Urban Atlas: Satellite Imagery for Urban Analysis

Ch Roshika[1], K Mamatha[2], Vijayalaxmi[3], Sk Tehazeeb[4], M Sandhya[5]
[1,2,3,4,5] UG Scholar - B Tech Information Technology BVRIT HYDERABAD College of Engineering for Women

***Abstract*—India's growth has been rapid, and we have seen a rapid shift in land use. Different taxonomies, such as houses, buildings, forests, roads, and public places, have changed a lot. Because of this rapid expansion, it is very important for us to find a good map structure and to watch and study these changes. So we came up with an idea to understand how rapidly growing cities are spreading and making better places. Our project UrbanAtlas, which is a smart system that can automatically recognise different taxonomies in a picture and make better decisions for our future. To differentiate them, we have used the YOLOv6 Segmentation model, which will separate the labels accurately based on 2 properties: threshold and overshadow. We have trained our model using LoveDa dataset, which has more than 45k pictures and contains both rural and urban pictures. We prepared our data, converted it into the appropriate format for building and testing the model, and then used Streamlit as a web platform to present the model's results. When tested, we can see some key metrics like precision, recall, validation percentage, and f1-score. this metrics help us determine how much our trained data set has learned.**

***Index Terms*—Key Terms: Urban growth, Semantic segmenta- tion, YOLOv6 segmentation, Threshold and overshadow proper- ties, Deep learning, Automated monitoring**

I. INTRODUCTION

Developing nations like India, where there is rapid urbanisa- tion, in cities like Hyderabad, Kolkata, and Mumbai, have been transformed from natural land to urban land till today [1]. to monitor all these changes, which is crucial for our sustainable development, resource allocation, and future effective city planning. For that, we have collected a large number of satellite images of a large-scale city [2]. This satellite imagery dataset is divided into 2 groups: one for urban areas and the other for rural areas. No city is completely urbanised; there might be at least a few small rural areas that cross through them. These pictures consist of buildings, rivers, oceans, seas, all kinds of water bodies, vegetative land for cultivation, and transport and road systems. Since there is a huge amount of data, it makes it difficult for a human to understand. Earlier, older methods relied on extracting the features from each tile, but there was a drawback, as each image is complex and poorly identifies all taxonomies [3]. Using deep learning methods like CNN can automatically detect spatial patterns from raw data and improve accuracy [4]. But our model used YOLO instead of CNN because it can detect and segment taxonomies properly, which offers both speed and precision [5]. Our project UrbanAtlas uses a newly released model called YOLOV26 to classify the land with minimal human effort [6]. It has only one automated workflow where it performs preprocessing, training, prediction, and visualisation. By doing this process, it supports better urban planning for people, making our environment balanced with equal trees and buildings, which will promote sustainable practices around the world.

II. EXISTING SYSTEM

Unlike older methods like CNN and U-Net, which needed manual feature extraction, deep learning automatically learns the patterns in those raw satellite images. And also, there are some struggles when object size varies, and if we have similar land categories, that will make our dataset unbalanced. But then YOLOv6 segmentation model is used for urban land-use for better classification, it is faster and gives more precise results.

III. PROPOSED SYSTEM

We collected a dataset called 'LoveDa' , which consists of Satellite images of a particular city, Hyderabad. Then, the initial processing started with data cleaning and standardising all images , so that we can put our data ready for further analysis. Since we found one of the best model, which is recently released, i.e., YOLOv26 nano, it is a super-fast AI model, which is used for automatic detection of objects, which are our taxonomies (buildings, forests, roads). Nano is a lightweight model, which is one of the versions of YOLOV26. It can process large areas very quickly without much GPU power. Then we will train YOLOV26 nano on our labelled satellite images, since it automatically does the testing process also, which reduces human effort. after that process there will be a folder of masked images, which are crucial for segmenting the objects, our model achieved accuracy of 94, and the predictions which it gave is accurate, the testing and training effiency is having low difference which means our model is considered as general, once, trained it is ready for deployment, which we have done using streamlit, then it automatically analyzes images and tell how much percent is of land use . and made urban growth tracking over time easier. The output isn't just maps—it's actionable insight for plan- ners and researchers. Next up: how we train and deploy the model.

