iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 12 **Issue:** XII **Month of publication:** December 2024

**DOI:** https://doi.org/10.22214/ijraset.2024.65527

# Predicting Illegal Forest Encroachment in Bamori Range, Guna, Madhya Pradesh Using Machine Learning

Akshaya Rathore
*DFO Guna Forest Division*

***Abstract: Illegal forest encroachment represents a critical challenge in the Bamori Range, Guna District, Madhya Pradesh. The conversion of forested land into agricultural areas disrupts biodiversity and threatens ecological stability. This study presents a predictive model, utilizing machine learning, to identify areas prone to illegal land-use changes. By employing spatial and temporal datasets within a Random Forest framework, the model achieved a high accuracy of 94% and an ROC AUC score of 0.9796. Key predictors, such as the "Rate of Spread" and "Neighbor Crop Count," emerged as significant drivers of encroachment, providing actionable insights for forest management. This approach offers a data-driven solution to enhance conservation strategies and mitigate environmental degradation.***

## I. INTRODUCTION

### *A. Background and Context*

Illegal conversion of forest lands into agricultural fields is a persistent and complex issue, particularly in this region. This encroachment disrupts ecological balance, and undermines conservation efforts. While remote sensing and geospatial analysis have been used extensively for monitoring land-use changes, their potential for proactive risk prediction remains underutilized. Incorporating predictive modeling can provide forest managers with the necessary foresight to address encroachment effectively.

### *B. Objective of the Study*

This study aims to create a data-driven model to forecast illegal land-use changes in future years. By using a Random Forest algorithm that integrates historical patterns and spatial data, this model provides forest management with a powerful tool to combat encroachment effectively.

### *C. Significance*

Accurate predictions can direct resources where they are most needed. Identifying areas prone to illegal activities can help forest managers implement targeted interventions, raise community awareness, and strengthen preventive measures.

## II. LITERATURE REVIEW

### *A. Prior Studies*

Previous research in India has often utilized remote sensing and GIS tools to track land-use changes. Machine learning models like Random Forests have shown potential for handling complex datasets. However, integrating temporal indicators and spatial context remains underexplored in predictive models for illegal encroachment

### *B. Addressing Gaps*

This study bridges the gap by combining spatial and temporal features, improving predictive accuracy and providing practical tools for forest management.

## III. DATA AND STUDY AREA

### *A. Study Area*

The Bamori Range, located in Guna District, Madhya Pradesh, is characterized by rich forest cover. However, ongoing encroachment for agricultural purposes poses a significant threat to this region.

### B. Data Sources

The land cover classification rasters used in this study were generated using a hyperparameter-tuned Random Forest model described in [Land Cover Classification in Bamori Range using Machine Learning Algorithms: A Comparative Study]. This model utilized Sentinel-2 imagery, Dynamic World probabilities, and spectral indices to classify land cover into five distinct classes. Each raster represents the land cover classes for the respective year, with consistent classification schemes across years, including classes for 'Forest', 'Crop', 'Crop fallow', 'Water bodies', and 'Built Up areas'.

## IV. METHODOLOGY

### A. Model Selection

A Random Forest classifier was chosen for its robustness, ability to handle complex datasets, and effectiveness in modeling non-linear relationships. Its reliability and resistance to overfitting made it ideal for predicting illegal forest encroachment.

### B. Feature Engineering

#### 1) Temporal Features

LULC Classes (2019–2022): Historical land cover data for each year to track patterns.

Change Indicators: Binary variables marking land use changes between consecutive years (2019–2022).

#### 2) Spatial Features

Neighbor Crop Count: Tracks the number of surrounding pixels classified as agricultural.

Directional Changes: Summarizes the spread of encroachment in neighboring pixels.

Rate of Spread: Counts how often a pixel's land class changed between 2019 and 2023

### C. Spatial Feature Calculation

Neighbor Identification: The eight surrounding pixels for each forested pixel were identified to capture influences from nearby areas.

Neighbor Crop Count: Calculated by analyzing land cover of neighboring pixels to assess agricultural pressure.

Directional Changes: Tracks encroachment trends in nearby pixels.

### D. Model Training and Validation

Data Splitting: A 70: 30 train-test split ensured balanced class representation.

