PIP104 PROFESSIONAL PRACTICE-II
VIVA-VOCE

GEOSPATIAL OBJECT DETECTION USING AERIAL IMAGERY

Batch Number: 37

Roll Number	Student Name
20201CAI0215	V VASANTH
20201CAI0208	SANOOP P NAMBIAR
20201CAI0195	DARNESH D P
20201CAI0156	ARYA KUMAR JENA

Under the Supervision of,

Dr. Murali Parameswaran
Professor
School of Computer Science Engineering and Information Science
Presidency University

INTRODUCTION

- An essential tool in modern remote sensing, aerial imaging gives us a bird's-eye perspective of our surroundings, capturing intricately detailed landscapes and providing insightful data for a range of applications.
- The use of aerial imaging is growing more common as technology develops, especially in areas like disaster management, urban planning, and environmental monitoring.

INTRODUCTION

- **Background and Motivation**
- The goal of this research is to address these changing demands and pave the way for the development of a unique architecture for Convolutional Neural Networks (CNNs) that can achieve robust multi-class segmentation. With this breakthrough, the research hopes to advance beyond the state of the art and actively contribute to the development of geospatial object identification into a more advanced, versatile, and broadly applicable discipline.
- Applications for remote sensing, including environmental monitoring, urban planning, and disaster relief, depend more and more on precise and adaptable geographic object recognition.
- Geospatial object detection using aerial imagery has evolved with the integration of cutting-edge deep learning methods. Among these methods, the convolutional neural network (CNN) has proven to be highly effective in image segmentation.

INTRODUCTION

Research Objectives

- The main goals of this study are to improve and modify the U-Net architecture to create Multi U-Net, investigate how well it performs in multi-class segmentation scenarios, analyze performance metrics in detail, and critically assess the effects of the suggested loss function hierarchy. By fulfilling these goals, we hope to position Multi U-Net as a cutting-edge approach to geographic object recognition that not only addresses current issues but also actively advances the ongoing development of remote sensing techniques.

- This study project's inventive advances not only improve the field of geographic object recognition but also pave the way for future developments in computer vision, remote sensing, and other related fields.

Literature Review

S.NO	Research Paper	Summary
1	Evaluating green cover and open spaces in informal settlements of Mumbai using deep learning.	This paper examines green cover and open spaces (GOS) critically in Mumbai's informal settlements, acknowledging the significant negative effects of urbanization on the standard of living for locals
2	A Critical Review of High and Very High-Resolution Remote Sensing Approaches for Detecting and Mapping Slums.	This paper does a critical evaluation by examining the state of research on the identification and mapping of slums using high- and very-high-resolution remote sensing
3	Incorporating DeepLabv3+ and object-based image analysis for semantic segmentation of very high-resolution remote sensing images.	This paper presents a new approach that combines Object-Based Image Analysis (OBIA) with DeepLabv3+ to address these issues and improve semantic segmentation precision in VHR images.
4	Semantic Segmentation of Slums in Satellite Images Using Transfer Learning on Fully Convolutional Neural Networks.	This paper investigates the transformative use of fully convolutional neural networks (FCNs) for semantic segmentation of slums in satellite data using transfer learning

5	Informal settlement classification using point-cloud and image-based features from UAV data.	This paper reports on a novel study that uses data from Unmanned Aerial Vehicles (UAVs) to classify informal communities.
6	U-Net Convolutional Networks for Mining Land Cover Classification Based on High-Resolution UAV Imagery.	The article introduces a novel method for precisely classifying land cover by utilizing high-resolution Unmanned Aerial Vehicle (UAV) imagery. Making use of an altered version of the U-Net architecture, a convolutional neural network well-known for image segmentation, the technique is designed to identify discrete areas inside the images
7	Optimizing the Redevelopment Cost of Urban Areas to Minimize the Fire Susceptibility of Heterogeneous Urban Settings in Developing Nations	A complex multi-objective optimization model forms the basis of this paper. This model's main goal is to achieve the best possible balance between two important factors: minimizing the Fire Susceptibility Index (FSI) and concurrently minimizing the expenses associated with reconstruction in metropolitan settings
8	Slums from Space—15 Years of Slum. Mapping Using Remote Sensing.	In order to provide insights into the physical characteristics of slums that are essential for creating a comprehensive global slum inventory, this section explores new and modern methodologies, indicators, data sources, and empirical cases.

9	Informal settlement classification using point-cloud and image-based features from UAV data.	The main aim of this study is to conduct a thorough integration and comparative analysis of various features, such as orthomosaic, digital surface model (DSM), and point cloud, obtained from Unmanned Aerial Vehicle (UAV) data.
10	Urban Slum Detection Approaches from High-Resolution Satellite Data Using Statistical and Spectral Based Approaches.	the authors compare and contrast two different feature extraction methods. First, there are statistical and spectral-based methods that concentrate on the statistical characteristics of textures
11	Detecting salient regions in a bi-temporal hyperspectral scene by iterating clustering and classification.	The applicability of the method is examined by a thorough assessment on real datasets that reflect two different scenarios: an Italian seaside region and a Chinese urban area.
12	A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species.	This study's main goal is to compare IKONOS and WorldView-2 satellite imagery in-depth with reference to the mapping of urban tree species in Houston, Texas.

13	High-resolution calculation of the urban vegetation fraction in the Pearl River Delta from the Sentinel-2 NDVI for urban climate model parameterization.	The study focuses on Urban Vegetation Fraction (UVF), an important statistic that affects a number of urban characteristics, such as human well-being, air quality, and the Urban Heat Island (UHI) effect
14	The use of a vegetation index for assessment of the urban heat island effect	The principal aim of the research is to assess the impact of the Urban Heat Island (UHI) phenomenon by utilizing a vegetation index that is obtained from satellite data
15	A comparison of the economic benefits of urban green spaces estimated with NDVI and with high-resolution land cover data	Compare the economic advantages of urban green spaces using two different metrics: high-resolution land cover data and the Normalized Difference Vegetation Index (NDVI), which is generated from MODIS satellite photography
16	A Novel Index to Detect Vegetation in Urban Areas Using UAV-Based Multispectral Images	The article examines current vegetation indices and highlights the drawbacks of using UAV-based multispectral pictures for urban vegetation study