IV. WORKING

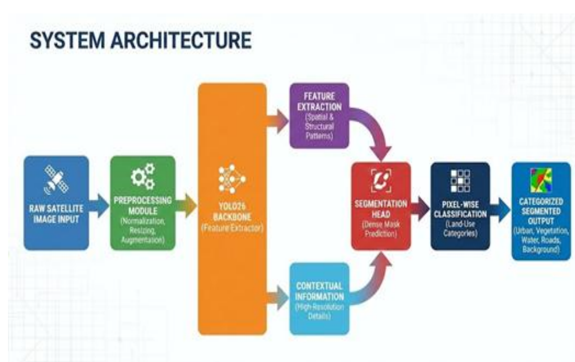
The system we proposed segments the usage of the land covered using the satellite images through a deep learning- based method called YOLO version 26, which is the YOLO26 segmentation model. We developed the complete system using Python, following a structured procedure that includes image Collection, preprocessing, model training, prediction and de- ployment through the web interface.

A. Problem Definition

Satellite images consist of huge amounts of geographical data, and manually identifying the usage of land, such as urban areas, vegetation, roads, and water bodies, is difficult and consumes a lot of time. Standard analytical approaches require high-level knowledge and cannot handle large-scale monitoring effectively, and also analysing it manually will reduce the accuracy, and there can be more errors, which can affect the reliability of the land-use classification. There can be delays in the analysis, which will again affect the decision- making during critical situations like disaster management and urban planning. So, there is a need of efficient and reliable approach to analyse satellite imagery accurately and quickly.

B. System Architecture

The Architecture begins with collecting the raw images of the satellite and given as input. These images consist of geographical data like land, water, vegetation, roads, buildings etc.Now before analysis images are prepared by normalising, resizing and augmentation. Then comes the feature extractor that is YOLOv6 this is the backbone of this system. It extracts important features from images and identifies the patterns such as edges, textures and structures. It captures the deeper spatial relationships in the image and helps us to understand how different regions are arranged. After this, the extracted features are passed for further processing, where both spatial and structural patterns are analysed in more detail. Along with this, the related details are also considered to capture the high-resolution details and understand the relationship between neighboring regions. This helps in improving the overall accuracy of the model. The processed features will be given to the segmentation head then dense mask prediction will be performed. In this stage, each pixel of the image will be assigned to a particular category based on its characteristics. After this, pixel-wise classification will be carried out in which every pixel is classified into categories. At last, the system produces a segmented output which is categorised and clearly represents the different regions in the image.



C. Dataset Description

We use LoveDA for training—a dataset made for high-res satellite segmentation, with both urban and rural scenes. It's got seven classes: background, building, road, water, barren land, forest, and agricultural land. Split into training and validation, formatted for YOLO.

For real-world performance, we tested on Hyderabad too, breaking images into tiles to handle high-res data. This mix lets us train with benchmarking and prove the model works on real city images.

D. Dataset Preparation and Organization

Images and masks sit in training and validation folders. Every image pairs with a mask that labels the land use class. This keeps things balanced and lets us check how well the model performs on fresh data.

E. Data Preprocessing

All images are resized, and annotation masks are converted for YOLO's format. We weed out bad samples and normal- ize pixel values. These steps help the model focus on real features—boundaries, textures—not random noise.

F. Model Training

So, here the YOLO based segementation model will learns to identify the different regions in the satellite images and classify them. The YOLO26 segmentation model is a pre trained model which has already learned general image fea- tures such as edges, shapes and textures so, we used this model as base model and during the training it is modified and some adjustments were made so that this model specially recognises the land usage patterns in satellite images. When the image is given as input this image passes through multiple convolutional layers where the feature extraction takes place here and the important features such as edges, patterns and spacial details will be extracted. As the data flows deeper into the layer more complex features also will be extracted.