Hyperparameter Tuning: Optimized parameters included:

n_estimators: 100

max_depth: 10

Class Weights: Balanced

Random State: 42

Threshold Adjustment: A 0.35 threshold maximized recall, ensuring most encroachment cases are flagged.

## V. RESULTS

### A. Model Performance

The Random Forest classifier achieved the following performance metrics on the test set:

Accuracy: 94%

Precision (Encroachment): 89%

Recall (Encroachment): 100%

F1 Score (Encroachment): 94%

ROC AUC Score: 0.9796

The model's ROC curve (in the appendix) demonstrates excellent discriminative ability.

Confusion Matrix

| | Predicted Not Encroachment | Predicted Encroachment |
|---|---|---|
| Actual No Encroachment | 581,004 | 77,869 |
| Actual Encroachment | 1,790 | 657,083 |

### B. Feature Importance

The analysis of feature importances revealed the following:

1) Rate of Spread: 44.21%
2) LULC_2022: 18.40%
3) Neighbor Crop Count: 15.91%
4) LULC_2020: 4.17%
5) LULC_2021: 3.76%

## VI. DISCUSSION

### A. Interpretation of Results

The model demonstrates strong predictive capabilities, identifying areas at risk of illegal encroachment by combining historical land cover changes and spatial context. Its high recall ensures that most potential encroachment cases are flagged, allowing forest managers to focus resources effectively.

### B. Limitations

Predictions rely on LULC data quality and may not account for unexpected factors like policy changes or natural disasters.

### C. Future Work

Enhancements include integrating socioeconomic data, and dynamically updating the model with new data.

## VII. CONCLUSION

This study demonstrates the power of machine learning in addressing environmental challenges. By predicting illegal encroachment in the Bamori Range, the Random Forest model empowers forest managers with actionable insights, contributing to long-term conservation efforts.

## REFERENCES

[1] Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26(1), 217222.
[2] Rawat, J. S., & Kumar, M. (2015). Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India. The Egyptian Journal of Remote Sensing and Space Science, 18(1), 7784.
[3] Joshi, P. K., Roy, P. S., Singh, S., Agrawal, S., & Yadav, D. (2016). Vegetation cover mapping in India using multitemporal IRS Wide Field Sensor (WiFS) data. Remote Sensing of Environment, 103(2), 190202.
[4] Patel, N. R., Kumar, S., & Shekhawat, R. S. (2020). Modeling land use land cover change using geospatial techniques and Markov chain analysis in a semiarid region of India. Environmental Earth Sciences, 79(19), 115.

