深層学習に基づく航空写真とデジタル地表モデルのエンドツーエンドの建物変化検出モデル

End-to-end Building Change Detection Model in Aerial Imagery and Digital Surface Model Based on Neural Networks

学籍番号　47-186839
氏　　名　連 欣蕾（Lian, Xinlei）
指導教員　柴崎 亮介 教授

Change detection is the process of identifying differentiations in the state of an objector phenomenon by observing multi-temporally. Essentially, it involves the ability to quantify temporal effects using data sets acquired at divergent time point. One of the major applications of remotely-sensed data obtained from Earth-orbiting satellites is change detection because of repetitive coverage at short intervals and consistent image quality. The process uses multitemporal datasets to qualitatively analyze the temporal effects of objects or phenomena and quantify the changes. Change detection studies with remote sensing data have been used in a wide variety of applications including land use and land cover, deforestation studies, natural disaster monitoring, as well as building damage assessments. The Due to the wide range of application scenarios including video surveillance, remote sensing, medical diagnosis and treatment, civil infrastructure, under- water sensing, driver assistance systems and so on, imagery-based change detection related research and algorithm development has remained to be an active research focus at remote sensing and computer vision domain in recent years.

Along with the increasing diversification of applications and data categories, the technique of change detection has advanced with ascending in the spatial resolution of remote sensing images and it has been used for applications at different spatial scales. Within the applications by applying multi-temporal remote sensing imagery to generate timely information on the earth’s natural environment and human activities, most of scholars lay emphasis on natural environment related ones. Nevertheless, the multi-temporal change detection of urban constructed environment, including building construction, traffic construction, urban facilities and other infrastructures, is significant for urban activities monitoring, real estate market mastery, resident’s mobility supervision and then whole city development promotion. Our study will be focusing on the application scenarios of urban construction change detection, consisting of three kinds: building new construction, demolishment as well as continuation, which is aiming at urban construction legitimacy

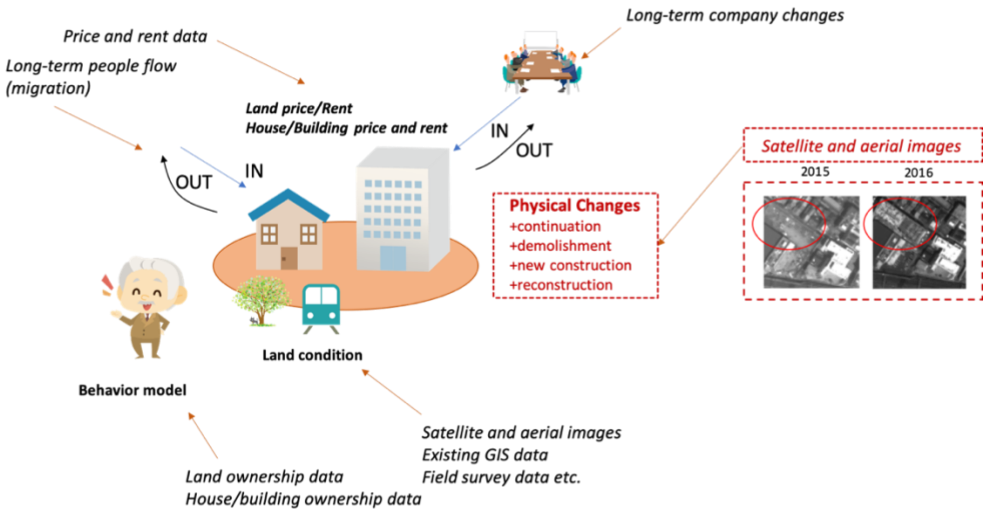

Figure 1 Research Background and Motivation of Building Change Detection

supervision and real estate commercial activity monitoring in a long term. The research background and motivation of our study is addressed in Figure 1.

