Smart Terrain Recognition System with Yolov8 Model

Anmol Deshratna Saxena (corresponding author)
SoCSE, Dr. Vishwanath Karad MIT World Peace University
Pune, Maharashtra, India
1032211624@mitwpu.edu.in

Aditya Gode
SoCSE, Dr. Vishwanath Karad MIT World Peace University
Pune, Maharashtra, India
1032211586@mitwpu.edu.in

Drishti Sinha
SoCSE, Dr. Vishwanath Karad MIT World Peace University
Pune, Maharashtra, India
1032210911@mitwpu.edu.in

Rashmi Phalnikar
SoCSE, Dr. Vishwanath Karad MIT World Peace University
Pune, Maharashtra, India
rashmi.phalnikar@mitwpu.edu.in

Ayush Thakre
*SoCSE,*Dr. Vishwanath Karad MIT World Peace University
Pune, Maharashtra, India
1032211517@mitwpu.edu.in

***Abstract*—**
Intelligent Terrain Identification for Agricultural and Environmental Uses Accurate terrain description is essential for effective agricultural and environmental planning. This work explores the precise and efficient terrain type detection using state-of-the-art computer vision techniques such as You Only Look Once (YOLO). Our objective is to combine image data with terrain features, such as plains, forests, mountains, and deserts, to develop a dependable system for forecasting landscape characteristics. This approach is less expensive and time-consuming than traditional survey methods, yet it provides insightful information that may be used to enhance agricultural practices, irrigation systems, and infrastructure development. Combining data fusion and model optimization ensures effective terrain recognition, enabling informed decision-making in various agricultural and environmental contexts.

Keywords: computer vision, data fusion, model optimization, remote sensing, land use planning, YOLO, terrain analysis, agricultural development, environmental management, and smart terrain recognition.

I INTRODUCTION

Accurately assessing topographical features is essential for maximizing land use and enhancing sustainability in the current era of agricultural development and environmental management. Time-consuming, labor-intensive, and expensive are the traditional methods of terrain analysis, which mostly rely on physical surveys and ground measurements. The swift progress of artificial intelligence(AI) and machinelearning technologies has led to the emergence of smart terrain identification systems as effective substitutes that provide increased accuracy and speed. These systems use computer vision models to enable real-time recognition and classification of various terrain types, including rocky, sandy, and plateau regions.

As a cutting-edge object detection framework renowned for its speed and accuracy, the "You Only Look Once" (YOLO) model is one of the most promising methods for smart terrain recognition. YOLO is excellent for applications in agriculture, where quick and precise land evaluation is essential. It can process high-resolution imagery and video streams to recognize and classify terrain elements in real-time. YOLO and other AI-driven terrain analysis techniques combined offer a comprehensive strategy for comprehending the features of the terrain and how they affect agricultural output.

Accurate terrain recognition has significant implications for agriculture, enabling better decisions regarding crop selection, irrigation planning, and infrastructure development. For instance, rocky or sandy terrains require distinct strategies for soil treatment, water retention, and crop growth, while plateau regions may present opportunities for specific types of agriculture or livestock management. By integrating smart terrain recognition systems into agricultural planning, stakeholders can optimize resource allocation, mitigate risks, and contribute to sustainable land development.

II. LITERATURE SURVEY

Planning for agriculture and the environment depends on accurate terrain identification as it affects resource distribution and land management. The time-consuming and expensive nature of traditional surveying techniques has sparked interest in computer vision innovations such as You Only Look Once (YOLO), which enable effective terrain analysis.

Numerous techniques for classifying terrain have been investigated, ranging from conventional remote sensing methods to machine learning methods. Aerial surveys are one example of a labor-intensive, weather-dependent traditional approach. On the other hand, huge datasets are processed effectively by machine learning, especially by Convolutional Neural Networks (CNNs), although they frequently demand a substantial amount of CPU power.

YOLO, which is a top real-time object identification algorithm that is renowned for processing photos quickly and accurately, which makes it perfect for classifying terrain. Research validates YOLO's capability for precisely detecting terrain types including plains and woods by demonstrating its efficacy in a variety of applications, such as urban planning and environmental monitoring.

Accurate terrain recognition is improved by combining several data sources. Comprehensive terrain models are produced by data fusion approaches that combine LiDAR and satellite photos. YOLO's adaptability to many environmental situations is further enhanced by model optimization procedures such hyperparameter tuning and transfer learning.

Agriculture and environmental management are significantly impacted by accurate terrain identification. With careful irrigation planning and crop selection based on topographical factors, it can improve agricultural operations . Furthermore, environmental hazards are reduced by well-informed infrastructure.

