LIM-CD: A LARGE-SCALE REMOTE SENSING CHANGE DETECTION DATASET FOR INCREMENTAL MONITORING

Hanchao Zhang [1], Ruiqian Zhang [1,*], Xiaogang Ning [1], Xiao Huang [2], You He[1],
Yixin Chen [1], Mingzhu Li [3], Wei Cui [4], Jiaming Wang [5]

[1] Institute of Photogrammetry and Remote Sensing, Chinese Academy of Surveying and Mapping, Beijing, China
[2] Department of Environmental Sciences, Emory University, Atlanta, GA, USA
[3] School of Surveying, Mapping and Geographic Information, Liaoning Technical University, Fuxin, Liaoning, China
[4] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei, China
[5] Hubei Key Laboratory of Intelligent Robot, Wuhan Institute of Technology, Wuhan, Hubei, China
E-mail address: zhangrq@casm.ac.cn

KEY WORDS: Change detection dataset, Remote sensing, Incremental monitoring, Large-scale, High resolution.

ABSTRACT:

In this paper, we introduce a new large-scale change detection dataset called LIM-CD, designed for training and evaluating change detection algorithms on high resolution remote sensing images. The dataset currently consists of 9,259 images with labels covering six construction land use change types (i.e., residential land, industrial land, commercial land, public facilities, transportation land, and special land). The image annotations contain not only newly added regions of construction land as change annotations but also auxiliary information about construction land present in pre-change image (image T1), which serves as secondary annotations. These annotations offer crucial information for incremental monitoring applications. The remote sensing images are carefully selected to cover a broad range of imaging variations, including different image sources, years, backgrounds, and terrain. Additionally, we have provided comprehensive metadata labels, which can serve as additional features to aid model training and optimization. To establish a baseline for future algorithm development, we applied seven widely used and state-of-the-art change detection algorithms to the LIM-CD dataset. We are confident that our dataset can serve as a valuable resource for the research community, enabling the development of more accurate and robust change detection models. More information about the project can be found at https://github.com/xiaoxiangAQ/LIM-CD-dataset.

1. INTRODUCTION

Since the first aerial imaging over Paris, France, in 1858, the observation of Earth's surface has entered a new era with the advent of various remote sensing technologies. With the rapid development of high-resolution remote sensing images, which capture detailed surface information on the Earth, efficient and rapid change detection based on such images has become feasible. As a result, numerous change detection methods have also been proposed to detect changed regions from high resolution remote sensing images, enabling their use in a wide range of applications that include land use surveys, urban planning, environmental monitoring, and urbanization research, among others.

In recent years, the accuracy of change detection has advanced significantly due to the proliferation of change detection datasets and deep Convolutional Neural Networks (CNNs). Deep learning network structures, including a variety of CNNs (Shi et al.,2022; Chen and Shi,2020; Ding et al.,2022; Zhang et al.,2020; Sakurada et al.,2020; Zhang et al., 2023) and networks based on transformers (Chen et al.,2022) and (Bandara and Patel,2022), have replaced the traditional change detection and become the mainstream approach in this field. Most deep learning based change detection methods utilize a data-driven strategy in which high-level semantic features of image change areas are obtained through large-scale data training and feature extraction. Therefore, large-scale datasets are of great significance for model construction in these methods.

To explore change detection method performance in various applications, researchers have released datasets from their research, including WHU building CD (Ji et al.,2019), LEVIR-CD (Chen and Shi,2020), DSIFN-CD (Zhang et al.,2020), SZTAKI (Benedek and Sziranyi,2009), and others. While these datasets offer remote sensing images from different regions with ground truth annotations of changed regions, they do not provide sufficient image metadata labels. This lead to limitations in dataset generalization and hinder the learning of large models. Moreover, the current datasets are either limited in scale or exhibit a relatively concentrated image collection area, which may not capture phenomena such as the same spectrum with different objects or the same object with different spectra, which are commonly encountered in practical applications. Consequently, many methods achieve comparable accuracy rates exceeding 90%, making it challenging to accurately assess model quality. Thus, there is an urgent need to establish a comprehensive and complex change detection dataset with diverse samples to address these limitations.

Despite the availability of numerous publicly available datasets, human visual interpretation, rather than automatic change detection techniques, is still widely used in various applications. The current study proposes that the primary impediment to the widespread implementation of machine learning models is the inherent variability in real-world data distributions. Maintaining a uniform distribution of existing public datasets and real-world data is challenging, requiring transfer learning or model fine-tuning for effective implementation. However, the limited metadata labels in training samples provided by contemporary public datasets, coupled with underdeveloped transfer learning and model fine-tuning techniques, hinder the practical application of automatic change detection methodology.

---

* Corresponding author

Therefore, providing additional image metadata label is essential for developing application-oriented change detection datasets.