After convolution, the image becomes maps and these maps contains where edges are and where the patterns are present so, instead of the raw image now model will see the important features. After this the model starts to predict and label each pixel as road, building, water etc. These predicted outputs are compared with the actual lables that is the ground truth masks to check the accuracy of the model. Here, we will calculate a loss function as difference between predicted output and actual ouput. Based on the loss value model will adjust its internal weights this process is called backpropogation where error is sent back through the layers and model will be correcting the error itself. This process is performed repeatedly for many iterations called epochs. In each epoch the model sees the trained images and gradually model becomes better at predicting the regions.

In our project the model is trained for 30 epochs with the image size 640 and the batch size 20. As the model is being trained the loss value will be reduced and accuracy increases. Finally the best version is saved as best.pt and it will be used later for the prediction and deployment.

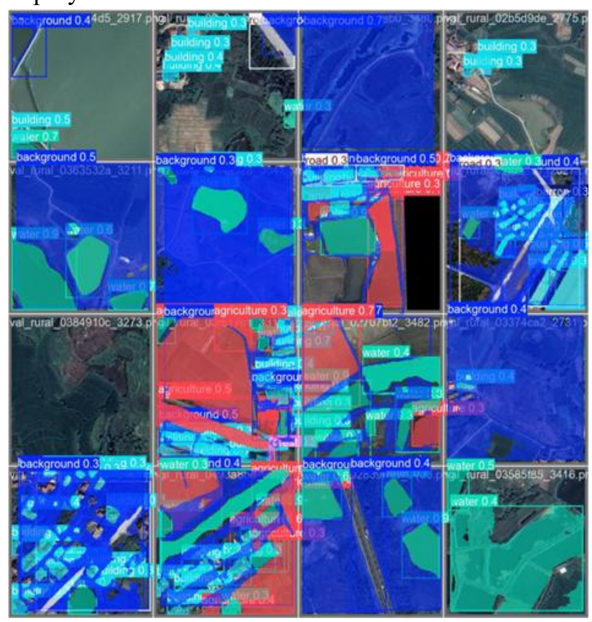


Fig. 1. Valid prediction of batch 20

G. Segmentation and Prediction Process

Once it's trained, the model's ready to label new images. Feed it a satellite photo—it spits out segmentation masks, tagging buildings, greenery, roads, water, and more. Way quicker than manual mapping.

H. Model Evaluation

After training the model we have to check how the model is performing. This can be done using the validation dataset and this dataset contains images that are not used during the training. To evaluate the performance we use few metrics and one of the main mertics is mAP (mean Average Precision) which tells how well the model is detecting the regions. Along with the mAP, precision and recall are also calculated.

