples.

### Overall Accuracy and Class Performance

All three models achieved high overall classification accuracy:

* SVM: 89,84%
* Random Forest: 91.75%
* 1D CNN: 95.26%

Table . F1-score comparison

|  |  |  |  |
| --- | --- | --- | --- |
| **Class** | **SVM F1-score** | **RF F1-score** | **CNN F1-score** |
| Water | 0.99 | 0.98 | 0.99 |
| Built Area | 0.94 | 0.95 | 0.98 |
| Vegetation | 0.59 | 0.63 | 0.77 |
| Others | 0.76 | 0.75 | 0.88 |

These results confirm that while **SVM** and **RF** are effective in cleanly separated spectral classes, they struggle with overlapping or mixed land covers, where **CNN** offers a substantial advantage. This class-wise evaluation aligns with the overall accuracy scores, reaffirming **CNN’s** robustness in both general and fine-grained classification tasks.

Among the three, the **1D CNN** model achieved the highest accuracy, showing its superior ability to capture complex spectral patterns in urban land cover. While both **SVM** and **RF** performed well in classifying distinct classes such as **Water** and **Built Area**, they were more prone to confusion in spectrally similar categories like **Vegetation** and **Others**.

### Model Performance Comparison

Table . Performance Comparison

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Accuracy (%)** | **Strengths** | **Weaknesses** |
| **SVM** | 89.84 | Robust with small training data; good at linear class boundaries | Lower generalization in complex or mixed land cover types |
| **RF** | 91.75 | Fast, interpretable, stable across classes | Misclassifies spectrally overlapping classes more frequently |
| **1D CNN** | 95.26 | Best performance; learns deep spectral features; handles class confusion | Requires more training time and computational resources |

# CONCLUSION

## Conclusion and Future work sugguest

This project successfully implemented and compared three pixel-based land cover classification approaches—Support Vector Machine, Random Forest, and a custom 1D Convolutional Neural Network—using Sentinel-2 satellite imagery over central Hanoi. Among the three models, the 1D CNN achieved the highest accuracy, demonstrating superior capability in handling complex spectral patterns and class overlaps common in urban environments.

The integration of spectral bands with NDVI significantly improved classification quality, particularly in distinguishing vegetation from bare or construction areas. The study highlighted the advantages of deep learning models in remote sensing applications, particularly in heterogeneous urban landscapes.

For future development, the following directions are proposed:

* Incorporate spatial context by adopting 2D CNNs or hybrid spatial-spectral architectures to improve classification of complex urban scenes.
* Expand to temporal analysis, enabling detection of land-use changes over time for dynamic monitoring of urbanization trends.
* Apply transfer learning to adapt the trained model for other cities with similar spectral characteristics, reducing training cost and effort.
* Integrate with GIS platforms to support visualization, analysis, and decision-making in real-world urban planning systems.
* By continuing to refine and expand this work, the project can serve as a foundation for scalable, intelligent land cover monitoring in Vietnam and beyond.

## Application of Segmentation Results

The land cover classification results generated in this study offer several practical applications across different domains:

* Urban Planning and Management: Accurate land use maps provide valuable input for city planners and policymakers to monitor urban growth, manage infrastructure development, and preserve green spaces. This supports sustainable urban planning in fast-developing cities like Hanoi.
* Environmental Monitoring: By detecting changes in vegetation cover, water bodies, and impervious surfaces, the results contribute to tracking environmental degradation, illegal land conversions, or deforestation within urban areas.
* Agricultural Analysis: Although the primary focus is urban classification, the methodology can be extended to peri-urban and rural areas to monitor crop patterns, classify agricultural zones, and support precision farming initiatives.
* Smart City Development: The classified data can be integrated into Geographic Information Systems (GIS) and decision support tools for intelligent land-use monitoring, enabling proactive management and data-driven governance.
* Disaster Risk Management: Knowing the exact location of built-up areas, vegetation, and water bodies is essential for flood risk assessment, emergency response planning, and climate resilience strategies.

# REFERENCES

Bibliography

[1]US Department of Commerce, National Oceanic and Atmospheric Administration, “What is remote sensing?,” *Noaa.gov*, 2020. <https://oceanservice.noaa.gov/facts/remotesensing.html>

[2]Wikipedia Contributors, “Remote sensing,” *Wikipedia*, Sep. 23, 2019. <https://en.wikipedia.org/wiki/Remote_sensing>

[3]M. A. Reddy, “Textbook Of Remote Sensing And Geographical Information Systems M Anji Reddy 3e 2008 Book : Free Download, Borrow, and Streaming : Internet Archive,” *Internet Archive*, 2018. <https://archive.org/details/TextbookOfRemoteSensingAndGeographicalInformationSystems/page/n39/mode/2up?view=theater> (accessed Jul. 06, 2025).

[4]S.-H. by Sinergise, “Normalized difference vegetation index,” *Sentinel-Hub custom scripts*. <https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/ndvi/>

[5]“S2 Products,” *Copernicus.eu*, 2016. <https://sentiwiki.copernicus.eu/web/s2-products#S2Products-Productanomaliesandfeatures> (accessed Jul. 06, 2025).

[6]Kamrrann, “GitHub - Kamrrann/Sentinel2\_Classification\_Using\_Deep-Convolutional-Neural-Network: This repository provides a comprehensive guide on performing Land Use and Land Cover (LULC) classification on Sentinel-2 multispectral images using finely-tuned Convolutional Neural Networks (CNNs). The entire workflow involves training an efficiently optimized CNN model and subsequently applying the pre-trained model to map LULC.,” *GitHub*, 2024. <https://github.com/Kamrrann/Sentinel2_Classification_Using_Deep-Convolutional-Neural-Network> (accessed Jul. 06, 2025).

[7]d-smit, “GitHub - d-smit/sentinel2-deep-learning: Image classification on Sentinel-2 satellite imagery.,” *GitHub*, 2019. <https://github.com/d-smit/sentinel2-deep-learning> (accessed Jul. 06, 2025).

[8]raoofnaushad, “GitHub - raoofnaushad/Land-Cover-Classification-using-Sentinel-2-Dataset: Application of deep learning on Satellite Imagery of Sentinel-2 satellite that move around the earth from June, 2015. This image patches can be trained and classified using transfer learning techniques.,” *GitHub*, 2015. <https://github.com/raoofnaushad/Land-Cover-Classification-using-Sentinel-2-Dataset> (accessed Jul. 06, 2025).

[9]J. R. Anderson, *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. 1976.

[10]A. Gowda Karegowda, Nandeesha, and Nagaraj, “International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Classification of Sentinel 2 Images using Customized Convolution Neural Networks,” *Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE*, vol. 2023, no. 1s, pp. 136–142.

[11]P. Thanh Noi and M. Kappas, “Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery,” *Sensors*, vol. 18, no. 1, p. 18, Dec. 2017, doi: <https://doi.org/10.3390/s18010018>.

[12]D. Phiri, M. Simwanda, S. Salekin, V. Nyirenda, Y. Murayama, and M. Ranagalage, “Sentinel-2 Data for Land Cover/Use Mapping: A Review,” *Remote Sensing*, vol. 12, no. 14, p. 2291, Jul. 2020, doi: <https://doi.org/10.3390/rs12142291>.

[13]C. Li, J. Wang, L. Wang, L. Hu, and P. Gong, “Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery,” *Remote Sensing*, vol. 6, no. 2, pp. 964–983, Jan. 2014, doi: <https://doi.org/10.3390/rs6020964>.