Our paper introduces a novel Large-scale Incremental Monitoring Change Detection dataset (LIM-CD), consists of 9,259 images of size 512×512 with a resolution ranging from 0.5-2 meters, providing not only newly added construction land regions as change annotations, but also auxiliary construction land information in the pre-change image (image T1) as secondary annotations for annual incremental monitoring change detection application. The secondary annotations mimic conventional incremental monitoring applications that aim to identify newly added suspected illegal patterns in non-construction land. The images were collected from a mosaic of 15 satellite sensors, carefully chosen to cover a wide range of imaging variations from 10 provinces across China, each with unique topographical features and different years of image acquisition. Due to the complexity and variability of the image sources comprising, automated change detection networks face significant challenges in this dataset. To aid in the development of knowledge-based network design, we provide additional comprehensive image information, called metadata labels, for large-scale modeling exploration in application-oriented change detection network experiments. The metadata labels comprise information that includes sensor type, acquisition year, region, and other pertinent details for each image. It is important to note that the metadata labels offer only categorical information with images from different years possessing varying numerical distinctions in the year metadata labels.

To assess the effectiveness of our proposed LIM-CD dataset, we compare it against the existing general change detection datasets and present their differences in Table I. In addition, we conduct experiments using seven baseline detectors on the LIM-CD dataset and analyze their performance, strengths, challenges, and limitations. Based on these analyses, we identify potential areas for future research. Our contribution can be summarized as follows:

(1) Our study presents the LIM-CD dataset, which to the best of our knowledge, is the first large-scale incremental monitoring change detection dataset that includes both change annotations and secondary annotations. The LIM-CD dataset is designed to support the development of algorithms that can learn a systematic understanding of the human-like monitoring process during routine monitoring.

(2) The selection of images in the LIM-CD dataset was conducted with care to ensure that it covers a wide range of imaging variations, such as different sources of images, acquisition years, backgrounds, and terrains. To further enhance the model training and optimization process, we provide additional comprehensive image label metadata labels that captures these variations and can be used as additional features.

(3) We also conducted a comprehensive evaluation of seven widely used and state-of-the-art change detection algorithms on the LIM-CD dataset, which can serve as a baseline for future algorithm development and comparison.

2. RELATED WORK

In this section, we outline the development for remote sensing change detection datasets and methods applied as reference for the current study, including change detection datasets and change detection algorithms.

2.1 Change Detection Datasets

The existing change detection datasets can be broadly classified into two categories (Sakurada et al.,2020): Binary Change Detection (BCD) datasets, where the sample label is binary indicating whether the region has changed or not, and Semantic Change Detection (SCD) datasets, where the sample label contains various types of changes. In general, BCD datasets mainly focus on detecting changes or changes of a certain type, and SCD datasets contain more information and generally include changes of all factors, which helps in network feature learning and extraction. However, considering the labeling of SCD datasets is time-consuming and labor-intensive, we mainly focus on BCD datasets in this paper.

As an earlier dataset, SZTAKI AirChange Benchmark Set (Benedek and Sziranyi,2009) contains seven aerial images captured between 2000 and 2005 and six images between 2000 and 2007. The sample images primarily consist of images captured in the same year, with similar quality, camera settings, seasons, and lighting conditions. Due to the limited amount of data, long image year span, relatively low image quality, and monotony in image season, this dataset is not suitable for complex and diverse scenes.

The AIST Building Change Detection (ABCD) dataset (Fujita et al.,2017) is specifically designed for developing and evaluating damage detection systems that can determine whether buildings have been washed away due to natural disasters, such as floods. The dataset is relatively simple and limited to a single application area, which may not be suitable for other types of change detection tasks.

The WHU building (Ji et al.,2019) comprises two aerial images, along with change vectors, change grid maps, and two corresponding building vectors. While the images have high resolution and large size, they cover a relatively simple area.

The GZCD dataset is composed of images acquired between 2006 and 2019, covering suburban areas of Guangzhou City, China. The dataset was collected using Google Earth service through the BIGEMAP software. Although the images contain various types of changes, including waters, roads, farmland, bare land, forests, buildings, ships, etc., the GZCD dataset focuses exclusively on building changes.

Lebedev-CD (Lebedev et al., 2018), also known as the CDD dataset in some studies, is a dataset that contains both synthetic and real images. This dataset's annotation is unique in that it focuses not only on common categories of changes (such as changes in regions with buildings, roads, and forests) but also on detailed objects such as cars and tanks.

LEVIR-CD (Chen and Shi,2020) dataset captured from 20 different regions in multiple cities in Texas, the United States, between 2002 and 2018. The dataset contains various types of buildings and includes changes in season, light, atmosphere, and sensor characteristics. However, the dataset also has some limitations, such as poor spectral information, limited geographic coverage.