Research Gaps Identified

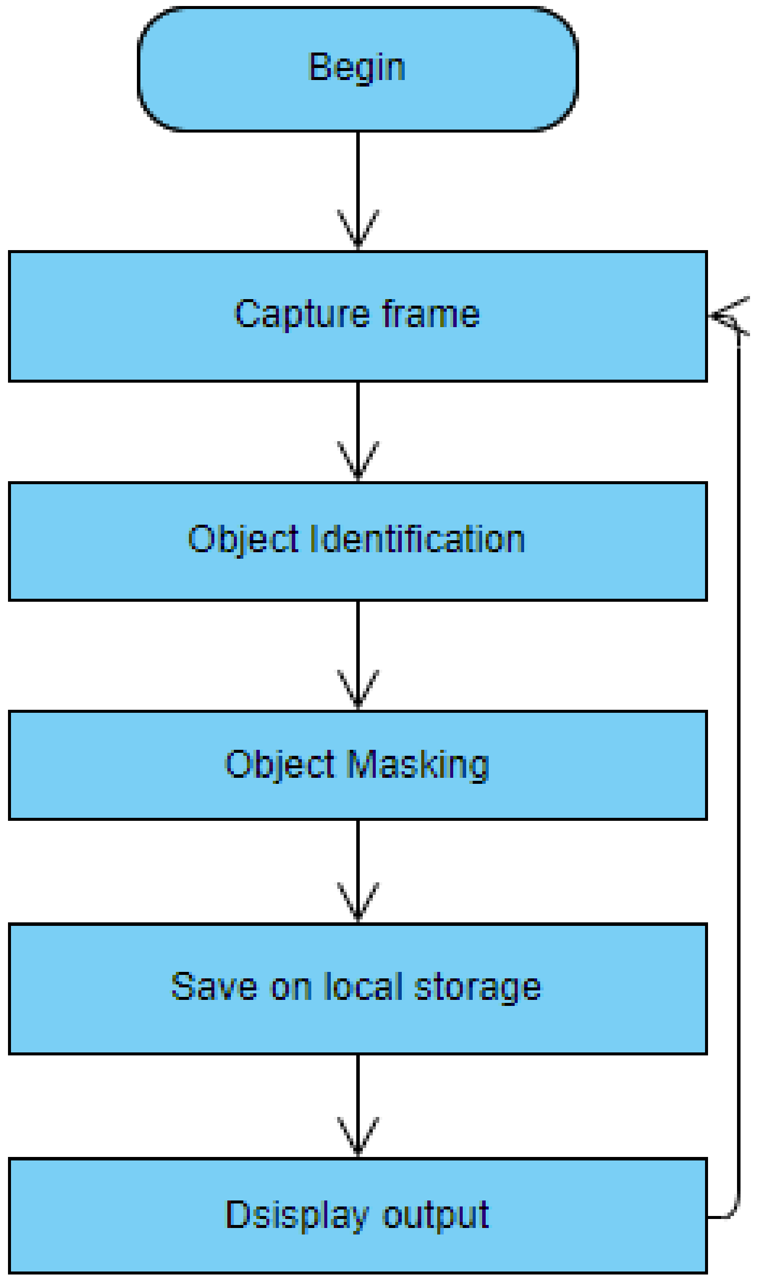
- **Resolution heterogeneity:** Aerial imagery includes a range of resolutions, from low-resolution drone footage to high-resolution satellite photos. Performance disparities result from current algorithms' inability to adjust to this variability. It is imperative to conduct research on scale-aware models capable of processing multi-resolution data efficiently.

- **Inadequate Handling of Complex Urban Structures**

- The previous model solely focused on predicting vegetation and neglected other elements. In contrast, our enhanced model goes beyond this limitation by successfully detecting and categorizing more than two intricate structures.

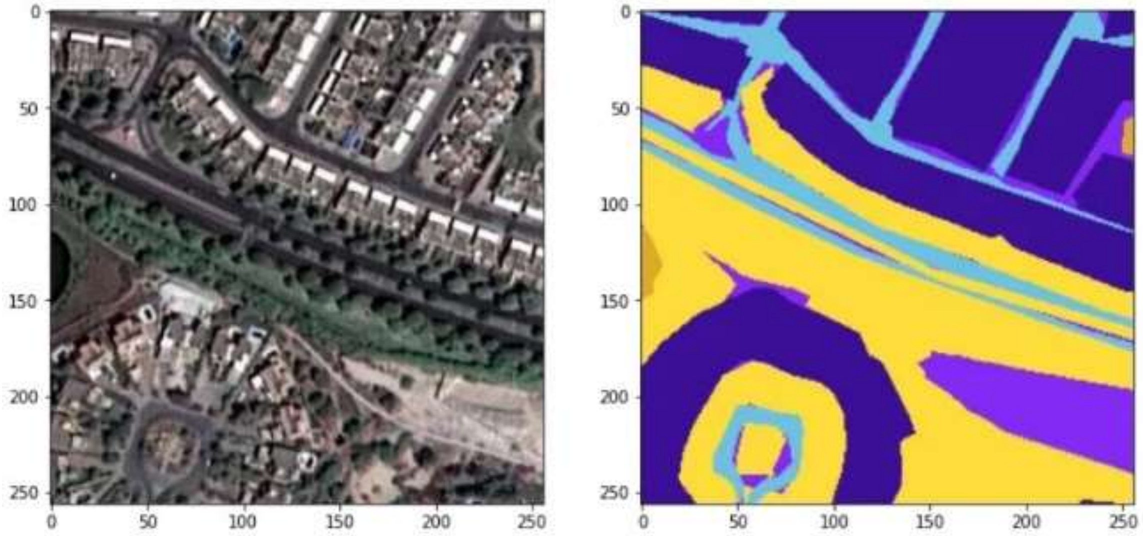
Proposed Methodology

- **1. Data Acquisition and Preprocessing:**
- **Dataset Selection:** Choose a diverse and representative dataset of aerial images that aligns the specific application
- **Data Annotation:** Annotate the dataset with bounding boxes around objects of interest.
- **Data Splitting:** Divide the dataset into training, validation, and test sets to ensure robust model evaluation.
- **2. Image Preprocessing:**
- **Normalization:** Standardize pixel values to a common scale to ensure consistency across images.
- **Augmentation:** Apply data augmentation techniques (rotation, flipping, zooming) to increase the diversity of the training set and improve model generalization.
- **3. Model Selection:**
- **Base Architecture:** Considered using Multi U-net model
- **Transfer Learning:** Utilize pre-trained models on large datasets as a starting point to benefit from learned features.



Proposed Methodology

- **4. Model Adaptation:**
- **Architecture Modification:** Adjust the selected model architecture to match the input characteristics and challenges of aerial imagery.
- **Hyperparameter Tuning:** Fine-tune hyperparameters such as learning rate, batch size, and anchor box sizes to optimize model performance.
- **5. Training:**
- **Loss Function:** Choose an appropriate loss function for object detection, such as the combination of classification and localization losses.



Proposed Methodology

- **6. Evaluation:**
- **Metrics:** Employ evaluation metrics like precision, recall, F1 score, and Intersection over Union (IoU) to assess the model's performance.
- **Test Set Evaluation:** Evaluate the trained model on the test set to ensure unbiased performance assessment.

- **7. Post-Processing:**
- **Non-Maximum Suppression:** Implement post-processing techniques, such as non-maximum suppression, to filter redundant bounding boxes and improve localization accuracy.