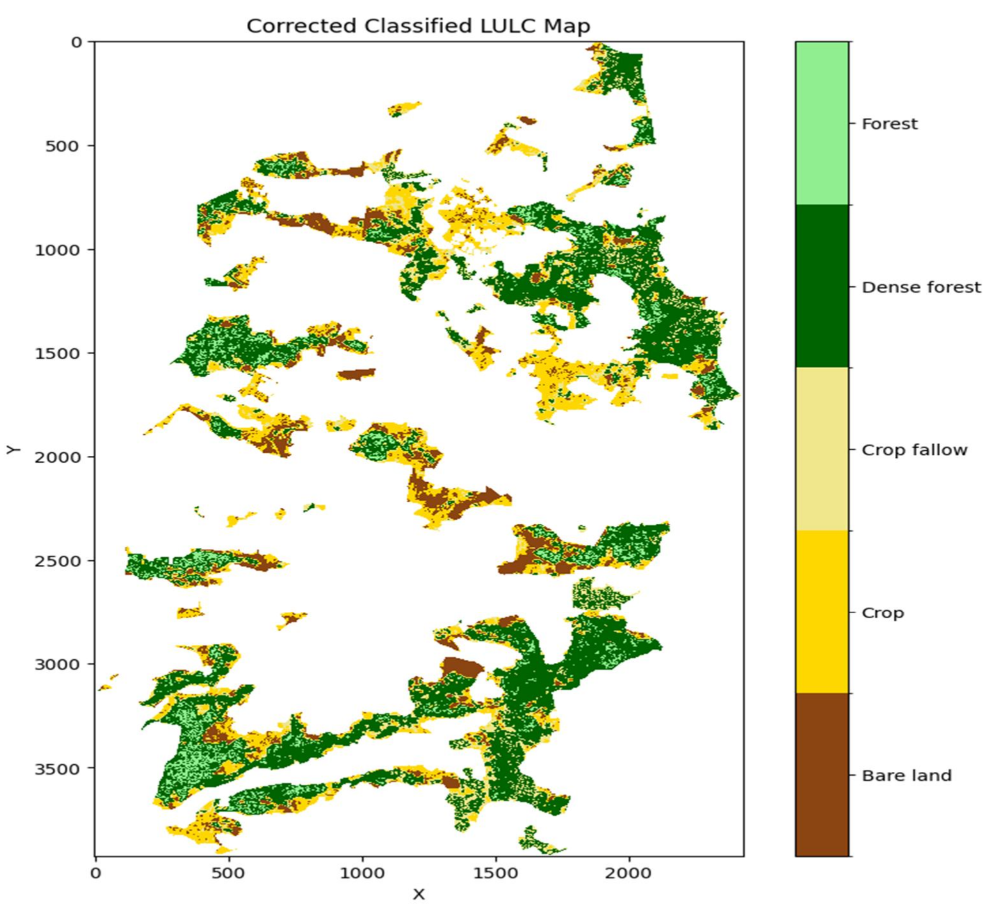
## APPENDIX

### A. Feature Definitions

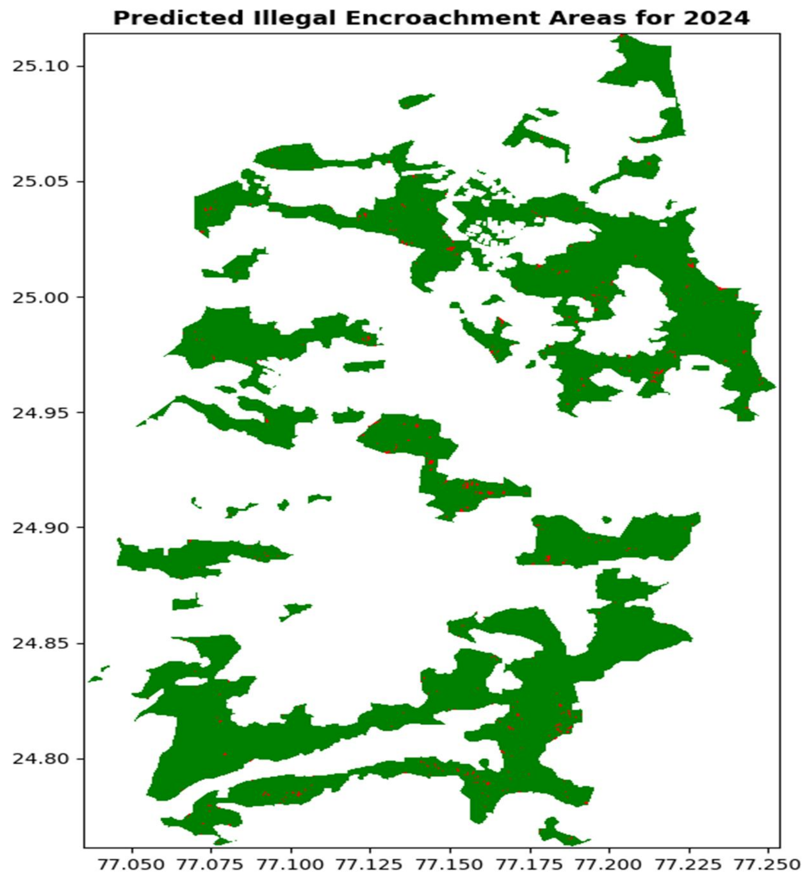
LULC Classes: Integer codes representing land cover types such as 'Forest', 'Crop', 'Crop fallow', etc.
Change Indicators: Binary values indicating whether a pixel's class changed between consecutive years.
Neighbor Crop Count: Number of neighboring pixels classified as 'Crop' or 'Crop fallow' in 2022.
Directional Changes: Sum of binary change indicators in neighboring pixels between 2021 and 2022.
Rate of Spread: Cumulative count of class changes for a pixel from 2019 to 2023.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue XII Dec 2024- Available at www.ijraset.com*