DSIFN-CD (Zhang et al.,2020) dataset is composed of multi-seasonal, multi-sensor images collected manually from Google Earth, covering a wide range of surface features from six cities in China, with no consideration given to change markers, seasonal changes, and changes in image brightness. Datasets with diverse types of changes are beneficial for training network

Table 1. Comparison of different BCD datasets

Dataset	Image count	Sample size	Resolution	Data Source	Regional Distribution	Metadata labels
SZTAKI (Benedek and Sziranyi,2009)	13	952×640	1.5m	Aerial dataset +Google Earth	Szada of Hungary	/
ABCD (Fujita et al.,2017)	4253	160×160	0.4m	Aerial dataset	Tohoku region of Japan	/
WHU building CD (Ji et al.,2019)	2	15354×32507	0.075m	Aerial dataset	Christchurch	/
GZCD (Peng et al.,2020)	19	1006×1168 ~ 4936×5224	0.55m	Google Earth	Guangzhou	/
Lebedev-CD (Lebedev et al., 2018)	16000	256×256	0.03-1m	Google Earth	/	/
LEVIR-CD (Chen and Shi,2020)	637	1024×1024	0.5m	Google Earth	Texas of the US	/
DSIFN-CD (Zhang et al.,2020)	3988	512×512	2m	Google Earth	6 Cities in China (Beijing, Chengdu, Shenzhen, Chongqing, Wuhan, Xi'an)	/
SYSU-CD (Shi et al.,2022)	20000	256×256	0.5m	Aerial dataset	Hong Kong	/
LIM-CD (Ours, 2023)	9,259	512×512	0.5m-2m	Mosaic images (15 satellites)	10 Provinces with different terrains in China	√

models; however, the challenge of improving accuracy limits their practical application.

SYSU-CD (Shi et al.,2022) is an aerial image dataset taken in Hong Kong between 2007 and 2014. The dataset contains changes from common urban, suburban, and vegetation, as well as high-rise, high-density building changes and Haikou-related changes. However, the dataset is limited by its concentration on small areas in Hong Kong, which presents a relatively simple environment.

In summary, the current BCD datasets are insufficient to address the diverse change detection scenarios. The limited diversity in change types in these datasets hinders their application to real-world needs. Moreover, the small size of these datasets makes them prone to over-fitting and results in poor model performance. The concentration of image areas in narrow ranges of seasonal years, sensors and terrain further complicate the development of model migration methods. Thus, there is a need for an incremental monitoring dataset that can continuously track changes over time. Furthermore, as most real-world applications involve complex backgrounds, multiple regions, and sensors, there is a pressing need to develop a large-scale multi-label incremental monitoring dataset.

2.2 Change Detection Algorithms

Remote sensing image change detection traditionally focuses on medium and low-resolution images with limited data volume. This approach predominantly relies on visual interpretation and direct comparative analysis of pre-and post-temporal images. However, with the advancement of image resolution and acquisition technology, conventional methods are insufficient to address the demands of large-scale and fine-grained change detection. Fortunately, deep learning methods possess robust feature learning and recognition classification capabilities (Khelifi and Mignotte,2020), and have been introduced into remote sensing change detection, achieving excellent results. This integration has greatly propelled the frontiers of research in remote sensing change detection.

CNNs based on deep learning have demonstrated excellent performance in various image processing tasks. As a powerful model, CNNs have also found widespread application in remote

sensing image change detection tasks. Many classic CNN networks (e.g., AlexNet, VGG, ResNet, UNet, etc.), as well as their improved version, have been proposed and effectively utilized for remote sensing change detection tasks (Shi et al.,2020). Daudt et al. (2018) proposed three modifications of the UNet model based on CNNs, called FC-EF, FC-Siam-conc and FC-Siam-diff respectively. Since then, numerous methods based on CNNs have been proposed, with a particular focus on attention models. For example, Shi et al. (2022) proposed a deep supervised attention metric-based network (DSAMNet) for pseudo-changes caused by external factors. Chen and Shi (2020) developed a Spatial-temporal attention-based network (STANet), utilizing basic spatial-temporal attention module (BAM) and pyramid spatial-temporal attention module (PAM). Ding et al. (2022) proposed the Bi-Temporal Semantic Reasoning (Bi-SRNet) change detection model that enhances the semantic consistency of the two-phase prediction map using CNN's cross-temporal semantic reasoning attention module. Zhang et al. (2020) developed a deeply supervised image fusion network (DSIFN) that includes a difference discrimination network (DDN) and deep feature extraction network (DFEN), and an attention module is added to fuse the features to improve the boundary integrity and internal compactness of the objects in the output change map.

Recently, the Transformer model has emerged as a popular choice for change detection due to its powerful semantic feature extraction ability, strong modeling ability, and efficient parallel computing ability. To further improve the modeling of context information in both space and time domains, recent studies have also introduced the Transformer model into the remote sensing change detection task. For instance, Chen et al. (2022) proposed a bitemporal image Transformer (BIT) that establishes context information more efficiently. Similarly, Changeformer (Bandara and Patel,2022) was developed based on the transformer-based network, where the Transformer encoder and the Multi-Layer Perception (MLP) decoder were used to perform change detection on a pair of co-registered remote sensing images.

However, deep learning-based change detection algorithms rely heavily on having sufficient data to provide support for their training, especially for transformer-based networks (Alexey Dosovitskiy,2020). Therefore, there is an urgent need for a large-scale, high-resolution change detection dataset to support the development of effective change detection algorithms.

3. DESIGN OF THE LIM-CD DATASET

The LIM-CD dataset comprises an extensive collection of image data from 15 different sensors, including Beijing-No.2 (BJ-2), Tianhui-No.1 (TH01), Ziyuan-No.1 (ZY1-02C), Ziyuan-No.3 (ZY3), Gaofen-No.1 (GF-1), Gaofen-No.2 (GF-2), Gaojing-No.1 (GJ-1), SPOT6, SPOT7, Kompsat-3A, Kompsat-3, PlanetLab, RapidEye, GeoEye-1 and Pleiades-1a/b. The dataset contains a total of 9259 images, including 6547 images in the training set, 1776 images in the validation set, and 936 images in the test set. All the images in the dataset have a resolution ranging from 0.5-2 meters with 512x512 pixels.