Along with the developing progression of remote sensing and photogrammetry, land cover change detection is not limited to the dataset deployment on low- and medium-resolution remote sensing images based on single spectral features. Ortho urban remote sensing images derived from high-resolution aerial imagery could accommodate adequate spectral features and detailed information for feature fusion and high-level feature extraction. Our study will execute experiments on the aerial ortho imagery dataset taken from part region Tokyo, Japan in 2015 and 2016 respectively, with the resolution of 0.16m. As traditional metropolitan area, Tokyo has typical urban texture, divergent types of urban objects and high building density, which will raise the difficulty and provide solid validation for this study simultaneously.

As a premise of change detection from radiance changes from spectral feature on imagery, disturbing factors from multi-temporal aerial imagery including misalignment of pixel, radiance error caused by illumination and atmosphere condition difference should be eliminated. Although current photogrammetry methodologies could be applicated to 2-D remote sensing imagery for rectification with sensor parameters, multi-temporal aerial or satellite imagery is not adequate to offer spectral and textual features for building change detection. Alongside recent development of Dense Image Matching (DIM) technology, the acquisition of 3-D point cloud and Digital Surface Model (DSM) has been generally realized, which could be combined with imagery, making building change detection more effective with greater spatial structure and texture information. Building change detection is a research domain, under both computer vision and remote sensing, which has developed for decades.

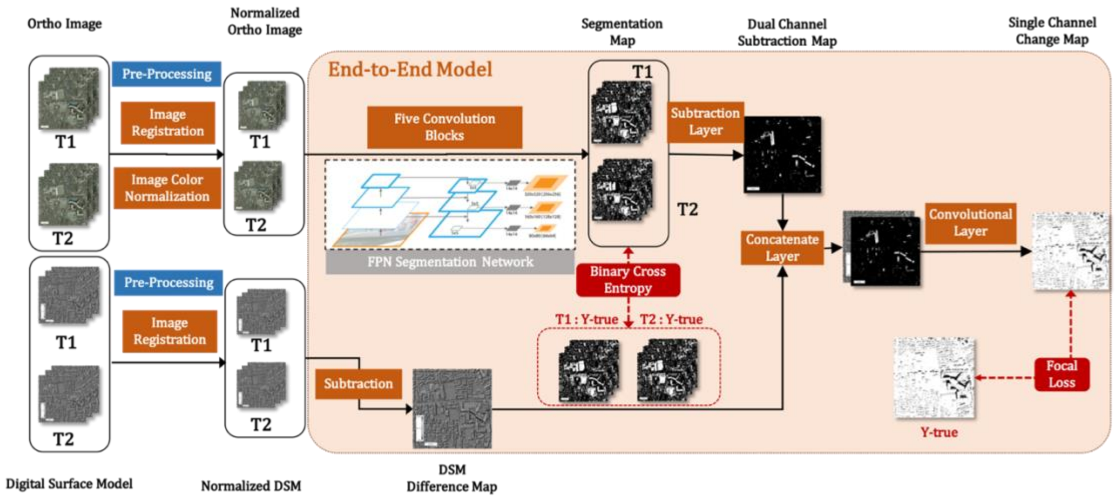

Figure 2 Overview Framework of Proposed End-to-End Change Detection Model

As acknowledged in existing relevant studies, the procedure of change detection could be divided to the following steps:

1. Data Acquisition: Differentiated by sensors and facilities, the data utilized for land cover change detection mainly comprises Visible Imagery (example: Aerial Imagery), Infrared Imagery as well as Multispectral, Hyperspectral and SAR Satellite Imagery, in which the last category is commonly utilized in practice.
2. Image Pre-processing: Likewise, depending on the sensor and data category, different image pre-processing techniques are applicated for better performance and higher efficiency in following change detection algorithm. Geometric registration, denoising and radiometric correction are basic processing methods for multi-temporal imagery data, avoiding inference factors like mis-alignment, image noise and radiance error caused by diverse perspective and illumination conditions when remote sensing equipment are working.
3. Change Detection Algorithm: Basically, the change detection algorithm could be categorized as pixel-based methods, object-based methods and spatial data mining methods, executed by traditional methodologies and learning-based methodologies.
4. Performance Evaluation and Accuracy Assessment: As for learning-based or model-based methods, traditional evaluation metrics including accuracy, confusion matrix, kappa coefficient and other indexes could be used for change detection algorithm performance evaluation. Also, some specialized indexes for a certain category is also frequently used.