The use of YOLO into terrain identification systems represents a significant development in environmental and agricultural planning. By integrating fast image processing with data fusion and model development, this technique offers a cost-effective solution for accurate terrain categorization. Subsequent investigations ought to enhance these methodologies and evaluate their suitability for other environments.

III. PROPOSED SYSTEM

The proposed system for intelligent terrain identification leverages cutting-edge technology, specifically the YOLO(You Only Look Once) model. At its core, the system is designed to analyze high-resolution satellite images, accurately identifying various terrain types such as plains, forests, mountains, and deserts. This real-time recognition capability allows agricultural and environmental planners to make informed decisions quickly, streamlining processes that would otherwise take much longer with traditional surveying methods.

To achieve this, the system begins with a well-structured dataset of satellite images, each labeled to indicate the specific terrain features present. The data undergoes thorough preprocessing to ensure it meets the requirements of the YOLO model, including resizing and normalization. This step is crucial for maintaining image quality while increasing efficiency. Additionally, data augmentation methods are applied to enhance the dataset, making the model more robust against variations in real-world conditions.

Once the data is prepared, the YOLOv8 model is curated using a carefully defined methodology that includes a composite loss function, optimizing the model's ability to detect and classify different terrain types. The training process is designed to ensure high accuracy, utilizing metrics like meanAveragePrecision to evaluate performance. After training, the model is capable of processing new satellite images, generating bounding boxes around detected terrain types, and providing insights that can directly inform agricultural practices and environmental management strategies.

In summary, this intelligent terrain identification system combines advanced image processing techniques with deep learning to deliver a powerful tool for stakeholders in agriculture and environmental management. By offering a faster, more accurate alternative to traditional methods, the system enhances decision-making capabilities and promotes sustainable land use practices.

IV. DATASET

The dataset used for this terrain recognition project consists of multiple combined satellite image datasets, each containing a diverse array of terrain types. These files contain satellite sensor-captured photos of a variety of environments, including plains, mountains, forests, and deserts. The data has been processed and organized in the YOLOv8 format, enabling efficient training for object detection and classification models.

The dataset is structured as follows:

- **Images**: High-definition satellite photos, each containing one or more different topographical kinds. To guarantee uniformity throughout the dataset, these photos undergo pre-processing to assure consistency in size and quality. Depending on the source dataset, the image resolutions range from 256x256 pixels to 1024x1024 pixels.

- **Labels**: Each image is paired with a corresponding label file that follows the YOLOv8 format. The labels identify terrain categories and provide bounding box coordinates around the areas of interest within the image. The structure consists of:
  - Class ID (identifying the kind of land cover: desert, mountains, forests, etc.)
  - The width, height, x, and y coordinates of the bounding box are normalized to the picture dimensions.

- **Categories**: There are four main terrain classifications in the dataset:

- Plains: Low-lying flat terrain, often covered with sparse flora.
- Mountains: Roughly topographically inclined areas that are frequently identified by peaks and ridges.
- Forests are regions with a high concentration of trees and dense vegetation.
- Deserts: Arid areas with little to no flora that are typified by sand dunes or rocky surfaces.

- **Sources of Data**: The dataset consists of both specially labeled data and publicly accessible satellite image archives. Important sources include curated collections from sites like Roboflow, which provide photos in YOLOv8 format, and datasets from geospatial repositories.

- **Data Splits**: To guarantee a balanced evaluation across all terrain types, the dataset is split into subsets for training, validation, and testing. 70%is for training, 20% validation, and 10% for testing the split ratio.

This dataset is appropriate for problems involving object detection and landscape categorisation using deep learning. Effective model training and fine-tuning are made possible by the YOLOv8 format's optimized labeling and image structuring. By merging different datasets, this collection delivers a strong and diversified training set, boosting model generalization across numerous satellite imagery sources and terrain types.

V. METHODOLOGY

In this project, we develop a terrain recognition system using different techniques, leveraging the YOLOv8 object detection model. The methodology includes preprocessing of the data, model selection, model architecture, training, and evaluation, with the ultimate goal of obtaining high classification and detection accuracy for various terrain types using satellite photos.

1. Data Preprocessing

Preprocessing raw satellite images is essential for compatibility with the YOLOv8 model. The images are resized to a standard input size of 640x640 pixels to balance detail and computational efficiency. Normalization follows, with pixel values are divided by 255 to bring them into the [0, 1] range, ensuring stability during training

Different techniques are to enhance model performance and prevent overfitting. This includes random flips (horizontal and vertical), rotations between -30° and 30°, and adjustments to contrast and brightness to accommodate varying conditions. Each image is paired with a corresponding label file in the YOLOv8 format, containing normalized bounding box coordinates and terrain class annotations, which supports effective training and evaluation.