3.1 Incremental monitoring annotations

The LIM-CD dataset presents two distinct label types, primary and secondary labels. The primary labels are encoded with a value of 255 and correspond to newly added construction land between two time periods T1 and T2, providing ground truth data that accurately reflect human activities and urban and rural development. Meanwhile, the secondary labels are encoded with a value of 128, indicating the extent of construction land present in the image at time T1, serving as an auxiliary input to reflect the construction situation at that specific time. To enable a comprehensive analysis of the urbanization process and human activities, both primary and secondary labels are jointly used to investigate and examine the newly added construction land between pairs of remote sensing images. Table 2 shows details regarding the land cover types and image examples.

3.2 Characteristics of the LIM-CD

Change detection in practical applications often involves complex backgrounds, multiple regions, and multiple sensors. To address these challenges, the developed LIM-CD dataset was designed with multiple labels and types, as multiple classification labels can facilitate knowledge learning during algorithm design. This approach allows for better analysis of the model's performance and problems, leading to model improvements and enabling the study of migration class methods in a more convenient manner.

3.2.1 Background Variation: As the application of deep learning in change detection continues to grow, datasets limited to a single region or small area can no longer meet the diverse needs of users. In response to the demand for large-scale research and application under the backdrop of big data and complexity, our dataset was designed to include images from over ten provinces in China, spanning Xinjiang, Guizhou, Sichuan, Hainan, Fujian, Jiangsu, Liaoning, Inner Mongolia, Shanxi, and Hebei. Figure 1 showcases images in the LIM-CD dataset from Liaoning (Northeast China), Hebei (North China), Inner Mongolia (Northwest China), Sichuan (Southwest China), Jiangsu (East China), and Hainan (South China), all of which are representative of distinct regions across China.

Table 2. Table of Land Use Classification for Construction

| Types	Composition	Examples	Image Examples (T1, T2, ground truth labels)
Residential land	Construction of residential buildings	Residential communities, Rural residential settlements, villa areas, etc	
Industrial land	Construction of industrial production and service facilities	Factories, warehouses, gas stations, parking lots, logistics parks, etc.	
Commercial land	Construction of commercial and service buildings	Shopping centers, supermarkets, hotels, restaurants, banks, hospitals, etc.	
Public facilities	Construction of public facilities	Schools, libraries, parks, squares, museums, sports arenas, clinics, etc	
Transportation land	Construction of transportation facilities	Airports, train stations, subway stations, highways, bridges, tunnels, etc	
Special land	Construction for special purposes	Such as water conservancy facilities, landscaping, cultural heritage preservation, etc	
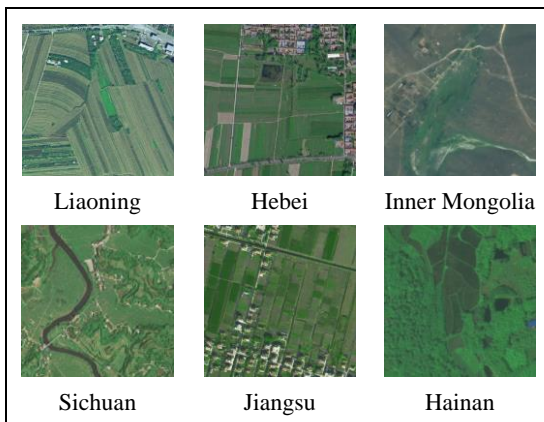
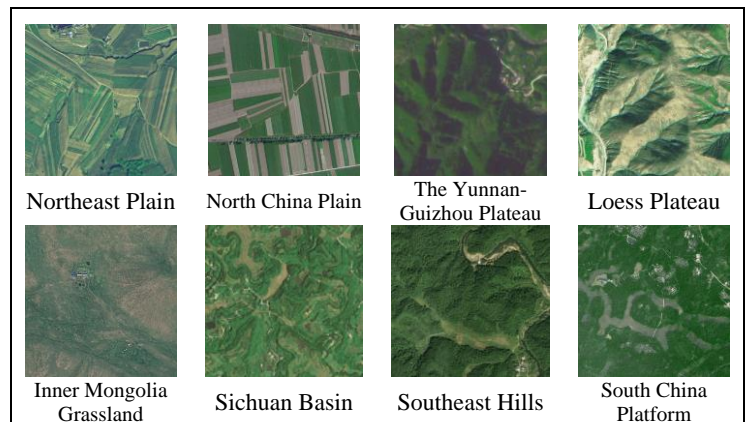**Figure 1**. Different region image examples in the dataset

Figure 2. Different terrain image examples in the dataset

3.2.2   Terrain Variation: Terrain is a crucial factor that impacts the precision of geometric correction, and the precision of geometric correction directly influences the accuracy of change detection. Inaccurate geometric correction leads to a high incidence of false changes. Consequently, diverse terrain data constitute a pivotal component of the change detection dataset. The LIM-CD dataset encompasses various terrains, such as plains, plateaus, grasslands, basins, hills, and platforms, as illustrated in Figure 2. Images with various terrains can facilitate the investigation of the algorithm's robustness in terrain alterations.