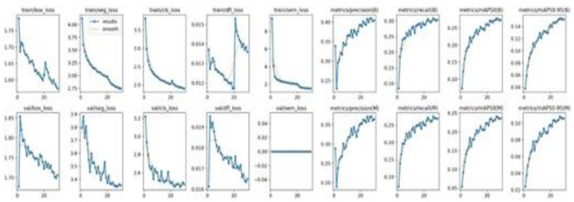


Fig. 2. Performance metrics

I. Web-Based Deployment

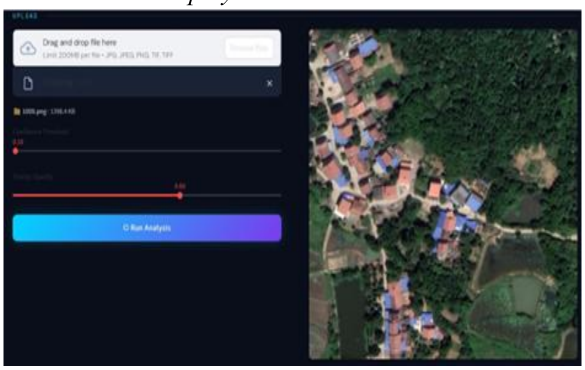


Fig. 3. Input

V. RESULTS

We evaluated the Urban Atlas system after training the YOLO26 nano segmentation model with the LoveDA dataset. Training proceeded smoothly, with loss values decreasing steadily, indicating stable learning and good convergence. The validation results demonstrated that the network effectively segmented key land-use types such as built-up areas, vegeta- tion, roads, and water bodies, all with clear spatial boundaries. These predicted masks align well with features in the satellite images. Figure 4 shows the original satellite image next to the model's segmented output for easy comparison. The masked areas highlight different urban features, allowing for quick interpretation of land distribution.

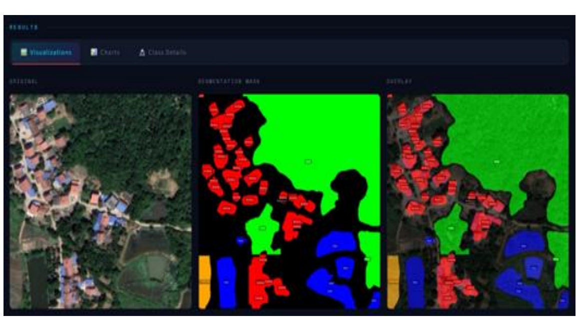


Fig. 4. Comparison between original satellite image and segmented output

We took the trained model and integrated it into a Streamlit web app, where users can upload their own satellite images and receive segmented outputs instantly in real time. Figure 5 zooms in on more prediction results, showing exactly how it classifies land-use regions into distinct segmentation masks. That setup really emphasises how practical it is for actual urban analysis. We even tested it on satellite image tiles from Hyderabad to see how it performs in real-world conditions. Sure enough, it identified the major land-use regions seam- lessly, proving it generalises well beyond the training data. Ultimately, the Urban Atlas framework provides effective, automated land-use segmentation, making it ideal for urban monitoring and planning tasks.

Fig. 5. Visualization of segmentation results across different urban regions

VI. CONCLUSION

After completion of our project, we have built a proper system that can analyse satellite images automatically and divide the taxonomies properly. By this, we can figure out how much land is used in that particular area. The model we used is YOLOV26 nano, which is a lightweight model and can segment automatically, which makes human work fast and reliable. It performs segmentation, which means it can segment different labels of the image and give a different colour to each taxonomy. The outputs of real satellite images will turn out to be annotated images. It not only works on the training dataset, but also for different cities. We have integrated using steamlit app. This helps users to upload any satellite images and get their images segmented instantly with a simple user interface.

ACKNOWLEDGMENT

The authors would like to thank their faculty guide and institution for the support and resources provided during the development of this project.

REFERENCES

[1] K. Chaurasia, R. Nandy, O. Pawar, R. R. Singh, and M. Ahire, "Semantic segmentation of high-resolution satellite images using deep learning," *Earth Science Informatics*, vol. 14, no. 4, pp. 2161–2170, 2021.

[2] Y. Xi, Y. Liu, T. Li, J. Ding, Y. Zhang, S. Tarkoma, Y. Li, and P. Hui, "A satellite imagery dataset for long-term sustainable development in united states cities," *Scientific Data*, vol. 10, no. 1, p. 866, 2023.

[3] J. Lv, Q. Shen, M. Lv, Y. Li, L. Shi, and P. Zhang, "Deep learning-based semantic segmentation of remote sensing images: a review," *Frontiers in Ecology and Evolution*, vol. 11, p. 1201125, 2023.

[4] I. Dimitrovski, V. Spasev, S. Loshkovska, and I. Kitanovski, "U-net en- semble for enhanced semantic segmentation in remote sensing imagery," *Remote Sensing*, vol. 16, no. 12, p. 2077, 2024.

[5] Ultralytics, "Ultralytics yolo," 2023.

[6] Ultralytics, "Yolo26 model documentation," 2024.