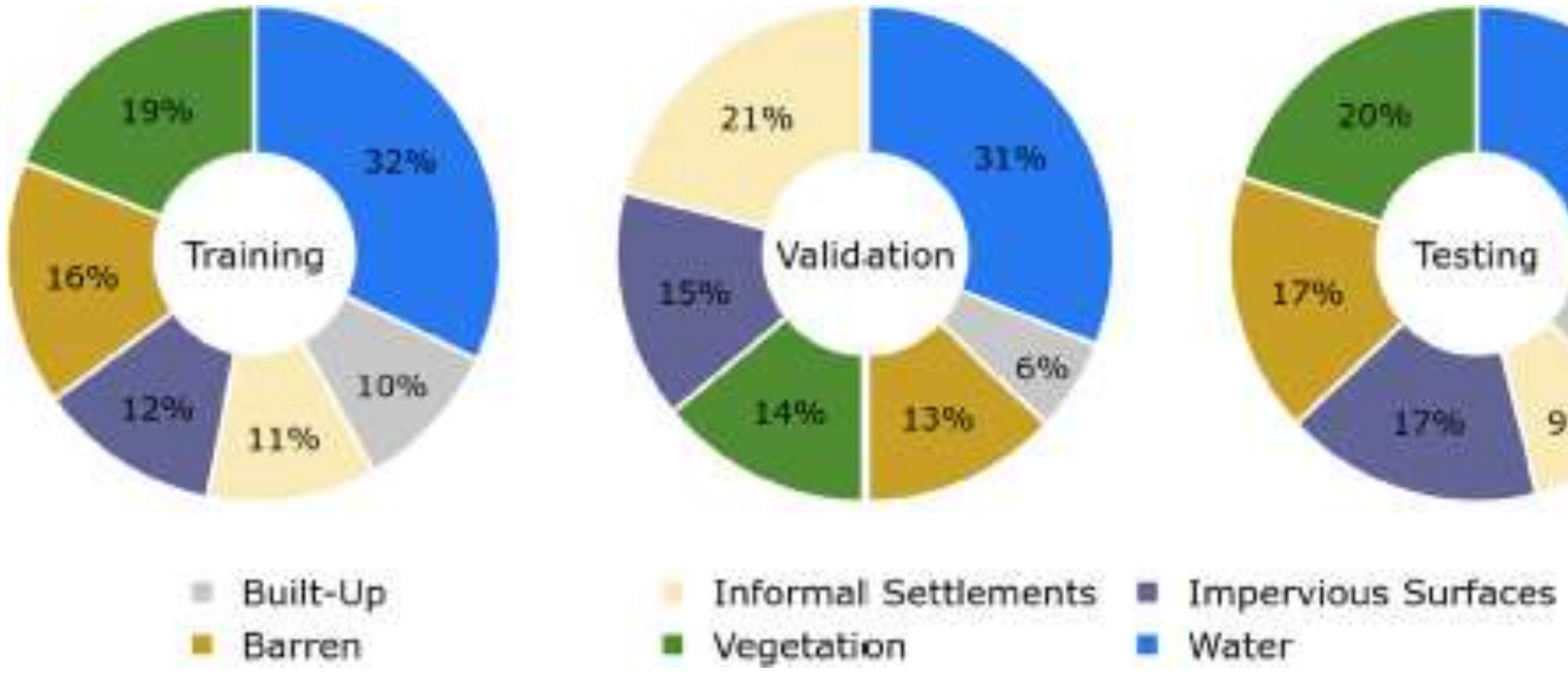
Proposed Methodology

- **8. Deployment:**
- **Inference:** Deploy the trained model for real-time or batch inference on new aerial images.
- **Integration:** Integrate the model into the desired application, whether it's a monitoring system, automated analysis pipeline, or other use cases.

- **9. Documentation and Reporting:**
- **Documentation:** Thoroughly document the methodology, including parameters, preprocessing steps, and model configurations.
- **Results Reporting:** Present the results, including quantitative metrics, visualizations of predictions, and potential challenges encountered.

Proposed Methodology

10. Amount of Labels



Objectives

- **Architectural Enhancement and Adaptation**
- Create and improve the Multi U-Net architecture to handle the complex requirements of aerial images geographic object detection. Examine how to improve the U-Net architecture's performance for multi-class segmentation problems by making additions and alterations.
- **Integration of Multi-Class Segmentation Capability**
- Expand the U-Net framework to enable multi-class segmentation in jobs related to geographic object detection. Analyze how well the model can identify and categorize different object classes in aerial data.