2. YOLOv8 Model Overview

The YouOnlyLookOnce v8 (YOLOv8) model is a one-stage object detection framework that divides an input image into an S×S grid and predicts bounding boxes and classprobabilities for every grid cell. Because YOLOv8 can identify many items in an image quickly and accurately, it is a good choice for tasks involving terrain detection where diverse terrains may coexist.

3. Model Architecture

YOLOv8 introduces a new neural network architecture optimized for fast object detection. Its key components are:

- **Backbone**: A CSP-Darknet53 version that uses the input images to extract key features. To prevent vanishing gradients and enhance feature propagation—both essential for learning a variety of terrain features—the backbone makes use of residual connections.

● **Neck**: The component used as the neck is the Path Aggregation Network (PANet). It improves the localisation and identification of terrain objects at various scales by combining features from the backbone's layers.

● **Head**: The detection head consists of three output layers, each responsible for detecting terrain objects at different scales (small, medium, large). The ability of the model to recognise different-sized terrains in satellite imagery is enhanced by this multi-scale detection.

4. Loss Function

The YOLOv8 model optimizes a composite loss function during training, comprising three key components:

● Bounding Box Loss: This loss uses Complete Intersectionover Union(CIoU) loss, an improvement on traditional IoU loss, to quantify the difference between the ground-truth boxes and the predicted bounding boxes. The formula for CIoU is:

$$\mathrm{CIoU} = 1 - \left(IoU - \frac{\rho^2(\mathbf{b}, \mathbf{b}^g)}{c^2} - \frac{v}{(1 - IoU) + v} \right)$$

Where:

● b and b^g are the predicted andground-truth boxes.
● ρ^2(b,b^g) is the squaredEuclideandistance between the center points of b and b^g.
● cis the diagonal length of the smallest enclosing box covering both b and b^g.
● v is a measure of aspect ratio consistency.

By including aspect ratio and distance in the loss function, CIoU enhances convergence and is crucial for precisely localising terrain regions.

● Objectness Loss: This component penalizes the model for predicting objects in grid cells where no terrain is present. Binary cross-entropy loss is used here:

$$L_{obj} = -\sum_{i=1}^{S^2} \left[p_i \log(\hat{p}_i) + (1 - p_i) \log(1 - \hat{p}_i) \right]$$

Where pi is the true objectnessscore, and pi cap is the predicted objectness score.

● **Classification Loss**: This is a standard cross-entropy loss used to classify terrain types within the bounding boxes:

$$L_{cls} = -\sum_{i=1}^{S^2} \sum_{c=1}^{C} p_i^c \log(\hat{p}_i^c)$$

Where pic is the true class label (terrain type) for grid cell i and class c, and pic cap is the predicted probability for class c.

The total loss function is a weighted sum of these three components:

$$L_{total} = \lambda_{box}L_{box} + \lambda_{obj}L_{obj} + \lambda_{cls}L_{cls}$$

where λbox, λobj, and λcls are weighting factors for the respective losses.

5. Training Procedure

The YOLOv8 model is trained using the Adam optimizer, known for its adaptability in handling noisy gradient updates. The following hyperparameters are used:

- **Learning Rate**: To avoid overfitting, the learning rate is initially set to 0.001 and is subsequently reduced over training via a cosine annealing schedule.
- **Batch Size**: To balance GPU memory efficiency and gradient estimation accuracy, a batch size of 16 is used.
- **Epochs**: The model is trained for100 epochs, with early stopping applied if the validation loss does not get better for 10 consecutive epochs.
- **Weight Decay**: To regularise the model and avoid overfitting, a weight decay i.e 0.0005 is employed.

6. Model Evaluation

The performance of the model is evaluated using the following metrics:

- **Mean Average Precision (mAP)**: The primary metric for evaluating the accuracy of the model. It measures the area under the precision-recall curve for each class and averages it over all classes. We report both mAP@0.5 and mAP@0.5:0.95.
- **Precision and Recall**: Precision calculates the proportion of rightly predicted terrain objects out of all predicted objects, the recall measures the proportion of correctly predicted objects out of all ground-truth objects. These are computed as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

- **F1 Score**: The harmonicmean(HM) of precision and recall is used to balance both metrics:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

7. Classification of Terrain and Post-Processing

The model produces bounding boxes with corresponding terrain class probabilities after detection. To eliminate unnecessary boxes and keep the predictions with the highest degree of confidence, Non-Maximum Suppression (NMS) is used. This guarantees the identification of only the most pertinent terrain objects.