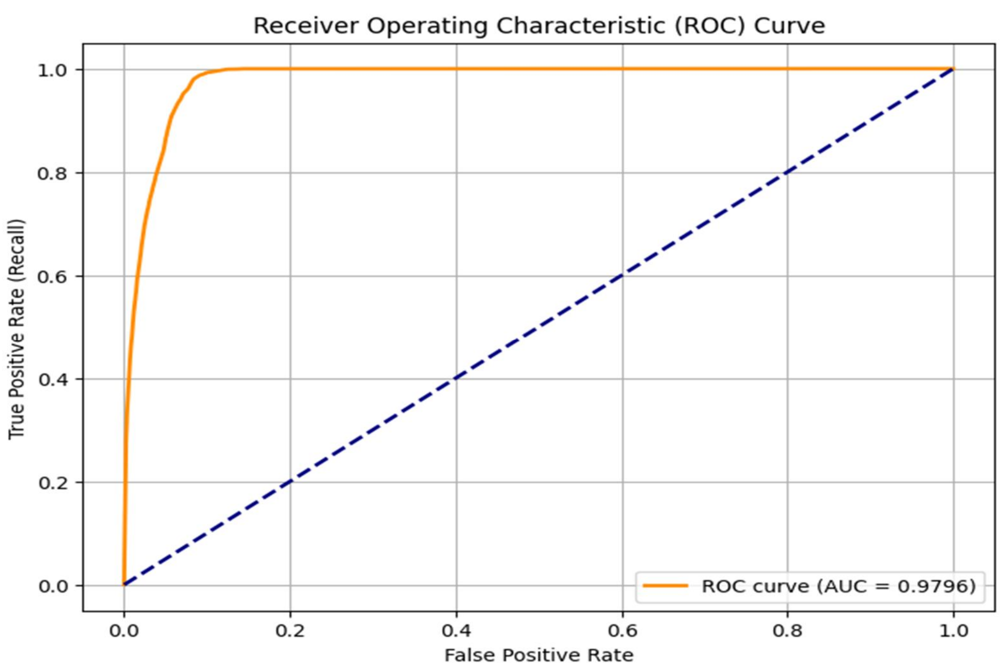
*1) Maps of the Study Area*

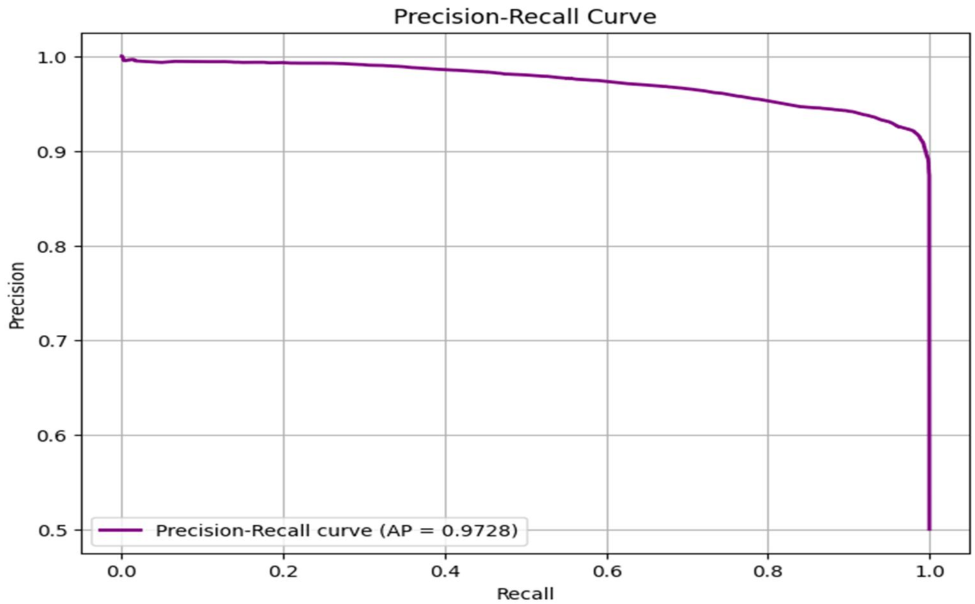
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue XII Dec 2024- Available at www.ijraset.com*