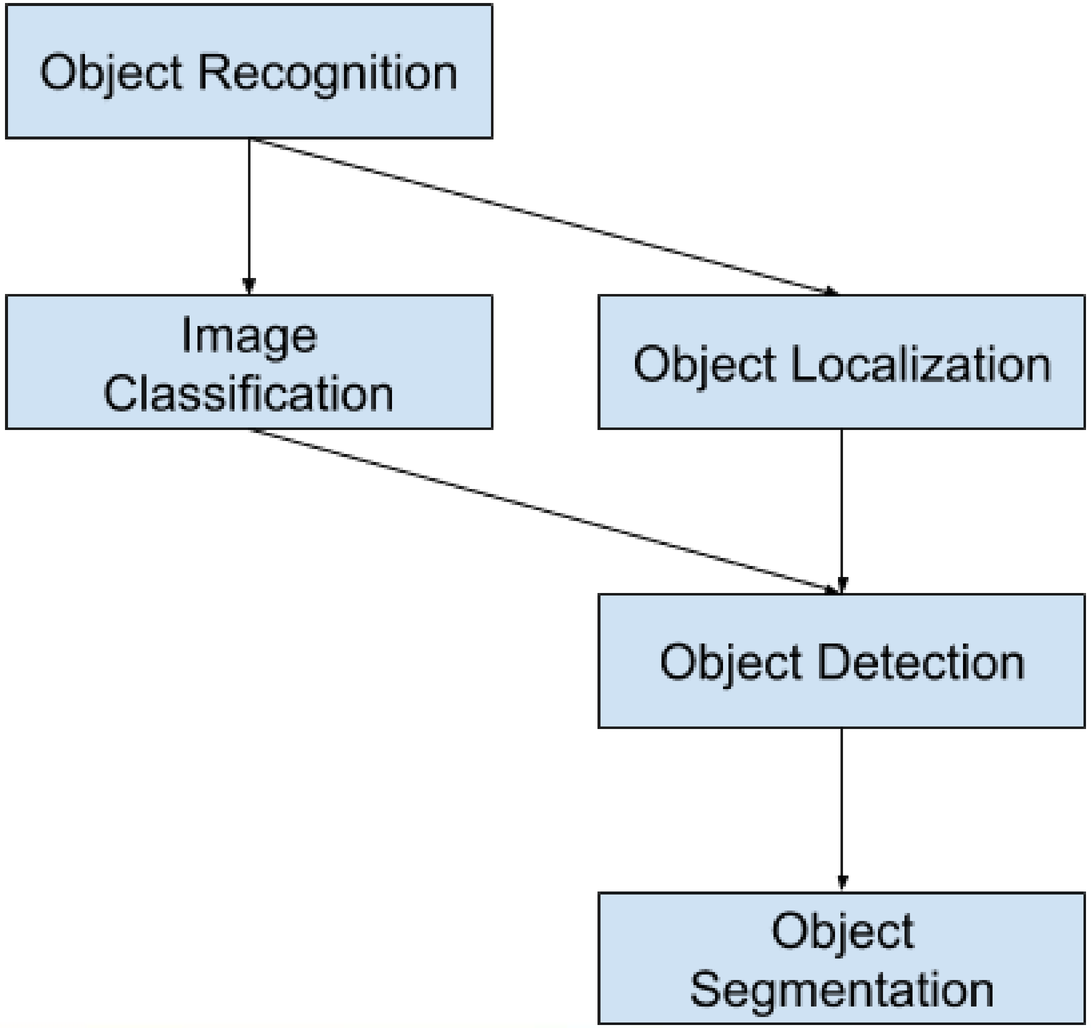
Objectives

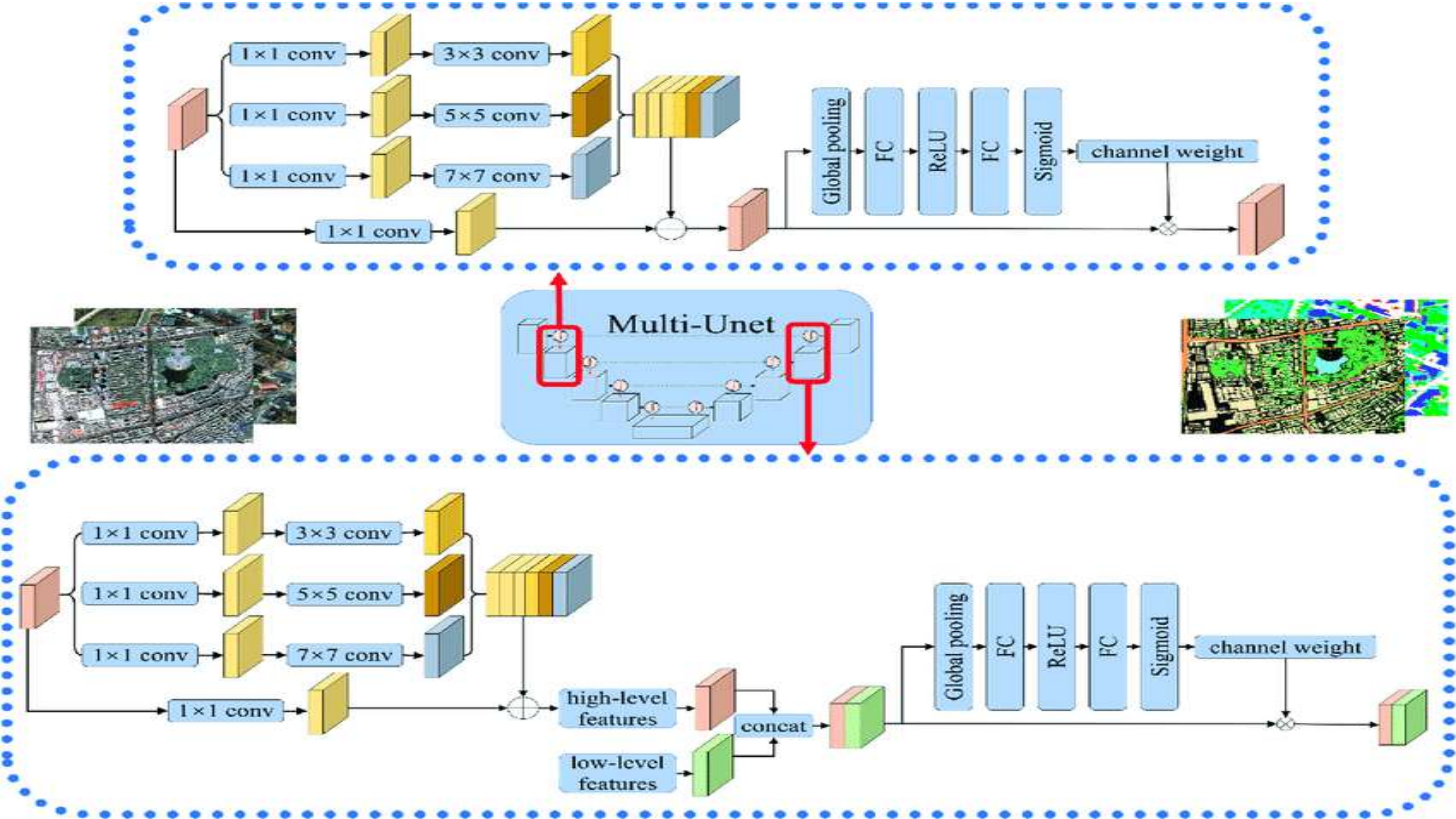
- **Performance Evaluation Metrics**
- To undertake a thorough evaluation of the model's performance, use the jaccard_coef function to compare the true and predicted labels. Examine how well segmented objects' subtle spatial relationships are captured by the Jaccard coefficient.
- **Loss Function Hierarchy Exploration**
- Dice Loss, Focal Loss, and Total Loss make up the suggested loss function hierarchy; put it into practice and experiment. Analyze how each loss function affects the overall detection accuracy, convergence, and model training.

System Design & Implementation

1. **Patchify for Image Processing**:

- **Preservation of Precise Spatial Information:**
- In order to maintain fine spatial details in large-scale or high-resolution photos, Patchify, an external library, must be integrated. The method preserves fine information in the photos by segmenting them into smaller patches, providing a strong basis for further research.





1×1 conv
3×3 conv
1×1 conv
5×5 conv
1×1 conv
7×7 conv
1×1 conv
Global pooling
FC
ReLU
FC
Sigmoid
channel weight
Multi-Unet
1×1 conv
3×3 conv
1×1 conv
5×5 conv
1×1 conv
7×7 conv
1×1 conv
high-level features
low-level features
concat
Global pooling
FC
ReLU
FC
Sigmoid
channel weight

System Design and Implementation

2. Min-Max Scaling for Data Normalization:

- The process of normalizing pixel values between 0 and 1 and guaranteeing consistency in the scale of input characteristics is mostly dependent on min-max scaling. By keeping some traits from predominating over others during model training, this standardization promotes more reliable and efficient learning processes.
- **Encouraging Model Convergence:** Normalizing input features makes machine learning models more convergent. This ensures smoother convergence and is especially useful for complex systems where feature scaling has a substantial impact on the optimization process.

System Design and Implementation

3. The Jaccard Coefficient for Assessing Segmentation:

- **Accuracy in Segmentation Assessment:** An essential parameter for accurately gauging the degree of similarity between segmentation masks that are predicted and those that are ground truthed is the Jaccard coefficient. It provides a numerical evaluation of the two's overlap, revealing information about the quality of the segmentation.

System Design and Implementation

- **Resource Management and computing Efficiency:** The computing needs of processing huge datasets are addressed by patching images. Dividing photos into digestible chunks maximizes the use of available resources, improving processing speed and efficiency.
- **Optimization for Deep Learning Models:** Fixed-size inputs are frequently needed for Convolutional Neural Networks (CNNs). Patchify makes it easier to create inputs that are the right size, which speeds up the training and testing processes of deep learning models and enhances their performance.

