Multitemporal Semantic Segmentation Applied to Land Use Classification

Yu Liang Lin[1], Tsung Han Wen[1], Tee Ann Teo[1]*

[1] National Yang Ming Chiao Tung University

This research explores the benefits of using temporal satellite imagery for land use classification. Compared to single satellite images, temporal satellite imagery provides richer information, and applying temporal data to land use classification can enhance classification accuracy. Garnot and Landrieu (2021) proposed a framework called U-TAE (U-Net with Temporal Attention Encoder) for temporal semantic segmentation of agricultural land types. This architecture is based on the U-Net and incorporates a temporal attention mechanism to process temporal features, thereby improving the ability to capture temporal feature information and structural patterns. The study utilized the PASTIS dataset based on Sentinel-2 satellite imagery, which effectively applies temporal semantic segmentation to agricultural land categories. Currently, land use surveys primarily rely on single-image classification, and the application of temporal data is not yet widespread. Therefore, this study utilizes temporal imagery from Formosat-2 (FS2) satellite to perform land use classification in Taiwan. The research employs FS2 four-band multispectral imagery in combination with panchromatic imagery and terrain features, spanning a period from 2012 to 2015, totaling 102 images. Training data includes existing land use maps as classification categories, supplemented with aerial base maps to generate accurate training area data. The classification categories are combined with FS2 natural color images and used to train the U-TAE deep learning framework. In addition to U-TAE, other architectures (Unet-3d, Unet-ConvLSTM, U-BiConvLSTM) are compared to demonstrate the generalizability of the U-TAE framework. The U-TAE framework demonstrates a notably higher MACRO IoU (0.41) and MACRO F1-score (0.53) in temporal semantic segmentation by pixel-wise comparison with alternative methods, namely 3D-Unet (MACRO IoU: 0.27, MACRO F1-score: 0.36), U-BiConvLSTM (MACRO IoU: 0.38, MACRO F1-score: 0.49), and Unet-ConvLSTM (MACRO IoU: 0.39, MACRO F1-score: 0.50). These evaluation metrics offer comprehensive insights into the model's effectiveness and its precise capacity to discern between various land use categories. The study encompasses nine distinct land use categories, including agriculture, forests, transportation, among others, with the primary objective of developing a highly proficient deep learning model tailored specifically for Formosat-2 image classification. These findings significantly contribute to the advancement of temporal satellite imagery's utilization for enhanced land use classification and monitoring endeavors.

Keywords: FORMOSAT-2, Temporal Semantic Segmentation, Land Use Classification