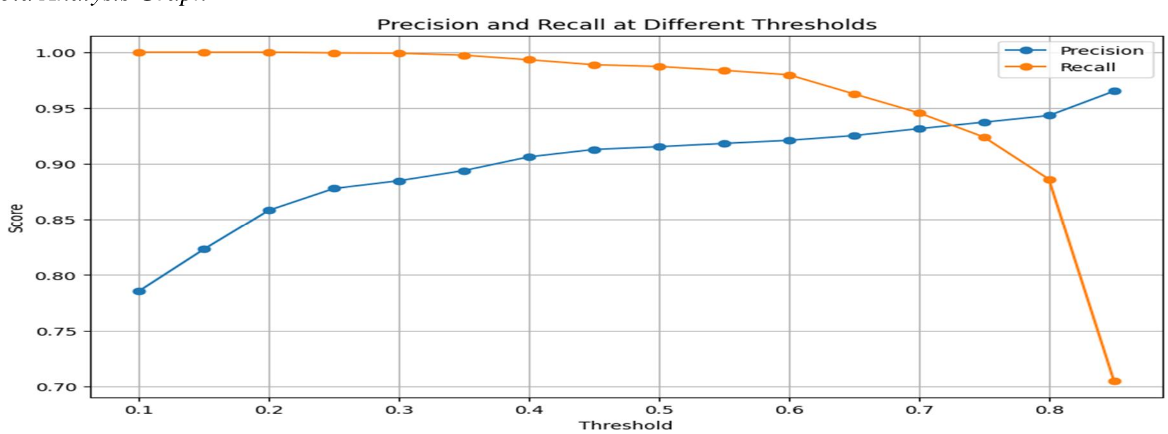

2) *Model Performance Graphs*



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue XII Dec 2024- Available at www.ijraset.com*