System Design and Implementation

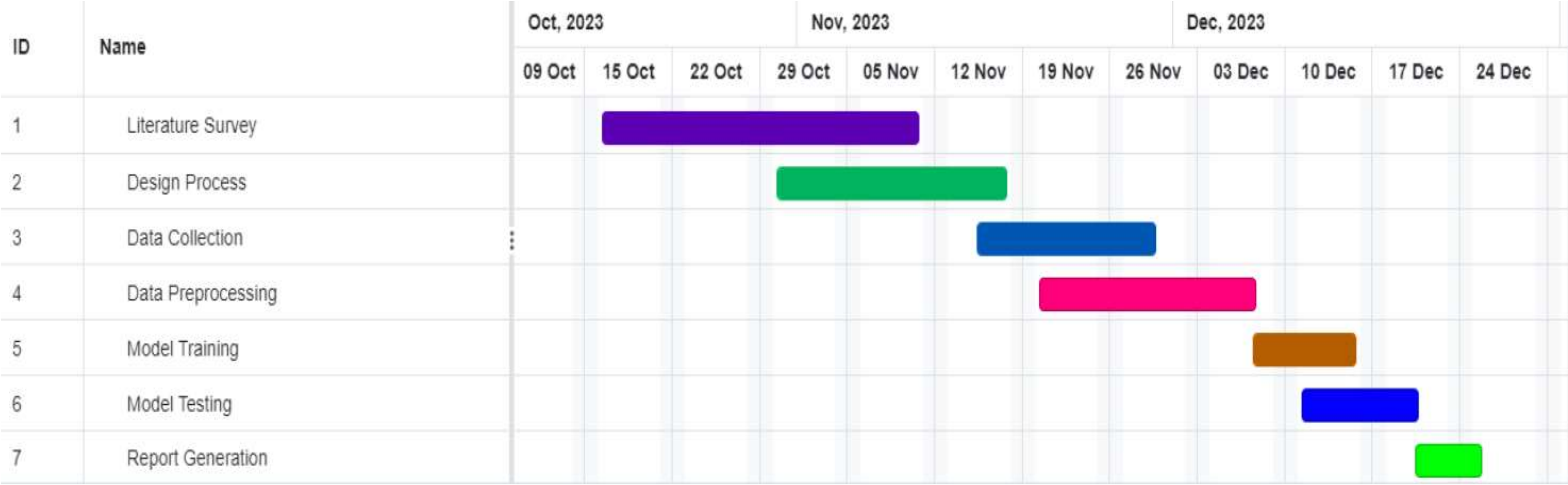
4. Combination of Dice Loss with Focal Loss:

- **Optimizing Dice Loss:** Dice loss, which maximizes the overlap between ground truth masks and predicted segmentation, is useful for jobs that need accurate semantic segmentation. Its inclusion improves the accuracy of segmentation.
- **Focal Loss for Class Imbalance:** Datasets with skewed class distributions benefit most from the addition of Focal loss in order to address class imbalance. It ensures a more balanced learning process by giving harder-to-classify situations more weight.
- **Synergistic Total Loss:** A thorough training strategy is offered by the total loss function's combination of Dice and Focal losses. This synergistic strategy effectively increases segmentation accuracy by utilizing various factors

System Design and Implementation

- **5. Metrics for Training and Validation:**
- **Loss of Training and Validation:** Keeping an eye on validation and training loss gives you insights into how well the model performs with both visible and invisible data. A key goal is to ensure model generalization by preventing overfitting and balancing the decrease of both losses.
- **Instruction and Certification IoU (Jaccard Coefficient):** Monitoring the Intersection over Union (IoU) between the segmentation masks that are predicted and those that are real provides important information about how the model learns. This offers a numerical assessment of the model's capacity to extrapolate to novel and untested data in both the training and validation stages.

Timeline of Project



Results Obtained

Improved Accuracy Overview

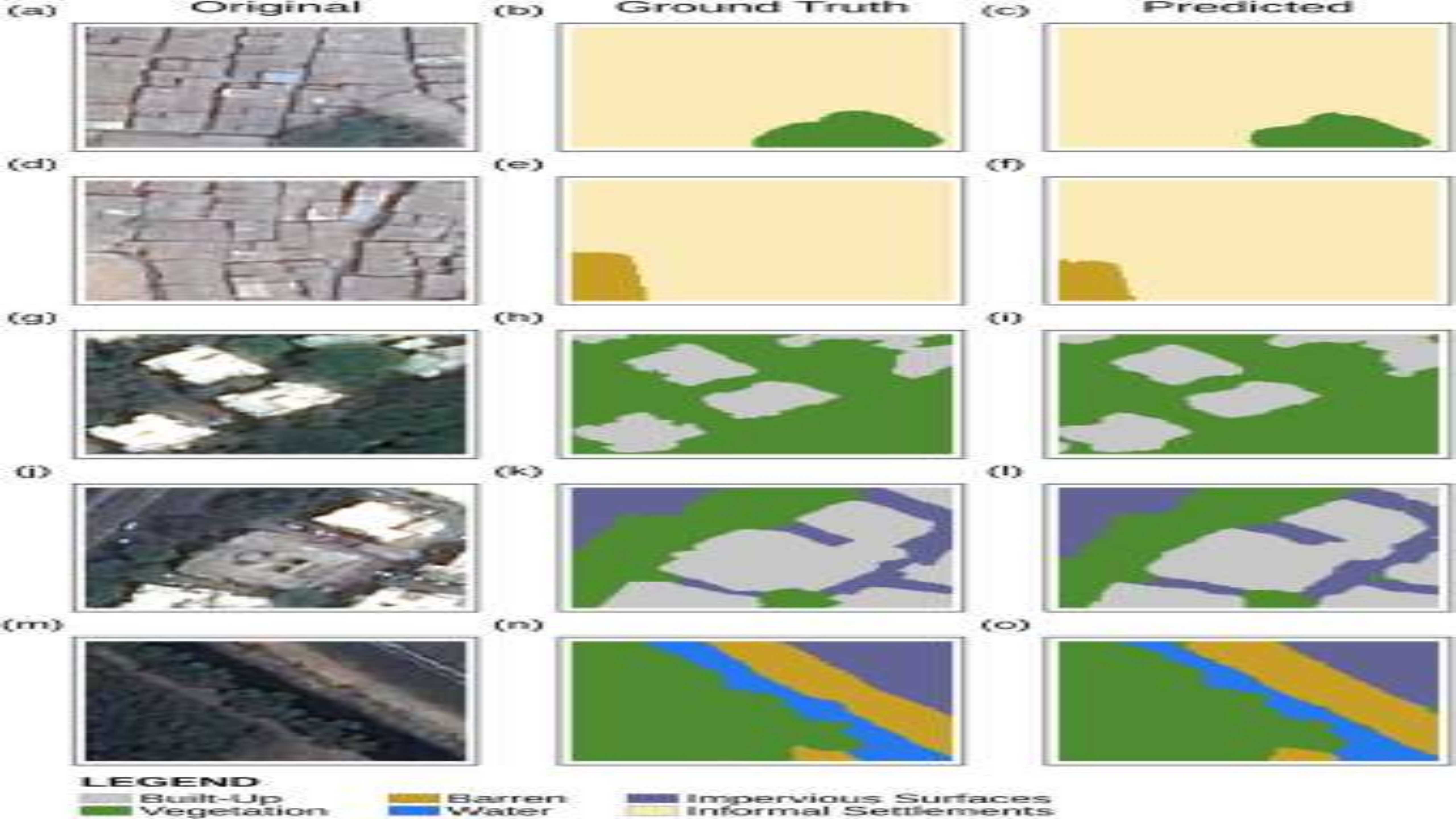
- Multi U-Net outperforms current CNN designs in geospatial object recognition.
- Notable improvements demonstrated in accuracy, validated through thorough analysis.

