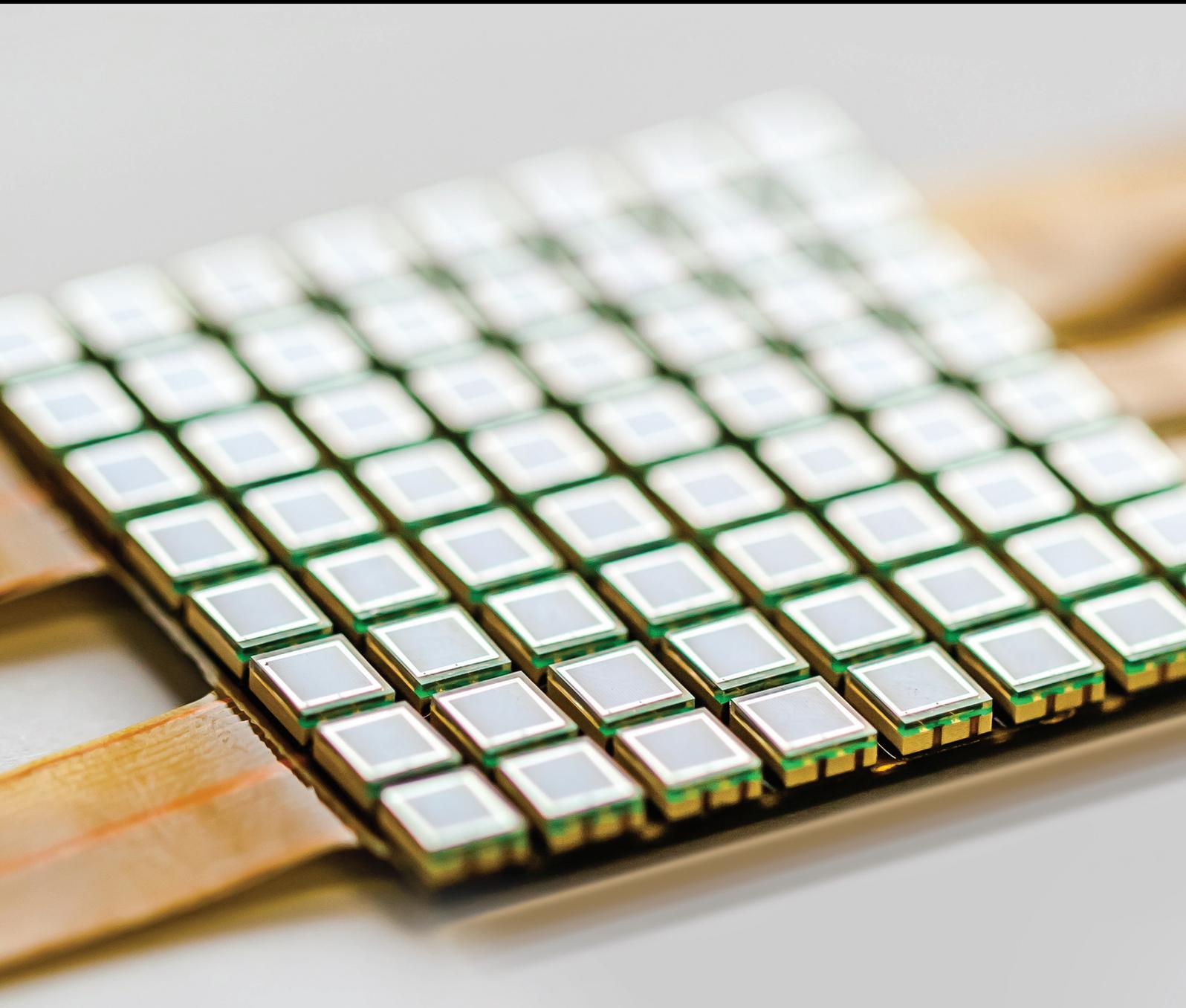


Systems and Sensors in Geoscience Applications

Special Issue Editor in Chief: Biswajeet Pradhan

Guest Editors: Hyung-Sup Jung and Ghassan Beydoun





Systems and Sensors in Geoscience Applications

Journal of Sensors

Systems and Sensors in Geoscience Applications

Special Issue Editor in Chief: Biswajeet Pradhan

Guest Editors: Hyung-Sup Jung and Ghassan Beydoun



Copyright © 2018 Hindawi. All rights reserved.

This is a special issue published in "Journal of Sensors." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

- Harith Ahmad, Malaysia
M. İlhan Akbaş, USA
Manuel Aleixandre, Spain
Bruno Andò, Italy
Constantin Apetrei, Romania
Marko Beko, Portugal
Fernando Benito-Lopez, Spain
Romeo Bernini, Italy
Shekhar Bhansali, USA
Wojtek J. Bock, Canada
Matthew Brodie, Australia
Paolo Bruschi, Italy
Belén Calvo, Spain
Stefania Campopiano, Italy
Domenico Caputo, Italy
Sara Casciati, Italy
Gabriele Cazzulani, Italy
Chi Chiu Chan, Singapore
Edmon Chehura, UK
Marvin H Cheng, USA
Nicola Cioffi, Italy
Mario Collotta, Italy
Marco Consales, Italy
Jesus Corres, Spain
Andrea Cusano, Italy
Antonello Cutolo, Italy
Dzung Dao, Australia
Luca De Stefano, Italy
Manel del Valle, Spain
Francesco Dell'Olio, Italy
Franz L. Dickert, Austria
Giovanni Diraco, Italy
Nicola Donato, Italy
Mauro Epifani, Italy
Abdelhamid Errachid, France
Stephane Evoy, Canada
Vittorio Ferrari, Italy
Luca Francioso, Italy
Manel Gasulla, Spain
Carmine Granata, Italy
Banshi D. Gupta, India
Mohammad Haider, USA
Clemens Heitzinger, Austria
- María del Carmen Horrillo, Spain
Evangelos Hristoforou, Greece
Syed K. Islam, USA
Stephen James, UK
Bruno C. Janegitz, Brazil
Hai-Feng Ji, USA
Sang Sub Kim, Republic of Korea
Antonio Lazaro, Spain
Laura M. Lechuga, Spain
Chengkuo Lee, Singapore
Chenzong Li, USA
Xinyu Liu, Canada
Eduard Llobet, Spain
Jaime Lloret, Spain
Yu-Lung Lo, Taiwan
Jesús Lozano, Spain
Oleg Lupan, Moldova
Frederick Mailly, France
Pawel Malinowski, Poland
Santiago Marco, Spain
Vincenzo Marletta, Italy
Carlos Marques, Portugal
Eugenio Martinelli, Italy
Antonio Martínez Olmos, Spain
Jose R. Martinez-De-Dios, Spain
Giuseppe Maruccio, Italy
Yasuko Y. Maruo, Japan
Mike McShane, USA
Fanli Meng, China
Carlos Michel, Mexico
Stephen. J. Mihailov, Canada
Heinz C. Neitzert, Italy
Calogero M. Oddo, Italy
Keat Ghee Ong, USA
M. Palaniswami, Australia
Alberto J. Palma, Spain
Lucio Pancheri, Italy
Roberto Paolesse, Italy
Giovanni Pau, Italy
Alain Pauly, France
Giorgio Pennazza, Italy
Michele Penza, Italy
Salvatore Pirozzi, Italy
- Antonina Pirrotta, Italy
Stavros Pissadakis, Greece
Stelios M. Potirakis, Greece
Biswajeet Pradhan, Malaysia
Valerie Renaudin, France
Armando Ricciardi, Italy
Christos Riziotis, Greece
M. Luz Rodriguez-Mendez, Spain
Jerome Rossignol, France
Carlos Ruiz, Spain
Ylias Sabri, Australia
Josep Samitier, Spain
José P. Santos, Spain
Isabel Sayago, Spain
Giorgio Sberveglieri, Italy
Andreas Schütze, Germany
Praveen K. Sekhar, USA
Sandra Sendra, Spain
Woosuck Shin, Japan
Pietro Siciliano, Italy
Vincenzo Spagnolo, Italy
Sachin K. Srivastava, India
Stefano Stassi, Italy
Vincenzo Stornelli, Italy
Weilian Su, USA
Tong Sun, UK
Salvatore Surdo, Italy
Raymond Swartz, USA
Hidekuni Takao, Japan
Guiyun Tian, UK
Suna Timur, Turkey
Vijay Tomer, USA
Abdellah Touhafi, Belgium
Aitor Urrutia, Spain
H. Vaisocherova - Lisalova, Czech Republic
Everardo Vargas-Rodriguez, Mexico
Xavier Vilanova, Spain
Luca Vollero, Italy
Tomasz Wandowski, Poland
Qihao Weng, USA
Qiang Wu, UK
Hai Xiao, USA
Chouki Zerrouki, France

Contents

Systems and Sensors in Geoscience Applications

Biswajeet Pradhan , Hyung-Sup Jung , and Ghassan Beydoun
Editorial (3 pages), Article ID 7242495, Volume 2018 (2018)

Deep Learning Approach for Building Detection Using LiDAR–Orthophoto Fusion

Faten Hamed Nahhas, Helmi Z. M. Shafri , Maher Ibrahim Sameen, Biswajeet Pradhan ,
and Shattri Mansor
Research Article (12 pages), Article ID 7212307, Volume 2018 (2018)

A High-Precision Control Scheme Based on Active Disturbance Rejection Control for a Three-Axis Inertially Stabilized Platform for Aerial Remote Sensing Applications

Xiangyang Zhou , Chao Yang, Beilei Zhao, Libo Zhao, and Zhuangsheng Zhu 
Research Article (9 pages), Article ID 7295852, Volume 2018 (2018)

Classification of Very High Resolution Aerial Photos Using Spectral-Spatial Convolutional Neural Networks

Maher Ibrahim Sameen, Biswajeet Pradhan , and Omar Saud Aziz
Research Article (12 pages), Article ID 7195432, Volume 2018 (2018)

Ocean Wave Information Retrieval Using Simulated Compact Polarized SAR from Radarsat-2

Xiaochen Wang , Yun Shao , Lu She , Wei Tian, Kun Li, and Long Liu
Research Article (12 pages), Article ID 1738014, Volume 2018 (2018)

A DLM-LSTM Framework for North-South Land Deformation Trend Analysis from Low-Cost GPS Sensor Time Series

Fangling Pu , Zhaozhuo Xu, Hongyu Chen, Xin Xu, and Nengcheng Chen 
Research Article (11 pages), Article ID 3054295, Volume 2018 (2018)

Wetland Change Detection Using Cross-Fused-Based and Normalized Difference Index Analysis on Multitemporal Landsat 8 OLI

Yan Gao , Zeyu Liang , Biao Wang , Yanlan Wu , and Penghai Wu 
Research Article (8 pages), Article ID 8130470, Volume 2018 (2018)

Development of Data Fusion Method Based on Topological Relationships Using IndoorGML Core Module

Junho Park , Dasol Ahn , and Jiyeong Lee 
Research Article (14 pages), Article ID 4094235, Volume 2018 (2018)

Research on Real-Time Supervisory System for Compaction Quality in Face Rockfill Dam Engineering

Shengxiang Huang, Wen Zhang , and Gen Wu 
Research Article (11 pages), Article ID 6487405, Volume 2018 (2018)

Performance Analysis of Network-RTK Techniques for Drone Navigation considering Ionospheric Conditions

Tae-Suk Bae  and Minho Kim
Research Article (8 pages), Article ID 5154697, Volume 2018 (2018)

Inversion Model of GPR Imaging Characteristics of Point Objects and Fracture Detection of Heritage Building

Gao Lv , Ning Li , Jie Yang, Xianchun Yao , Dexiu Hu, and Rixuan Pang
Research Article (10 pages), Article ID 3095427, Volume 2018 (2018)

Evaluation and Analysis of Dam Operating Status Using One Clock-Synchronized Dual-Antenna Receiver

Yunlong Zhang , Songlin Yang, Jiankun Liu, Dongwei Qiu , Xiaoyan Luo, and Jianhong Fang
Research Article (12 pages), Article ID 9135630, Volume 2018 (2018)

Tile-Based Semisupervised Classification of Large-Scale VHR Remote Sensing Images

Haikel Alhichri , Essam Othman , Mansour Zuair , Nassim Ammour, and Yakoub Bazi 
Research Article (14 pages), Article ID 6257810, Volume 2018 (2018)

Editorial

Systems and Sensors in Geoscience Applications

Biswajeet Pradhan ¹, **Hyung-Sup Jung** ², and **Ghassan Beydoun**¹

¹Research Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and Information Technology, University of Technology Sydney, NSW 2007, Australia

²Dept. of Geoinformatics, Lab. for Advanced Remote Sensing (LARS), University of Seoul, 90 Jeonnong-dong, Dongdaemun-gu, Seoul 130-743, Republic of Korea

Correspondence should be addressed to Biswajeet Pradhan; biswajeet24@gmail.com

Received 11 July 2018; Accepted 11 July 2018; Published 23 September 2018

Copyright © 2018 Biswajeet Pradhan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The science branch of geosciences encompasses various disciplines such as geology, geophysics, geotechnology, and computer simulations dealing with the planet Earth and aiming at understanding the Earth's history and its future evolution. Through fieldwork, Earth observations from space, modeling, and theoretical studies and geoscience applications provide essential information to further understand the developments in the Earth sciences and relevant environmental engineering aspects.

Various systems and sensors such as visible imaging, synthetic aperture radar (SAR), global navigation satellite system (GNSS), and laser scanning (LiDAR) are continually developing to acquire a broad range of data types about the Earth surface supporting new numerous applications. Major advances in geoscience applications have recently occurred, particularly with the development of new sensors of high accuracy and efficient technology, computing performance, new theories, and technical methods. Furthermore, based on the latest research in geoscience, more advanced sensors and both theoretical and technical methods offer new insights into solving difficult and multidiscipline challenges. Overall, those new developments in geoscience lead to a more systematic investigation using advanced geospatial technologies.

This special issue is aimed at bringing researchers to contribute in various methods for geoscience applications using Earth observation sensors and systems. New ways of data acquisition, data integration strategies, and processing with innovative methods and modeling continue to improve our

understanding of the landscape elements and their interactions with the climate and the environment.

This special issue, which had opened for 10 months, is particularly dedicated to publish a set of high quality papers dealing with the up-to-date systems and sensor applications of state-of-the-art techniques and tools in various disciplines of geosciences and civil engineering to solve real-world problems.

An article by X. Wang et al. demonstrated about the compact polarized synthetic aperture radar (CPSAR) and investigated its capability using Souyris' and Nord's algorithms. Nord's algorithm was adapted to retrieve ocean wave information. The result from their study showed that Nord's algorithm has a better convergence ability than Souyris'. Ocean wave slope spectrum and other parameters of main wave were further calculated. Their finding also revealed that the buoy output and CP SAR-based wave field information clearly showed a good agreement.

F. Pu et al. presented an interesting case study on landslide displacement trend analysis using a multisensor: space-borne remote sensing methods such as synthetic aperture radar and wireless low-cost global positioning systems (GPS). They developed a framework of a dynamic linear model based on long short-term memory (DLM-LSTM) to extract and predict north-south land deformation trends from one meter accuracy GPS receivers. Authors used a Kalman filter to calculate the deformation trend with submeter-level accuracy. Their designed framework could be used in broader application such as geological disaster

monitoring, earthquake deformation studies, and terrestrial mass movements.

G. Lv et al. applied an inversion model as a very fast and effective method to detect the defect in the building heritage using a ground-penetrating radar (GPR). They assessed image characteristics of hyperbolic curves with different depths and radii using finite-difference time-domain method (FDTD). Subsequently, they went on to establish inversion models of buried depth and radius of the point object with better accuracy. Their results demonstrated an important application of GPR to detect hidden defects in civil engineering.

An article presented by J. Park et al. proposed a fusion method by using IndoorGML core module to integrate and analyze various datasets and data formats. They proposed a novel geospatial data fusion model and the topological relation-based data fusion model (TRDFM), using topological relations among spatial objects to integrate different geospatial datasets and different data formats. They implemented the TRDFM in a geospatial application system to execute better information-based services without the need for reformatting the data or geometric data exchange. Authors used a 3D GIS software to describe the concept of the proposed TRDFM method.

In another paper, T.-S. Bae and M. Kim analyzed the effects of the ionospheric conditions on the GNSS Network-based Real-time Kinematic (NRTK), as well as the possibility of applying the mobile NRTK to drone navigation for mapping. Over the years, the NRTK systems are predominantly used for precision positioning in many fields such as surveying and agriculture mapping purpose. The success rate usually depends on the local environment and the ionospheric condition. Their results demonstrated that even though a submeter accuracy could be achieved, it is still important to understand the process of dealing with ionospheric disturbances.

A study by H. Alhichri et al. described an efficient tile-based semisupervised classification scheme for large-scale very high-resolution (VHR) remote sensing (RS) images where typical pixel-based classification methods are not feasible. Authors proposed a pretrained convolutional neural network (CNN) to extract descriptive features from each tile. They proposed a novel model that exploits the spectral as well as the spatial relationship. Finally, they tested the method with several experiments to check the classification accuracy. One of the most promising findings of their work is that, even with less than ten training samples per class, their proposed method performs extremely well in terms of classification accuracy.

S. Huang et al. described that their aim was to develop a supervisory system based on Global Navigation Satellite System (GNSS) technology, wireless data communication, internet of things technology, and computer technology. The basic goal of their system is to supervise the real-time roller compaction parameters of the working surface including rolling track, rolling times, rolling speed, thickness, and smoothness. Further, the authors described the conventional method used for quality control. The feasibility and robustness of their developed supervisory system were illustrated

in a case study in the face rockfill dam of Shuibuya project in China. Their findings showed that their system indeed provides a new and effective method of process control that could be used for better construction quality.

M. I. Sameen et al. in their article described the classification of very high-resolution aerial photos using spectral-spatial convolutional neural networks. Their network is characterised by a convolution layer, a kernel of size 3×3 , pooling size of 2×2 , normalization, dropout, and a layer with Softmax activation. According to their results, the proposed model is effective with overall accuracy and Kappa coefficient of 0.973 and 0.967, respectively. Their proposed model demonstrated robustness especially for high-resolution aerial photo classification provided if the parameters are carefully selected.

F. H. Nahhas et al. applied the state-of-the-art deep learning approach for building detection from aerial orthophotos. Authors utilized an object-based image analysis (OBIA) through their proposed model for creating objects, feature fusion, and autoencoder-based dimensionality reduction which was used for object classification. The architecture optimized for a grid search method and the sensitivity and parameters were also analyzed. Their results showed that the detection accuracy was 86.06% to 86.19% in the working area and 77.92% to 78.26% in the testing area. Finally, their results revealed that the use of an autoencoder in deep learning models could in fact improve the accuracy of recognition of buildings in fused data.

X. Zhou et al. described in their article the improvement of the accuracy of an inertial stabilized platform in the remote sensing applications using high-precision control scheme. They have used active disturbance rejection control (ADRC) to suppress the effects of disturbance on accuracy. According to the results of their study, ADRC has better capacity in disturbance rejection than a CPID controller through which the friction disturbance will be weakened and the stabilization accuracy can be improved.

Y. Gao et al. assessed wetland change detection using cross-fused- (CF-) based and normalized difference index (NDI) on multitemporal Landsat 8 images. Their main aim was to quantify the changes in wetland cover by using an image-to-image comparison change detection method. Normalized difference vegetation index (NDVI) and normalized difference water index (NDWI) were used to enhance the information on vegetation and water, respectively. Their proposed method performed well with reduced error and an overall accuracy of 97% and 93% for the interannual and seasonal datasets, respectively. The proposed method can be used as an effective tool for wetland change detection.

Y. Zhang et al. proposed an improved GNSS scheme using a finite element model to simulate the deformation. Further, the dam displacement warning standard was determined based on surface displacement and safety factor. Finally, the operation status was evaluated using the BDS/GPS deformation monitoring system. Their new proposed scheme was found to be effective and feasible.

In conclusion, we expect that this special issue provides important and new scientific pieces of evidence in sensors, tools, technical and theoretical methods, and computer

simulations for obtaining and processing datasets and knowledge discovery in geosciences and civil engineering applications and that it makes a practical progress for improving geospatial information systems.

Acknowledgments

We express our great appreciation to all authors for their excellent contributions and reviewers for their valuable help. The Lead Guest Editor would like to thank the two Guest Editors for their dedicated cooperation.

*Biswajeet Pradhan
Hyung-Sup Jung
Ghassan Beydoun*

Research Article

Deep Learning Approach for Building Detection Using LiDAR–Orthophoto Fusion

Faten Hamed Nahhas,¹ Helmi Z. M. Shafri^{1,2}, Maher Ibrahim Sameen,³ Biswajeet Pradhan³ and Shattri Mansor^{1,2}

¹Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia

²Geospatial Information Science Research Center (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia

³Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW, Australia

Correspondence should be addressed to Helmi Z. M. Shafri; helmi@upm.edu.my

Received 21 November 2017; Accepted 26 June 2018; Published 5 August 2018

Academic Editor: Antonio Lazaro

Copyright © 2018 Faten Hamed Nahhas et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper reports on a building detection approach based on deep learning (DL) using the fusion of Light Detection and Ranging (LiDAR) data and orthophotos. The proposed method utilized object-based analysis to create objects, a feature-level fusion, an autoencoder-based dimensionality reduction to transform low-level features into compressed features, and a convolutional neural network (CNN) to transform compressed features into high-level features, which were used to classify objects into buildings and background. The proposed architecture was optimized for the grid search method, and its sensitivity to hyperparameters was analyzed and discussed. The proposed model was evaluated on two datasets selected from an urban area with different building types. Results show that the dimensionality reduction by the autoencoder approach from 21 features to 10 features can improve detection accuracy from 86.06% to 86.19% in the working area and from 77.92% to 78.26% in the testing area. The sensitivity analysis also shows that the selection of the hyperparameter values of the model significantly affects detection accuracy. The best hyperparameters of the model are 128 filters in the CNN model, the Adamax optimizer, 10 units in the fully connected layer of the CNN model, a batch size of 8, and a dropout of 0.2. These hyperparameters are critical to improving the generalization capacity of the model. Furthermore, comparison experiments with the support vector machine (SVM) show that the proposed model with or without dimensionality reduction outperforms the SVM models in the working area. However, the SVM model achieves better accuracy in the testing area than the proposed model without dimensionality reduction. This study generally shows that the use of an autoencoder in DL models can improve the accuracy of building recognition in fused LiDAR–orthophoto data.

1. Introduction

Buildings are a fundamental element in forming a city and are essential for urban mapping [1]. The extraction of accurate building objects from remote sensing data has become an interesting topic and has received increasing attention in recent years. Building information is important in several geospatial applications, such as urban planning, risk and damage assessment of natural hazards, 3D city modeling,

and environmental sciences. Building objects can be delineated from many data sources, such as satellite images, aerial photos, radar images, and laser scanning data. In particular, Light Detection and Ranging (LiDAR) offers an accurate and efficient approach for obtaining elevation data, which can be used to extract ground objects, such as buildings [2]. The advantages of using LiDAR over traditional photogrammetry include the capability to collect high-density point clouds at a relatively short time, high vertical accuracy,

and low cost. However, the accurate extraction of buildings in urban areas with precise boundaries is a difficult task due to the presence of nearby objects, such as trees, which frequently have the same elevations as buildings. Therefore, the fusion of LiDAR point clouds and aerial images can be an important step toward improving the quality of building detection.

Numerous methods have been proposed for building detection in the past decades by using LiDAR data and by fusing other remote sensing data with LiDAR data to improve accuracy and quality. Li et al. [3] proposed a series of novel algorithms for detecting building boundaries from the fusion of LiDAR and high-resolution images. Their results indicate that the fusion of LiDAR and high-resolution images is a promising approach for the accurate detection of building boundaries (correctness = 98% and completeness = 95%). Li et al. [4] proposed an improved building extraction method based on the fusion of optical imagery and LiDAR data. The aforementioned method comprises four steps: filtering, building detection, wall point removal, and roof patch detection. Their results suggest that the proposed method can automatically extract building objects with complex shapes. Saeidi et al. [5] also applied a data-driven method based on Dempster–Shafer theory to fuse LiDAR and SPOT (Satellite Pour l’Observation de la Terre) data for building extraction. These researchers examined the potential of slope and height information extracted from the LiDAR-based digital elevation model (DEM) and digital surface model (DSM), as well as from the normalized difference vegetation index (NDVI) created from SPOT images. Their results show that NDVI/normalized DSM (nDSM) fusion performs better than NDVI/slope for building extraction.

Uzar and Yastikli [6] developed an automatic building detection method based on LiDAR data and aerial photographs. This method includes segmentation and classification with object-based image analysis. The accuracy assessment shows an overall accuracy of approximately 93%, a completeness of 96.73%, and a correctness of 95.02% for building extraction. Uzar [7] developed an automatic approach for building extraction based on multisensor data (LiDAR and aerial photographs) and rule-based classification. He applied fuzzy classification to improve building extraction results. His method achieved a completeness of 81.71% and a correctness of 87.64% based on a comparison between the extracted buildings and reference data. Furthermore, Awrangjeb et al. [8] proposed an automatic building detection technique using LiDAR data and multispectral imagery. They utilized the normalized difference vegetation index to separate the buildings from trees and extract the residential buildings in the area. Awrangjeb et al. [9] also developed a building detection technique for complex scenes. In their method, a rule-based procedure was established to utilize the normalized digital surface model extracted from LiDAR data for the task in hand effectively.

