

Urban Functional Zone Change Detection via Multi-Source Temporal Fusion of Street View and Remote Sensing Imagery

Chenghan Yang^a, Hong Fang^a, Shanchuan Guo^a, Pengfei Tang^a, Zilong Xia^a, Xingang Zhang^a, Peijun Du^{a*}

^a School of Geography and Ocean Science, Nanjing University, Nanjing, 210023, China.

OBJECTIVES

This study seeks to develop a deep learning-based method for Urban Functional Zone (UFZ) change detection through the integration of multi-temporal Street View (SV) and Remote Sensing (RS) imagery. The objective is to overcome the limitations of traditional approaches, such as manual surveys and RS imagery, which often fail to capture fine-grained or ground-level details, by introducing the Multi-Source Temporal Fusion (MSTF) model. The MSTF aims to dynamically fuse SV and RS features via a Multi-Source Fusion (MSF) module and to capture temporal dynamics through a Multi-Temporal Fusion (MTF) module. An additional goal is to improve detection accuracy by aggregating SV classifications at the UFZ level. Ultimately, the study intends to establish a robust framework for multi-source data integration, thereby enhancing UFZ analysis to support improved land use optimization and urban management.

METHOD

The proposed method, illustrated in Fig. 1, integrates multi-temporal RS and SV data for highly accurate UFZs change detection. For each period (T1, T2), SV images are systematically classified, and SV category frequencies are calculated to create detailed statistical maps. The Multi-Source Temporal Fusion Model (MSTFM) combines a Multi-Source Fusion Module (MSFM) and a Multi-Temporal Fusion Module (MTFM). In MSTFM, CNNs extract features from RS imagery at T1, T2, and their difference image. The MSFM uses an attention mechanism to effectively fuse RS and SV data, while the MTFM applies self-attention to thoroughly integrate temporal features, emphasizing relevant time steps. A dense classification module then processes the fused features for final change detection and classification results.

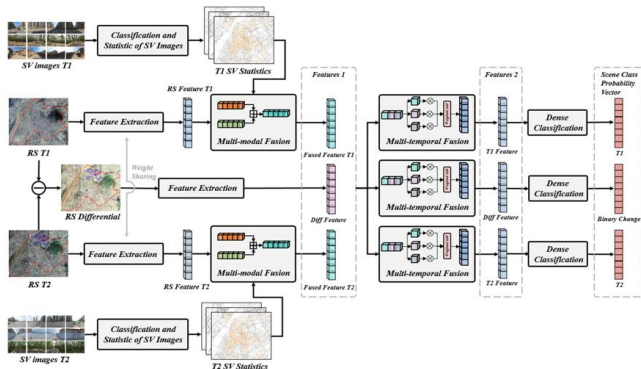


Fig. 2 Overall framework

MAJOR RESULT

As depicted in Fig. 2, the proposed method excelled in Nanjing UFZs classification and change detection (T1 to T2), aligning closely with ground truth for vegetation, residential, and industrial zones. Unlike VGG16 and ResNet18, which faltered in spatial consistency for commercial and public service areas, and DSIFN and FC-EF, which lacked boundary precision, the proposed model delivered accurate delineations. Its binary change detection robustly captured transitions, minimizing errors, affirming its efficacy for urban land use monitoring.

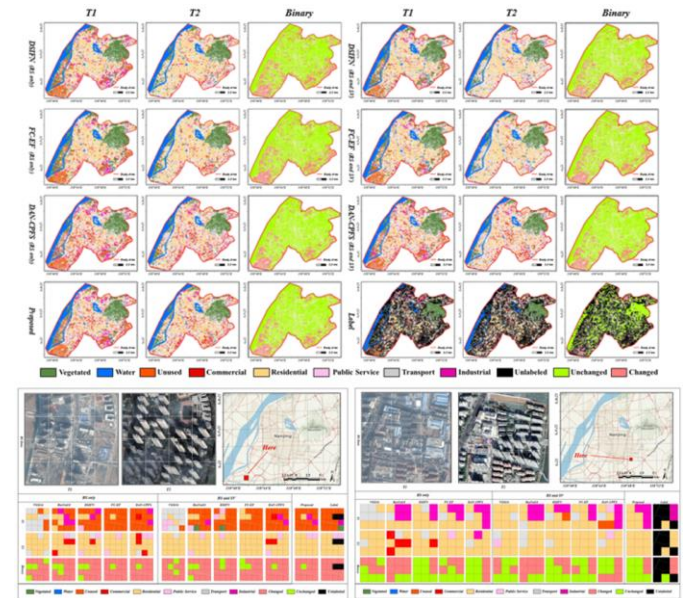


Fig. 2 UFZ maps and two examples in Nanjing

Two areas were examined: Hexi (Example A), transitioning from mixed unused/industrial/residential to residential with a school, and Daming Road (Example B), shifting from industrial to residential. Per Fig. 2, in Example A, RS-only models (e.g., VGG16) misclassified T1 zones, while DSIFN and FC-EF misidentified areas as transport. In T2, the proposed method accurately detected the residential and public service shift, with binary change maps confirming precise transition detection, leveraging SV data for enhanced urban change analysis.

PUBLICATION

Yang Chenghan, Fang Hong, Guo Shanchuan, Tang Pengfei, Xia Zilong, Zhang Xingang, Du Peijun*. Urban Functional Zone Change Detection via Multi-Source Temporal Fusion of Street View and Remote Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*. (Under Review, IF=7.5)