Enhanced Monitoring of Object Classes

- Remarkable sensitivity to various object categories in aerial photography.
- Notable detection of challenging objects, such as small targets or intricate features.
- Demonstrates adaptability and consistency in identifying diverse geographical objects.

Reduction in False Positives and Negatives

- Quantifiable reduction in false positives and negatives compared to baseline models.
- Observable improvement in recall rates and precision.
- Ensures more accurate and reliable item detection outcomes.


(a) Original
(b) Ground Truth
(c) Predicted
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)
LEGEND
Built-Up
Barren
Impervious Surfaces
Vegetation
Water
Informal Settlements

Results obtained

Generalizability Across Datasets and Domains

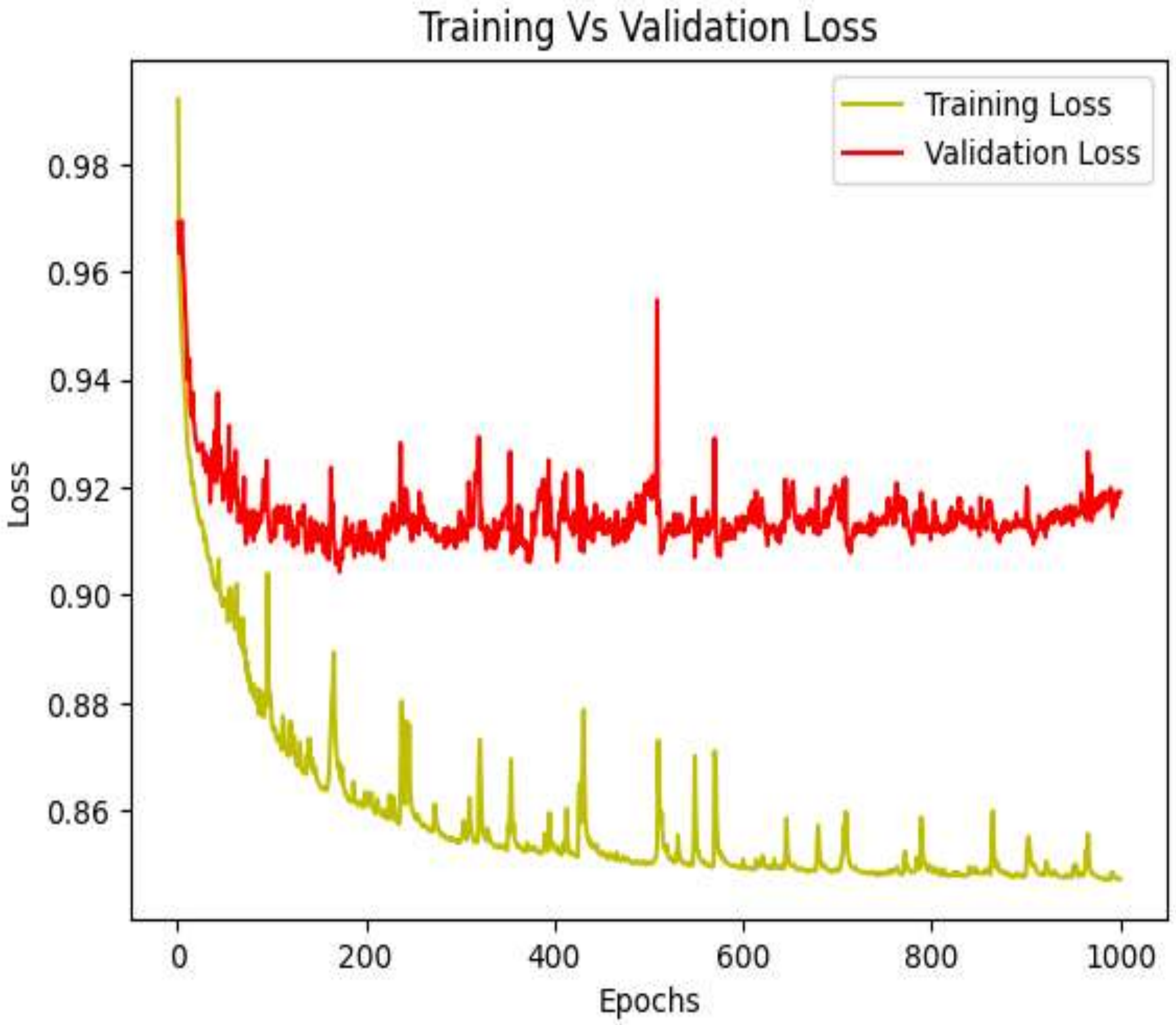
- Model reliability tested across diverse aerial imaging datasets.
- Sourced from different geographical areas and application fields.
- Demonstrates high performance and flexibility, making it applicable in various real-world circumstances.