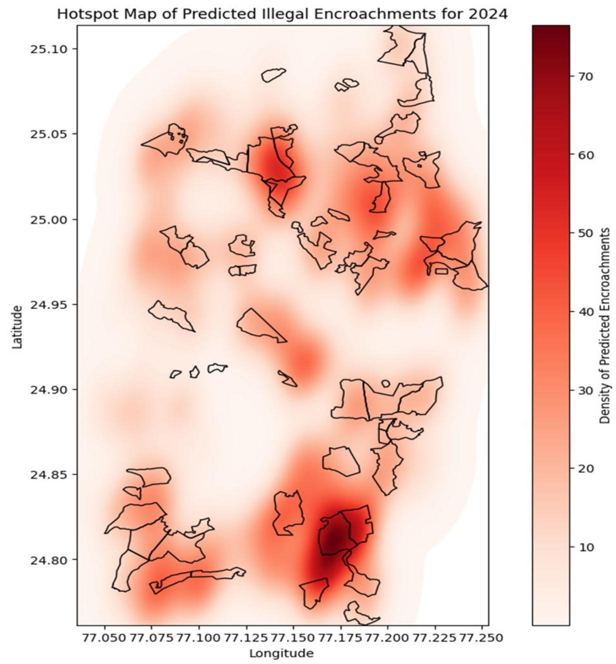


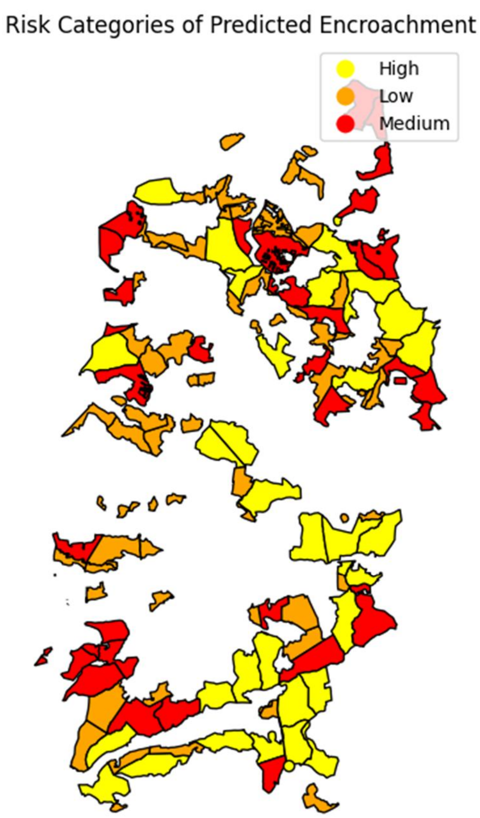

*3) Threshold Analysis Graph*

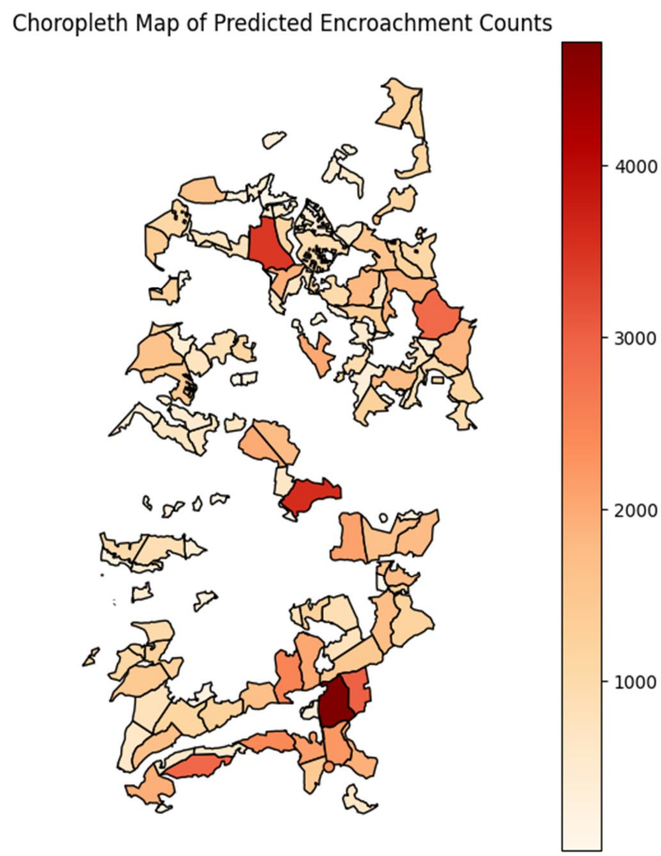


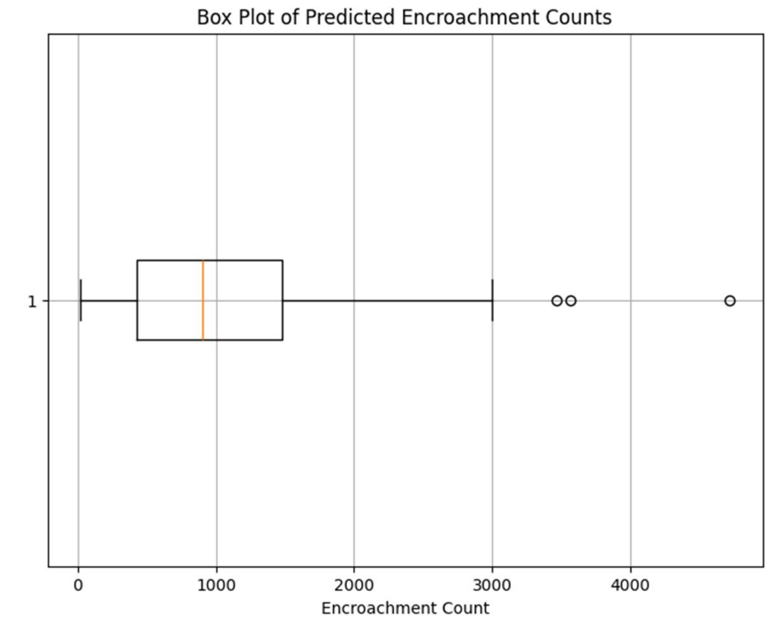

| | Threshold | Precision | Recall | F1-Score |
|---|---|---|---|---|
| 0 | 0.10 | 0.785563 | 1.000000 | 0.879905 |
| 1 | 0.15 | 0.823217 | 1.000000 | 0.903038 |
| 2 | 0.20 | 0.858472 | 1.000000 | 0.923847 |
| 3 | 0.25 | 0.877993 | 0.999334 | 0.934742 |
| 4 | 0.30 | 0.884954 | 0.999059 | 0.938551 |
| 5 | 0.35 | 0.894049 | 0.997283 | 0.942849 |
| 6 | 0.40 | 0.906314 | 0.993190 | 0.947765 |
| 7 | 0.45 | 0.912932 | 0.988738 | 0.949324 |
| 8 | 0.50 | 0.915430 | 0.987198 | 0.949961 |
| 9 | 0.55 | 0.918327 | 0.983713 | 0.949896 |
| 10 | 0.60 | 0.921103 | 0.979746 | 0.949520 |