Recently, Wang et al. [10] presented an automatic method for building boundary extraction from LiDAR data. This method includes height-based segmentation, shape recognition by shape indices, and boundary reconstruction

using Hough transformation and a sequential linking technique. Their findings show that the proposed method can achieve accurate extraction of building boundaries at rates of 97%, 85%, and 92% for three LiDAR datasets with different scene complexities. Prerna and Singh [11] assessed a building detection method based on the segmentation of LiDAR and high-resolution photographs. These researchers determined that an object-based-oriented classification yielded the best accuracy ($R^2 = 0.86$) compared with using only LiDAR. Zhao et al. [12] presented a building extraction method using LiDAR data and connected operators. Their results demonstrate that the proposed method performs effectively. The efficient and average offset values of simple and complex building boundaries are 0.2 m to 0.4 m and 0.3 m to 0.6 m, respectively. Tomljenovic et al. [13] applied object-based analysis for building extraction from LiDAR data. Their obtained results exhibit high accuracies for the initial study area and on the International Society for Photogrammetry and Remote Sensing benchmark without any modification.

Tomljenovic et al. [2] reviewed building extraction methods based on LiDAR data. Their analysis shows that the main limitations of current building detection methods are their application to wide-area datasets and the lack of transferability studies and measures. Other challenges in building detection from LiDAR include point cloud sparsity, high spectral variability, differences of urban objects, surrounding complexity, and data misalignment [14]. Gilani et al. [14] proposed a methodology that extracts and regularizes buildings using features from LiDAR data and orthoimagery to overcome some of the aforementioned limitations. Their results demonstrate the robustness of their approach. However, this method is affected by the registration error between LiDAR data and orthoimagery, which requires a further validation on different datasets. The lack of transferability of current methods is mainly due to the use of rule-based classification.

Therefore, the current paper reports on a building detection method based on the fusion of LiDAR data and orthophotos using a deep learning (DL) approach. At present, DL has gone beyond multilevel perceptrons and comprises a collection of techniques and computational methods for building compassable differentiable architecture. In particular, this study develops a framework based on an autoencoder to reduce feature dimensionality and a convolutional neural network (CNN) to distinguish building objects from non-building objects after segmentation is performed on LiDAR and orthoimage data.

2. Methodology

This section describes the proposed model and explains its components that have been designed to detect buildings from LiDAR and orthophotos based on a DL approach. It describes the overall workflow, data preprocessing and preparation, feature extraction through multiresolution and spectral difference segmentations, feature fusion and abstraction using autoencoders and CNN, and building detection that applies fully connected layers with sigmoid activation to the final layer.

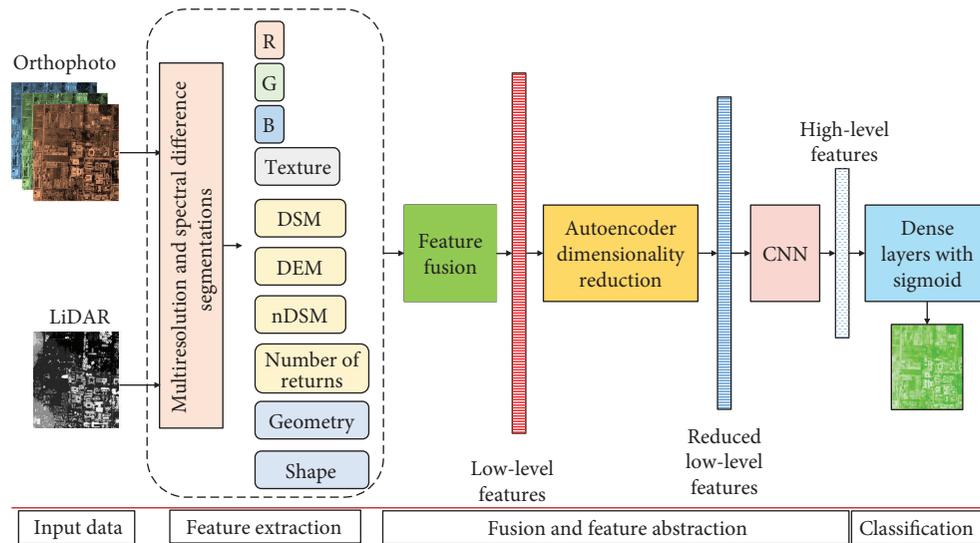


FIGURE 1: Architecture of the proposed building detection method using DL and LiDAR-orthophoto fusion.

2.1. Overall Architecture. This study proposes a DL model for detecting buildings from fused LiDAR and orthophoto data. The overall workflow of this model is presented in Figure 1. The proposed pipeline encompasses four main components: preprocessing and preparation of input data, feature extraction, fusion and feature abstraction, and classification. The first component, that is, data preparation, includes the geometric correction and registration of LiDAR point clouds with orthophotos. The point clouds were filtered to create DSM, DEM, and nDSM samples. DSM was created by interpolating point clouds using the inverse weighted distance (IDW) method. DEM was created by filtering non-ground points using the multiscale curvature algorithm of ArcGIS [15]. nDSM was created by subtracting DEM from DSM. The LiDAR-derived DSM, DEM, nDSM; number of returns; and orthophoto bands (i.e., red, green, and blue) were then composited at 0.3m spatial resolution and prepared for segmentation.

The second component, that is, feature extraction, was implemented to extract the spectral and texture features from the orthophotos and DSM, DEM, number of returns, and geometry and shape features from LiDAR data. The third component includes feature fusion and abstraction using an autoencoder DL model to reduce features and a CNN model to transform low-level features into high-level features. The last component adopted fully connected layers and a sigmoid layer to classify image objects into background and buildings. The details of these processing steps are explained in the following sections.

2.2. Feature Extraction. A total of 21 features, including spectral, shape, textural, and LiDAR-based features, were initially extracted to detect building objects in the LiDAR and orthophoto data. Spectral features were used to evaluate the mean pixel values in the orthophoto bands. The shape features refer to the geometric information of meaningful objects, which is calculated from the pixels that form these objects. An accurate segmentation of the map is necessary to ensure

the successful use of these features. Texture features were also derived from the Haralick texture features based on the gray-level cooccurrence matrix (GLCM) or the gray-level difference vector. Alternatively, the LiDAR-based features were used to describe the topography and height of objects.

The low-level features (Table 1) were calculated based on the image objects created via multiresolution and spectral difference segmentations. The features extracted from the LiDAR data and orthophoto were fused at the feature level. The features were then reduced by applying an autoencoder-based dimensionality reduction approach. The reduced low-level features were then fed into the CNN model to extract the high-level features for classification. The following sections describe the aforementioned processes.

2.3. Fusion and Feature Abstraction. Building detection and description are important steps in reconstructing building objects. The former refers to the process of identifying building objects among other objects [20], whereas the latter refers to the process of delineating the geometric boundary of building objects to describe their geometry and extract information as attributes linked to the objects in a geographic information system (GIS). On the one hand, orthophotos have a significant capacity in spatial resolution and exhibit strong reflectance around building boundaries. However, the spectral similarity of different ground objects generates difficulties in extracting buildings from orthophotos. On the other hand, extracting building edges with height discontinuity is difficult in LiDAR due to the relatively small footprint size of the laser beam and disadvantageous backscattering from illuminated targets [20]. Thus, the fusion of orthophotos and LiDAR can improve the accuracy of building detection and description processes.

Data fusion is defined as the process of using or combining data from multiple sources to form a new dataset and accomplish a particular objective [21]. The three fusion levels that can combine data from different sources are classified as pixel, feature, and decision fusions [22]. The present study

TABLE 1: Extracted features from the LiDAR and orthophoto data [16–19].

Data source	Feature group	Feature	Description
	Spectral	Mean red	Average value of the pixels that cover the segment in the red band
		Mean green	Average value of the pixels that cover the segment in the green band
		Mean blue	Average value of the pixels that cover the segment in the blue band
		GLCM angular	$\sum_{i,j=0}^{N-1} P_{i,j}^2$
		GLCM contrast	$\sum_{i,j=0}^{N-1} P_{i,j} i - j ^2$
		GLCM correlation	$\sum_{i,j=0}^{N-1} [((i - \mu_i)(j - \mu_j))(\sigma_i^2 + \sigma_j^2)^{-1/2}]$
Orthophoto	Texture	GLCM dissimilarity	$\sum_{i,j=0}^{N-1} P_{i,j} i - j $
		GLCM entropy	$\sum_{i,j=0}^{N-1} P_{i,j} (-\ln P_{i,j})$
	GLCM homogeneity	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{(1 + (i - j)^2)}$	
	GLCM mean	$f_{\mu_i} = \mu_i \sum_{i,j=0}^{N-1} i(P_{i,j}), f_{\mu_j} = \mu_j \sum_{i,j=0}^{N-1} j(P_{i,j})$	
	GLCM variance	$\sum_{i,j=0}^{N-1} P_{i,j} (i - \mu)^2$	
LiDAR	Shape	Area	Total area of segment without holes
		Compactness	Ratio of the area of a polygon to the area of a circle with the same perimeter
		Density	Distribution in space of the pixels of an image object
		Length/width	Length-width ratio of the envelope rectangle
		Rectangular fit	Goodness of a building that fits into a rectangle
		Roundness	Area of the segment to the square of the maximum diameter of the referred segment
		Shape index	Border length of the segment divided by four times the square root of its area
	LiDAR	DEM	Digital elevation model
		DSM	Digital surface model
		nDSM	Object height by subtracting DEM from DSM

In the equations above, i is the row number of the cooccurrence matrix, j is the column number of the cooccurrence matrix, and $P_{i,j}$ is the normalized value in cell i, j ($P_{i,j} = V_{i,j} / \sum_{i,j=0}^{N-1} V_{i,j}$), where $V_{i,j}$ is the value in cell i, j of the cooccurrence matrix and N is the number of rows or columns of the cooccurrence matrix.

adopts the feature level because building detection and description with object-based analysis are easier and more efficient. Orthophoto features (e.g., spectral and textural features) and LiDAR features (e.g., DSM, DEM, nDSM, and spatial features) are combined to form low-level features for building detection (Table 1).

Many features that are related to spectral, textural, topographical, and shape groups can be extracted from orthophotos and LiDAR data. The use of many features can cause overfitting, particularly when the training samples are relatively small. The other disadvantages of using a large number of features are noise, redundant information, and increasing computing time. The current study introduces an autoencoder-based approach that reduces feature space dimensionality and improves low-level features by

transforming them into fewer features (i.e., reduced low-level features) to address the aforementioned issue. The transformed features are expected to be more informative than the raw features and to improve the performance of the overall methodology workflow of building detection. A CNN model is also developed to select the relevant features for detecting buildings and to transform the reduced low-level features into high-level features by applying a set of convolution and pooling operations. The process of reducing (or abstracting) low-level features by using the autoencoder and CNN models is described in the following sections.

2.3.1. Autoencoders. Autoencoders (Figure 2) are neural networks that attempt to reconstruct their inputs without using labels (unsupervised); they have two logical parts, that is, the

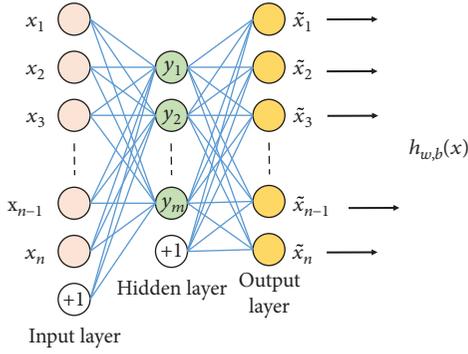


FIGURE 2: Simple structure of an autoencoder (adapted from [24]).

encoder and the decoder [23]. The former comprises the network layers that create a hidden representation of the input data, whereas the latter comprises the network layers that take the hidden representation from the encoder and create an output that is similar to the input data of the encoder. Thus, the last layer in autoencoder networks has the same size as the input of the first input layer. This process allows the network to learn features regarding the input data and regularization parameters. Hidden representation can be smaller than the input data; hence, the major benefit of autoencoders is dimensionality reduction.

Autoencoders adopt the backpropagation algorithm for training [23]. In an autoencoder, the output $h_{w,b}(x) = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)^T$ is equal to the input $x = (x_1, x_2, \dots, x_n)^T$.

$$\begin{aligned} h_{w,b}(x) &= g(f(x)) \approx x, \\ J(W, b; x, y) &= \frac{1}{2} \|h_{w,b}(x) - y\|^2, \end{aligned} \quad (1)$$

where x is an input that belongs to the n -dimensional space, y is a new representation that belongs to the m -dimensional space, and J is the reconstruction error.

A standard autoencoder comprises three layers. The first layer to the second layer amounts to an encoder f , and the second layer to the third layer amounts to a decoder g . Then, the algorithm minimizes J by adjusting the parameters in the encoder and the decoder to obtain the reconstructed input. The number of hidden layer nodes m is restricted to less than the number of original input nodes n to utilize the autoencoder as a dimensionality reduction algorithm.

The proposed autoencoder architecture includes 21 input features, 3 hidden layers with 128, 50, and 30 nodes, and a central layer with dimensions 5, 10, and 15 that are evaluated iteratively. Several hidden layers and their associated number of nodes were selected using a grid search method [25] and evaluated based on the similarity between the input and reconstructed data measured via the mean squared error. The network was trained through the Adamax optimization method [26] with its default parameters in Keras (TensorFlow backend) [27] and a batch size of 32. A sparsity constraint (L1 activity regularizer) was also added to the encoded representations to avoid overfitting and reduce model complexity.

2.3.2. CNNs. CNN [28] is a method that simulates a multi-layer structure of the human brain. It can extract the features of input data from a low to a high layer incrementally to improve classification or prediction processes. It abstracts the relationships among data and improves optimization performance with a reduction in training parameters. The structure of CNN consists of three layers that can be described as the convolution, subsample (pooling), and fully connected layers (Figure 3).

The proposed CNN architecture is encompassed by two stacked feature stages. Each stage contains a convolution layer followed by a pooling layer. A 2D convolution with 128 filters and maximum pooling were used. The high-level features were produced by flattening the 2D features estimated via the convolution and pooling operations. The network was also trained using the Adamax optimization method with a batch size of 8. Once the high-level features were obtained, a fully dense layer with 10 nodes and a dropout rate of 0.2 were used to classify features into building or background classes. The trained CNN model was then adopted to predict the class of test data, and the outputs were utilized to create the final building maps in GIS. The CNN network was optimized using the grid search method, which is explained in the next section.

2.3.3. Optimization Procedure. The optimization of hyperparameters is a crucial step in developing an efficient object detection model through DL methods, which are easy to use. Optimization can improve the overall performance, prediction accuracy, and generalization capacity of models, particularly when they are used to predict unseen data. The current study utilizes the grid search method to determine the optimal hyperparameters among specific search spaces of the CNN model. The grid search typically identifies a better set of hyperparameters than a manual search within the same amount of time. The optimized parameters, their search spaces, and their determined optimal values are shown in Table 2. Five hyperparameters, namely, the optimizer, number of filters, number of hidden units of the dense layer, dropout rate, and batch size, were optimized. The search spaces of the hyperparameters (excluding the optimizer) were manually selected after several random experiments.

3. Results and Discussion

This section describes the experimental datasets, the results of building detection with the accuracy assessment, and the sensitivity analysis of the proposed model. The proposed model was developed in Python using Google's TensorFlow library. It was then implemented in a personal computer with an Intel® Core i7 at 2.00 GHz and a memory (RAM) of 16 GB.

3.1. Experimental Datasets. The proposed building detection model was evaluated on two datasets (i.e., working and testing) selected from the Universiti Putra Malaysia campus located in the state of Selangor, Malaysia (Figure 4). The selected areas are geographically located between latitudes $7^{\circ}11'00''\text{E}$ and $7^{\circ}14'00''\text{E}$ and longitudes $3^{\circ}00'00''\text{N}$ and

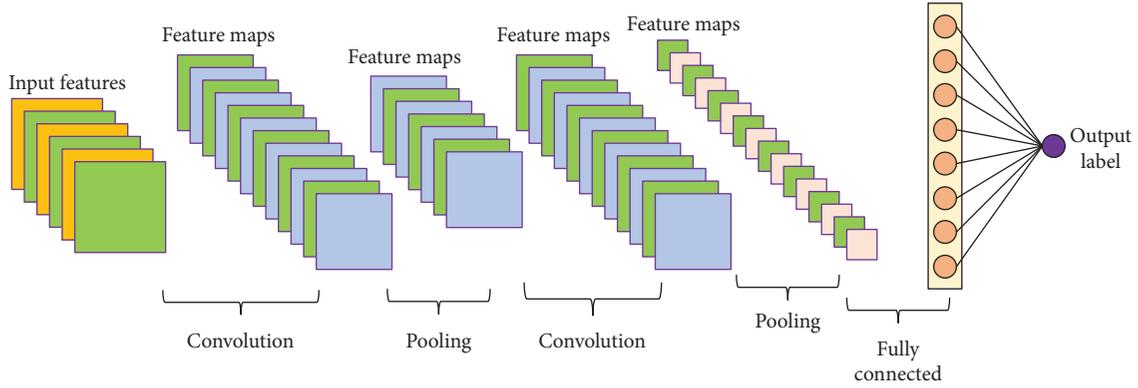


FIGURE 3: Typical CNN architecture.

TABLE 2: Optimization of CNN hyperparameters.

Parameter	Search space	Optimal value
Optimizer	(rmsprop, adam, nadam, Adamax, Adadelata, sgd)	Adamax
Number of filters	(16, 32, 64, 128)	128
Number of hidden nodes	(3, 10, 50, 100)	10
Dropout rate	(0, 0.2, 0.3, 0.5)	0.2
Batch size	(4, 8, 16, 32, 64)	8

$3^{\circ}40'00''N$ of the Kertau RSO Malaya coordinate system. The areas were selected because they include a mixture of urban features, such as asphalt roads, trees, dense vegetation, water bodies, and buildings. The buildings have different roofing materials, shapes, sizes, and heights.

The LiDAR data used in this study was obtained with a laser scanning system (Riegl LM Q5600 and Camera Hassleblad 39 Mp) on March 8, 2015. The systems had a scanning angle of 60° and a camera angle of $\pm 30^{\circ}$. The average point density of the LiDAR data was 4 points/m² with an average point space of 0.43 m. Overall, the LiDAR data contained 9.24 million points in the working and testing areas. The minimum and maximum elevations are in the working area at 37.65 m and 79.83 m, respectively. The elevations in the testing area range from 36.86 m to 100.36 m. Three different products were derived from the raw LiDAR point clouds, namely, DEM, DSM, and height feature or nDSM. Furthermore, the laser scanning system also collected RGB images along the point clouds. The spatial resolution of the collected orthophotos is 13 cm.

DSM was derived with IDW interpolation at 0.5 m spatial resolution. Meanwhile, DEM was derived using an ArcGIS filtering algorithm called multiscale curvature classification (MCC) [15]. The validations of this filtering method exhibit improvement in removing understory vegetation, which addresses topological differences across scales [15]. The other advantages of this approach include a built-in function in ArcGIS software, which makes its implementation easy and enables its integration into an automatic processing pipeline. The MCC algorithm filters LiDAR point clouds by classifying LiDAR returns as ground and nonground points. This

algorithm combines curvature filtering with a scale component and variable curvature tolerance [15]. MCC then interpolates a surface at different resolutions through the thin-plate spline method, and points are classified based on a progressive curvature threshold parameter (0.78 in this study). Other LiDAR data filtering methods are presented in the works of [29, 30].

3.2. Results of Building Detection. The image objects created via multiresolution segmentation were classified into buildings and backgrounds using the proposed DL model. Classification was applied with the complete set of features (21) and the best number of features obtained by the autoencoders (10 features). The building detection results are shown in Figure 5. Figure 5(a) shows the buildings detected by the model in the working area without reducing the dimensionality of the input features. The total number of buildings detected is 2808, which is higher by 8% than the real number in the reference dataset. The reason for this misclassification is mainly due to noise, which leads to small objects being incorrectly detected as buildings. With regard to the geometry of the detected buildings, Figure 5(c) shows that the detected buildings were affected by nearby objects, such as roads and trees. These objects create a problem in accurately describing the buildings. For example, a single building is composed of several objects, which cannot offer accurate building counting in the study area. Additional nearby objects attached to the detected buildings also create an issue in describing building objects, such as estimating their roofing geometry, floor area, and even their height. By contrast, the results of the model with reduced features show better building detection with less misclassification and better boundary delineation (Figures 5(b) and 5(d)). The number of buildings calculated using this method is 281, which is 0.86% lower than the reference number of buildings. Reducing the number of features using the autoencoder model may contribute to the removal of features that create overfitting in the model and offer better building detection results. Figure 5(d) shows an example of how reducing the features used for building detection can also contribute to improving the boundary delineation of objects. This property is extremely useful in counting the buildings in the study area with better accuracy. Furthermore, building detection with

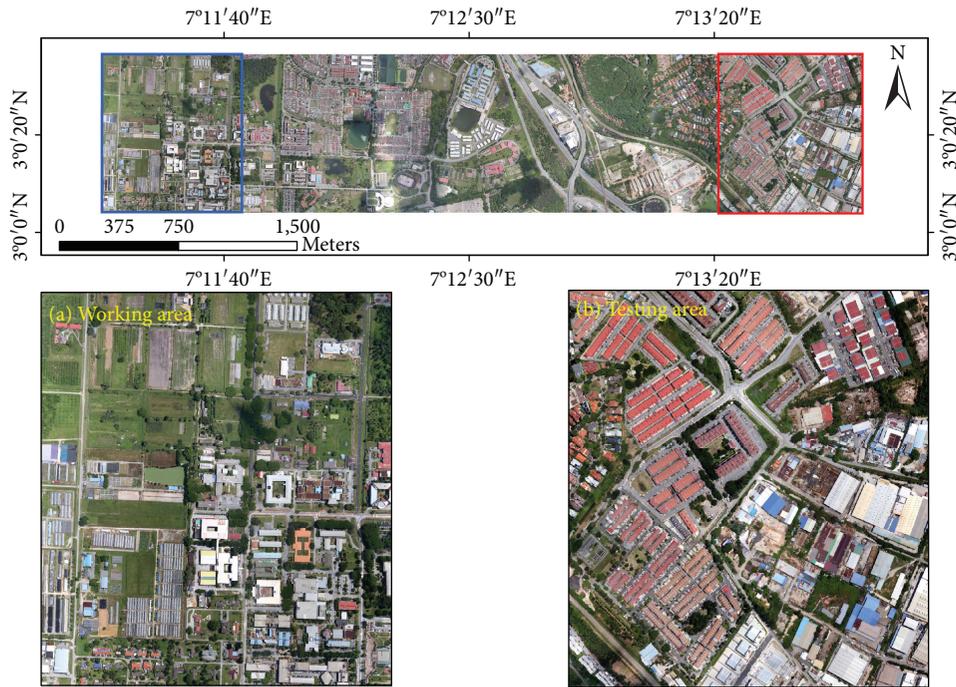


FIGURE 4: Location of the study area and experimental datasets: (a) working area and (b) testing area.

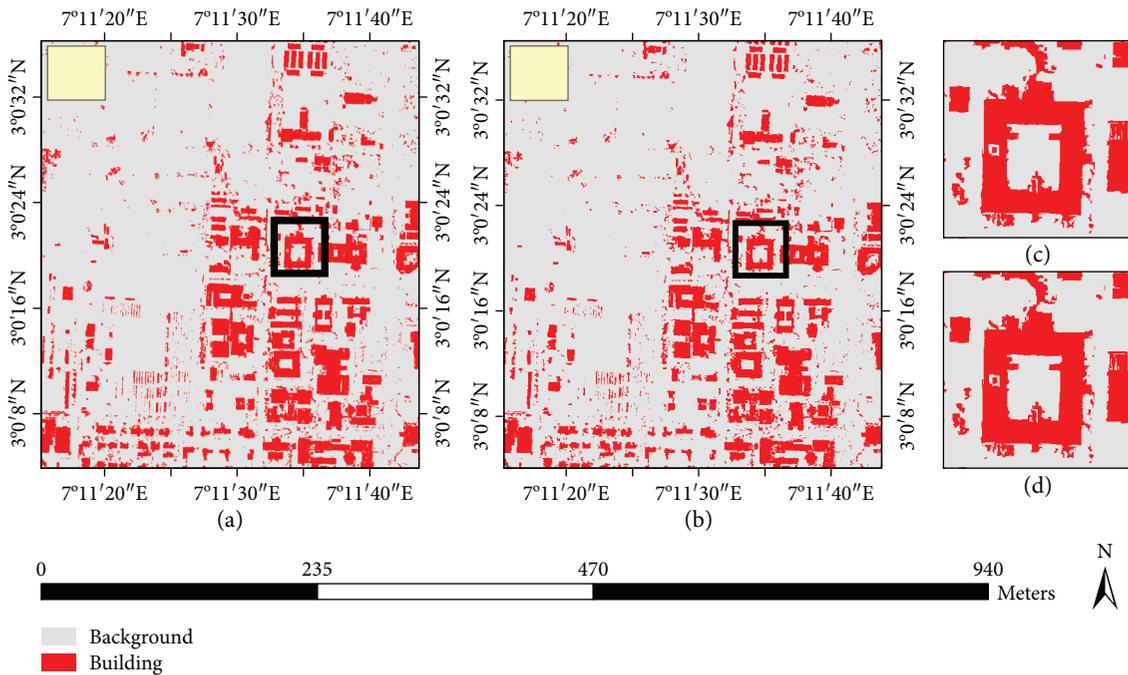


FIGURE 5: Results of building detection using the proposed method in the working area: (a) without feature dimensionality reduction, (b) with dimensionality reduction (10 features), (c) example of a detected building through a complete set of input features, and (d) example of building detection after feature reduction by the autoencoder dimensionality reduction approach.

accurate boundary can calculate several spatial and geometric attributes of the objects with high precision. The model outputs that apply autoencoders for feature fusion and abstraction allow exporting of building information in the study area that can be useful for decision-making and urban planning, among other applications.