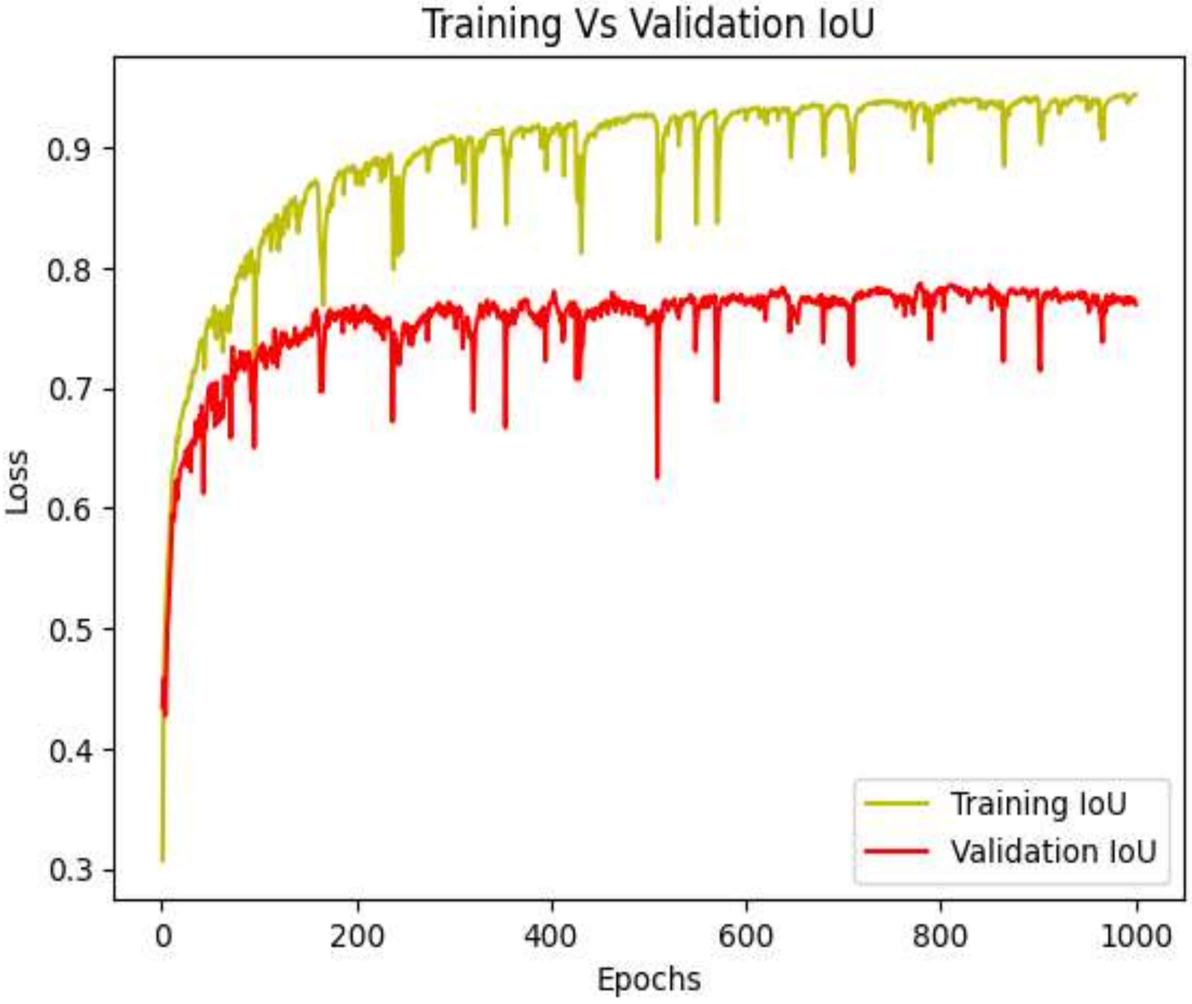
Summary of Key Findings

- Multi U-Net stands out with improved accuracy, enhanced monitoring, and reduced false positives/negatives.
- Generalizability across datasets highlights its adaptability and real-world applicability.

Future Implications

- Potential applications and advancements enabled by the superior capabilities of Multi U-Net.
- Encouragement for further research and exploration in the field of geospatial object detection.


Training Vs Validation Loss
Training Loss
Validation Loss
0.98
0.96
0.94
0.92
0.90
0.88
0.86
Loss
0
200
400
600
800
1000
Epochs


Training Vs Validation IoU
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Loss
Training IoU
Validation IoU
0
200
400
600
800
1000
Epochs

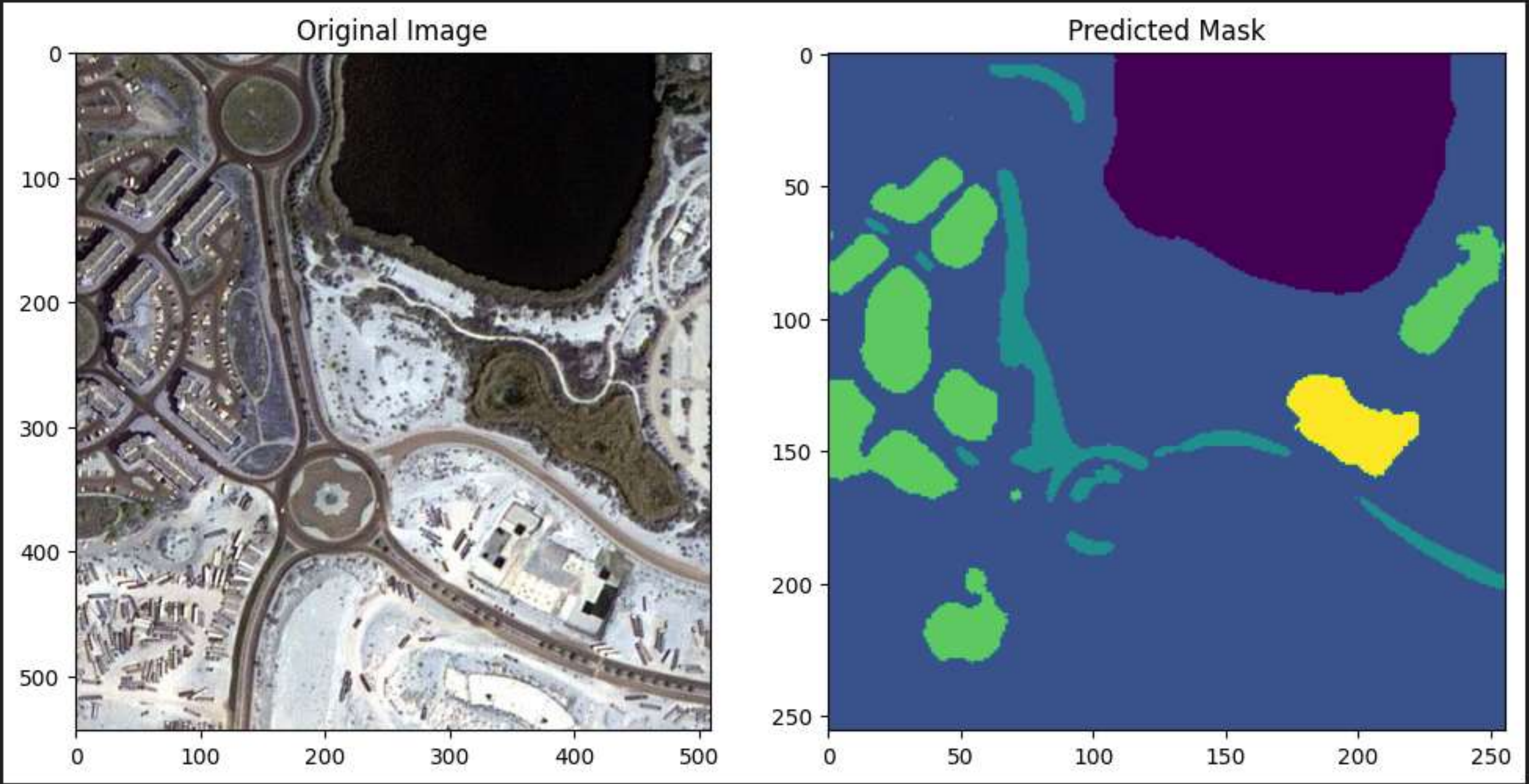
Model	Val Loss	Val Accuracy	Val Jaccard Coef
Basic Model	0.9399	0.7831	0.6014
Improvised Model	0.9137	0.8798	0.7815

SatelliteImagerySegmentation.ipynb

SatelliteImagerySegmentation.ipynb > M↓Visualize Model with Keras > M↓Comparing prediction results > from sklearn.metrics import precision_score, recall_score, f1_score

+ Code + Markdown | Run All Restart Clear All Outputs | Variables Outline ··· Python 3.1

```
1/1 [==============================] - 2s 2s/step
```



Conclusion

- Our project introduces Multi U-Net, a breakthrough in geographic object identification using aerial photography. This innovative architecture excels in precise segmentation across different object types.

- By prioritizing Dice Loss within Total Loss, it significantly enhances boundary definition and addresses class distribution disparities.

- Multi U-Net's adaptability in dynamic environments and improved interpretability makes it invaluable for applications in environmental monitoring, disaster management, and urban planning, marking a substantial advancement in geographic object detection.