| 11 | 0.65 | 0.925362 | 0.962651 | 0.943638 |
| 12 | 0.70 | 0.931548 | 0.945709 | 0.938575 |
| 13 | 0.75 | 0.937321 | 0.924038 | 0.930633 |
| 14 | 0.80 | 0.943265 | 0.886301 | 0.913896 |
| 15 | 0.85 | 0.965063 | 0.705001 | 0.814783 |

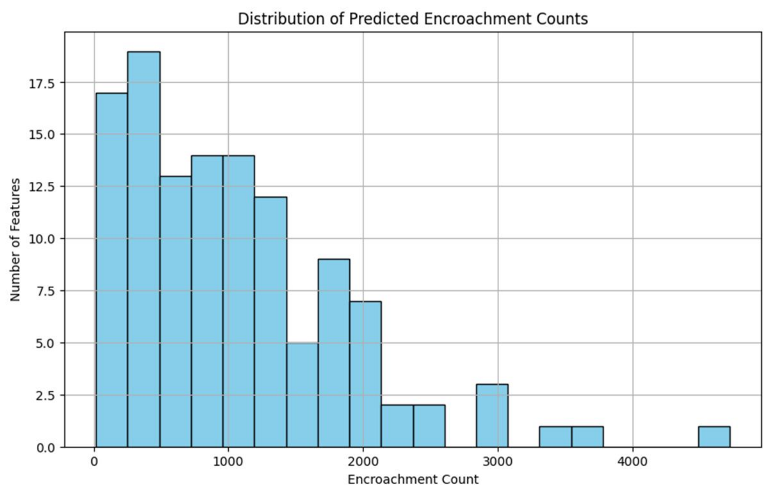
Sample Predictions Overlay

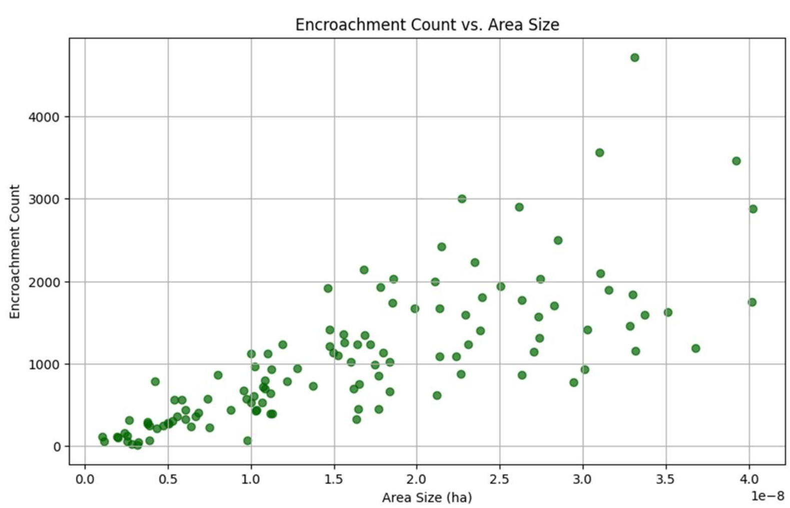



Box Plot of Predicted Encroachment Counts
1
0
1000
2000
3000
4000
Encroachment Count


Distribution of Predicted Encroachment Counts
Number of Features
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
0
1000
2000
3000
4000
Encroachment Count


Encroachment Count vs. Area Size
Encroachment Count
0
1000
2000
3000
4000
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1e−8
Area Size (ha)