Furthermore, the proposed model was also used to detect buildings in the testing area, and the results are shown in Figure 6. The model was applied with and without feature reduction. Figure 6(a) shows the buildings in the testing area obtained by the model without using autoencoders. The number of buildings in this map is 1029, which is 4.47%

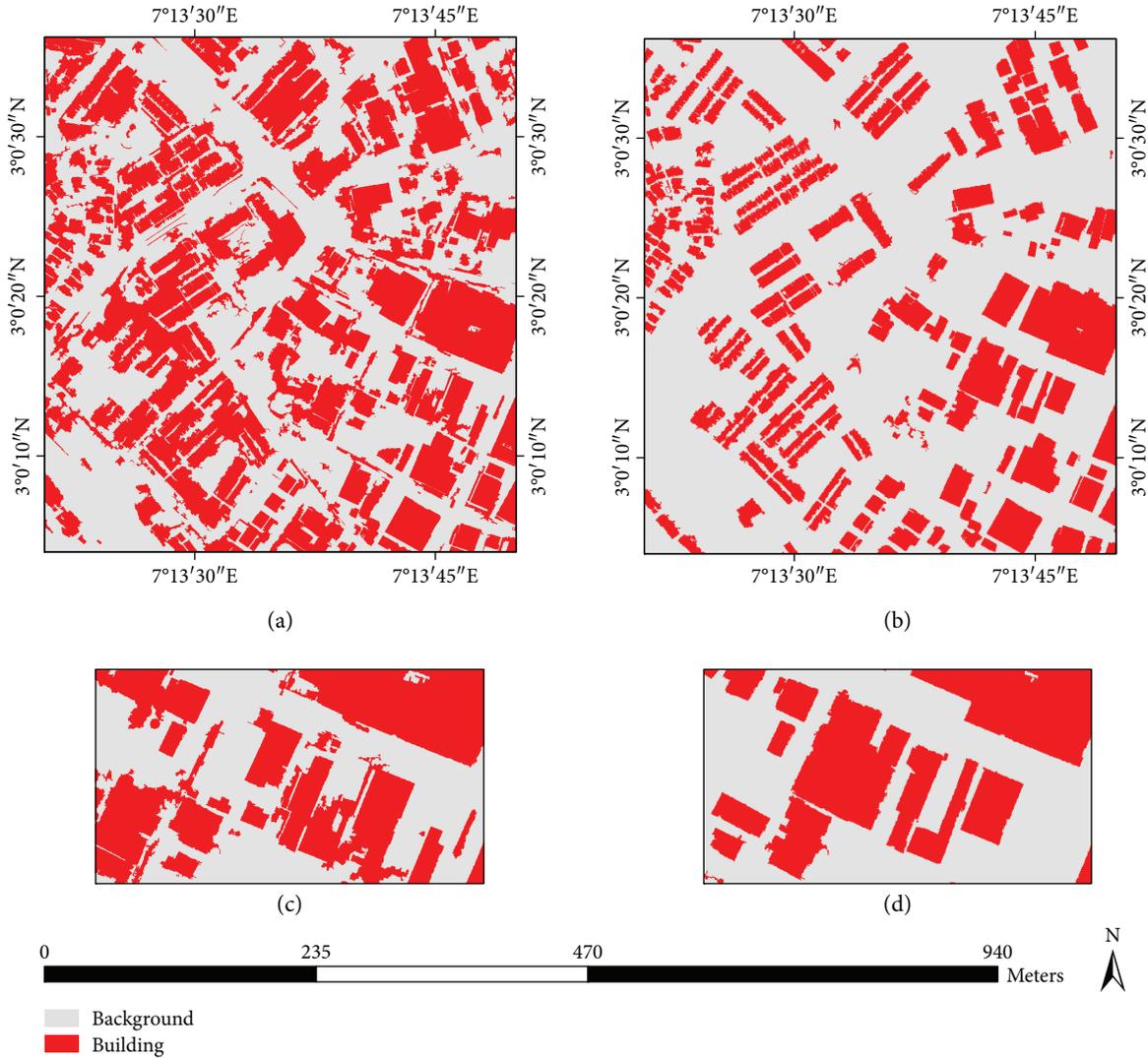


FIGURE 6: Results of building detection using the proposed method in the testing area: (a) without feature dimensionality reduction, (b) with dimensionality reduction (10 features), (c) example of a detected building through a complete set of input features, and (d) example of building detection after feature reduction by the autoencoder dimensionality reduction approach.

higher than the ground truth number. The geometry of the detected objects also shows noisy boundaries and additional nearby objects, such as trees combined with building objects (Figure 6(c)). The noisy boundaries limit the applications of the produced building map due to insufficient accuracy regarding counting and geometry. By contrast, the model that uses autoencoders presents better results (Figures 6(b) and 6(d)). The number of buildings calculated in the map is 256 (1.11% higher than the ground truth). Similarly, the results indicate that reducing the number of features by using the autoencoder approach can improve building detection accuracy and its boundary delineation.

3.3. Sensitivity Analysis. The proposed model has several hyperparameters with significant effects on the accuracy of building detection from the fusion of LiDAR–orthophoto data. Thus, this section presents a sensitivity analysis of these hyperparameters.

TABLE 3: Effects of dimensionality reduction on building detection accuracy.

Number of features	Accuracy (%)	
	Working area	Testing area
21	86.06	77.92
15	85.90	76.71
10	86.19	81.86
5	84.77	78.26

3.3.1. Effects of Dimensionality Reduction. Autoencoders can reduce the dimensionality of input features to a lower number of features by specifying the dimension of its central layer. Table 3 shows the different experiments applied to detect buildings in the input data with different dimensions of the middle layer of the autoencoder model. The dimensions explored were 15, 10, and 5. The model with the complete set of features achieved an accuracy of 86.06% and

77.92% in the working and testing areas, respectively. After reducing the number of features to 15, the model detected the buildings in the areas with accuracies of slightly less than those using the complete set of features. The accuracy in the working and testing areas with 15 features were 85.90% and 76.71%, respectively. In addition, the use of 10 features offered the best results in the working and testing areas with overall accuracies of 86.19% and 81.86%, respectively. By contrast, when the number of features was reduced to 5, the overall accuracy of building detection decreased by 1.29% in the working area and slightly improved in the testing area compared with using the complete set of features.

Autoencoders learn a compressed representation of the input; thus, using the transformed features instead of the complete set of features can reduce noise and redundant information in the features. Although the use of autoencoders in DL models can exhibit lower performance in the training data, better generalization power can still be attained. Moreover, reducing the number of features improves the computing performance of the model while keeping accuracy as high as possible. The use of autoencoders can be more efficient than the multiclass recognition problem or the one-class classification problem for building detection. The main reason for this finding is that detecting one feature type frequently requires relatively fewer significant features than using many features, wherein some of the features may be irrelevant to the task. In the case of multiclass recognition problems, features that are irrelevant to a specific class may be significant for others, and vice versa.

3.3.2. Effects of the CNN Model. The CNN model has several hyperparameters, such as the number of filters, the optimizer, the number of hidden units in the fully connected layer, batch size, and dropout rate. The selection of hyperparameter values significantly affects detection accuracy; therefore, the parameters were carefully analyzed and optimized. Figure 7 shows the results of the sensitivity analysis of these parameters evaluated based on the 10-fold cross-validation accuracy achieved for building detection in the testing area. With regard to the number of filters, the results show that the best number of filters is 128, which achieves an accuracy of 81.86%. The lowest accuracy (15.5%) was obtained with 64 filters. The analysis also shows that the best optimizer is Adam, which realized an accuracy of 81.41% and is significantly better than other methods. By contrast, the number of hidden units in the dense layer has disregarded effects. The highest accuracy (81.86%) was attained by using 10 units or 100 units. The use of 3 units and 50 units obtained slightly lower accuracy (81.61%). The use of a lower number of units in the fully connected layer improves the computing performance of the model; therefore, the optimal value of these parameters is regarded as 10. Furthermore, the sensitivity analysis results show that the best batch size is 8, which achieved an accuracy of 81.86%. The use of a batch size of 4 also attained a slightly similar accuracy (81.32%). However, the use of batch sizes larger than 8 shows a reduction in accuracy of nearly 50%. Finally, the analysis indicates that the dropout rate can have direct effects on the accuracy of building recognition. The best dropout rate is 0.2, which achieved

an accuracy of 81.73%. The combination of the best parameter values is considered the best set of parameters and thus is used to produce the final maps (Figures 5 and 6).

3.4. Comparison with Support Vector Machine (SVM). The proposed model was compared with the traditional machine learning method of SVM. Table 4 shows the accuracy assessment of the different methods applied to detect buildings in the working and testing areas. The results of the comparison experiments in the working area show that the best accuracy (86.19%) was obtained using the proposed model with a lower number of features selected by the autoencoder model. The proposed model without dimensionality reduction also obtained higher accuracy than the SVM models. The results show that the SVM model can achieve relatively good accuracy when its hyperparameters are optimized. However, the SVM model with default parameters can attain the lowest accuracy (76.56%). The experiments in the testing area similarly show that the best accuracy (81.86%) can be obtained using the proposed model with dimensionality reduction. However, the SVM model with optimized hyperparameters outperforms the proposed model without dimensionality reduction. The accuracy of the SVM model with optimization is 79.27%, whereas the DL model without using autoencoders for dimensionality reduction achieved 77.92% building detection accuracy. The SVM model with default parameters obtained the lowest accuracy (74.11%).

Figure 8 presents an example of the building detection results from the testing area for the proposed model and SVM method. Figure 8(a) shows the study subset that contains different building types with various geometric and roofing characteristics. The results of the proposed model without dimensionality reduction are presented in Figure 8(b), whereas those with dimensionality reduction are presented in Figure 8(c). The results that used autoencoders are more accurate, with less noise in nearby building boundaries. For example, the results show that the proposed model with a lower number of features can obtain results that are more precise with regard to building geometry. The buildings in Figure 8(b) were combined, and the model could not detect the features between the buildings. By contrast, the results of the proposed model (with 10 features) present better building separation compared with those using the complete set of features. Furthermore, using the transformed features instead of the original set of features can better distinguish between buildings and nearby trees. The SVM models present relatively similar results. However, the optimized SVM exhibits better detection accuracy and less misclassification between buildings and nearby trees. Moreover, the results of the optimized SVM show better building separation as highlighted by the green circles in Figure 8(d). Overall, the accuracy assessment and visual interpretation of the classification results show that the proposed model is more accurate than the SVM model.

4. Conclusion

This study developed a DL approach based on autoencoders and CNN models to detect buildings in a fused LiDAR–

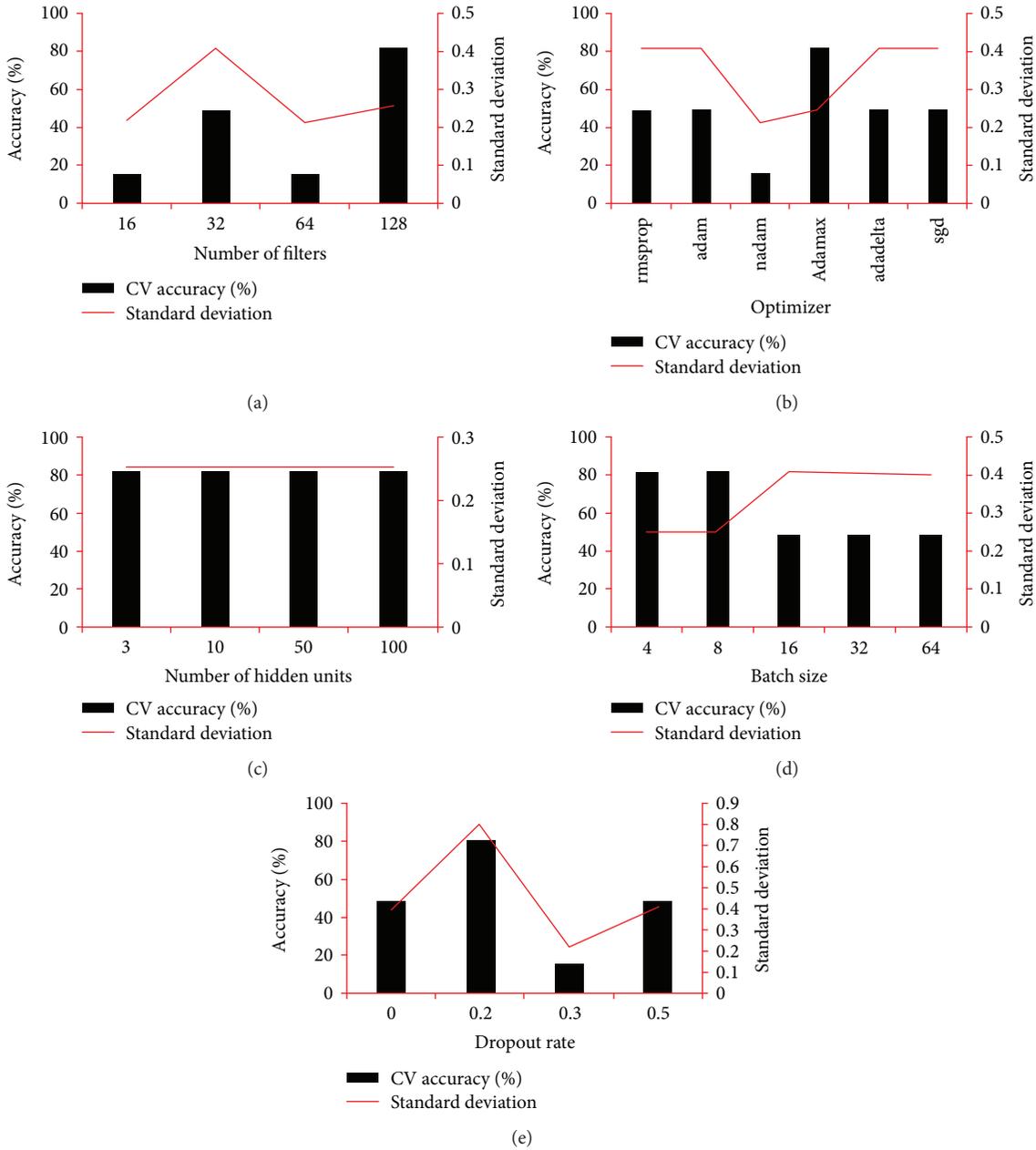


FIGURE 7: Effects of CNN hyperparameters on building detection accuracy.

TABLE 4: Accuracy assessment of the testing area.

Dataset	Model	Properties	Accuracy (%)
Working area	Proposed model	Without dimensionality reduction	86.06
	Proposed model	With dimensionality reduction	86.19
	SVM	Without optimization	76.56
	SVM	With optimization	82.34
Testing area	Proposed model	Without dimensionality reduction	77.92
	Proposed model	With dimensionality reduction	81.86
	SVM	Without optimization	74.11
	SVM	With optimization	79.27

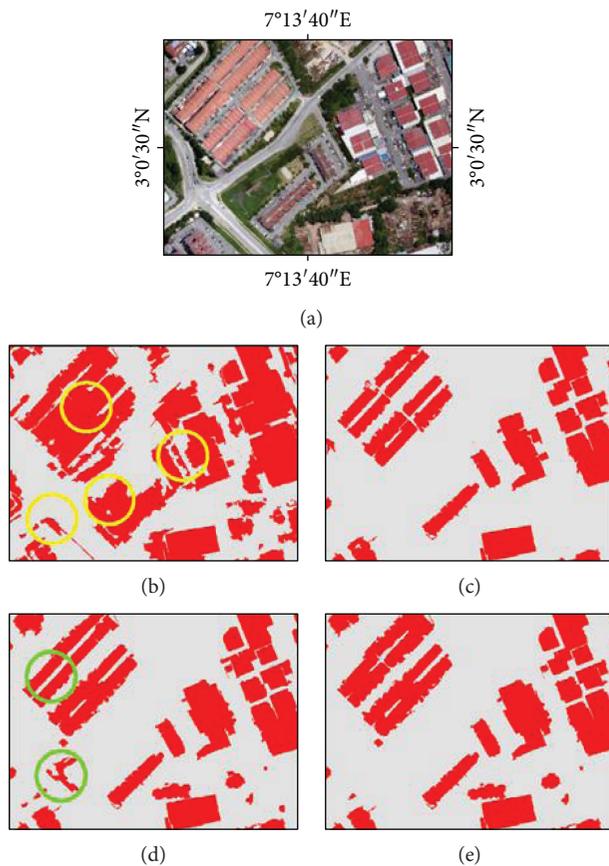


FIGURE 8: Examples of the results of the proposed model and the SVM method: (a) orthophoto of the area, (b) results of the proposed model without dimensionality reduction, (c) results of the proposed model with dimensionality reduction, (d) SVM results without optimization, and (e) SVM results with optimization.

orthophoto dataset. The proposed architecture includes multiresolution and spectral difference segmentations to create objects by grouping the image pixels according to their shape and spectral properties. A total of 21 features from spectral, textural, LiDAR, and spatial features were identified for building detection. These low-level features were then fused at the feature level and compressed into 10 features using the autoencoder model. The compressed features were transformed into high-level features, which were then used to classify the objects into buildings and nonbuildings. The main advantages of applying such architecture to building detection include automatic feature selection and removal of redundant features for improved building detection in datasets.

The main findings of the study suggest that using autoencoders as a dimensionality reduction step can improve the accuracy of building recognition and improve the computing performance of the model. The proposed model achieved the best accuracy of 86.19% in the working area and 81.86% in the testing area. The comparative study shows that the proposed model outperforms the SVM model in the working and testing areas. Furthermore, the sensitivity analysis indicates that the hyperparameters of the DL model and SVM method should be fine-tuned to obtain better accuracy levels

in building detection. Although the proposed method that was determined to be useful for building detection achieves better results than the SVM model, several points still have to be considered in the future. Further research should be performed to improve the proposed model for large-scale building mapping and testing. Future studies should also test whether using satellite images instead of orthophotos can improve accuracy or will only increase the cost of data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank Universiti Putra Malaysia (UPM) (Grant GP-IPS/2016/9491800) for funding the project.

References

- [1] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny, "Automatic building extraction from DEMs using an object approach and application to the 3D-city modeling," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 3, pp. 365–381, 2008.
- [2] I. Tomljenovic, B. Höfle, D. Tiede, and T. Blaschke, "Building extraction from airborne laser scanning data: an analysis of the state of the art," *Remote Sensing*, vol. 7, no. 4, pp. 3826–3862, 2015.
- [3] Y. Li, H. Wu, R. An, H. Xu, Q. He, and J. Xu, "An improved building boundary extraction algorithm based on fusion of optical imagery and LiDAR data," *Optik - International Journal for Light and Electron Optics*, vol. 124, no. 22, pp. 5357–5362, 2013.
- [4] H. Li, C. Zhong, X. Hu, L. Xiao, and X. Huang, "New methodologies for precise building boundary extraction from LiDAR data and high resolution image," *Sensor Review*, vol. 33, no. 2, pp. 157–165, 2013.
- [5] V. Saeidi, B. Pradhan, M. O. Idrees, and Z. Abd Latif, "Fusion of airborne LiDAR with multispectral spot 5 image for enhancement of feature extraction using Dempster-Shafer theory," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6017–6025, 2014.
- [6] M. Uzar and N. Yastikli, "Automatic building extraction using LiDAR and aerial photographs," *Boletim de Ciências Geodésicas*, vol. 19, no. 2, pp. 153–171, 2013.
- [7] M. Uzar, "Automatic building extraction with multi-sensor data using rule-based classification," *European Journal of Remote Sensing*, vol. 47, no. 1, pp. 1–18, 2014.
- [8] M. Awrangjeb, M. Ravanbakhsh, and C. S. Fraser, "Automatic detection of residential buildings using LiDAR data and multispectral imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 5, pp. 457–467, 2010.
- [9] M. Awrangjeb, C. Zhang, and C. S. Fraser, "Building detection in complex scenes thorough effective separation of buildings from trees," *Photogrammetric Engineering & Remote Sensing*, vol. 78, no. 7, pp. 729–745, 2012.
- [10] R. Wang, Y. Hu, H. Wu, and J. Wang, "Automatic extraction of building boundaries using aerial LiDAR data," *Journal of Applied Remote Sensing*, vol. 10, no. 1, article 016022, 2016.

- [11] R. Prerna and C. K. Singh, "Evaluation of LiDAR and image segmentation based classification techniques for automatic building footprint extraction for a segment of Atlantic County, New Jersey," *Geocarto International*, vol. 31, no. 6, pp. 694–713, 2016.
- [12] Z. Zhao, Y. Duan, Y. Zhang, and R. Cao, "Extracting buildings from and regularizing boundaries in airborne lidar data using connected operators," *International Journal of Remote Sensing*, vol. 37, no. 4, pp. 889–912, 2016.
- [13] I. Tomljenovic, D. Tiede, and T. Blaschke, "A building extraction approach for airborne laser scanner data utilizing the object based image analysis paradigm," *International Journal of Applied Earth Observation and Geoinformation*, vol. 52, pp. 137–148, 2016.
- [14] S. Gilani, M. Awrangzeb, and G. Lu, "An automatic building extraction and regularisation technique using LiDAR point cloud data and orthoimage," *Remote Sensing*, vol. 8, no. 3, p. 258, 2016.
- [15] J. S. Evans and A. T. Hudak, "A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 4, pp. 1029–1038, 2007.
- [16] R. M. Haralick and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 610–621, 1973.
- [17] L. Ma, T. Fu, T. Blaschke et al., "Evaluation of feature selection methods for object-based land cover mapping of unmanned aerial vehicle imagery using random forest and support vector machine classifiers," *ISPRS International Journal of Geo-Information*, vol. 6, no. 2, p. 51, 2017.
- [18] M. A. Aguilar, R. Vicente, F. J. Aguilar, A. Fernández, and M. M. Saldaña, "Optimizing object-based classification in urban environments using very high resolution GeoEye-1 imagery," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I-7, pp. 99–104, 2012.
- [19] K. Charoenjit, P. Zuddas, P. Allemand, S. Pattanakiat, and K. Pachana, "Estimation of biomass and carbon stock in Para rubber plantations using object-based classification from Thaichote satellite data in Eastern Thailand," *Journal of Applied Remote Sensing*, vol. 9, no. 1, article 096072, 2015.
- [20] G. Sohn and I. Dowman, "Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 1, pp. 43–63, 2007.
- [21] M. I. Sameen, F. H. Nahhas, F. H. Buraihi, B. Pradhan, and A. R. B. M. Shariff, "A refined classification approach by integrating Landsat Operational Land Imager (OLI) and RADARSAT-2 imagery for land-use and land-cover mapping in a tropical area," *International Journal of Remote Sensing*, vol. 37, no. 10, pp. 2358–2375, 2016.
- [22] J. Zhang, "Multi-source remote sensing data fusion: status and trends," *International Journal of Image and Data Fusion*, vol. 1, no. 1, pp. 5–24, 2010.
- [23] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [24] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232–242, 2016.
- [25] G. Hinton, "A practical guide to training restricted Boltzmann machines," *Momentum*, vol. 9, p. 926, 2010.
- [26] D. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2014, <https://arxiv.org/abs/1412.6980>.
- [27] F. Chollet, "Keras," 2015, <https://keras.io>.
- [28] Y. LeCun, B. Boser, J. S. Denker et al., "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.
- [29] D. Mongus and B. Žalik, "Parameter-free ground filtering of LiDAR data for automatic DTM generation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 67, pp. 1–12, 2012.
- [30] Z. Chen, B. Devereux, B. Gao, and G. Amable, "Upward-fusion urban DTM generating method using airborne Lidar data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 72, pp. 121–130, 2012.

Research Article

A High-Precision Control Scheme Based on Active Disturbance Rejection Control for a Three-Axis Inertially Stabilized Platform for Aerial Remote Sensing Applications

Xiangyang Zhou ^{1,2}, Chao Yang,¹ Beilei Zhao,¹ Libo Zhao,² and Zhuangsheng Zhu ¹

¹School of Instrumentation Science and Opto-electronics Engineering, Beihang University (BUAA), Beijing 100191, China

²State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Correspondence should be addressed to Xiangyang Zhou; xyzhou@buaa.edu.cn and Zhuangsheng Zhu; zszhu@buaa.edu.cn

Received 1 December 2017; Revised 12 February 2018; Accepted 5 April 2018; Published 17 July 2018

Academic Editor: Biswajeet Pradhan

Copyright © 2018 Xiangyang Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a high-precision control scheme based on active disturbance rejection control (ADRC) to improve the stabilization accuracy of an inertially stabilized platform (ISP) for aerial remote sensing applications. The ADRC controller is designed to suppress the effects of the disturbance on the stabilization accuracy that consists of a tracking differentiator, a nonlinear state error feedback, and an extended state observer. By the ADRC controller, the effects of both the internal uncertain dynamics and the external multisource disturbances on the system output are compensated as a total disturbance in real time. The disturbance rejection ability of the ADRC is analyzed by simulations. To verify the method, the experiments are conducted. The results show that compared with the conventional PID controller, the ADRC has excellent capability in disturbance rejection, by which the effect of main friction disturbance on the control system can be weakened seriously and the stabilization accuracy of the ISP is improved significantly.

1. Introduction

It is of crucial importance for an aerial remote sensing system to obtain the high-resolution images under a nonideal multisource disturbance environment. The inertially stabilized platform (ISP) is a key component for an aerial remote sensing system, which is used to hold and control the line of sight (LOS) of the imaging sensors keeping steady relative to the tracking object in the inertial space [1, 2]. If the high-resolution images of the imaging sensor were obtained, the ISP should have the capability to effectively isolate the attitude changes of aircraft in three axes; besides, it is required to suppress the multisource disturbances whatever inside or outside of the aircraft body so that the high control precision can be gotten. Therefore, the most critical performance metric for the ISP is disturbance rejection.