3.2.3   Time Variation: Time is a fundamental measure of change, and image data acquired at different times constitute the fundamental component of change detection. Time variation in the dataset necessitates a more robust algorithm to ensure accuracy and reliability. The LIM-CD dataset includes image data spanning from 2014 to 2018, employing an incremental monitoring system. In addition to "changed" and "unchanged " labels, the dataset includes the secondary labels indicating the previous year's building area, which facilitates the extraction of building changes in non-building areas, making it more suitable for numerous applications.

3.2.4   Resolution Variation: The spatial resolution of remote sensing images is one of the essential indicators for evaluating sensor performance and remote sensing information and serves as a critical basis for identifying the shape and size of ground objects. Different spatial resolutions may capture distinct data features of the same area. The designed LIM-CD dataset offers image data with multiple resolutions. In practical applications, the algorithm developed using the LIM-CD dataset can be utilized effectively for detecting changes in images of varying resolutions.

3.3 Statistical characteristics

To enhance users' understanding of the LIM-CD dataset, we conducted pixel-based label statistics for the dataset. The dataset is segmented into three subsets: the test set, validation set, and training set. We have computed the pixel counts for the pixels in primary labels, secondary labels and invariant area, within each subset. The statistical findings are depicted in Figure 3. In addition, we also classified the images in the dataset by region and time, and counted the number of pixels included in each category, including pixels in the primary labels, the unchanged regions and the secondary labels. The statistical results are shown in Figures 4 and 5.
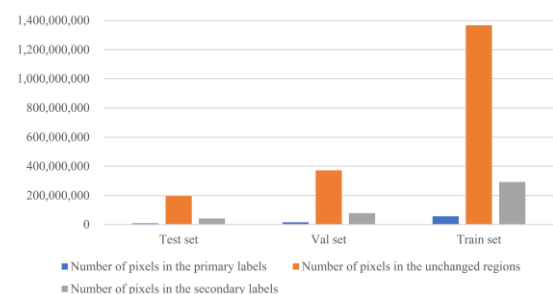

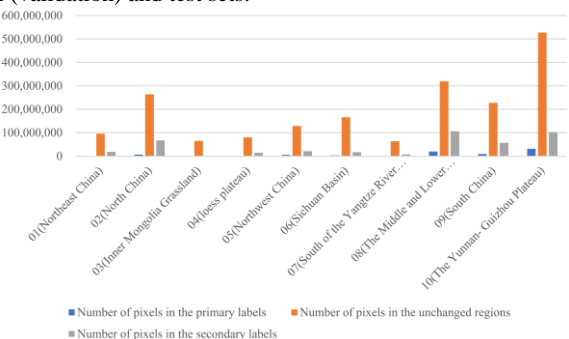**Figure 3.** Statistics of the number of pixels contained in train, val (validation) and test sets.

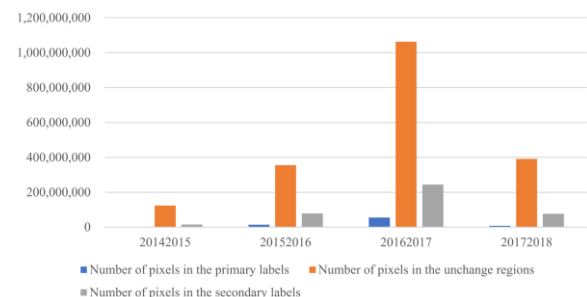**Figure 4.** The number of pixels of different types by region.

Figure 5. The number of pixels of different types by region.

4.   EVALUATION RESULTS

To validate the efficacy of change detection using the proposed LIM-CD dataset and provide reference evaluation results for researchers using the dataset, we retrained and evaluated different baseline detectors in this section.

4.1 Baseline Algorithms

In our baseline experiments, we selected seven widely used, state-of-the-art baseline change detectors, including FC-EF (Daudt et al.,2018), FC-Siam-conc (Daudt et al.,2018), FC-Siam-diff(Daudt et al.,2018), ISNet (Cheng et al.,2022), SNUNet-CD (Fang et al.,2022), BIT CD (Chen et al.,2022), and Changeformer (Bandara and Patel,2022). Among them, FC-EF, FC-Siam-conc, and FC-Siam-diff are considered classic deep learning baseline networks. ISNet and SNUNet-CD leverage CNN and attention models to achieve state-of-the-art change detection results. BIT CD and Changeformer introduce Transformer models, offering novel development ideas to the change detection field.

FC-X algorithms: The FC-EF algorithm is a modification of the UNet model that uses four max pooling and four upsampling layers. The FC-Siam-conc algorithm employs a fully convolutional siamese connection approach, where the encoder is divided into two branches with identical structure and weight sharing. The FC-Siam-diff algorithm employs a full convolution twin difference calculation method, where the encoder adopts a twin network structure and skip connections are used to process the absolute value of the difference between the encoders. These algorithms are frequently employed as benchmark experiments for various change detection datasets due to their simplicity and ease of training.