With the given background and motivation above, the main objective of this study is to generate change map classified into three classes including new construction,

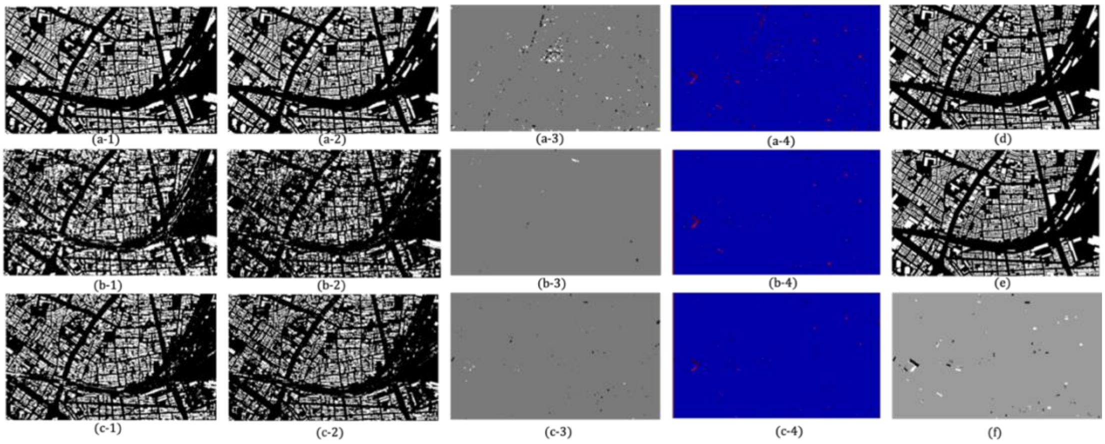

Figure 3 Experiment Results of Baselines and Proposed Method

demolishment and continuation by end-to-end model based on Feature Pyramid Network (FPN) with ortho aerial urban imagery and digital surface models (DSM) of the same target area. Specifically, the overall framework of our proposed end-to-end, multi-input and multi-output change detection model is shown in figure 2.

The results are shown in figure3, consisting of (a-1) 2015 Segmentation Result of Baselin1 (a-2) 2016 Segmentation Result of Baselin1 (a-3) Change Map Result of Baseline1(a-4) Change Map Visual Evaluation of Baseline2 (b-1) 2015 Segmentation Result of Baselin2 (b-2) 2016 Segmentation Result of Baselin2 (b-3) Change Detection Result of Baseline2 (b-4) Change Map Visual Evaluation of Baseline2 (c-1) 2015 Segmentation Result of Proposed Method (c-2) 2016 Segmentation Result of Proposed Method (c-3) Change Detection Result of Proposed Method (c-4) Change Map Visual Evaluation of Proposed Method (d) 2015 Segmentation Map Ground Truth (d) 2016 Segmentation Map Ground Truth (d) Change Map Ground Truth

For the building semantic segmentation results, the baseline1 had much better performance than the other two because of the model loss concentrating on the segmentation task only. Nevertheless, the performance of post classification methods heavily relies on the capability of segmentation model and much noises arise due to the misalignment though the segmentation has high accuracy.

Our method behaves relatively better than the baseline2 in segmentation results. Despite the two end-to-end models also distribute the optimizing concentration to change map by hyper losses, the DSM introduction also contribute to exact construction features by importing elevation details information.

For the change detection results, by visual inspection, the result of baseline1 generates severe noises due to the sample misalignment and inadequate morphological filtering in spite of the manual fine-tuning. Two end-to-end models have much better performances and our proposed method increased the ratio of True Positive and True Negative.