Generally, the ISP is placed between the aircraft and imaging sensor served as a physical interface. It is a principal issue for the control system of the ISP to minimize the effects

of disturbances introduced on the ISP [1]. In [2], a proportional integral control scheme is proposed to realize zero steady-state error of angular output when facing disturbances from carrier. In [3], a hybrid adaptive fuzzy proportional-integral-derivative (PID) controller is developed for a servo system with nonlinear property and uncertainties. In [4], a new current loop inside rate loop is particularly designed to suppress the influences of voltage fluctuation and motor back electromotive force on control precision. In [5], the model uncertainties and unmeasurable disturbances existing objectively are solved by introducing internal model control for a two-axis ISP system. In [6], an approach to compensate disturbing torque based on acceleration feed-forward for the three-axis ISP is proposed. In [7], a parameter identification method for LuGre model based on multilevel coordinate search algorithm was presented. In [8], a compound control scheme based on both of the model reference adaptive control (MRAC) and PID control methods is proposed.

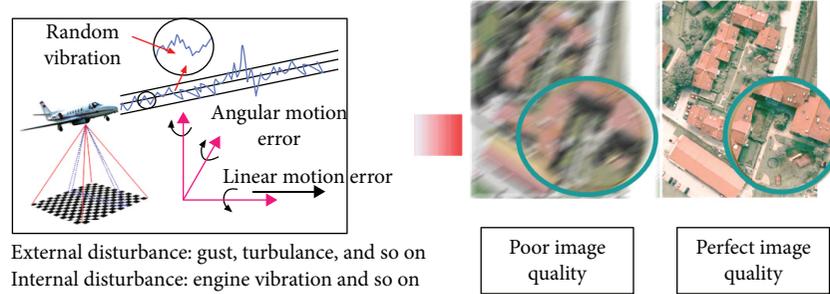


FIGURE 1: Schematic diagram: effect of the ISP on improving the images quality in an aerial remote sensing system.

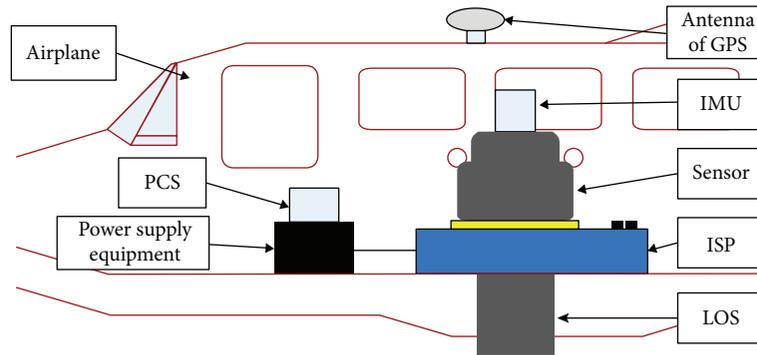


FIGURE 2: Schematic diagram of an aerial remote sensing system.

The active disturbance rejection control (ADRC) was proposed to deal with the nonlinear systems, where both uncertain dynamics and disturbances exist. The ADRC is composed of three components, that is, a tracking differentiator (TD), a nonlinear state error feedback (NLSEF), and an extended state observer (ESO). The key in the ADRC is to online estimate the total uncertainty, which lumps the internal uncertain dynamics and the external disturbances by the ESO. The ADRC can solve quickly and effectively the problems of nonlinear, strong coupling, and so on [9, 10]. The ADRC not only can improve the dynamic performance of the system but also can achieve good adaptability and robustness [11, 12]. A large number of simulation experiments show that the ADRC algorithm is simple to detect and compensate the disturbance of the control object automatically [13, 14]. This design methodology requires very little information about the plant dynamics, is very easy to tune, and has a very good disturbance rejection capability. In [15], a feedback linearization based the ESO control law is proposed for the trajectory tracking control of a flexible joint robotic system. In [16], the effectiveness of the ESO in the estimation of the states and uncertainties is demonstrated. In [17], an ADRC for the antenna pointing control of a large flexible satellite system is designed, by which the high pointing accuracy and rotation speed are achieved.

In this paper, to improve the ability of disturbance rejection of an aerial three-axis ISP, a disturbance rejection scheme based on the ADRC is proposed. To verify the method, the simulations and experiments are carried out, respectively. The rest of this paper is arranged as follows.

Section 2 first introduces the background of this work. Section 3 explains the ADRC principle and algorithm that includes three parts: the TD, the NLSEF, and the ESO. Section 4 demonstrates the simulation analysis, including ADRC controller as well as the PID controller. The ADRC experiment is in Section 5, including the static base experiment and the moving base experiment. In the end, a brief conclusion is given in Section 6.

2. Background

2.1. Aerial Remote Sensing System. Figure 1 shows a schematic diagram to illustrate the important effect of the ISP on improving the image quality in an aerial remote sensing system. In general, an aerial remote sensing system consists of an aircraft vehicle, an imaging sensor, a three-axis ISP, and a position and orientation system (POS). During the work, the ISP is used as a physical interface to separate the imaging sensor and the aircraft. Depending on the ISP, the serious influences of various disturbances on the imaging sensors are effectively suppressed, by which the LOS can be kept stable all along and the high-resolution images are gotten eventually.

2.2. The Role of Three-Axis ISP System. Figure 2 shows the schematic diagram of an aerial remote sensing system. As shown in this figure, during application, the ISP is installed on the floor of the airplane. Meanwhile, the imaging sensor is installed on the inner azimuth gimbal of the ISP, and the POS is installed on the top of the imaging

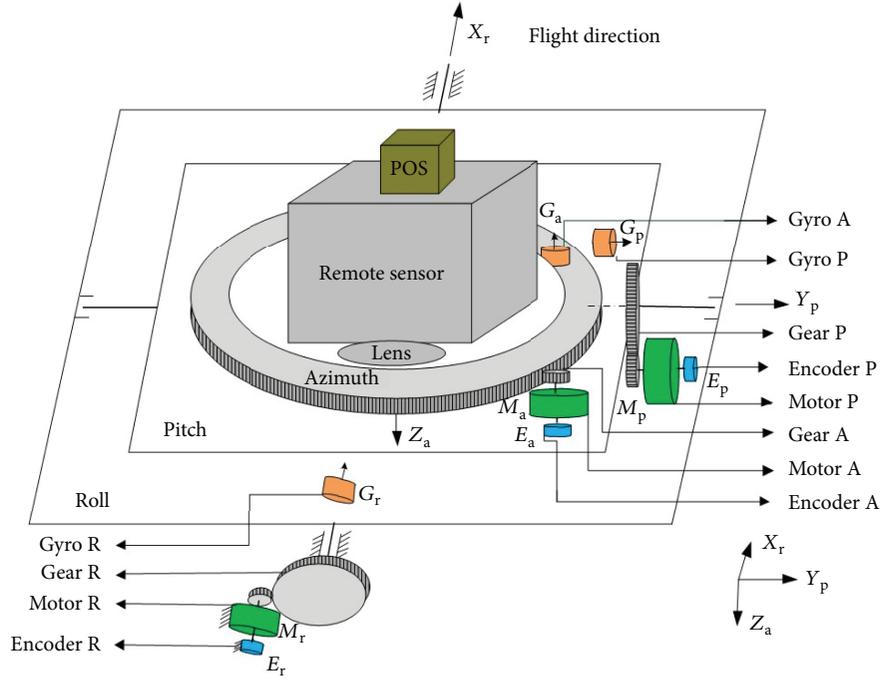


FIGURE 3: Schematic diagram of the three-axis ISP's principle.

sensor. When the aircraft rotates or jitters, the high-precision attitude reference information is measured by the POS in real time, which is then provided to the ISP control system to make the LOS of imaging sensor to achieve the high stabilization precision.

2.3. Working Principle of Three-Axis ISP System. Figure 3 shows the schematic diagram of the three-axis ISP's principle. As shown in the figure, the ISP consists of three gimbals, which are azimuth gimbal (A-gimbal), pitch gimbal (P-gimbal), and roll gimbal (R-gimbal) from the outside to the inside. Three rate gyros installed on the three different gimbals are used to measure inertial angular rate of P-gimbal, R-gimbal, and A-gimbal, which are used to improve the stabilization precision through compensating the difference between the rate command input and the angular rate of the gimbal. On the other aspects, three photoelectric encoders installed on different gimbals are used to measure the relative angular position changes between two adjacent gimbals. Three servo motors installed on different gimbals are used to drive R-gimbal, P-gimbal, and A-gimbal rotating inversely to compensate the attitude changes to three gimbals in real time.

3. Design of the ADRC Controller

3.1. Tracking Differentiator (TD). For a second-order system [18], we get

$$\begin{aligned} \dot{z}_1 &= z_2, \\ \dot{z}_2 &= f(z_1, z_2), \end{aligned} \quad (1)$$

if the solution of (1) satisfies

$$\begin{aligned} \lim_{t \rightarrow \infty} z_1(t) &= 0, \\ \lim_{t \rightarrow \infty} z_2(t) &= 0. \end{aligned} \quad (2)$$

So, for any boundary integrable variable, $v(t)$, and any constants, $T > 0$ and $r > 0$, the solution $x_1(t)$ of the system

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= r^2 \cdot f\left(x_1 - v(t), \frac{x_2}{r}\right) \end{aligned} \quad (3)$$

can meet the following relationship:

$$\lim_{r \rightarrow \infty} \int_0^T |x_1(r, t) - v(t)| dt = 0. \quad (4)$$

Therefore, assuming that a function $f(z_1, z_2)$ can be found to make (1) asymptotically stable at the origin, a non-linear tracking differentiator of $v(t)$ in the form of (3) can be constructed, in which the same function $f(z_1, z_2)$ is used [19].

Kasim and Riyadh [20] have made some investigations on traditional structures and essential properties of the non-linear TD. A kind of second-order nonlinear TD based on second-order bang-bang switch system has been proposed:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -r \cdot \text{sign}\left(x_1 - v_0(t) + \frac{x_2 |x_2|}{2r}\right), \end{aligned} \quad (5)$$

where x_1 is the desired trajectory and x_2 is its derivative. Note that the parameter r is an application dependent and it is set

accordingly to speed up or slow down the transient profile. Then, x_2 is denoted as the TD of $v_0(t)$.

In order to avoid chattering near the origin, changing the sign function to linear saturation function sat , then the modified TD is represented by [21–23]

$$\text{sat}(x, \delta) = \begin{cases} \text{sign}(x), & |x| \geq \delta, \\ x, & |x| < \delta. \end{cases} \quad (6)$$

The TD gets the differentiated signal by computing the integration of the relational viable. So it can efficiently weaken the effect of noise, while the classical differential component enlarges it and the differentiated signal obtained by this way is far from the real one. In some sense, the differential component of the TD is feasible physically, while differential components of the classical methods are not feasible physically [21].

3.2. Nonlinear State Error Feedback (NLSEF). For the first-order system [22],

$$\dot{x} = w(x, t) + u. \quad (7)$$

$w(x, t)$ is the sum of the perturbations of the system, and linear feedback is used for the control object:

$$u = -kx, \quad k > 0. \quad (8)$$

Then, the closed-loop system can be expressed:

$$\dot{x} = -kx + w(x, t). \quad (9)$$

The closed-loop equation of the finishing system is obtained [23]:

$$\frac{1}{2} \frac{dx^2}{dt} = -kx^2 + xw = -k \left(x - \frac{w}{2k} \right)^2 + \frac{w^2}{4k}. \quad (10)$$

If the state quantity x converges, then $dx^2/dt < 0$ is needed, that is to say, it needs to be satisfied:

$$\left| x - \frac{w}{2k} \right| > \frac{|w|}{2k}. \quad (11)$$

It is easy to get

$$|x| > \frac{|w|}{k}. \quad (12)$$

As long as $|x| > |w|/k$, there must be $dx^2/dt < 0$. The state parameter x converges, and the system disturbance $w(x, t)$ can also be well suppressed. But with the convergence of the state quantity x , it will gradually approach the range of $|x| \leq |w|/k$, which leads to the suppression of the disturbance of the system and the steady-state error of the system.

The ADRC reduces the steady-state error of the system by adopting appropriate nonlinear functions. According to the different error states, the nonlinear function can make the feedback gain appropriate increase and decrease, so as to reduce the steady-state error of the system, and effectively avoid the disadvantages caused by adopting the fixed feedback gain.

Nonlinear feedback is used to control objects:

$$u = -k|x|^\alpha \text{sign}(x), \quad 0 < \alpha < 1. \quad (13)$$

Then, the controlled object is converted into a closed loop:

$$\dot{x} = w(x, t) - k|x|^\alpha \text{sign}(x). \quad (14)$$

Then get

$$\frac{1}{1+\alpha} \frac{d|x|^{1+\alpha}}{dt} = -k \left(|x|^\alpha \text{sign}(x) - \frac{w(x, t)}{2k} \right)^2 + k \left(\frac{w(x, t)}{2k} \right)^2. \quad (15)$$

When $d|x|^{1+\alpha}/dt < 0$, the state x is convergent, which requires

$$|x|^\alpha \text{sign}(x) - \frac{w(x, t)}{2k} > \frac{w(x, t)}{2k}. \quad (16)$$

Then get

$$|x| > \left(\frac{|w(x, t)|}{k} \right)^{1/\alpha}. \quad (17)$$

According to the analysis of the upper form, when $|x| > (|w(x, t)|/k)^{1/\alpha}$, the state x converges, and with the further convergence, state x can converge to the range of $|x| \leq (|w(x, t)|/k)^{1/\alpha}$, so the steady-state error of the system will not be greater than $(|w(x, t)|/k)^{1/\alpha}$.

It is easier to know the linear feedback obtained the maximum steady-state error $|w(x, t)|/k$ compared with $(|w(x, t)|/k)^{1/\alpha}$ that when $0 < \alpha < 1$ and $k > w(x, t)$, there are

$$\left(\frac{|w(x, t)|}{k} \right)^{1/\alpha} \ll \frac{|w(x, t)|}{k}. \quad (18)$$

For systems subject to uncertainties, the steady-state error obtained by using nonlinear feedback is much smaller than that by using linear feedback. In addition, in nonlinear feedback, if the α value is changed slightly, the $(|w(x, t)|/k)^{1/\alpha}$ can rapidly approach zero.

The TD generates the arranged transition process and its derivatives. The outputs of ESO estimate the states of the control system. By comparing the difference between the outputs of TD and those of ESO, the NLSEF control law $u_0(t)$ is used to drive the state trajectory to the desired reference signal. Its mathematic expression is given as [24]

$$\varepsilon_i = v_i - z_i \quad (i = 1, 2, \dots, n). \quad (19)$$

Thus, the nonlinear combination of system state error feedback is formed:

$$u_0(t) = \sum_{i=1}^n \beta_i \text{fal}(\varepsilon_i, \alpha_i, \delta), \quad (20)$$

where α_i, β_i ($i = 1, 2, \dots, n$), and δ are adjustable parameters.

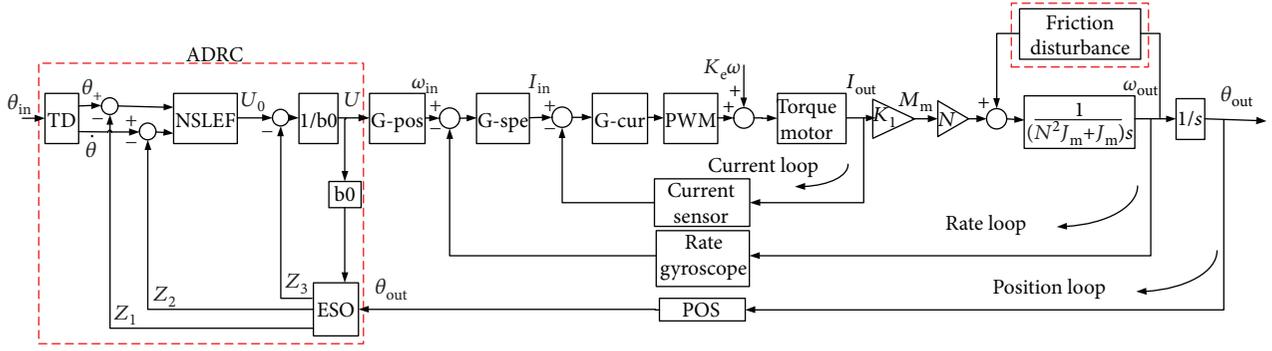


FIGURE 4: Block diagram of the ADRC-based three-loop compound controller.

$$\text{fal}(e) = \begin{cases} |e|^\alpha \text{sign}(e), & |e| > c, \\ \frac{e}{c^{1-\alpha}}, & |e| \leq c, \end{cases} \quad (21)$$

where $\text{fal}(e)$ is defined as a modified exponential gain function [25].

3.3. Extended State Observer (ESO). The performance of the ESO is very important to the performance of the controller [18, 20]. The state observer is used to determine the internal state of the system by observing the external state of the system in different situations.

If the function $f(x_1, x_2)$ is not known and the above description is not valid, the way to expand the state observer is to expand $f(x_1, x_2)$ into a new state variable x_3 :

$$\begin{aligned} \dot{x}_3 &= f(x_1(t), x_2(t)), \\ \dot{x}_3 &= w(t). \end{aligned} \quad (22)$$

In this way, nonlinear objects can be regarded as new linear objects [25]:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_3 + bu, \\ \dot{x}_3 &= w(t), \\ y &= x_1. \end{aligned} \quad (23)$$

Then, the state observer is established:

$$\begin{aligned} e_1 &= z_1 - y, \\ \dot{z}_1 &= z_2 - \beta_{01}e_1, \\ \dot{z}_2 &= z_3 - \beta_{02}|e_1|^{1/2} \text{sign}(e_1) + bu, \\ \dot{z}_3 &= -\beta_{03}|e_1|^{1/4} \text{sign}(e_1). \end{aligned} \quad (24)$$

In this way, the provided parameters β_{01} , β_{02} , and β_{03} are selected properly, the state variables of the system can be estimated in real time by z_1 , z_2 , and z_3 , and the real-time action amount of the acceleration $f(x_1, x_2)$ of the system is

measured. Equation (25) can not only achieve good control effect but also avoid high-frequency tremor.

$$\text{fal}(e, \alpha, \delta) = \begin{cases} \frac{e}{\delta^{\alpha-1}}, & |e| \leq \delta, \\ |e|^\alpha \text{sign}(e), & |e| > \delta, \end{cases} \quad (25)$$

where δ is the length of the linear section.

$$\begin{aligned} z_1(t) &\longrightarrow x_1(t), \\ z_2(t) &\longrightarrow x_2(t), \\ z_3(t) &\longrightarrow f(x_1(t), x_2(t)). \end{aligned} \quad (26)$$

If $f(x_1, x_2)$ contains time t and unknown disturbances $w(t)$, it is available [17]:

$$z_3 \longrightarrow x_3 = f[x_1(t), x_2(t), w(t)]. \quad (27)$$

The real-time action of acceleration acting on the system can be obtained:

$$a(t) = f[x_1(t), x_2(t), t, w(t)]. \quad (28)$$

Thus, (28) can become the extended state observer of the system. The ESO is a dynamic process that does not take advantage of system parameters and specific models but uses the information between inputs and outputs of the system to estimate the system state variables directly.

3.4. The ADRC-Based Three-Loop Compound Controller.

Figure 4 shows the block diagram of the ADRC-based three-loop compound controller. As shown in this figure, the blocks of G-cur, G-spe, and G-pos indicate the controllers from the inner loop to the outer loop, which are the current loop, rate loop, and position loop, respectively; the PWM block indicates the pulse width modulation, which is used for the current amplification to drive the torque motor; the symbol L , R , K_t , and N are the inductance of a torque motor, resistance, motor torque coefficient, and transition ratio from the torque motor to the gimbal, respectively. Besides, J_m and J_t indicate the moment of inertia of the motor and the moment of the gimbal inertia along the rotation axis, respectively.

As shown in this figure, in order to improve the control accuracy of the three-axis ISP system, the ADRC controller

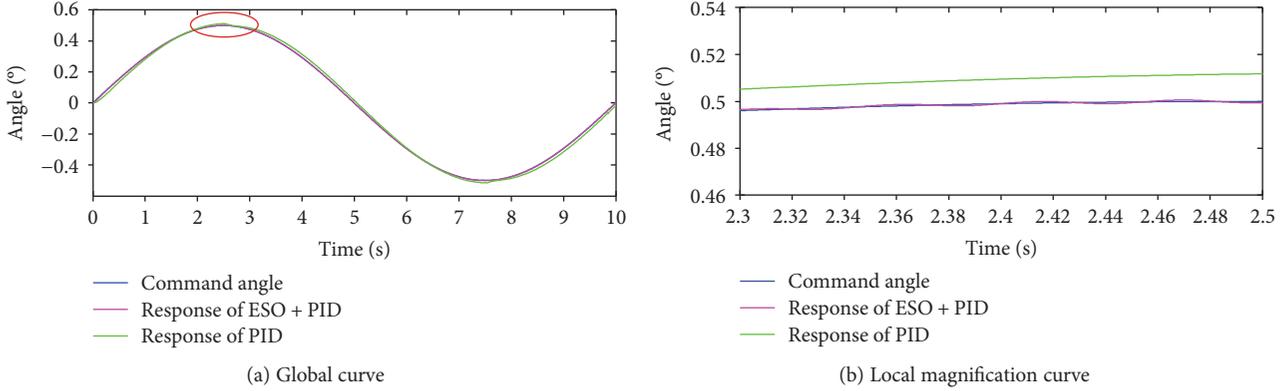


FIGURE 5: PID and ESO/PD sine angle position feedback contrast diagram.

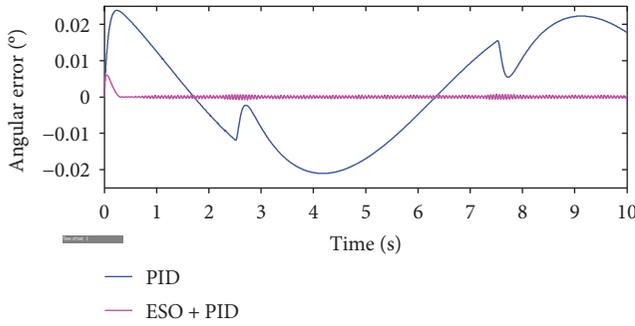


FIGURE 6: PID and ESO/PD angle position error contrast diagram.

is designed. One of the inputs of the ADRC is the tracking instruction of angular position of the ISP control system, θ_{in} , and another input of the ADRC, θ_{out} , is the real angular position of the ISP measured by the POS. The control amount u calculated by the ADRC is acted on the ISP position loop.

In Figure 4, the generalized inputs, θ and $\dot{\theta}$, are first obtained by the TD. The generalized outputs of the ISP control system, z_1 and z_2 , are obtained by the ESO. z_1 is the angle value measured by POS in the position loop of the three-axis ISP system, z_2 is the derivative of the angle, and z_3 is the compensatory amount of the ESO to the disturbance in the ISP control system. The compared errors between the generalized inputs and the generalized outputs are inputted to the NLSEF to generate the control amount u_0 , which are compared with z_3 to generate the control amount u . The control amount u after compensation is acted on the position loop of the ISP. Thus, the disturbance rejection ability and the control accuracy of the ISP control system are improved.

4. Simulation Analysis

4.1. Angular Position Response to Step Input. In order to verify the validity of ADRC algorithm, MATLAB/Simulink simulation is firstly carried out. The main parameters of the NLSEF control law are $\{\beta_1, \beta_2, \alpha_3, \alpha_4, \delta_2\}$. Normally, the value of α_3 is from 0.5 to 0.75, and the value of α_4 is from 1.25 to 1.5. The value of δ_2 is the same as the integral step

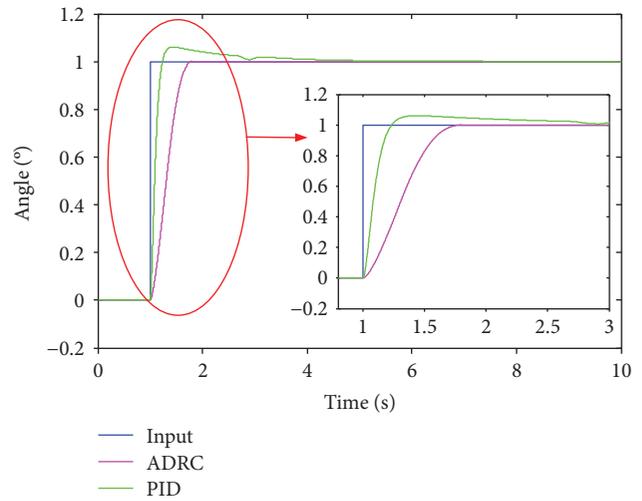


FIGURE 7: Comparison of position feedback between step angle of conventional PID and ADRC controller.

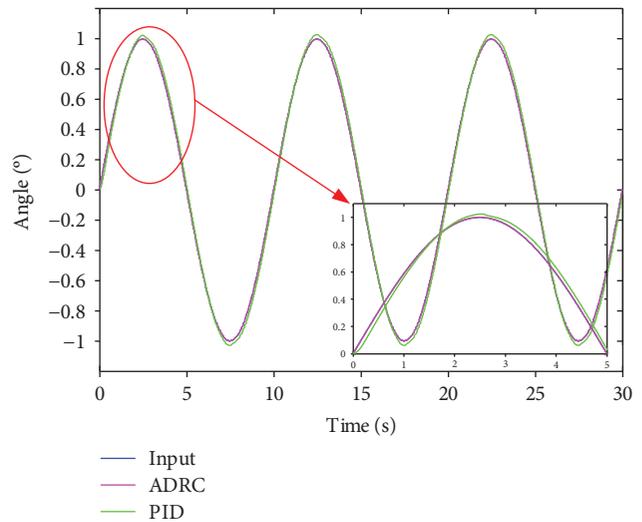


FIGURE 8: Comparison of position feedback to sinusoidal signal between conventional PID and ADRC controller.