ISNet and SNUNet-CD: The ISNet technique is dedicated to enhancing feature refinement by proposing deep learning networks that deliver improved separability. This approach incorporates not only margin maximization but also targeted arrangement of attention mechanisms. The SNUNet-CD method leverages a hybrid approach that combines Siamese networks and NestedUNet, resulting in exceptional results across multiple datasets. This approach addresses the challenge of lost localization information in the deeper layers of neural networks through dense information transfer between the decoder and decoder. Additionally, an integrated channel attention module (ECAM) is employed for deep monitoring.

BIT CD and Changeformer: The BIT-CD conducts a bitemporal image Transformer (BIT) to efficiently and effectively model contexts within the spatial-temporal domain. The Changeformer network unifies the hierarchically structured Transformer encoder with Multi-Layer Perception (MLP) decoder in a Siamese network architecture to efficiently render multi-scale long-range details required for accurate change detection. These two algorithms both rely on Transformer architectures.

4.2 Implementation Details

In model training and testing, the operating system used in experiments is Ubuntu 22.04. All experiments were run on a desktop workstation with an Intel(R) Xeon(R) Gold 5222 CPU @ 3.80GHz, and an NVIDIA GeForce RTX 3090 GPU with 24G memory, and implemented based on the PyTorch platform.

For a fair comparison, we conducted experiments without data augmentation and resized the input images to the same scale (512×512). Preprocessing techniques such as random Gaussian noise, random rotation, and scale cropping were employed on the training set to enhance the model's robustness. The batch size and epoch number in all experiments were set to 8 and 100. Given the significant differences in training strategies between CNNs and transformers, we used distinct parameter settings for

each methodology to ensure adherence to the original literature. Specifically, for FC-EF, FC-Siam-conc, FC-Siam-diff, ISNet, and SNUNet-CD, we employed AdamW optimizer with CosineAnnealingLR learning rate adjustment strategy and used a hybrid loss function (Fang et al.,2022). The initial learning rate was set to 4e-4. In contrast, the Transformer methods experiments on BIT CD and Changeformer used 0.01 as the initial learning rate with linear decreasing (slope=1) learning rate adjustment strategy on SGD optimizer. These two methods used a cross-entropy loss function.

4.3 Evaluation Metrics

To evaluate the performance of these models, we used commonly employed evaluation indexes in change detection studies, including intersection over union (IOU), overall accuracy (Acc), Precision, Recall, F1score, and kappa coefficient (Kappa). These metrics are computed using the following formulae:

$$IOU = \frac{TP}{TP+FP+FN} \qquad Recall = \frac{TP}{TP+FP}$$

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \qquad F1score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$Precision = \frac{TP}{TP+FN} \qquad Kappa = \frac{p_o - p_e}{1 - p_e}$$

where *TP*, *TN*, *FP*, and *FN* denote true positive, true negative, false positive, and false negative in CD predictions, respectively. The Kappa coefficient is used for consistency testing and can also be used to measure classification accuracy.

4.4 Results and Analysis

4.4.1 Quantitative Results and Analysis: Based on the quantitative results in Table 3, we observe that existing binary change detection techniques achieve 54%-64% change detection results (as measured by the F1) in the LIM-CD dataset. These results suggest that current methods face challenges in accurately detecting changed regions in our dataset. Additionally, we note that the accuracy metrics of the various methods exhibit a variance of approximately 10%, indicating a considerable heterogeneity in the effectiveness of these methods.

Among the three traditional deep learning methods for change detection, the FC-Siam-conc approach exhibits inferior performance compared to FC-Siam-diff and FC-EF in all evaluation metrics. Specifically, FC-EF outperforms FC-Siam-conc in terms of accuracy, with an F1 score of over 59%. Among the most recent CNN-based change detection techniques, the SUNET approach demonstrates superior performance, achieving an F1 score of 64.24%, which is more than 3% higher than other methods. It is worth noting that even when trained for a shorter duration of 50 epochs, the SUNET method still achieves higher accuracy levels compared to other approaches.

The performance of Transformer-based change detection methods, specifically Changeformer, has shown unsatisfactory performance in our experiments. This shortcoming could be attributed to the focus of Changeformer on the acquisition of intricate information, which may not be well-suited for complex remote sensing images in the LIM-CD dataset. Nonetheless, considering the advancements of the Transformer methods in other domains, it is plausible that future methods with more refined training strategies and deeper Transformer structures may yield a significant enhancement in the accuracy of the LIM-CD dataset.