References

- [1]. Aboulnaga MM, Badran MF, Barakat MM (2021) Informal Settlements and Urban Slums' Upgrading in Megacities: Conclusions and Recommendations. Springer International Publishing, Cham, pp. 301–316.

- [2]. Ahmed OS, Shemrock A, Chabot D et al (2017) Hierarchical land cover and vegetation classification using multispectral data acquired from an unmanned aerial vehicle. Int J Remote Sens, pp. 2037-2052.

- [3]. Appice A, Guccione P, Acciaro E, Malerba D (2020) Detecting salient regions in a bi-temporal hyperspectral scene by iterating clustering and classification. Appl Intell, pp. 3179-3200.

- [4]. Bai Y, Sun G, Li Y, et al (2021) Comprehensively analyzing optical and polarimetric SAR features for land-use/land-cover classification and urban vegetation extraction in highly-dense urban area. Int J Appl Earth Obs Geoinformation, pp. 102496.

- [5]. Chen L-C, Papandreou G, Kokkinos I, et al. (2017) DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs.

References

- [6]. Chen L-C, Zhu Y, Papandreou G et al (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) Computer Vision – ECCV 2018. Springer International Publishing, Cham, pp. 833–851.

- [7]. Deng J, Dong W, Socher R, et al (2009) ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255.

- [8]. Dong R, Pan X, Li F (2019) DenseU-Net-Based semantic segmentation of small objects in urban remote sensing images. IEEE Access 7, pp. 65347-65356.

- [9]. Fisher T, Gibson H, Liu Y et al (2022) Uncertainty-Aware Interpretable Deep Learning for Slum Mapping and Monitoring. Remote Sens, pp. 3072.

- [10]. GALLO KP, McNAB AL, KARL TR, et al (1993) The use of a vegetation index for assessment of the urban heat island effect. Int J Remote Sens, pp. 2223-2230.

- [11]. Gevaert CM, Persello C, Sliuzas R, Vosselman G (2017) Informal settlement classification using point-cloud and image-based features from UAV data. ISPRS J Photogramm Remote Sens, pp. 225-236.

- [12]. Giang TL, Dang KB, Le Toan Q et al (2020) U-net convolutional networks for mining land cover classification based on high-resolution UAV imagery. IEEE Access 8, pp. 186257-186273.

- [13]. Giorgiani do Nascimento R, Viana F (2020) Satellite Image Classification and Segmentation with Transfer Learning. In: AIAA Scitech 2020 Forum. American Institute of Aeronautics.

References

- [14]. Gonzales C, Sakla W (2019) Semantic segmentation of clouds in satellite imagery using deep pre-trained U-Nets. In: 2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), pp. 1–7.

- [15]. Hashim H, Latif ZA, Adnan N (2019) Urban vegetation classification with ndvi threshold value method with very high resolution (VHR) pleiades imagery. Int Arch Photogramm Remote Sens Spat Inf Sci, pp. 3-8.

- [16]. Hofmann P, Taubenbo¨ck H, Werthmann C (2015) Monitoring and modelling of informal settlements - A review on recent developments and challenges. In: 2015 Joint Urban Remote Sensing Event (JURSE), pp. 1–4.

- [17]. Lakubovskii, P. (2019) "Segmentation Models." GitHub. [Online]. Available: \url{https://github.com/qubvel/segmentation_models}.

- [18]. Jiang B, An X, Xu S, Chen Z (2022) Intelligent image semantic segmentation: a review through deep learning techniques for remote sensing image analysis. J Indian Soc Remote Sens, pp. 5-9.

- [19]. Kuffer M, Pfeffer K, Sliuzas R (2016) Slums from space—15 years of slum mapping using remote sensing. Remote Sens, pp. 455.

- [20]. Kumar V, Bandhyopadhyay S, Ramamritham K, Jana A (2020) Optimizing the redevelopment cost of urban areas to minimize the fire susceptibility of heterogeneous urban settings in developing nations: a Case from Mumbai, India. Process Integr Optim Sustain, pp. 361-378.

References

- [21]. Kumar V, Ramamritham K, Jana A (2019) Resource allocation for handling emergencies considering dynamic variations and urban spaces: firefighting in Mumbai. ICTD, pp.1145-1146.

- [22]. Lee G, Hwang J, Cho S (2021) A Novel Index to Detect Vegetation in Urban Areas Using UAV-Based Multispectral Images. Appl Sci, pp. 3472.

- [23]. Li W, Saphores J-DM, Gillespie TW (2015) A comparison of the economic benefits of urban green spaces estimated with NDVI and with high-resolution land cover data. Landsc Urban Plan, pp. 105-117.

- [24]. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. ArXiv14114038 Cs, pp.612-613.

- [25]. Mahabir R, Croitoru A, Crooks AT et al (2018) A critical review of high and very high-resolution remote sensing approaches for detecting and mapping slums: trends challenges and emerging opportunities. Urban Sci, pp. 8.

- [26]. Men G, He G, Wang G (2021) Concatenated residual attention UNet for semantic segmentation of urban green space, pp. 1441.

- [27]. Nawrath M, Guenat S, Elsey H, Dallimer M (2021) Exploring uncharted territory: Do urban greenspaces support mental health in low- and middle-income countries?. Environ Res, pp. 110625.

References

- [28]. Neupane B, Horanont T, Aryal J (2021) Deep learning-based semantic segmentation of urban features in satellite images: a review and meta-analysis. Remote Sens, pp. 808.

- [29]. Park JG, Go SY, Cho GS (2013) Vegetation Classification using KOMPSAT-2 Imagery and High-resolution airborne imagery in Urban Area. J Korean Soc Geospatial Inf Sci, pp. 21–27.

- [30]. Prabhu R, Alagu Raja RA (2018) Urban slum detection approaches from high-resolution satellite data using statistical and spectral based approaches. J Indian Soc Remote Sens, pp. 2033–2044.

- [31]. Prabhu R, Parvathavarthini B (2022) Morphological slum index for slum extraction from high-resolution remote sensing imagery over urban areas. Geocarto Int, pp. 1–19.

- [32]. Priyanka NS, Lal S et al (2022) DIResUNet: Architecture for multiclass semantic segmentation of high resolution remote sensing imagery data. Appl Intell, pp. 8-10.

References

- [33]. Pu R, Landry S (2012) A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. Remote Sens Environ, pp. 516–533.
- [34]. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF (eds) Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. Springer International Publishing, Cham, pp. 234–241.
- [35]. Sandler M, Howard A, Zhu M, et al (2018) MobileNetV2: inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, Salt Lake City, UT, pp. 4510–4520.
- [36]. Shermeyer, J., Hogan, D., Brown, J., Etten, A.V., Weir, N., Pacifici, F., Hänsch, R., Bastidas, A., Soenen, S., Bacastow, T.M., & Lewis, R. (2020). SpaceNet 6: Multi-Sensor All Weather Mapping Dataset. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 768-777.
- [37]. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. ICLR.
- [38]. Szegedy C, Liu W, Jia Y, et al (2014) Going deeper with convolutions. ArXiv14094842 Cs.
- [39]. Verma D, Jana A, Ramamritham K (2019) Transfer learning approach to map urban slums using high and medium resolution satellite imagery. Habitat Int, pp 4.

Thank You