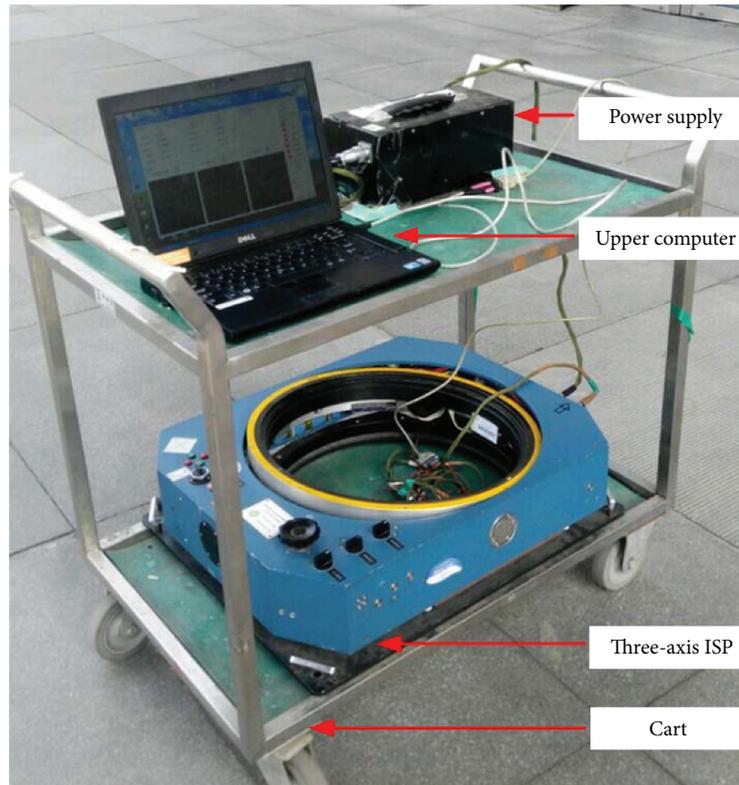


FIGURE 9: Picture of moving base experiment system.

length as 0.01 generally. The parameters tuned by the experimental method are as follows: $\beta_1 = 6.0$ and $\beta_2 = 1.5$.

In the simulation, the pitch gimbal system is chosen as the control object. When the given input angle instruction is $\theta_{id} = 0.5 \sin(\pi \cdot n \cdot T_s)$ deg, $T_d = 0$, the disturbance resistance performance of the PID and the ESO/PD controller is simulated by adding the disturbance model; the results are shown in Figure 5.

In Figure 5(a), the PID and the ESO/PD composite controllers can track the sinusoidal signals in different position precision. In the local magnification diagram of Figure 5(b), we can see that the PID has approximately 0.01° errors in tracking the sinusoidal instruction, but the ESO/PD compound controller tracking curve almost coincided with the sinusoidal instruction. Therefore, the ESO/PD controller has strong disturbance resistance ability, which can effectively suppress the disturbance effect, and its tracking error is almost zero.

Under the influence of disturbance, the comparison of angular position error between the PID and the ESO/PD controller is shown in Figure 6. From Figure 6, we can see that the angular error of PID fluctuates slowly from -0.02° to 0.025° . When the disturbance amplitude is large, the tracking error curve of the PID will increase correspondingly, while the angular error of the ESO/PD composite controller is stable at 0° . Therefore, the ESO/PD controller has better anti-interference ability than the PID controller.

Figure 7 shows the comparison of position feedback to step angle between conventional PID and ADRC controller.

When the track reference is a step signal with magnitude 1° which lasts for 10 s, the ADRC controller has a good dynamic performance compared to the conventional PID controller. As shown in this figure, compared to the conventional PID controller, the ADRC controller has nearly no any overshoot and can more easily reach stabilization state with much shorter adjusting time.

4.2. Angular Position Response to Sinusoidal Input. Figure 8 shows the comparison of position feedback to sinusoidal signal between conventional PID and ADRC controller. When the track reference is sinusoidal signal with amplitude 1° and frequency 0.1 Hz, the ADRC controller can achieve better disturbance rejection performance compared to the conventional PID controller. As shown in this figure, compared to the conventional PID controller, the tracking curve of the ADRC controller coincides almost with the sine command.

5. Experimental Verification

In order to verify the validity and practicability of the ADRC, the experiments are carried out. Figure 9 shows the picture of moving base experiment system. In the experiments, the three-axis ISP is placed on a moving cart. The experimental system includes a three-axis ISP, a data acquisition and processing system (upper computer), and a 28 V DC power supply. The weight of the ISP is 40 kg and its maximum support load is 80 kg. The maximum leveling rotation

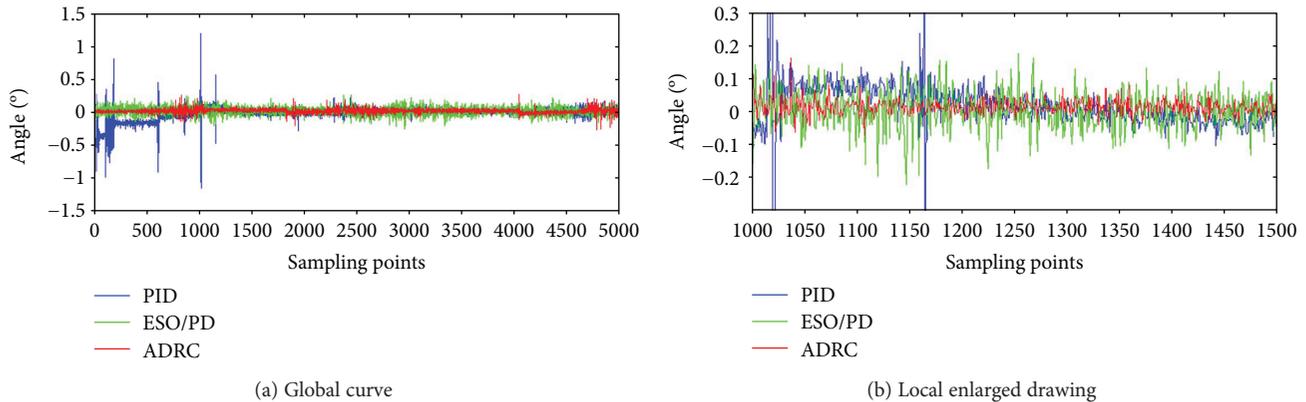


FIGURE 10: Comparison of angle error of dynamic cart experiment.

angle range of the ISP is $\pm 5^\circ$; the maximum heading rotation angle range is $\pm 25^\circ$.

Set the tracking reference angle of the roll gimbal system as 0° ; sampling points is 5000, and the sampling frequency is 50 Hz. As shown in Figure 10, in leveling experiments, the results obtained by three different methods are as follows: for the PID controller, the RMS (root mean square) error is 0.2673° ; for the ESO/PD controller, the RMS error is 0.1315° , which is decreased up to 50.8% compared to the PID controller; for the ADRC controller, the RMS error is 0.0826° , which are decreased up to 69.1% and 37.2% compared to the PID controller and the ESO/PD controller, respectively. The results show that the ADRC control method has better control precision and interference suppression capability compared to the PID and the ESO/PD.

6. Conclusion

This paper investigates the high-precision control method for a three-axis inertially stabilized platform (ISP). Based on the analysis, a disturbance rejection control scheme based on the active disturbance rejection control (ADRC) is proposed to realize high-precision position servo control, which is robust to the multisource disturbances. Then, the numerical simulations and the experimental validations are conducted to analyze and verify the scheme, respectively. The results show that the ADRC control method can get a better control accuracy and have strong disturbance rejection ability than the conventional PID methods. Under the moving base experimental condition, the stabilization precision of the ISP obtained by the ADRC controller is improved up to 69.1% compared to the conventional PID control method.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This project is supported in part by the National Natural Science Foundation of China (Grant nos. 51775017, 51375036), by the Beijing Natural Science Foundation

(Grant no. 3182021), and by the Open Research Fund of the State Key Laboratory for Manufacturing Systems Engineering (sklms2018005).

References

- [1] M. K. Masten, "Inertially stabilized platforms for optical imaging systems," *IEEE Control Systems Magazine*, vol. 28, no. 1, pp. 47–64, 2008.
- [2] Y. Han, Y. Lu, and H. Qiu, "An improved control scheme of gyro stabilization electro-optical platform," in *2007 IEEE International Conference on Control and Automation*, pp. 346–351, Guangzhou, China, 2007.
- [3] W. Ji, Q. Li, and B. Xu, "Design study of adaptive fuzzy PID controller for LOS stabilized system," in *Sixth International Conference on Intelligent Systems Design and Applications*, pp. 336–341, Jinan, China, 2006.
- [4] X. Zhou, Y. Jia, Q. Zhao, and R. Yu, "Experimental validation of a compound control scheme for a two-axis inertially stabilized platform with multi-sensors in an unmanned helicopter-based airborne power line inspection system," *Sensors*, vol. 16, no. 3, p. 366, 2016.
- [5] X. Zhou, H. Zhang, and R. Yu, "Decoupling control for two-axis inertially stabilized platform based on an inverse system and internal model control," *Mechatronics*, vol. 24, no. 8, pp. 1203–1213, 2014.
- [6] Q. Mu, G. Liu, M. Zhong, and Z. Chu, "Imbalance torque compensation for three-axis inertially stabilized platform using acceleration feedforward," in *2012 8th IEEE International Symposium on Instrumentation and Control Technology (ISICT) Proceedings*, pp. 157–160, London, UK, 2012.
- [7] Z. Li, D. Fan, and S. Fan, "LuGre-Model-based friction compensation in direct-drive inertially stabilized platforms," *IFAC Proceedings Volumes*, vol. 46, no. 5, pp. 636–642, 2013.
- [8] X. Zhou, C. Yang, and T. Cai, "A model reference adaptive control/PID compound scheme on disturbance rejection for an aerial inertially stabilized platform," *Journal of Sensors*, vol. 2016, Article ID 7964727, 11 pages, 2016.
- [9] W. Xue and Y. Huang, "Comparison of the DOB based control, a special kind of PID control and ADRC," in *Proceedings of the 2011 American Control Conference*, pp. 4373–4379, San Francisco, CA, USA, 2011.
- [10] W. Xue, W. Bai, S. Yang, K. Song, Y. Huang, and H. Xie, "ADRC with adaptive extended state observer and its application to air–fuel ratio control in gasoline engines," *IEEE*

- Transactions on Industrial Electronics*, vol. 62, no. 9, pp. 5847–5857, 2015.
- [11] Y. Huang and J. Han, “Analysis and design for the second order nonlinear continuous extended states observer,” *Chinese Science Bulletin*, vol. 45, no. 21, pp. 1938–1944, 2000.
- [12] Y. X. Su, C. H. Zheng, and B. Y. Duan, “Automatic disturbances rejection controller for precise motion control of permanent-magnet synchronous motors,” *IEEE Transactions on Industrial Electronics*, vol. 52, no. 3, pp. 814–823, 2005.
- [13] W. Xue and Y. Huang, “On performance analysis of ADRC for a class of MIMO lower-triangular nonlinear uncertain systems,” *ISA Transactions*, vol. 53, no. 4, pp. 955–962, 2014.
- [14] Q. Zheng, L. Dong, D. H. Lee, and Z. Gao, “Active disturbance rejection control for MEMS gyroscopes,” *IEEE Transactions on Control Systems Technology*, vol. 17, no. 6, pp. 1432–1438, 2009.
- [15] S. E. Talole, J. P. Kolhe, and S. B. Phadke, “Extended-state-observer-based control of flexible-joint system with experimental validation,” *IEEE Transactions on Industrial Electronics*, vol. 57, no. 4, pp. 1411–1419, 2010.
- [16] C. Zhao and Y. Huang, “ADRC based input disturbance rejection for minimum-phase plants with unknown orders and/or uncertain relative degrees,” *Journal of Systems Science and Complexity*, vol. 25, no. 4, pp. 625–640, 2012.
- [17] S. Li, X. Yang, and D. Yang, “Active disturbance rejection control for high pointing accuracy and rotation speed,” *Automatica*, vol. 45, no. 8, pp. 1854–1860, 2009.
- [18] Y. Tang, Y. Wu, M. Wu, X. Hu, and L. Shen, “Nonlinear tracking-differentiator for velocity determination using carrier phase measurements,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 4, pp. 716–725, 2009.
- [19] P. Wang and K. Zhang, “Research on line-of-sight rate extraction of strapdown seeker,” in *Proceedings of the 33rd Chinese Control Conference*, pp. 859–863, Nanjing, China, 2014.
- [20] I. Kasim and W. Riyadh, “On the improved nonlinear tracking differentiator based nonlinear PID controller design,” *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 10, 2016.
- [21] H. Huang, L. Wu, J. Han, G. Feng, and Y. Lin, “A new synthesis method for unit coordinated control system in thermal power plant - ADRC control scheme,” in *2004 International Conference on Power System Technology, 2004. PowerCon 2004*, pp. 133–138, Singapore, 2004.
- [22] J. Han, “Nonlinear state error feedback control law—NLSEF,” *Control and Decision*, vol. 3, 1995.
- [23] J. Han, “From PID technique to active disturbances rejection control technique,” *Control Engineering of China*, vol. 9, pp. 13–18, 2002.
- [24] W. Jun, Y. J. Wang, and H. Jian, “Artificial bee colony algorithm based auto-disturbance rejection control for rehabilitation robotic arm driven by PM-TS actuator,” in *2012 Proceedings of International Conference on Modelling, Identification and Control*, pp. 802–807, Wuhan, Hubei, China, 2012.
- [25] Y. Bing, Z. H. Tian, and S. J. Shi, “Novel fault diagnosis scheme for HVDC system via ESO,” *High Voltage Engineering*, vol. 11, 2007.

Research Article

Classification of Very High Resolution Aerial Photos Using Spectral-Spatial Convolutional Neural Networks

Maher Ibrahim Sameen, Biswajeet Pradhan , and Omar Saud Aziz

School of Systems, Management and Leadership, Faculty of Engineering and Information Technology, University of Technology Sydney, Building 11, Level 06, 81 Broadway, P.O. Box 123, Ultimo, NSW 2007, Australia

Correspondence should be addressed to Biswajeet Pradhan; biswajeet24@gmail.com

Received 3 March 2018; Revised 17 April 2018; Accepted 6 May 2018; Published 26 June 2018

Academic Editor: Paolo Bruschi

Copyright © 2018 Maher Ibrahim Sameen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Classification of aerial photographs relying purely on spectral content is a challenging topic in remote sensing. A convolutional neural network (CNN) was developed to classify aerial photographs into seven land cover classes such as building, grassland, dense vegetation, waterbody, barren land, road, and shadow. The classifier utilized spectral and spatial contents of the data to maximize the accuracy of the classification process. CNN was trained from scratch with manually created ground truth samples. The architecture of the network comprised of a single convolution layer of 32 filters and a kernel size of 3×3 , pooling size of 2×2 , batch normalization, dropout, and a dense layer with Softmax activation. The design of the architecture and its hyperparameters were selected via sensitivity analysis and validation accuracy. The results showed that the proposed model could be effective for classifying the aerial photographs. The overall accuracy and Kappa coefficient of the best model were 0.973 and 0.967, respectively. In addition, the sensitivity analysis suggested that the use of dropout and batch normalization technique in CNN is essential to improve the generalization performance of the model. The CNN model without the techniques above achieved the worse performance, with an overall accuracy and Kappa of 0.932 and 0.922, respectively. This research shows that CNN-based models are robust for land cover classification using aerial photographs. However, the architecture and hyperparameters of these models should be carefully selected and optimized.

1. Introduction

Classifying remote sensing data (especially orthophotos of three bands—red, green, blue (RGB)) with traditional methods is a challenge even though some methods in literature have produced excellent results [1, 2]. The main reason behind is that remote sensing datasets have high intra- and interclass variability and the amount of labeled data is much smaller as compared to the total size of the dataset [3]. On the other hand, the recent advances in deep learning methods like convolutional neural networks (CNNs) have shown promising results in remote sensing image classification especially hyperspectral image classification [4–6]. The advantages of deep learning methods include learning high-order features from the data that are often useful than the raw pixels for classifying the image into some predefined labels. Other advantages of these methods are spatial learning

of contextual information from data via feature pooling from a local spatial neighborhood [3].

There are several methods and algorithms that have been adopted by many researchers to efficiently classify a very high-resolution aerial photo and produce accurate land cover maps. Methods such as object-based image analysis (or OBIA) was mostly investigated because of its advantage in very high-resolution image processing via spectral and spatial features. In a recent paper, Hsieh et al. [7] applied aerial photo classification by combining OBIA with decision tree using texture, shape, and spectral feature. Their results achieved an accuracy of 78.20% and a Kappa coefficient of 0.7597. Vogels et al. [8] combined OBIA with random forest classification with texture, slope, shape, neighbor, and spectral information to produce classification maps for agricultural areas. They have tested their algorithm on two datasets, and the results showed the employed methodology

to be effective with accuracies of 90% and 96% for the two study areas, respectively. On the other hand, a novel model was presented by Meng et al. [9], where they applied OBIA to improve vegetation classification based on aerial photos and global positioning systems. Results illustrated a significant improvement in classification accuracy that increased from 83.98% to 96.12% in overall accuracy and from 0.7806 to 0.947 in the Kappa value. Furthermore, Juel et al. [10] showed that random forest with the use of a digital elevation model could achieve relatively high performance for vegetation mapping. In a most recent paper, Wu et al. [2] developed a model based on a comparison between pixel-based decision tree and object-based SVM to classify aerial photos. The object-based support vector machine (SVM) had higher accuracy than that of the pixel-based decision tree. Albert et al. [11] developed classifiers based on conditional random fields and pixel-based analysis to classify aerial photos. Their results showed that such techniques are beneficial for land cover classes covering large, homogeneous areas.

2. Related Works

The success of CNN in the fields like computer vision, language modeling, and speech recognition has motivated the remote sensing scientists to apply it in image classification. There are several works that have been done on CNN for remote sensing image classification [12–15]. This section briefly explains some of these works highlighting their findings and their limitations.

Sun et al. [16] proposed an automated model for feature extraction and classification with classification refinement by combining random forest and CNN. Their combined model could perform well (86.9%) and obtained higher accuracy than the single models. Akar [1] developed a model based on rotation forest and OBIA to classify aerial photos. Results were compared to gentle AdaBoost, and their experiments suggested that their method performed better than the other method with 92.52% and 91.29% accuracies, respectively. Bergado et al. [17] developed deep learning algorithms based on CNN for aerial photo classification in high-resolution urban areas. They used data from optical bands, digital surface models, and ground truth maps. The results showed that CNN is very effective in learning discriminative contextual features leading to accurate classified maps and outperforming traditional classification methods based on the extraction of textural features. Scott et al. [13] applied CNN to produce land cover maps from high-resolution images. Other researchers such as Cheng et al. [12] used CNN as a classification algorithm for scene understanding from aerial imagery. Furthermore, Sherrah [14] and Yao et al. [15] used CNN for semantic classification of aerial images.

This research investigates the development of a CNN model with regularization techniques such as dropout and batch normalization for classifying aerial orthophotos into general land cover classes (e.g., road, building, waterbody, grassland, barren land, shadow, and dense vegetation). The main objective of the research is to run several experiments exploring the impacts of CNN architectures and

hyperparameters on the accuracy of land cover classification using aerial photos. The aim is to understand the behaviours of the CNN model concerning its architecture design and hyperparameters to produce models with high generalization capacity.

3. Methodology

This section presents the dataset, preprocessing, and the methodology of the proposed CNN model including the network architecture and training procedure.

3.1. Dataset and Preprocessing

3.1.1. Dataset. To implement the current research, a pilot area was identified based on the diversity of the land cover of the area. The study area is located in Selangor, Malaysia (Figure 1).

3.1.2. Preprocessing

(1) *Geometric Calibration.* Since the orthophoto was captured by an airborne laser scanning (LiDAR) system, it was essential to calibrate it geometrically to correct the geometric errors. In this step, the data was corrected based on ground control points (GCPs) collected from the field (Figure 2). There were 34 GCPs identified from clearly identifiable points (i.e., road intersections, corners, and power lines). The geometric correction was done in ArcGIS 10.5 software. The steps of geometric correction included identification of transformation points in the orthophoto, application of the least square transformation, and calculation of the accuracy of the process. The selected points were uniformly distributed in the area. After that, the least square method (Kardoulas et al., 1996) was applied to estimate the coefficients, which are essential for the geometric transformation process. After the least square solution, the polynomial equations were used to solve for X , Y coordinates of GCPs and to determine the residuals and RMS errors between the source X , Y coordinates and the retransformed X , Y coordinates.

(2) *Normalization.* Since the aerial orthophotos have integer digital values and initial weights of the CNN model are randomly selected within 0-1, a z -score normalization was applied to pixel values of the orthophotos to avoid abnormal gradients. This step is essential as it improves the progress of the activation and the gradient descent optimization (LeCun et al., 2012).

$$X' = \frac{(X/\max) - \mu}{\sigma}, \quad (1)$$

where \max is the maximum pixel value in the image, μ and σ are the mean and standard deviation of X/\max , respectively, and X' is normalized data.

3.2. The Proposed Approach

3.2.1. Overview. An orthophoto is composed of $m \times n \times d$ digital values, where m , n , and d are the image width, length, and depth, respectively. The goal of a classification model is

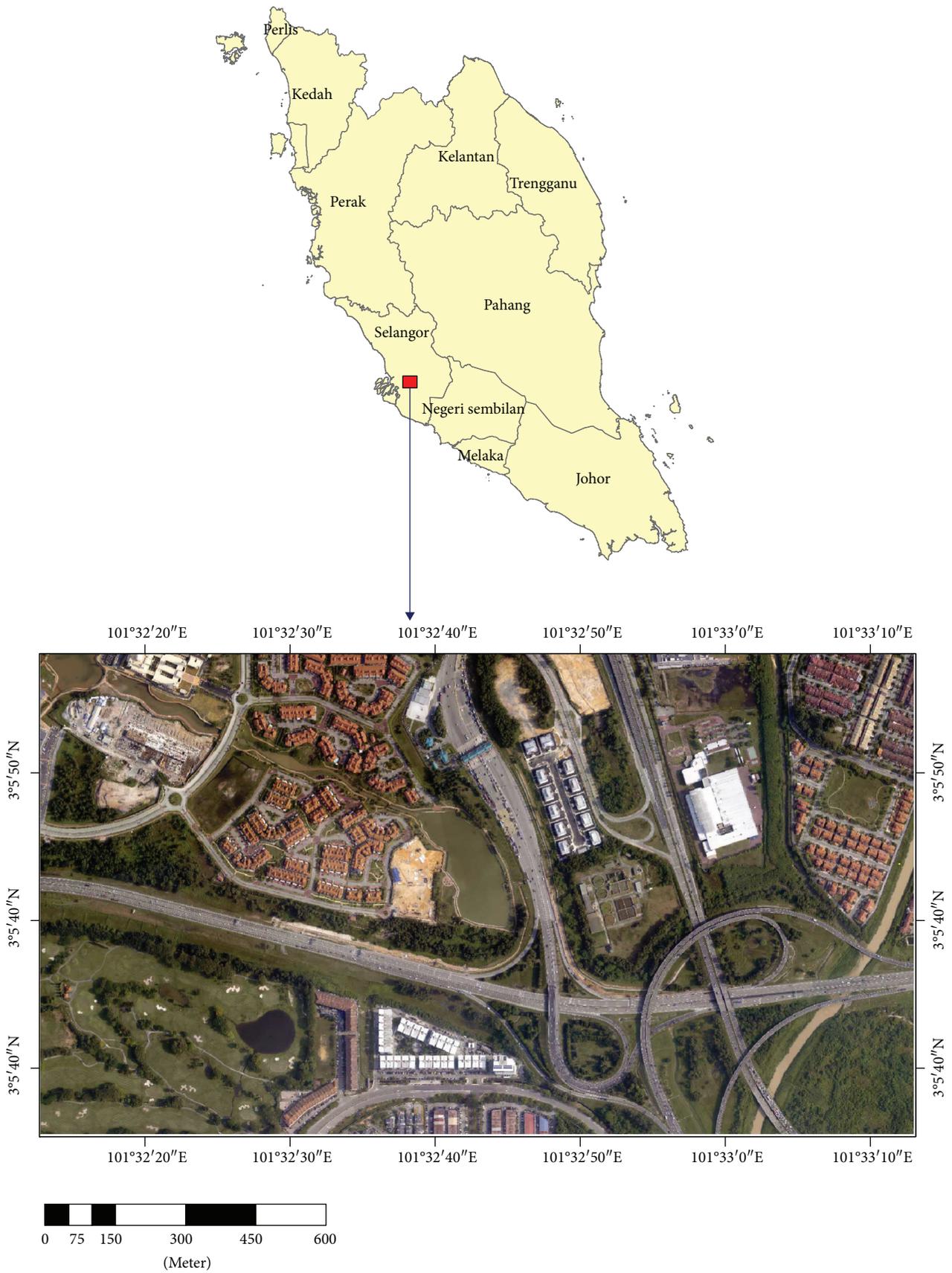


FIGURE 1: The study area location map.