Table 3. Change detection results on the LIM-CD Dataset

| Method	IOU(%)	Acc(%)	Recall(%)	Precision(%)	F1 score(%)	Kappa(%)
FCEF(Daudt et al.,2018)	42.06	97.15	54.47	64.87	59.22	57.75
FC-Siam-diff (Daudt et al.,2018)	41.38	97.18	52.41	66.29	58.54	57.09
FC-Siam-conc (Daudt et al.,2018)	37.30	97.00	46.92	64.54	54.34	52.83
BIT CD (Chen et al.,2022)	43.40	97.47	51.05	74.34	60.53	59.27
Changeformer (Bandara and Patel,2022)	38.22	97.21	45.36	70.84	55.31	53.94
ISNet (Cheng et al.,2022)	42.80	97.22	54.63	66.41	59.95	58.52
SUNET_EP50 (Fang et al.,2022)	45.98	97.50	55.99	72.01	63.00	61.72
SUNET (Fang et al.,2022)	47.31	97.58	57.19	73.27	64.24	63.00
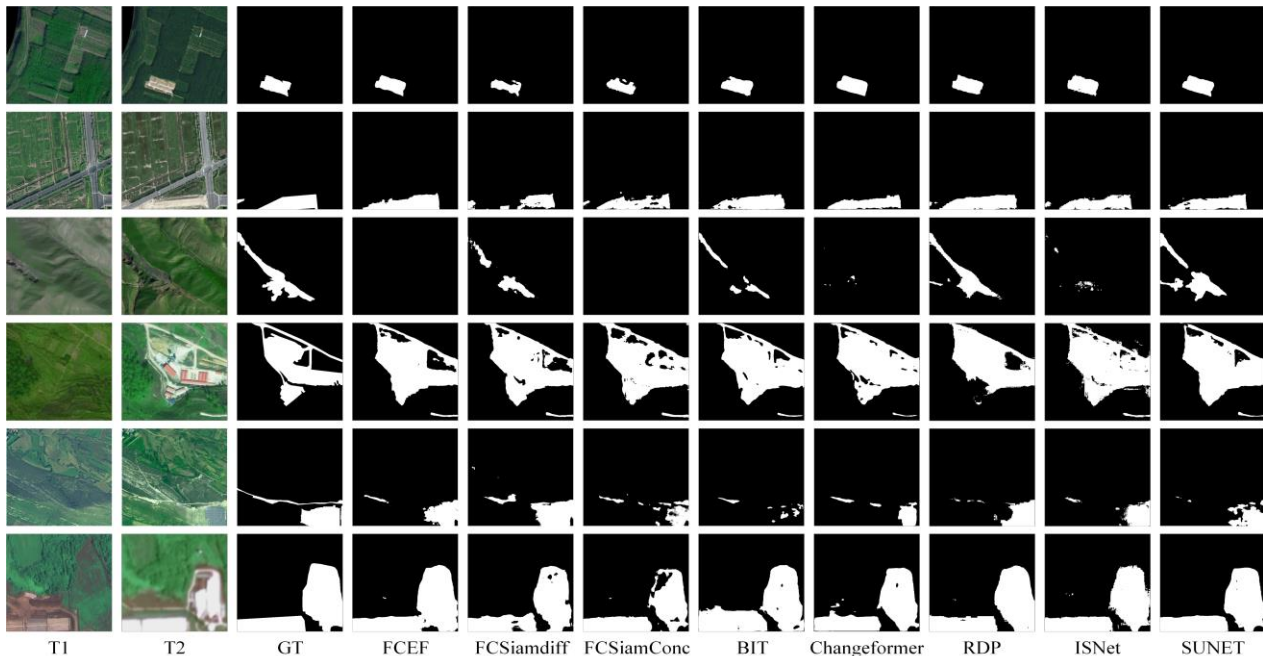
Figure 6. Change detection results. The T1, T2 images and ground truth labels are shown in the first three columns. The next few columns are the experimental results of the corresponding method in the last row.

The results presented in Table 3 highlight the complexity of the LIM-CD dataset and showcase the potential of deep learning's flexible network structures in change detection. However, the noticeable discrepancies in accuracy across different network designs indicate that existing deep learning-based change detection methods still require further optimization to handle the challenges posed by intricate image conditions. Further research may focus on the development of more refined deep learning architectures, optimization strategies, and training methodologies to enhance the accuracy of change detection in remote sensing images.

4.4.2 Qualitative Results and Analysis: We also conducted qualitative analysis on the change detection results, where randomly selected experimental results are presented in Figure 6. We observe that the alterations in the first and third rows were primarily due to the modification of vegetation coverage to push-fill. In the first row, the substantial disparity between the T1 image vegetation traits and the push-fill traits in the T2 image enabled the detection of a comprehensive range of changes using various methods. However, in the third row, the minimal disparity in features between the two images led to limited detection of the altered region by FC-EF and FC-Siam-Conc, while Changeformer and ISNet methods exhibited partial detection of localized changes. The identification of recently

added roads in rows two, four, and five demonstrated that the effectiveness of road detection may be attributed to the narrow width of road segments. In rows four and six, the detection outcomes were generally satisfactory, owing to the distinct construction features of post-temporal images. Nonetheless, certain false changes and gaps were still detected.