8. Implementation

The model is used as a terrain recognition system after it has been trained and assessed. The system receives fresh satellite images as input, applies the YOLOv8 model to process them, and outputs bounding boxes surrounding identified terrain types along with associated class labels.

VI. RESULT AND ANALYSIS

The training results for the terrain recognition model demonstrate significant progress over ten epochs, highlighted by a dramatic decrease in training loss from 0.81296 in the first epoch to just 0.10993 by the tenth. This decline indicates learning from the training data, as evidenced by the consistent Top-1 accuracy of 1 throughout the training process, showcasing the model's reliability in correctly identifying the top predicted terrain class. Meanwhile, the Top-5 accuracy increases from 0.54123 to 0.69345, reflecting the model's improving ability to recognize a broader range of terrain types, although there is still room for enhancement in ranking these predictions more accurately.

Validation loss, which starts at 1.1882 and decreases to 1.0356, suggests that the model is generalizing well to unseen data, indicating it is not only memorizing the training examples. However, the relatively high validation loss compared to training loss hints at potential overfitting, signaling a need for further tuning. Strategies such as regularization techniques or adjustments to learning rates could be beneficial in optimizing performance. Overall, these results indicate a promising trajectory for the terrain recognition system, with clear pathways for improving accuracy and reliability in practical applications.

VII. CONCLUSION

Our findings demonstrate, in summary, that terrain type recognition for agricultural and environmental applications is significantly improved by the You Only Look Once (YOLO) method. We developed a low-cost technology that offers insightful data and saves time and money compared to traditional surveys by fusing picture data with terrain parameters. The revolutionary potential of YOLO in landscape management and decision-making is shown by this strategy, which improves agricultural techniques and provides guidance for infrastructure development.

REFERENCES

[1] Li, Y., Zhao, Z., Luo, Y., & Qiu, Z. (2020). Real-time pattern-recognition of GPR images with YOLO v3 implemented by TensorFlow. *Sensors, 20*(22), 6476.

[2] Cui, B., Liang, L., Ji, B., Zhang, L., Zhao, L., Zhang, K., ... & Créput, J. C. (2024). Exploring the YOLO-FT deep learning algorithm for UAV-based smart agriculture detection in communication networks. *IEEE Transactions on Network and Service Management*.

[3] Paul, A., Machavaram, R., Kumar, D., & Nagar, H. (2024). Smart solutions for capsicum harvesting: Unleashing the power of YOLO for detection, segmentation, growth stage classification, counting, and real-time mobile identification. *Computers and Electronics in Agriculture, 219,* 108832.

[4] Likith, S., Reddy, B. R., & Reddy, K. S. (2021, December). A smart system for detection and classification of pests using YOLO and CNN techniques. In *2021 International Conference on Computational Performance Evaluation (ComPE)* (pp. 049-052). IEEE.

[5] Li, J., Li, J., Zhao, X., Su, X., & Wu, W. (2023). Lightweight detection networks for tea bud on complex agricultural environment via improved YOLO v4. *Computers and Electronics in Agriculture, 211,* 107955.

[6] Chen, C., Zheng, Z., Xu, T., Guo, S., Feng, S., Yao, W., & Lan, Y. (2023). YOLO-based UAV technology: A review of the research and its applications. *Drones, 7*(3), 190.

[7] Badgujar, C. M., Poulose, A., & Gan, H. (2024). Agricultural object detection with You Only Look Once (YOLO) algorithm: A bibliometric and systematic literature review. *Computers and Electronics in Agriculture, 223,* 109090.

[8] Latha, R. S.,Sreekanth,G. R., Rajadevi, R., Nivetha, S. K., Kumar, K. A., Akash, V.,& Anbarasu, P. (2022, January). Fruits andvegetables recognitionusing YOLO. In *2022 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-6). IEEE.

[9] Junos, M. H., Mohd Khairuddin, A. S., Thannirmalai, S., & Dahari, M. (2021). An optimized YOLO-based object detection model for crop harvesting system. *IET Image Processing, 15*(9), 2112-2125.

[10] Horng, G. J., Liu, M. X., & Chen, C. C. (2019). The smart image recognition mechanism for crop harvesting system in intelligent agriculture. *IEEE Sensors Journal, 20*(5), 2766-2781.