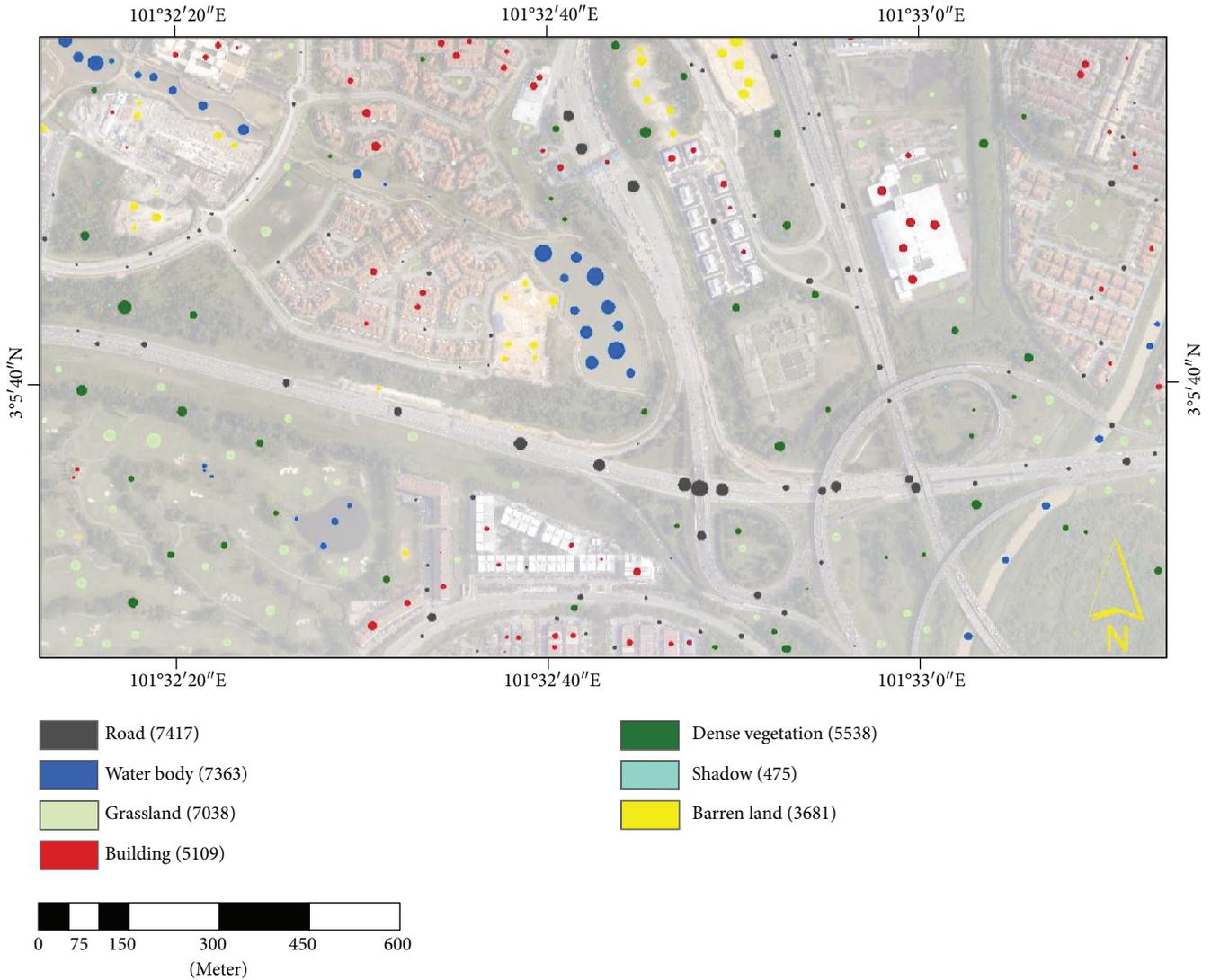


FIGURE 2: The ground truth samples over the study area, which were manually selected for seven land cover classes, for example, road, water body, grassland, building, dense vegetation, shadow, and barren land. The number in the brackets indicates the number of pixels in each class.

to assign a label to each pixel in the image given a set of training examples with their ground truth labels. In general, the common classification methods utilize the spectral information (image pixels across different bands) to achieve that goal. In addition, some of other techniques such as object-based image analysis (OBIA) segment the input image into several homogeneous contiguous groups before classification. This method uses additional features like spatial, shape, and texture to boost the classification performance of the classifier. However, both the methods, pixel-based and OBIA have several challenges like speckle noise in the first method and segmentation optimization in OBIA. Furthermore, both methods require careful feature engineering and band selection to obtain high accuracy of classification. More recently, classification methods using image patches and deep learning algorithms have been proposed to overcome the above challenges. Among the common methods is CNN. As a result, this study has proposed a classification method that is based on CNN and spectral-spatial feature learning for classifying

very high-resolution aerial orthophotos. The following sections describe the proposed model and its components including the basics of CNN, the network architecture, and the training methodology.

The pseudocode of the proposed classification model is presented in Algorithm 1. We developed the CNN model in the current study by running several experiments with different configurations. Then, we designed the ultimate model with best hyperparameters and architecture based on some statistical accuracy metrics such as overall accuracy, Kappa index, and per-class accuracies.

3.2.2. Basics of CNN. Convolutional neural networks (CNNs) or ConvNets are a type of artificial neural networks that simulate the human vision cortical system by using local receptive field and shared weights. It was introduced by LeCun and his colleagues [18]. Figure 3 shows a typical CNN with convolutional max pooling operations. CNN is suitable for analyzing images, videos, or data in the form of

```

Algorithm 1: CNN for orthophoto classification
Input: RGB image ( $I$ ) captured by the aerial remote sensing system, training/testing samples ( $D$ )
Output: Land cover classification map with seven classes ( $O$ )
 $I, D, O$ 
Preprocessing (Section 3.1.2):
calibrate  $I$  using the available 34 GCPs
normalize pixel values using Eq. 1
Classification (CNN) (Section 3.2.2 and Section 3.2.3):
for Patch_x_axis:
  initialize sum = 0
  for Patch_y_axis:
    calculate dot product(Patch, Filter)
    result_convolution ( $x, y$ ) = Dot product
for Patch_x_axis:
  for Patch_y_axis:
    calculate Max (Patch)
  result_maxpool ( $x, y$ ) = Dot product
update  $F = \max(0, x)$ 
  result_cnn_model = trained model
Prediction:
  apply the trained model to the whole image and get  $O$ 
Mapping:
  get the results of prediction
  reshape the predicted values to the original image shape
  convert the array to image and write it on the hard disk

```

ALGORITHM 1: The pseudocode of the proposed CNN developed for land cover mapping using aerial images.

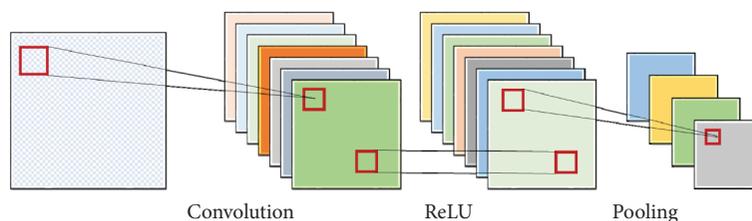


FIGURE 3: Illustration of typical layers of a CNN.

n -dimensional arrays that have a spatial component. This unique property makes them suitable for remote sensing image classification as well. A typical architecture of CNN consists of a series of layers such as convolution, pooling, fully connected (i.e., dense), and logistic regression/Softmax. However, additional layers like dropout and batch normalization also can be added to avoid overfitting and improve the generalization of these models. The last layer depends on the type of the problem, where for binary classification problems, a logistic regression (sigmoid) layer is often used. Instead, for multiclass classification problems, a Softmax layer is used. Each layer has its operation and is aimed in these models. For example, the convolutional layers are aimed at constructing feature maps via convolutional filters that can learn high-level features that allow taking advantage of the image properties. The output of these layers then passes through a nonlinearity such as a ReLU (rectified linear unit). Local groups of values in array data are often highly correlated, and local statistics of images are invariant to

location [19]. In addition, pooling layers (or subsampling) are used to merge semantically similar features into one. The most common method of subsampling computes the maximum of a local patch of units in feature maps. Other pooling operations are averaging max pooling and stochastic pooling. In general, several convolutional and subsampling layers are stacked, followed by dense layers and a Softmax or a logistic regression layer to predict the label of each pixel in the image.

3.2.3. Network Architecture. The architecture of the CNN model was built with a single convolutional layer followed by a max pooling operation, batch normalization, and two dense layer classifiers (Figure 4). This architecture yielded 3527 total parameters where 96 parameters are not trainable. The convolutional kernels were kept as 3×3 , and the pooling size in the max pooling layer was kept at 2×2 . Dropout was performed in the convolutional layer and the first dense layer with a drop probability of 0.5 to avoid overfitting. The

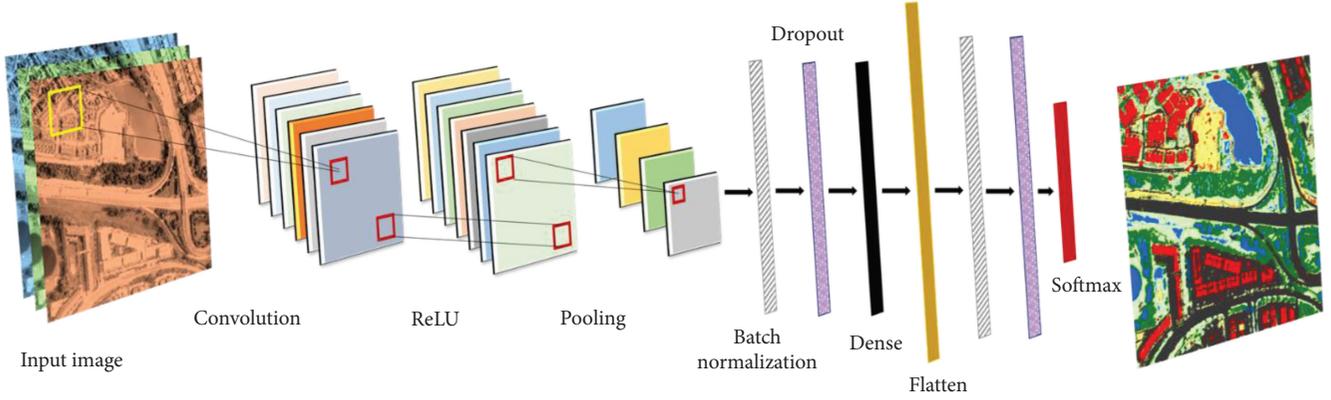


FIGURE 4: The architecture of the proposed CNN for aerial orthophoto classification.

TABLE 1: The summary of the CNN model layers.

Layer (type)	Output shape	Number of parameters
Input	(None, 3, 7, 7)	0
2D convolution	(None, 1, 5, 32)	2048
Max pooling	(None, 1, 2, 16)	0
Batch normalization	(None, 1, 2, 16)	64
Dropout	(None, 1, 2, 16)	0
Flatten	(None, 32)	0
Dense	(None, 32)	1056
Batch normalization	(None, 32)	128
Dropout	(None, 32)	0
Dense (Softmax)	(None, 7)	231

minibatch of stochastic gradient descent (SGD) was set to 32 images. Under the framework of Keras with Tensorflow backend, the whole process was run on a CPU Core i7 2.6 GHz and memory ram (RAM) of 16 GB. In the experiments, 60% of the total samples were randomly chosen for training, and the rest were chosen for testing, and overall accuracy (OA), average accuracy (AA), Kappa coefficient (κ), and per-class accuracy (PA) are used to evaluate the performance of the CNN classification method (Congalton and Green, 2008). The summary of the model's layers is shown in Table 1.

3.2.4. Training the Model. The CNN model was trained with backpropagation algorithm and stochastic gradient descent (SGD). It uses the minibatch's backpropagation error to approximate the error of all the training samples, which accelerates the cycle of the weight update with smaller back propagation error to speed up the convergence of the whole model. The optimization was run to reduce the loss function (J) (i.e., categorical cross entropy) of CNN expressed as the following:

$$J(X', W, b, \theta) = -\frac{1}{N} \left[\sum_{i=1}^N \sum_{j=1}^k 1\{y^i = t\} \cdot y_t^i \right], \quad (2)$$

where X' is normalized features, W and b are parameters of

CNN, θ is the parameters of Softmax layer, N is the number of samples, k is the number of land cover classes, $y^i = (y_1^i, y_2^i, \dots, y_k^i)$ is the prediction vector geo by the Softmax classifier (3), and y_t^i represents the possibility of the i th sample label being t and is computed by (3).

$$y_t^i = \frac{\exp(\theta_t^T c)}{\sum_{j=1}^k \exp(\theta_j^T c)}. \quad (3)$$

During back propagation, (4) are adapted to update W and b in every layer, where λ is the momentum which help accelerate SGD by adding a fraction of the update value of the past time step to the current update value, α is the learning rate, ∇W and ∇b are the gradients of $J(\cdot)$ with respect to W and b , respectively, and t just stands for the number of epoch during SGD:

$$\begin{aligned} W_{t+1} &= W_t - \lambda V_t - \alpha \nabla W, \\ b_{t+1} &= b_t - \lambda U_t - \alpha \nabla b. \end{aligned} \quad (4)$$

3.2.5. Evaluation. This study uses several statistical accuracy measures to evaluate different models and compare them under various experimental configurations. These metrics are overall accuracy (OA), average accuracy (AA), per-class accuracy (PA), and Kappa index (κ). They are calculated using the following equations [20]:

$$\begin{aligned} OA &= \frac{\sum D_{ii}}{N}, \\ AA &= \frac{\sum_1^m PA_m}{m}, \\ PA &= \frac{D_{ij}}{R_i}, \\ \kappa &= \frac{N \sum_{i,j=1}^m D_{ij} - \sum_{i,j=1}^m R_i \cdot C_j}{N^2 - \sum_{i,j=1}^m R_i \cdot C_j}, \end{aligned} \quad (5)$$

where $\sum D_{ii}$ is the total number of correctly classified pixels, N is total number of pixels in the error matrix, m is the

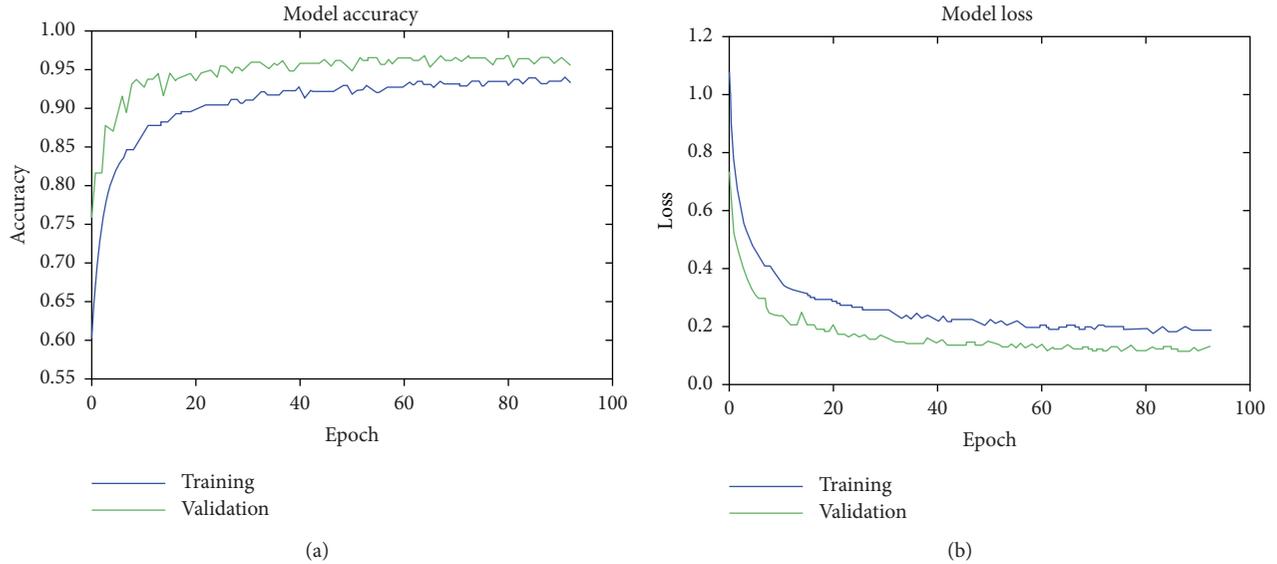


FIGURE 5: Performance of the CNN model with optimum parameters set, (a) model accuracy and (b) model loss for 93 epochs (early stopping).

TABLE 2: PA of the CNN model.

Class	PA
Road	0.971
Waterbody	0.944
Grassland	0.972
Building	0.995
Dense vegetation	0.999
Shadow	0.894
Barren land	0.980

number of classes, D_{ij} is the number of correctly classified pixels in row i (in the diagonal cell), R_i is the total number of pixels in row i , and C_j is the total number of pixels in column j .

4. Experimental Results

4.1. Performance of the Proposed Model

4.1.1. CNN with Dropout and Batch Normalization. Figure 5 shows the accuracy performance of the CNN model with dropout and batch normalization for 93 epochs on both training and validation datasets. The increment in model accuracy and reduction in model loss over time indicates that the model has learned useful features to classify the image pixels into the different class labels. The fluctuations in the accuracy from one epoch to another are because of using dropout that yielded a slightly different model at each epoch. The OA, AA, κ of this model on validation dataset was 0.973, 0.965, 0.967, respectively. In addition, Table 2 shows the per-class accuracy (PA) achieved by the model. The results suggest that the CNN model could

classify almost all the classes with relatively high accuracy. The minimum accuracy was 0.894 for the shadow class. While examining the confusion matrix (Table 3), the results indicate that several (~11) samples of this class were misclassified as dense vegetation affecting its PA. The confusion matrix also shows that there were several samples of water body class misclassified as grassland.

4.1.2. CNN Model with Other Configurations. The CNN model was also trained without dropout and batch normalization to see their impacts on the accuracy of the classification map. Table 4 summarizes the results of comparing CNN models with different configurations (i.e., CNN + dropout + batch normalization, CNN + dropout, CNN + batch normalization, and CNN). The results suggest that the use of dropout and batch normalization could improve the accuracy (OA, AA, and κ) of the classification by almost 4%. The use of batch normalization slightly performed better (OA = 0.964, AA = 0.956, κ = 0.961) than just using dropout (OA = 0.958, AA = 0.956, κ = 0.954). Nevertheless, the use of either dropout or batch normalization could improve the accuracy of the classification compared to not using any of these techniques with the CNN model. The CNN model without these techniques achieved the following accuracies: OA = 0.932, AA = 0.922, κ = 0.922 indicating the importance of such regularization methods for aerial orthophoto classification. The classified maps produced by these methods are shown in Figure 6. Furthermore, the performance plot (Figure 7) of the CNN model without dropout and batch normalization shows that this model overfits the training data and performs worse when applied to new data. Overall, the experimental results on both training and validation data sets infer that the proposed CNN architecture is a robust and efficient model, while the use of dropout and batch normalization

TABLE 3: The confusion matrix calculated for the CNN model.

	Road	Waterbody	Grassland	Building	Dense vegetation	Shadow	Barren land
Road	1474	0	0	23	0	0	21
Water body	0	1463	85	0	0	1	0
Grassland	0	10	1323	0	27	0	0
Building	4	0	0	991	0	0	0
Dense vegetation	0	0	0	0	1070	1	0
Shadow	0	0	0	0	11	93	0
Barren land	6	0	0	8	0	0	716

TABLE 4: Performance of CNN model with different configurations.

Model	OA	AA	κ
CNN + dropout + batch normalization	0.973	0.965	0.967
CNN + dropout	0.958	0.956	0.954
CNN + batch normalization	0.964	0.956	0.961
CNN	0.932	0.922	0.922

techniques as regularization methods is essential to obtain high accuracy of classification for the entire area rather than just predicting the labels of the training samples.

4.2. Sensitivity Analysis. The performance of CNN while classifying orthophotos is highly dependent on its architecture and hyperparameters. Thus, the sensitivity analysis could serve as an essential step in finding a good set of parameters and architecture configurations in addition to an understanding of the model behavior. Figure 8 shows the impact of different parameters (e.g., number of convolutional filters, activation function, drop probability, optimizer, batch size, and patch size) on the validation accuracy of CNN.

For convolutional filters, the sensitivity analysis shows that larger number of filters can lead to an increase in performance. However, use of larger number of filters can increase training time and overfit the training data if the model is not regularized properly. Thus, this parameter was set to 32 as an optimal setting and not exploring a larger number of filters. With this configuration, the model could achieve the following accuracies: OA=0.956, AA=0.945, and κ =0.947. In addition, this analysis shows that the activation function “ReLU” outperformed the other two functions (“Sigmoid” and “ELU”). By using this activation, the CNN model could achieve an OA of 0.956 higher than the second best activation “Sigmoid” by ~4.4%. ReLU also facilitates faster training and reduced likelihood of vanishing gradient. The experiments on drop probability showed that different parametric values can improve the performance of CNN depending on the accuracy metric. For example, results showed that the use of drop probability as 0.2 could optimize the model for OA and κ , where the model achieved an OA and κ of 0.975, 0.970, respectively. However, drop probability of 0.3 could perform better than the value of 0.2 for this parameter

regarding AA. Furthermore, performances of CNN with different optimizers have been investigated, and the results indicated that “Adam” could be effective in training compared to other optimizers. The highest OA (0.975) and κ (0.970) were achieved by the CNN model that was trained with “Adam.” However, when the optimizer “Nadam” was used to train CNN, the model could achieve the highest AA (0.974). The worst performance of CNN (OA=0.945, AA=0.949, and κ =0.934) was found to be when the model was trained with SGD. Moreover, the efficiency of CNN was compared with different batch sizes such as 4, 8, 16, 32, and 64. The batch size of 32 was found the best considering OA (0.975) and κ (0.970), while the batch size of 64 achieved the highest AA (0.975).

Another important parameter in the proposed CNN is the patch size, which is the local neighborhood area that forms with the size ($n \times n$). The advantage of using patch-based learning for orthophoto classification is sourced from the benefits of spectral and spatial information of the data that can improve the accuracy compared to just using the individual pixels (only spectral information). To understand this parameter and find its suboptimal value, several experiments were conducted with different patch sizes ($n = [3, 5, 7, 9, 11, 13]$). The statistical analysis in terms of model accuracy indicates that using larger n yields higher accuracy (Figure 8). However, when analyzing the classification map visually, the use of larger n reduces the spatial quality of the features in the classification map (Figure 9). As a result, we considered $n = 7$ as an effective value for this parameter as it achieved relatively high accuracy measured by OA, AA, and κ as well as high spatial quality features.

4.3. Training Time Analysis. The computing performance of the CNN model was dependent on the use of dropout and batch normalization layers in the network architecture in addition to other hyperparameters such as a number of convolutional filters and image patch size. Table 5 shows the training time of the CNN model with different configurations. When early stopping was applied, the training of CNN with dropout and batch normalization took about 124 seconds on a CPU. Removing the batch normalization from the architecture yielded a training time of 150 seconds, whereas CNN with dropout took 75 seconds to be trained. The CNN model without the use of dropout and batch normalization took the shortest time (58.4 seconds)

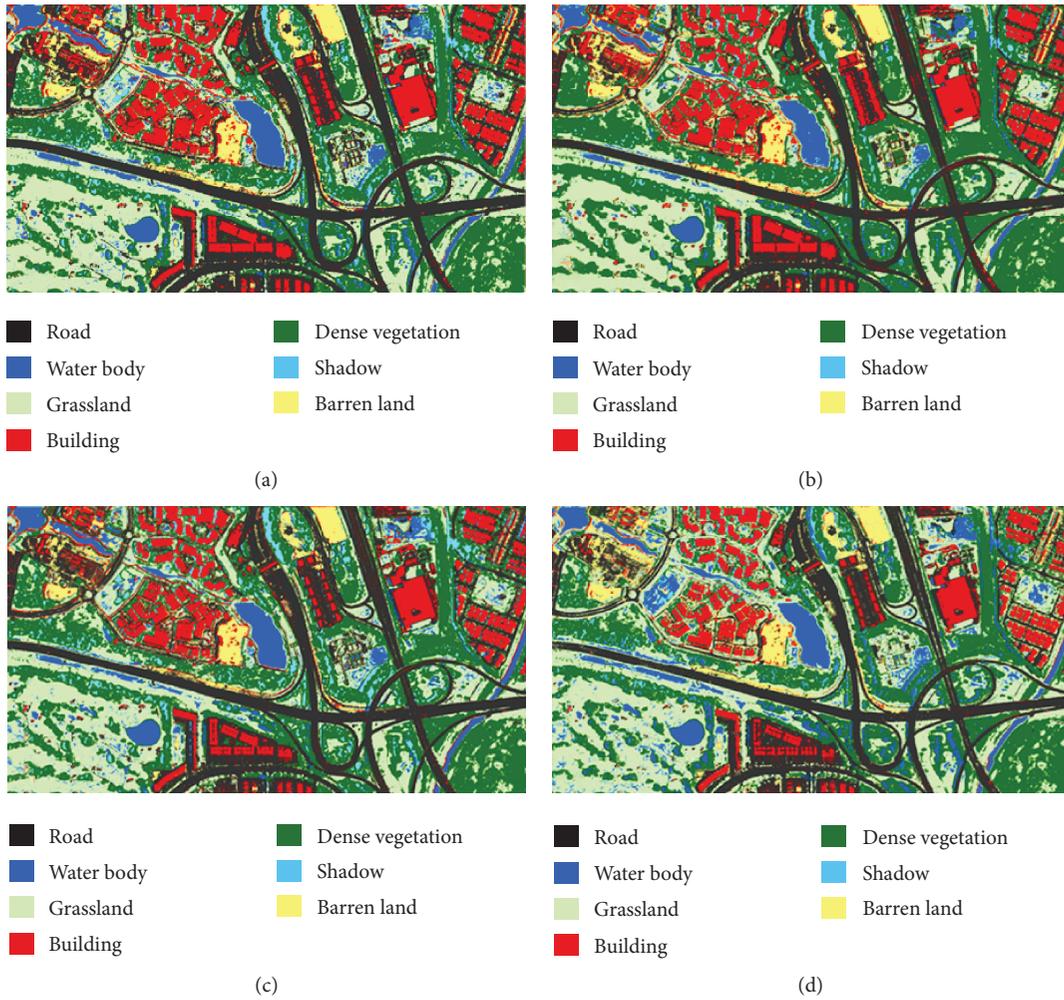


FIGURE 6: Classification maps produced by CNN models, (a) CNN + dropout + batch normalization, (b) CNN + dropout, (c) CNN + batch normalization, and (d) CNN.