Overall, it can be concluded that among the three conventional change detection techniques, FC-Siam-diff exhibits greater stability in diverse scenarios, whereas the other two methods are struggle to identify changes occurring in areas with subtle differences. Comparative analysis of the two Transformer methods indicates that Changeformer exhibits higher sensitivity towards intricate information, thereby resulting in a closer approximation of the actual boundary of the detected change area. However, its inability to detect certain changes results in relatively low performance. Among the latest CNN-based change detection techniques, ISNet prioritizes the boundary problem. Nonetheless, experimental findings indicate that this method generates a considerable amount of noise. Additionally, SUNET demonstrates false changes and missed detection issues. These results highlight the persistent challenges faced by various methods when dealing with complex and diverse datasets like our developed LIM-CD dataset. It is important to note that current methodologies do not incorporate metadata

labels as a reference due to the lack of such information in existing datasets. However, the LIM-CD dataset presented in this study is annotated with ample metadata labels, providing an opportunity to leverage this information to improve the efficacy and performance of change detection networks. Therefore, we suggest that utilizing hint information is a promising strategy for further experiments and improving the performance of change detection techniques.

5. CONCLUSION

In this study, we present LIM-CD, a large-scale dataset of high resolution remote sensing images specifically designed for incremental monitoring change detection tasks. The dataset consists of 9,259 images annotated with labels covering six construction land use change types. The annotations include newly added construction land regions as change annotations and auxiliary construction land information in image T1 as secondary annotations, providing crucial information for incremental monitoring applications. The remote sensing images were carefully selected to cover all possible imaging variations, including different backgrounds, terrain, years and image sources. We also offer additional image label metadata labels, which can serve as additional features to aid in model training and optimization. To evaluate the effectiveness of the LIM-CD dataset, we conducted experiments using widely used and state-of-the-art change detection algorithms. The qualitative and quantitative experimental results demonstrate that the LIM-CD dataset still presents many challenges for existing change detection methods. Considering the low accuracy of existing methods, we encourage researchers to conduct further in-depth explorations guided by metadata labels. The results of this study can serve as a baseline for future algorithm development, and we believe that the LIM-CD dataset can be a valuable resource for researchers in the field of change detection.

ACKNOWLEDGEMENTS

REFERENCES

Alexey Dosovitskiy, L. B. A. K. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ICLR 2021 / ArXiv*. doi: 10.48550/arXiv.2010.11929

Bandara, W. G. C., & Patel, V. M. (2022). A Transformer-based Siamese Network for Change Detection. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 207-210.

Benedek, C., & Sziranyi, T. (2009). Change Detection in Optical Aerial Images by a Multilayer Conditional Mixed Markov Model. *IEEE Transactions on Geoscience and Remote Sensing*, 47(10), 3416-3430.

Chen, H., Qi, Z., Shi, Z. (2022). Remote Sensing Image Change Detection With Transformers. *IEEE transactions on geoscience and remote sensing*, 60, 1-14.

Chen, H., & Shi, Z. (2020). A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection. *Remote Sensing*, 12(10), 1662.

Cheng, G., Wang, G., J. Han. (2022). ISNet: Towards Improving Separability for Remote Sensing Image Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*. 60, 1-11.

Daudt, R. C., Saux, B. L. Boulch, A. (2018). Fully Convolutional Siamese Networks for Change Detection. *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, PP, 4063-4067.

Ding, L., Guo, H., Liu, S., Mou, L., Zhang, J., Bruzzone, L. (2022). Bi-Temporal Semantic Reasoning for the Semantic Change Detection in HR Remote Sensing Images. *IEEE transactions on geoscience and remote sensing*, 60, 1-14.

Fang, S., Li, K., Shao, J., Li, Z. (2022). SNUNet-CD: A Densely Connected Siamese Network for Change Detection of VHR Images. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.

Fujita, A., Sakurada, K., Imaizumi, T., Ito, R., Hikosaka, S., Nakamura, R. (2017). Damage detection from aerial images via convolutional neural networks. *The IAPR International Conference on Machine Vision Applications (MVA)*, pp. 5–8.

Ji, S., Wei, S., Lu, M. (2019). Fully Convolutional Networks for Multisource Building Extraction From an Open Aerial and Satellite Imagery Dataset. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), 574-586.

Khelifi, L., & Mignotte, M. (2020). Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis. *IEEE Access*, 8, 126385-126400.

Lebedev, M.A., Vizilter, Yu.V., Vygolov, O.V., Knyaz, V.A. Rubis, A.Y. (2018). Change Detection in Remote Sensing Images Using Conditional Adversarial Networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2, pp, 565-571.

Peng D., Bruzzone L., Zhang Y., Guang H., Huang X. (2020). SemiCDNet: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7): 5891-5906.

Sakurada, K., Shibuya, M., Wang, W. (2020). Weakly Supervised Silhouette-based Semantic Scene Change Detection. *IEEE International Conference on Robotics and Automation (ICRA)*, pp, 6861-6867.

Shi, Q., Liu, M., Li, S., Liu, X., Wang, F., Zhang, L. (2022). A Deeply Supervised Attention Metric-Based Network and an Open Aerial Image Dataset for Remote Sensing Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-16.

Shi, W., Zhang, M., Zhang, R., Chen, S., Zhan, Z. (2020). Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges. *Remote Sensing*, 12(10), 1688.

Zhang, C., Yue, P., Tapete, D., Jiang, L., Shangguan, B., Huang, L., Liu, G. (2020). A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166, 183-200.

Zhang R., Zhang H., Ning X., Huang X., Wang J., Cui W. (2023). Global-aware siamese network for change detection on remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 199, 61-72.