# Aerial Image Semantic Segmentation

Chinatip Lawansuk

Department of Electrical Engineering and Computer Science

National Taipei University of Technology

Email: t112998405@ntut.edu.tw

*Abstract*—**Aerial image semantic segmentation is a crucial task in remote sensing and geographic information systems (GIS). This report provides an overview of the methods and techniques used in semantic segmentation of aerial images, discusses the challenges involved, and highlights recent advancements in the field. The report also covers the evaluation metrics commonly used to assess the performance of segmentation models.**

## I. INTRODUCTION

Surveying land usage is critical for sustainable development, environmental conservation, and efficient resource management. Traditional land surveying techniques, while foundational, are often limited by their scalability, accuracy, and reliance on extensive human intervention. The advent of aerial imagery has revolutionised this field, offering high-resolution and comprehensive coverage of vast areas, significantly enhancing the efficiency and accuracy of land surveys. However, interpreting these images using traditional algorithms, such as manual interpretation, supervised and unsupervised classification, edge detection, and segmentation, poses challenges including high labour costs, human error, and sensitivity to data variability.

In recent years, deep learning, particularly convolutional neural networks (CNNs), has emerged as a powerful tool for image segmentation. Among these, the U-Net architecture stands out for its ability to combine high-level contextual information with precise localisation, making it particularly effective for complex image segmentation tasks. U-Net's symmetric encoder-decoder structure, augmented with skip connections, facilitates the retention of fine-grained details, which is crucial for accurately delineating land features in aerial imagery. Enhanced variants such as Deep U-Net further improve segmentation accuracy by incorporating deeper networks and advanced techniques like attention mechanisms and multi-scale processing.

This work explores the application of U-Net and its deep variants in the segmentation of aerial images, demonstrating how these architectures address the limitations of traditional methods. Additionally, we focus on the critical aspects of data preparation and preprocessing, which are essential for improving the performance of deep learning models. Proper data preparation, including tasks such as noise reduction, image enhancement, and augmentation, ensures that the models are trained on high-quality datasets, thereby enhancing their robustness and generalisation capabilities.

By leveraging the strengths of deep learning and thorough data preparation, this approach aims to enhance the precision, scalability, and efficiency of land use surveys, ultimately contributing to better-informed decision-making in urban planning, environmental monitoring, and resource management.

## II. DATASET

describe where dataset from

### A. Data Preprocessing

*1) Crop:* why crop

*2) split:* why split

also published

## III. MODEL ARCHITECTURE

Normal method, Architecture

*Specifications and Constraints*

### A. Architectural Design

*1) Multiple Submodel Single Channel:*

*2) Single Model Multiple Channel:*

### B. Activation Functions

*1) ReLu:*

*2) LeakyReLu:*

*3) Tanh:*

## IV. METRICS AND EVALUATION

Several challenges are associated with aerial image semantic segmentation:

### A. Accuracy

### B. Mean Intersect Over Union (MeanIoU)

### C. RMSE

## V. RESULTS

### A. Performance on Training

### B. Performance Comparison

## VI. APPLICATION

### A. Application with Real Data from Google Maps

## VII. CONCLUSION

Aerial image semantic segmentation is a vital task with numerous applications. While traditional methods have laid the groundwork, deep learning-based approaches have brought significant improvements. Despite the challenges, ongoing research continues to push the boundaries of what is possible in this field.