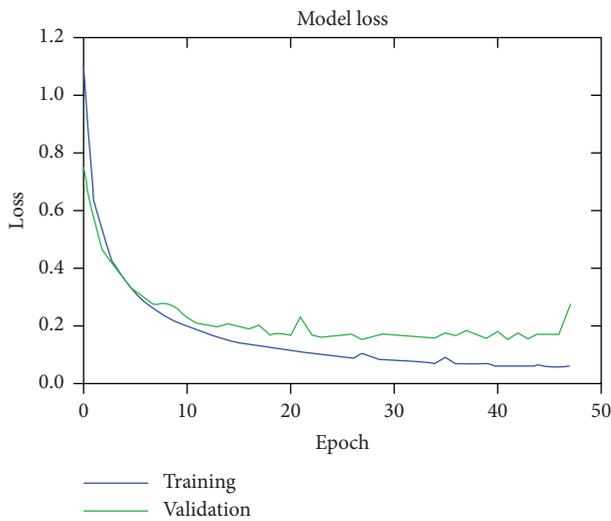


FIGURE 7: The loss of the CNN model without dropout and batch normalization.

to be trained. On the other hand, when the model was trained with 200 epochs without early stopping, the model (CNN + dropout + batch normalization) took about 230 seconds longer than that with early stopping by 106 seconds. In addition, the other models (CNN + dropout, CNN + batch normalization, and CNN) were also required a longer time to train as it was expected due to more number of epochs run. Overall, the computing performance of the proposed model is efficient for the investigated data. However, for larger datasets, the training of such models will require longer time, and as a result, graphical processing units will be essential.

5. Conclusion

In this paper, a classification model based on CNN and spectral-spatial feature learning has been proposed for aerial photographs. With the utilization of advanced regularization techniques such as dropout and batch normalization, the proposed model could balance generalization ability and training efficiency. Use of such methods to improve the CNN model along with other techniques like preprocessing

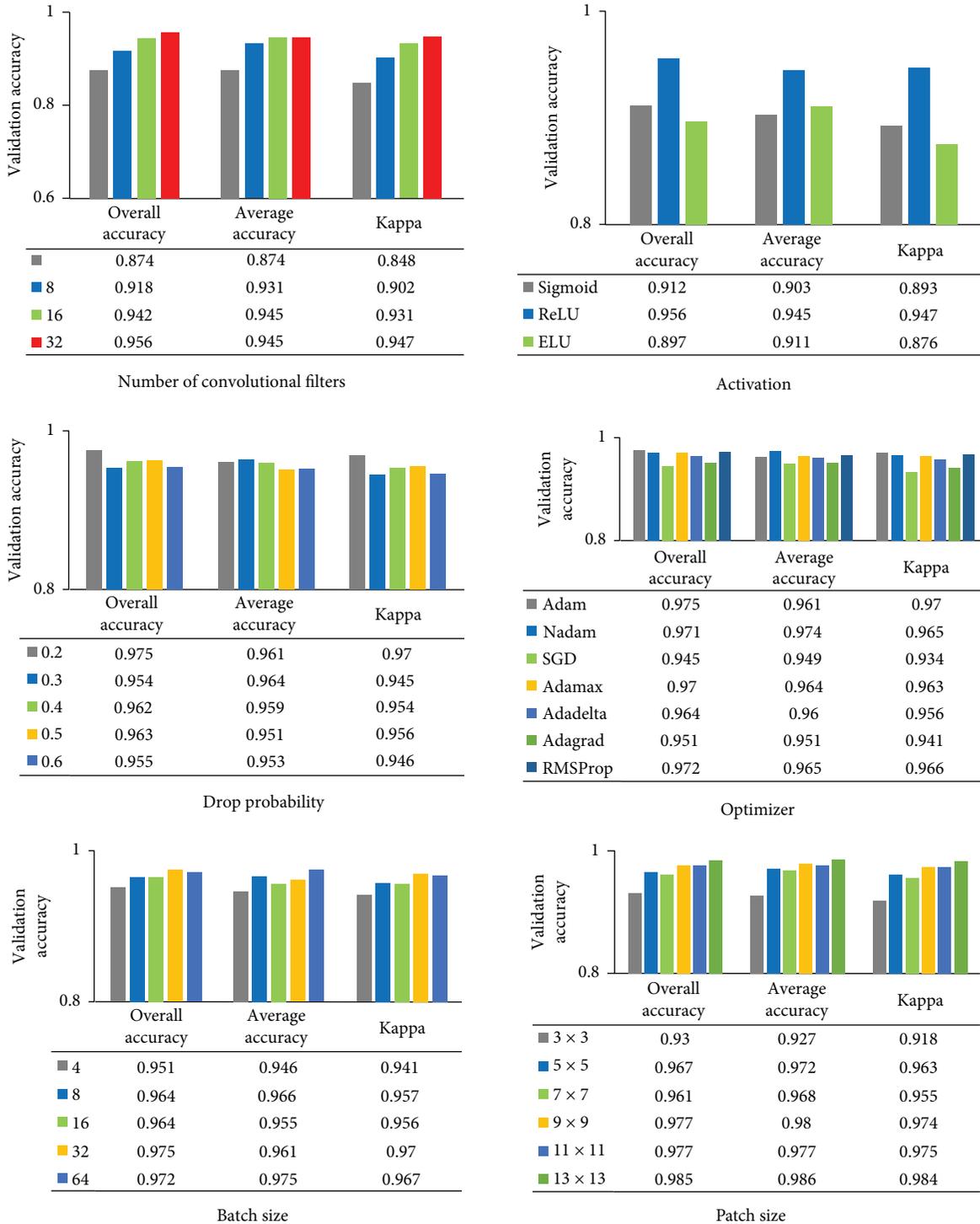


FIGURE 8: The influence of hyperparameters, the number of convolutional filters, activation function, drop probability, optimizer, batch size, and patch size.

(geometric calibration and feature normalization) and sensitivity analysis could make these models robust for classifying the given dataset. The CNN model acts as a feature extractor, and a classifier could be trained end-to-end given training samples. The network architecture can effectively handle the inter- and intraclass complexity inside the scene. The best model achieved OA = 0.973, AA = 0.965, and $\kappa = 0.967$

outperforming the traditional CNN model by ~4% in all the accuracy indicators. The short training time (124 seconds) confirmed the robustness of the proposed model for small and medium scale remote sensing datasets. The future work should focus on scaling this architecture for large remote sensing datasets and other data sources such as satellite images and laser scanning point clouds.

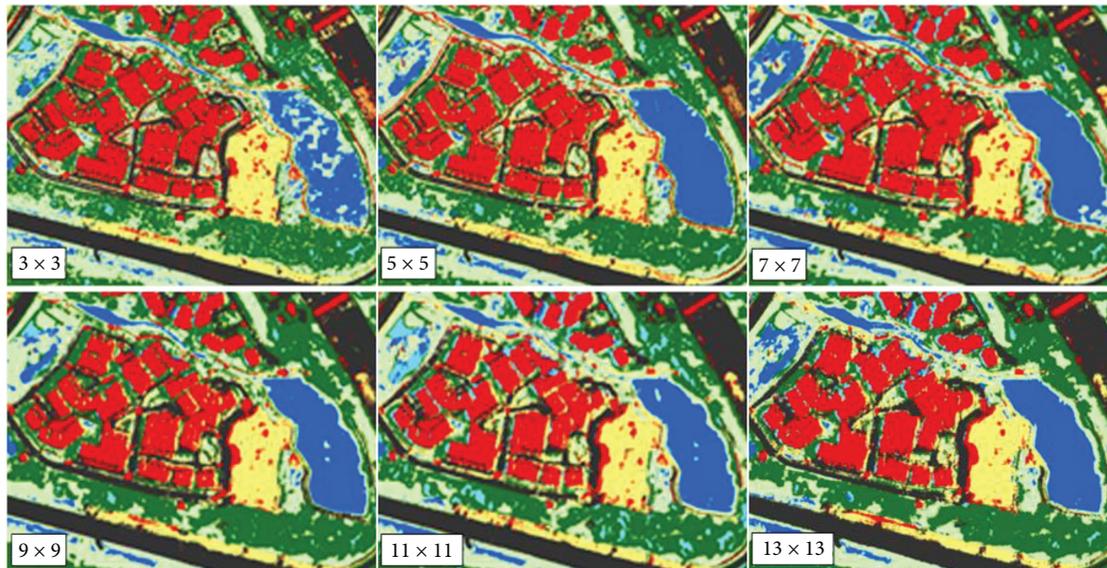


FIGURE 9: Effects of patch size on the quality of classified maps.

TABLE 5: The training time in seconds of CNN with different configurations for 200 epochs.

Model	Time (seconds)—with early stopping	Time (seconds)—full training
CNN + dropout + batch normalization	124	230
CNN + dropout	150	168
CNN + batch normalization	75	219
CNN	58.4	158

Data Availability

These data were used from a research project lead by Professor Biswajeet Pradhan. Very high resolution aerial photos were used in this research. The data can be made available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Ö. Akar, "The Rotation Forest algorithm and object-based classification method for land use mapping through UAV images," *Geocarto International*, vol. 33, no. 5, pp. 538–553, 2017.
- [2] Q. Wu, R. Zhong, W. Zhao, H. Fu, and K. Song, "A comparison of pixel-based decision tree and object-based Support Vector Machine methods for land-cover classification based on aerial images and airborne lidar data," *International Journal of Remote Sensing*, vol. 38, no. 23, pp. 7176–7195, 2017.
- [3] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, "DeepSat: a learning framework for satellite imagery," in *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, p. 37, New York, NY, USA, November 2015.
- [4] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2094–2107, 2014.
- [5] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 844–853, 2017.
- [6] W. Zhao and S. Du, "Spectral–spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4544–4554, 2016.
- [7] Y. T. Hsieh, C. T. Chen, and J. C. Chen, "Applying object-based image analysis and knowledge-based classification to ADS-40 digital aerial photographs to facilitate complex forest land cover classification," *Journal of Applied Remote Sensing*, vol. 11, no. 1, article 015001, 2017.
- [8] M. F. A. Vogels, S. M. De Jong, G. Sterk, and E. A. Addink, "Agricultural cropland mapping using black-and-white aerial photography, object-based image analysis, and random forests," *International Journal of Applied Earth Observation and Geoinformation*, vol. 54, pp. 114–123, 2017.
- [9] X. Meng, N. Shang, X. Zhang et al., "Photogrammetric UAV mapping of terrain under dense coastal vegetation: an object-oriented classification ensemble algorithm for classification and terrain correction," *Remote Sensing*, vol. 9, no. 11, p. 1187, 2017.
- [10] A. Juel, G. B. Groom, J. C. Svenning, and R. Ejrnaes, "Spatial application of random forest models for fine-scale coastal vegetation classification using object based analysis of aerial orthophoto and DEM data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 42, pp. 106–114, 2015.

- [11] L. Albert, F. Rottensteiner, and C. Heipke, "A higher order conditional random field model for simultaneous classification of land cover and land use," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 63–80, 2017.
- [12] G. Cheng, C. Ma, P. Zhou, X. Yao, and J. Han, "Scene classification of high resolution remote sensing images using convolutional neural networks," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 767–770, Beijing, China, July 2016.
- [13] G. J. Scott, M. R. England, W. A. Starms, R. A. Marcum, and C. H. Davis, "Training deep convolutional neural networks for land-cover classification of high-resolution imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 4, pp. 549–553, 2017.
- [14] J. Sherrah, "Fully convolutional networks for dense semantic labelling of high-resolution aerial imagery," 2016, <http://arxiv.org/abs/1606.02585>.
- [15] W. Yao, P. Poleswki, and P. Krzystek, "Classification of urban aerial data based on pixel labelling with deep convolutional neural networks and logistic regression," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B7, pp. 405–410, 2016.
- [16] X. Sun, X. Lin, S. Shen, and Z. Hu, "High-resolution remote sensing data classification over urban areas using random forest ensemble and fully connected conditional random field," *ISPRS International Journal of Geo-Information*, vol. 6, no. 8, p. 245, 2017.
- [17] J. R. Bergado, C. Persello, and C. Gevaert, "A deep learning approach to the classification of sub-decimetre resolution aerial images," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1516–1519, Beijing, China, July 2016.
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [20] A. Bogoliubova and P. Tymków, "Accuracy assessment of automatic image processing for land cover classification of St. Petersburg protected area," *Acta Scientiarum Polonorum. Geodesia et Descriptio Terrarum*, vol. 13, no. 1-2, 2014.

Research Article

Ocean Wave Information Retrieval Using Simulated Compact Polarized SAR from Radarsat-2

Xiaochen Wang^{1,2,3}, Yun Shao^{1,2,3,4}, Lu She^{2,3}, Wei Tian^{1,2,3}, Kun Li^{1,2,3}
and Long Liu^{1,2,3}

¹Laboratory of Target Microwave Properties, Deqing Academy of Satellite Applications, Huzhou 313200, China

²Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, 1 Beichen Road, Beijing 100101, China

³University of Chinese Academy of Sciences, 19A Yuquan Road, Beijing 100049, China

⁴State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, 1 Beichen Road, Beijing 100101, China

Correspondence should be addressed to Lu She; shelu_whu@163.com

Received 17 October 2017; Accepted 28 March 2018; Published 7 June 2018

Academic Editor: Biswajeet Pradhan

Copyright © 2018 Xiaochen Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The main objective of this paper is to demonstrate the capability of compact polarized (CP) synthetic aperture radar (SAR) to retrieve ocean wave field parameters. Souyris' and Nord's algorithms are used to carry out the reconstruction of CP SAR pseudo quad-polarized data for the ocean surface under both the circular transmit linear receive (CTLR) and $\pi/4$ mode. The results show that, for the CP reconstruction, Nord's algorithm has a better convergence ability than Souyris'. In addition, the investigation of the reconstruction accuracy shows that the CTLR mode is superior to the $\pi/4$ mode, in terms of ocean surface reconstruction. It is, therefore, concluded that the reconstructed parameters of CP CTLR mode data by Nord's algorithm adapt to retrieve ocean wave information. The ocean wave slope spectrum and other main wave parameters are also calculated from reconstructed CP data and compared with measurements from in situ National Data Buoy Center (NDBC) matched-up buoys. Comparison of CP SAR-based wave field information with buoy outputs also shows good agreement in the case of dominate wave height, wave direction, and wave period, with biases of 0.36 m, 17.96°, and 0.88 s, respectively.

1. Introduction

Ocean waves are the most important ocean surface features in the fields of ocean physics and coastal engineering. Ocean surface parameters, such as wave height, wave direction, and wave length, are important for sea-ice monitoring, ship detection, marine pollution forecasting, and in other areas [1, 2]. The wave slope spectrum, which provides information about the 2-D directional distribution of the wave energy density, is important to ocean wave information retrieval. In recent years, remote sensing as an efficient tool for earth surface monitoring gradually plays a key role in ocean information retrieval. Of all remote sensing sensors, SAR systems are the active satellite sensors with a resolution high enough to provide information about the ocean wave spectrum and

so these systems play a key role in the measurement of ocean surface waves [3–5]. In contrast to in situ buoys, SAR can operate in all weather conditions or in areas far from the shore and can measure ocean waves across the globe continuously [6]. In addition to improving the understanding of the properties of the ocean, SAR observations can be used to develop the forecast accuracy of the operational wave model. It is, therefore, important to investigate the capability of SAR to carry out ocean surface monitoring, especially in terms of ocean wave retrieval.

In recent decades, satellites such as ERS-1, ERS-2, Envisat, Radarsat-1/2, Sentinel-1 A/B, and Chinese GF-3 have carried on many SAR systems for use in ocean surface monitoring. Several wave retrieval algorithms have also been developed for ocean information retrieval. For example, K.

Hasselmann and S. Hasselmann [7, 8] proposed the Max Planck Institute (MPI) algorithm based on the closed nonlinear transformation relation between the ocean spectrum and the SAR image spectrum; the output from the wave model (WAM) is input as a first-guess spectrum in this method. Mastenbroek and de Valk [9] developed a semiparametric retrieval algorithm (SPRA) based on Hasselmann's nonlinear theory [7] by separating the wind-driven wave and swell—the wind-wave information is provided by synchronous scatterometer measurements and the swell information is derived from the rest of the SAR signal. Wang [10] proposed a semiempirical algorithm which has the objective to estimate the wave height from Envisat ASAR wave mode imagery without any prior knowledge. Shao et al. [11] develop a semiempirical function for significant wave height (H_s) and mean wave period (T_{mw}) retrieval from C-band VV polarization Sentinel-1 SAR. With the development of polarization technology, quad-polarized SAR has also been applied in ocean surface monitoring. Zhang et al. [12] present a synergistic method to retrieve both ocean surface wave and wind fields from spaceborne quad-polarization SAR, with consideration to the nonlinear mapping relationship between ocean wave spectra and SAR image spectra, in order to synergistically retrieve wind fields and wave directional spectra. Schuler and Lee [13, 14] proposed an algorithm for retrieving ocean wave spectra using quad-polarized SAR imagery and also investigated the results using L- and P-band SAR data. However, this method is only suitable for the L and P bands. By modifying the polarization orientation angle, He et al. [15, 16] developed a new retrieval algorithm so that it could be applied to C-band quad-polarized SAR data. He et al.'s method was also validated using Radarsat-2 quad-polarized SAR data by Zhang et al. [17, 18].

In addition, although the quad-polarized SAR can provide abundant polarization information to develop the process of ocean information retrieval, the limited swath width has also become the disadvantage of quad-polarized SAR for large-scale monitoring. For overcoming the limitation of CP SAR, a new type of SAR sensor called compact p(CP) SAR has been invented, which provides a new way to measure ocean waves and its potential for use in ocean surface monitoring has gradually become clear [19]. Compared with quad-polarized SAR, CP SAR operates at lower power but still has the capability to reconstruct quad-polarized data, thus overcoming the disadvantages of the narrow swath and high power of quad-polarized SAR and the insufficient polarized information provided by single- or dual-polarized SAR [20, 21]. CP SAR can, in fact, be interpreted as a compromise between dual-polarized and quad-polarized SAR so that swath width is increased and the operating power is reduced at the same time. To date, India's RISAT-1, launched in April 2012, and the Japanese ALOS-2 SAR satellite, launched in May 2014, have carried CP SAR. In the future, the planned Radarsat-2 constellation of satellites will also support CP SAR. Although CP SAR is carried on existing SAR platforms, there is still no well-calibrated data that can be used in research, especially in the case of ocean wave measurements. It is, therefore, important to study the potential of CP SAR for the retrieval of ocean wave information. In a

previous study, Nunziata et al. [22] investigated sea-surface scattering with a polarized model and exploited the hybrid polarized (HP) mode to interpret ocean surface polarization parameters. Yin et al. [23] proposed a new extended X-Bragg model and also used C-band quad-polarized SAR data acquired by both SIR-C/X-SAR and Radarsat-2 to emulate the HP and $\pi/4$ CP modes. Geldsetzer et al. [24] proposed that the dependence between incidence angle, wind speed, and wind direction of CP SAR parameters are presented for open water. Selected 20 CP parameters are related to C-band geophysical model function (CMOD) output. CP parameters are simulated from polarized Radarsat-2 data and emulate data available on the pending Radarsat Constellation Mission (RCM). Li et al. [25] present a new empirical reconstruction algorithm for the modified N based on a data set of more than 2000 Radarsat-2 quad-polarization images and collocated buoy observations. It is proven that cross-polarization facilitates wind speed retrieval and improves the accuracy of the results, especially with respect to high wind speeds. In summary, the capability of CP SAR to reconstruct pseudo quad-pol data and ocean wind retrieval has been investigated in detailed; however, its potential for ocean surface reconstruction has not been discussed yet, let alone wave information retrieval.

In this paper, two quad-polarized SAR images are used to reconstruct the CP pseudo quad-polarized data and demonstrate the potential of CP SAR for ocean wave information retrieval. The data set used in this paper is illustrated in Section 2. In Section 3, we briefly introduced the CP theory and the ocean wave retrieval method. In Section 4, we investigate two CP algorithms used for ocean surface reconstruction and evaluate the optimum CP mode for ocean surface reconstruction. The FP and CP ocean wave retrieval results are compared with NDBC buoy data. Conclusions are given in Section 5.

2. Data Sets

2.1. SAR Data. Two scenes of Radarsat-2 quad-polarized image are used in this study. Both of two image acquisition parameters are listed in Table 1, and the study areas are shown in Figures 1(a) and 1(b). HH-, HV-, and VV-polarized single look complex image (SLC) SAR images are also shown in Figures 1(a) and 1(b).

2.2. Buoy Data. In order to validate the CP SAR wave retrieval results, NDBC buoy data available from the National Oceanic and Atmospheric Administration (NOAA) were collected along with the SAR imagery. The main wave parameters were extracted from the original measurements and are listed in Table 2. Among all the wave measurement data, the dominant wave height (DWH), the dominant wave period (DWP), and dominant wave direction (DWD) were used for validating the retrieval results.

Because of limited collection of NDBC buoy, only one buoy was collocated with SAR imagery of scene 2. So the imagery of scene 1 was performed to CP reconstruction and ocean slope spectrum retrieval methods, and scene 2 was added to validate the method.

TABLE 1: Radar parameters of C-band Radarsat-2 quad-polarized data used for this study.

Scene ID	Image file name	Observation time	Incidence angle	Scene position	
				Latitude	Longitude
1	RS2-SLC-FQ3-ASC-31-Mar-2008_18.10-PDS_00058910	31 Mar. 2008 18:10	20.8°–22.8°	35.96°N	5.65°W
2	RS2-SLC-FQ9-ASC-09-Apr-2008_02.01-PDS_00058900	09 Apr. 2008 02:01	28.0°–29.8°	37.82°N	122.46°W

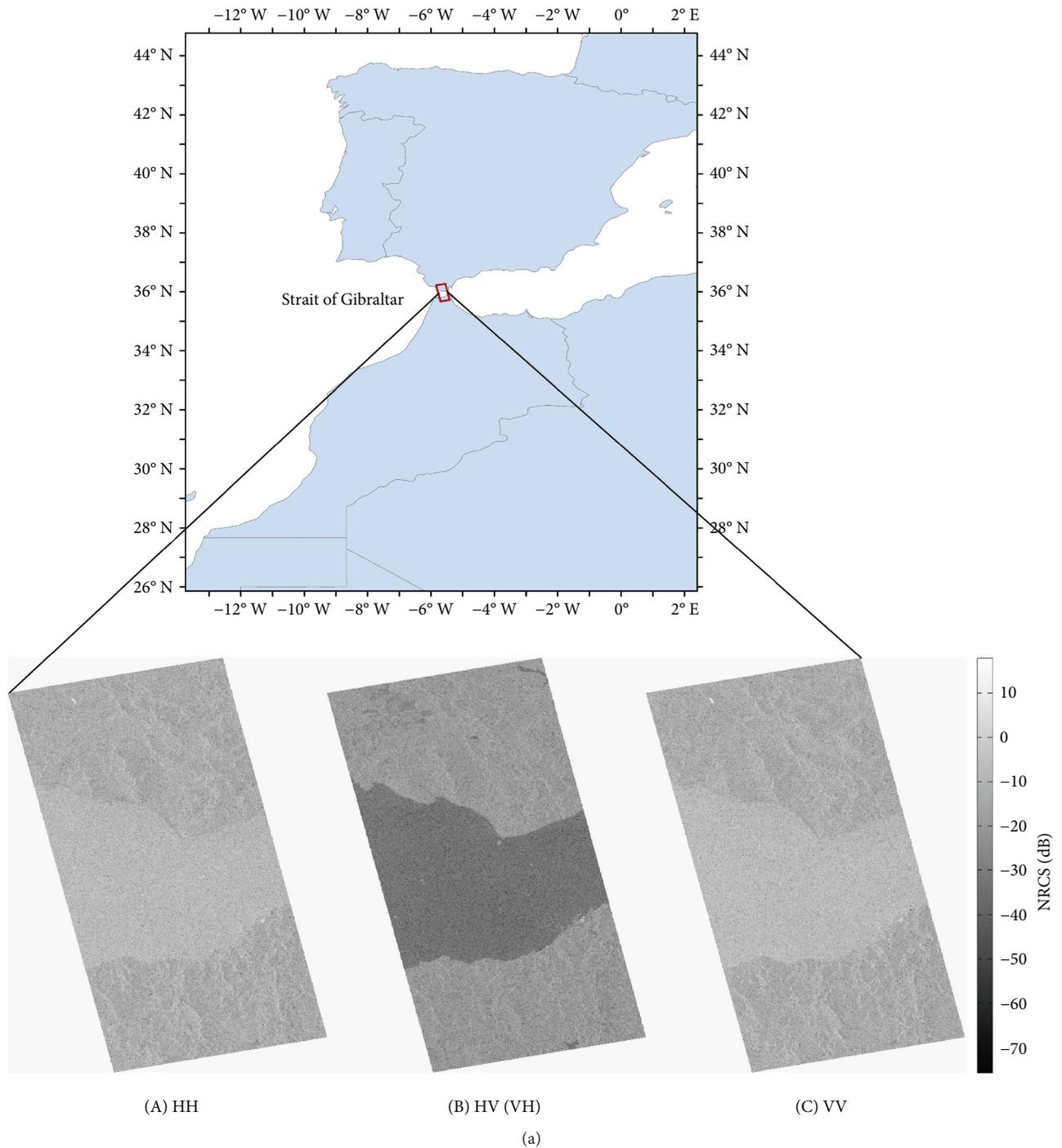


FIGURE 1: Continued.

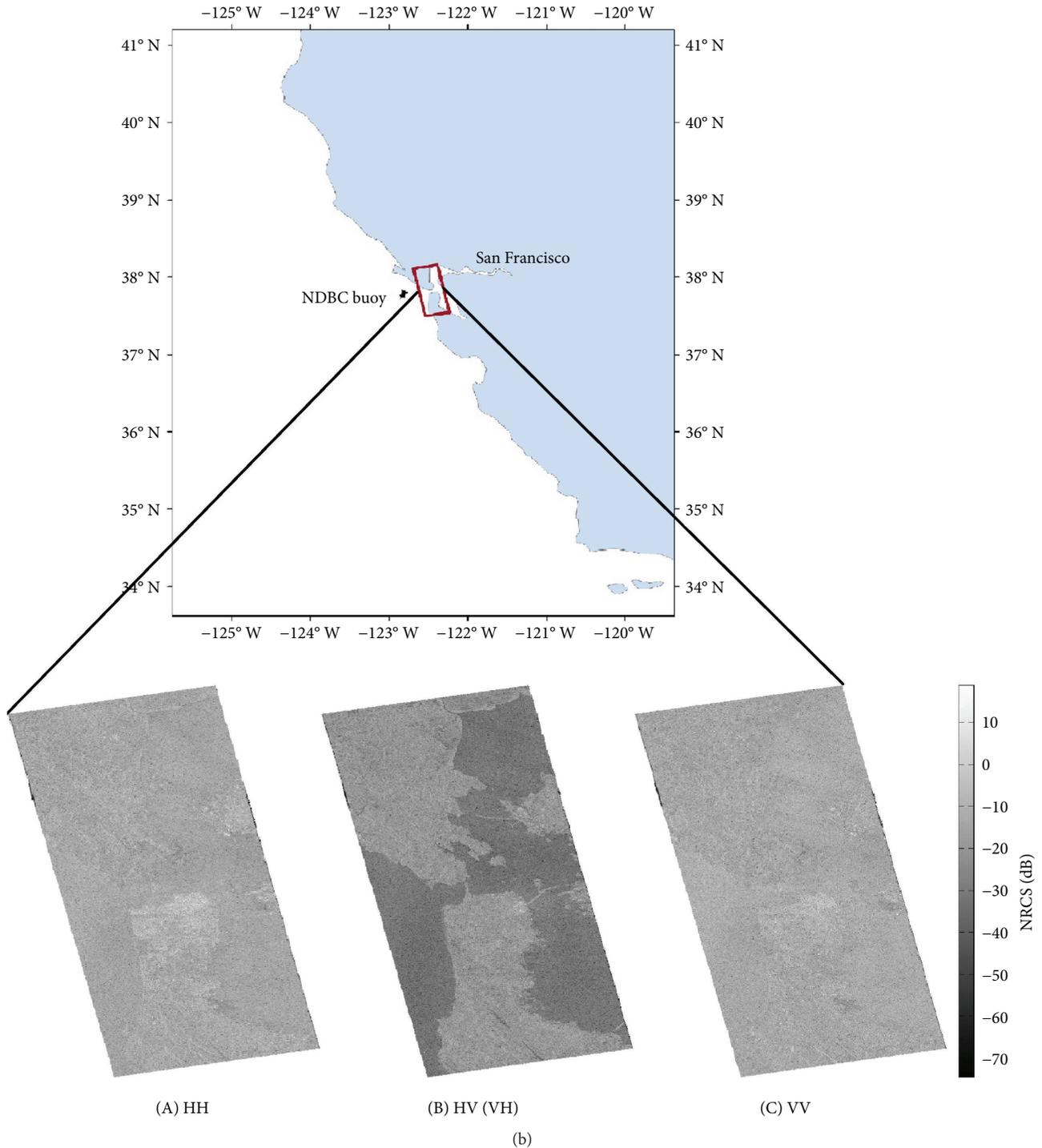


FIGURE 1: (a) The study area of scene 1: the coast near Strait of Gibraltar. The Radarsat-2 quad-polarized SLC SAR image was acquired on 2008/4/1. (A–C) correspond to the HH, HV (VH), and VV polarization channels, respectively. (b) The study area of scene 2: the coast near San Francisco. The Radarsat-2 quad-polarized SLC SAR image was acquired on 2008/4/9. The radar image and NDBC buoys are also marked. (A–C) correspond to the HH, HV (VH), and VV polarization channels, respectively.

3. Theory and Method

3.1. CP Reconstruction. CP theory is a newly developed method that can produce pseudo quad-polarized data from dual-polarized SAR data. The polarized information can be reconstructed in a satisfied precision with the decrease

of sensor operation power [26, 27]. The CP has three main modes: the CTRLR, dual circular polarized (DCP), and $\pi/4$ modes.

3.1.1. CTRLR Mode. In the CTRLR mode, the microwaves are transmitted in left-handed or right-handed circular

TABLE 2: Standard meteorological record from an NDBC in situ buoy.

Buoy ID	46237
Date	2008/4/9 01:51
Location	37.786 N 122.634 W
Water depth (m)	16.8
Dominant wave height (m)	2.18
Dominant wave period (s)	14.29
Average wave period (s)	7.03
Mean wave direction (°)	246

polarization and received in HH and VV polarization. The scattering vector $\vec{\mathbf{k}}_{\text{CTLR}}$ is expressed as

$$\vec{\mathbf{k}}_{\text{CTLR}} = \frac{[\mathbf{S}_{\text{hh}} - i\mathbf{S}_{\text{hv}} - i\mathbf{S}_{\text{vv}} + \mathbf{S}_{\text{hv}}]^T}{2}. \quad (1)$$

The relevant Hermitian covariance matrix is

$$\begin{aligned} \mathbf{C}_{\text{CTLR}} &= \left\langle \vec{\mathbf{k}}_{\text{CTLR}} \vec{\mathbf{k}}_{\text{CTLR}}^{*T} \right\rangle \\ &= \frac{1}{2} \begin{bmatrix} \langle |\mathbf{S}_{\text{hh}}|^2 \rangle & \langle i(\mathbf{S}_{\text{hh}} \mathbf{S}_{\text{vv}}^*) \rangle \\ \langle -i(\mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hh}}^*) \rangle & \langle |\mathbf{S}_{\text{vv}}|^2 \rangle \end{bmatrix} \\ &\quad + \frac{1}{2} \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \\ &\quad + \frac{1}{2} \begin{bmatrix} -2\Im(\langle \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{hv}}^* \rangle) & \langle \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{hv}}^* \rangle + \langle \mathbf{S}_{\text{hv}} \mathbf{S}_{\text{vv}}^* \rangle \\ \langle \mathbf{S}_{\text{hh}}^* \mathbf{S}_{\text{hv}} \rangle + \langle \mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hv}}^* \rangle & 2\Im(\langle \mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hv}}^* \rangle) \end{bmatrix}, \end{aligned} \quad (2)$$

where T stands for matrix transpose operations, $*$ stands for matrix conjugate operations, $\langle \rangle$ stands for average operations, \Re is the real part of complex number, \mathbf{S}_{pq} is polarization scattering matrix and p, q are horizontal or vertical polarized channel, i is the imaginary unit, and \Im is the imaginary part of complex number. It is also noticeable that this matrix can be simplified based on the assumption that the reflection is symmetrical.

$$\begin{aligned} \mathbf{C}_{\text{CTLR}} &= \left\langle \vec{\mathbf{k}}_{\text{CTLR}} \vec{\mathbf{k}}_{\text{CTLR}}^{*T} \right\rangle \\ &= \frac{1}{2} \begin{bmatrix} \langle |\mathbf{S}_{\text{hh}}|^2 \rangle & \langle i(\mathbf{S}_{\text{hh}} \mathbf{S}_{\text{vv}}^*) \rangle \\ \langle -i(\mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hh}}^*) \rangle & \langle |\mathbf{S}_{\text{vv}}|^2 \rangle \end{bmatrix} \\ &\quad + \frac{1}{2} \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}. \end{aligned} \quad (3)$$

In order to reconstruct a pseudo quad-polarized covariance matrix for the CP CTLR mode, $\langle |\mathbf{S}_{\text{hv}}|^2 \rangle$ also needs to be iteratively solved. This covariance matrix can be expressed as

$$\begin{aligned} &[\mathbf{C}_{\text{CTLR}}]_{\text{pseudo-quad}} \\ &= \begin{bmatrix} C_{11} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 & iC_{12} + \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \\ 0 & 2\langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 \\ (iC_{12} + \langle |\mathbf{S}_{\text{hv}}|^2 \rangle)^* & 0 & C_{22} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \end{bmatrix}. \end{aligned} \quad (4)$$

It is noticed that C_{ij} in (4) is element from CP 2×2 covariance matrix.

3.1.2. DCP Mode. In DCP mode, the microwaves are transmitted in right-handed circular polarization and received in left-handed or right-handed circular polarization. The scattering vector $\vec{\mathbf{k}}_{\text{DCP}}$ is expressed as

$$\vec{\mathbf{k}}_{\text{DCP}} = \frac{[\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}} + i2\mathbf{S}_{\text{hv}} i(\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}})]^T}{2}. \quad (5)$$

The relevant Hermitian covariance matrix is

$$\begin{aligned} \mathbf{C}_{\text{DCP}} &= \left\langle \vec{\mathbf{k}}_{\text{DCP}} \vec{\mathbf{k}}_{\text{DCP}}^{*T} \right\rangle \\ &= \frac{1}{4} \begin{bmatrix} \langle |\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}}|^2 \rangle & \langle -i(\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}})(\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}})^* \rangle \\ \langle -i(\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}})(\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}})^* \rangle & \langle |\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}}|^2 \rangle \end{bmatrix} \\ &\quad + \frac{1}{4} \begin{bmatrix} 4\langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 \\ 0 & 0 \end{bmatrix} \\ &\quad + \frac{1}{4} \begin{bmatrix} 4\Im(\langle (\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}}) \cdot \mathbf{S}_{\text{hv}}^* \rangle) & 2\langle (\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}})^* \rangle \\ 2\langle (\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}}) \cdot \mathbf{S}_{\text{hv}} \rangle & 0 \end{bmatrix}. \end{aligned} \quad (6)$$

It should be noted that the last term can be set to zero based on the assumption of reflectional symmetry and that the covariance matrix can be simplified to

$$\begin{aligned} \mathbf{C}_{\text{DCP}} &= \left\langle \vec{\mathbf{k}}_{\text{DCP}} \vec{\mathbf{k}}_{\text{DCP}}^{*T} \right\rangle \\ &= \frac{1}{4} \begin{bmatrix} \langle |\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}}|^2 \rangle & \langle -i(\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}})(\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}})^* \rangle \\ \langle -i(\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}})(\mathbf{S}_{\text{hh}} - \mathbf{S}_{\text{vv}})^* \rangle & \langle |\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{vv}}|^2 \rangle \end{bmatrix} \\ &\quad + \frac{1}{4} \begin{bmatrix} 4\langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \quad (7)$$

In order to reconstruct a pseudo quad-polarized covariance matrix for the CP DCP mode, $\langle |\mathbf{S}_{\text{hv}}|^2 \rangle$ needs to be iteratively solved. This covariance matrix can be expressed as [28]

$$\begin{aligned}
& [\mathbf{C}_{\text{DCP}}]_{\text{pseudo-quad}} \\
&= \begin{bmatrix} C_{11} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 & C_{12} + \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \\ 0 & 2\langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 \\ (C_{12} + \langle |\mathbf{S}_{\text{hv}}|^2 \rangle)^* & 0 & C_{22} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \end{bmatrix}. \quad (8)
\end{aligned}$$

It is noticed that C_{ij} in (8) is element from CP 2×2 covariance matrix.

3.1.3. $\pi/4$ Mode. In $\pi/4$ mode, the microwaves are transmitted in linear polarization with a 45° inclination and received in HH and VV polarization. The scattering vector $\vec{\mathbf{k}}_{\pi/4}$ is expressed as

$$\vec{\mathbf{k}}_{\pi/4} = \frac{[\mathbf{S}_{\text{hh}} + \mathbf{S}_{\text{hv}} \mathbf{S}_{\text{vv}} + \mathbf{S}_{\text{hv}}]^T}{\sqrt{2}}. \quad (9)$$

The relevant Hermitian covariance matrix is

$$\begin{aligned}
\mathbf{C}_{\pi/4} &= \left\langle \vec{\mathbf{k}}_{\pi/4} \vec{\mathbf{k}}_{\pi/4}^{*T} \right\rangle \\
&= \frac{1}{2} \begin{bmatrix} \langle |\mathbf{S}_{\text{hh}}|^2 \rangle & \langle \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{vv}}^* \rangle \\ \langle \mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hh}}^* \rangle & \langle |\mathbf{S}_{\text{vv}}|^2 \rangle \end{bmatrix} + \frac{1}{2} \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\
&+ \frac{1}{2} \begin{bmatrix} 2\Re(\langle \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{hv}}^* \rangle) & \langle \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{hv}}^* \rangle + \langle \mathbf{S}_{\text{hv}} \mathbf{S}_{\text{vv}}^* \rangle \\ \langle \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{hv}}^* \rangle + \langle \mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hv}}^* \rangle & 2\Re(\langle \mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hv}}^* \rangle) \end{bmatrix}. \quad (10)
\end{aligned}$$

It should be noted that the last term can be set to zero based on the assumption of reflectional symmetry and that the covariance matrix can be simplified to

$$\begin{aligned}
\mathbf{C}_{\pi/4} &= \left\langle \vec{\mathbf{k}}_{\pi/4} \vec{\mathbf{k}}_{\pi/4}^{*T} \right\rangle = \frac{1}{2} \begin{bmatrix} \langle |\mathbf{S}_{\text{hh}}|^2 \rangle & \langle \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{vv}}^* \rangle \\ \langle \mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hh}}^* \rangle & \langle |\mathbf{S}_{\text{vv}}|^2 \rangle \end{bmatrix} \\
&+ \frac{1}{2} \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \quad (11)
\end{aligned}$$

In order to reconstruct the pseudo quad-polarized covariance matrix for the CP $\pi/4$ mode, $\langle |\mathbf{S}_{\text{hv}}|^2 \rangle$ needs to be iteratively solved. The pseudo quad-polarized covariance matrix for the $\pi/4$ mode can be expressed as [28]

$$\begin{aligned}
& [\mathbf{C}_{\pi/4}]_{\text{pseudo-quad}} \\
&= \begin{bmatrix} C_{11} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 & C_{12} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \\ 0 & 2\langle |\mathbf{S}_{\text{hv}}|^2 \rangle & 0 \\ (C_{12} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle)^* & 0 & C_{22} - \langle |\mathbf{S}_{\text{hv}}|^2 \rangle \end{bmatrix}. \quad (12)
\end{aligned}$$

It is noticed that C_{ij} in (12) is element from CP 2×2 covariance matrix.

It should also be noted that the scattering vectors $\vec{\mathbf{k}}_{\text{CTRL}}$ and $\vec{\mathbf{k}}_{\text{DCP}}$ are related by

$$\vec{\mathbf{k}}_{\text{DCP}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \vec{\mathbf{k}}_{\text{CTRL}}. \quad (13)$$

The polarization information for the CTRL and DCP modes will be similar because of the above linear relationship, which means that the CP reconstruction result for these two modes will be the same [29, 30].

3.2. Wave Spectrum Retrieval Algorithm. For the linearly polarized backscattering cross section, the quad-polarized covariance matrix is used for the ocean surface wave polarization signatures, as follows [31]:

$$\begin{aligned}
\mathbf{C} &= \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \\
&= \begin{bmatrix} |\mathbf{S}_{\text{hh}}|^2 & \sqrt{2}(\mathbf{S}_{\text{hh}} \mathbf{S}_{\text{hv}}^*) & \mathbf{S}_{\text{hh}} \mathbf{S}_{\text{vv}}^* \\ \sqrt{2}(\mathbf{S}_{\text{hv}} \mathbf{S}_{\text{hh}}^*) & 2|\mathbf{S}_{\text{hv}}|^2 & \sqrt{2}(\mathbf{S}_{\text{hv}} \mathbf{S}_{\text{vv}}^*) \\ \mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hh}}^* & \sqrt{2}(\mathbf{S}_{\text{vv}} \mathbf{S}_{\text{hv}}^*) & |\mathbf{S}_{\text{vv}}|^2 \end{bmatrix}. \quad (14)
\end{aligned}$$

It is noticed that C_{ij} in (14) is element from FP 3×3 covariance matrix. Because the maximum contribution to the wave slope is obtained using only linear polarization, here, we set the ellipticity to zero and so the linearly polarized backscattering cross section at any polarization orientation angle can be derived as [14, 15]:

$$\begin{aligned}
\sigma(0, \varphi) &= \frac{1}{4} (\sigma_{\text{hh}} + \sigma_{\text{vv}}) [1 + \cos^2(2\varphi)] \\
&+ \frac{1}{2} (\sigma_{\text{hh}} - \sigma_{\text{vv}}) \cos 2\varphi + \sigma_{\text{hv}} \\
&+ \frac{1}{2} \Re(\sigma_{\text{hhvv}}) \sin^2(2\varphi), \quad (15)
\end{aligned}$$

where σ_{hh} , σ_{hv} , and σ_{vv} are the linearly polarized backscattering cross sections for the horizontal, cross, and vertical polarizations, respectively. σ_{hhvv} is the correlation between the horizontal and vertical polarizations, which can be directly derived from the covariance matrix.

In the linear modulation theory of ocean waves, the ocean surface height can be regarded as a superposition of sine waves and the radar backscattering cross section expressed as [16, 32, 33]

$$\xi(r, t) = \sum_k \xi_k \exp [i(k \cdot \mathbf{r} - \omega t)] + \text{cc}, \quad (16)$$

$$\sigma_{\text{pp}}(r, t) = \overline{\sigma_{\text{pp}}} \left(1 + \left\{ \sum_k T_{\text{PPK}}^{\text{R}} \xi_k \exp [i(k \cdot \mathbf{r} - \omega t)] + \text{cc} \right\} \right), \quad (17)$$

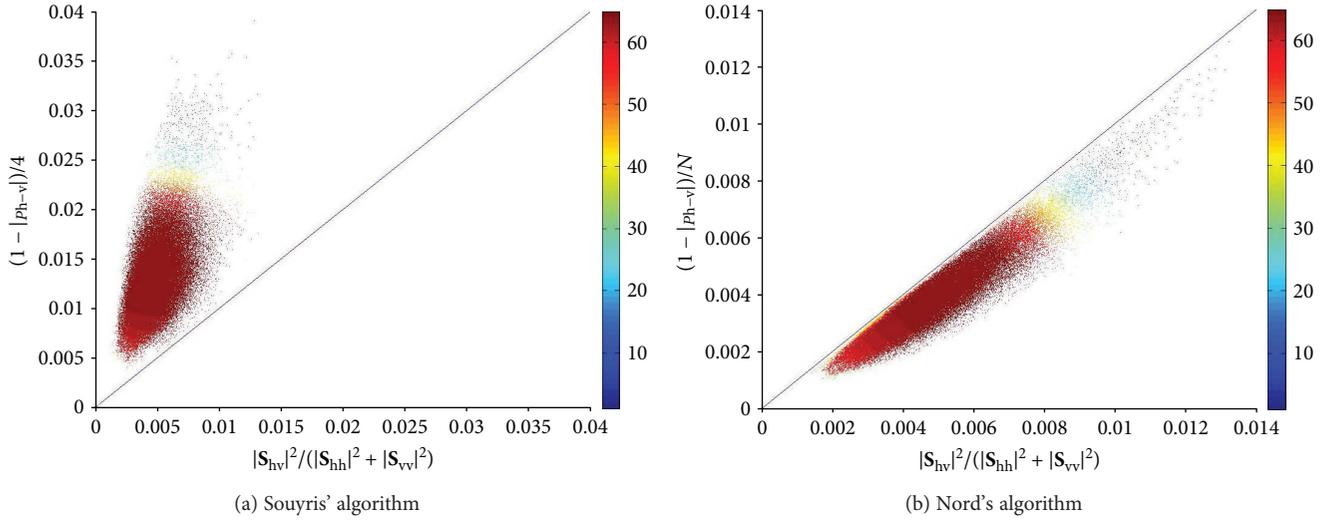


FIGURE 2: Scatter plot between $|S_{hv}|^2/(|S_{hh}|^2 + |S_{vv}|^2)$ and $(1 - \rho_{h-v})/N$. The value of N in (a) is 4 whereas in (b) it is calculated as $|S_{hh} - S_{vv}|^2/|S_{hv}|^2$. The convergence capability is clearly illustrated.

where ξ is the ocean surface height, k is the wavenumber, \mathbf{r} is the wave space vector, ω is the wave angular frequency, and cc is the complex conjugate of the series. T_{PPK}^R is the modulation transfer function for SAR, which includes the tilt modulation, hydrodynamic modulation, polarization angle modulation, range and azimuthal direction modulation, and velocity bunching modulation. In all modulation functions, only the tilt and polarization orientation angle modulation depend on the radar polarization [16]. Moreover, the wave slope induced by the polarization orientation angle modulation can be neglected for horizontal or vertical polarization [16]. By (18), the terms that are not sensitive to the polarization vanish and the equation can be simplified to [16]

$$\begin{aligned} \frac{\Delta\sigma_{vv}}{\sigma_{vv}} - \frac{\Delta\sigma_{hh}}{\sigma_{hh}} &= \sum_k (T_{vv}^t - T_{hh}^t) \xi_k \exp[i(k \cdot \mathbf{r} - \omega t)] + c_1, \\ \frac{\Delta\sigma_{\varphi\varphi}}{\sigma_{\varphi\varphi}} - \frac{\Delta\sigma_{vv}}{\sigma_{vv}} &= \sum_k (T_{\varphi\varphi}^t - T_{vv}^t + T_{\varphi\varphi}^p) \xi_k \exp[i(k \cdot \mathbf{r} - \omega t)] + c_2, \end{aligned} \quad (18)$$

where c_1 and c_2 are the complex conjugates of series, respectively; T_{vv}^t and T_{hh}^t are given by [16]

$$T_{vv}^t = ik_x \frac{4 - 0.5(1 - \sin^2\theta)}{\tan\theta(1 - \sin^2\theta)}, \quad (19)$$

$$T_{hh}^t = ik_x \frac{4 - 0.5(1 + \sin^2\theta)}{\tan\theta(1 + \sin^2\theta)}, \quad (20)$$

where k_x is the radar incidence wavenumber and θ is the incidence angle. $T_{\varphi\varphi}^t$ and $T_{\varphi\varphi}^p$ are given by (16) in [16]. By consolidating the equations, the straightforward algebraic calculation can be summarized as [16]

$$\begin{aligned} \frac{\Delta\sigma_{vv}}{\sigma_{vv}} - \frac{\Delta\sigma_{hh}}{\sigma_{hh}} &= -\frac{8 \tan\theta}{1 + \sin^2\theta} \frac{\delta\xi}{\delta x}, \\ \frac{\Delta\sigma_{\varphi\varphi}}{\sigma_{\varphi\varphi}} - \frac{\Delta\sigma_{vv}}{\sigma_{vv}} &= A \frac{\delta\xi}{\delta x} + B \frac{\delta\xi}{\delta y}, \end{aligned} \quad (21)$$

where $\delta\xi/\delta x$ and $\delta\xi/\delta y$ are the range and azimuthal wave slope, respectively. Parameters A and B are also given by (8) [16]. With an appropriate choice of φ , the range wave slope and azimuthal wave slope can be directly obtained. The wave slope spectrum and the other wave parameters can be estimated from wave slope.

4. Results and Discussion

4.1. Comparison of Reconstruction Algorithms. Besides the reflectional symmetry hypothesis, the relation between the copolarized and cross-polarized responses is the main condition limiting the solution of the CP reconstruction equations [21, 34].

$$\frac{|S_{hv}|^2}{|S_{hh}|^2 + |S_{vv}|^2} \approx \frac{1 - \rho_{h-v}}{N}. \quad (22)$$

Although both Souyris et al.'s [21] and Nord et al.'s [34] algorithms are used to reconstruct pseudo quad-polarized data, they differ in their definitions of N . Souyris et al. use 4 as the value of N , while Nord et al. calculate the value as

$$N = \frac{|S_{hh} - S_{vv}|^2}{|S_{hv}|^2}. \quad (23)$$

We, therefore, used the two reconstruction algorithms to calculate the slope rate of curve between $|S_{hv}|^2/(|S_{hh}|^2 + |S_{vv}|^2)$ and $(1 - \rho_{h-v}/N)$ (Figure 2). It is important to note that N in Figure 2(a) is equal to 4 while in Figure 2(b) N is given

TABLE 3: Convergence capability for CP reconstruction as found using Souyris' and Nord's algorithms.

Image ID	Total number of pixels	Number of nonconvergent pixels			
		CTLR_Souyris	$\pi/4$ _Souyris	CTLR_Nord	$\pi/4$ _Nord
1	6,220,800	7651	11,341	3962	1612
2	4,176,000	173	344	41	81

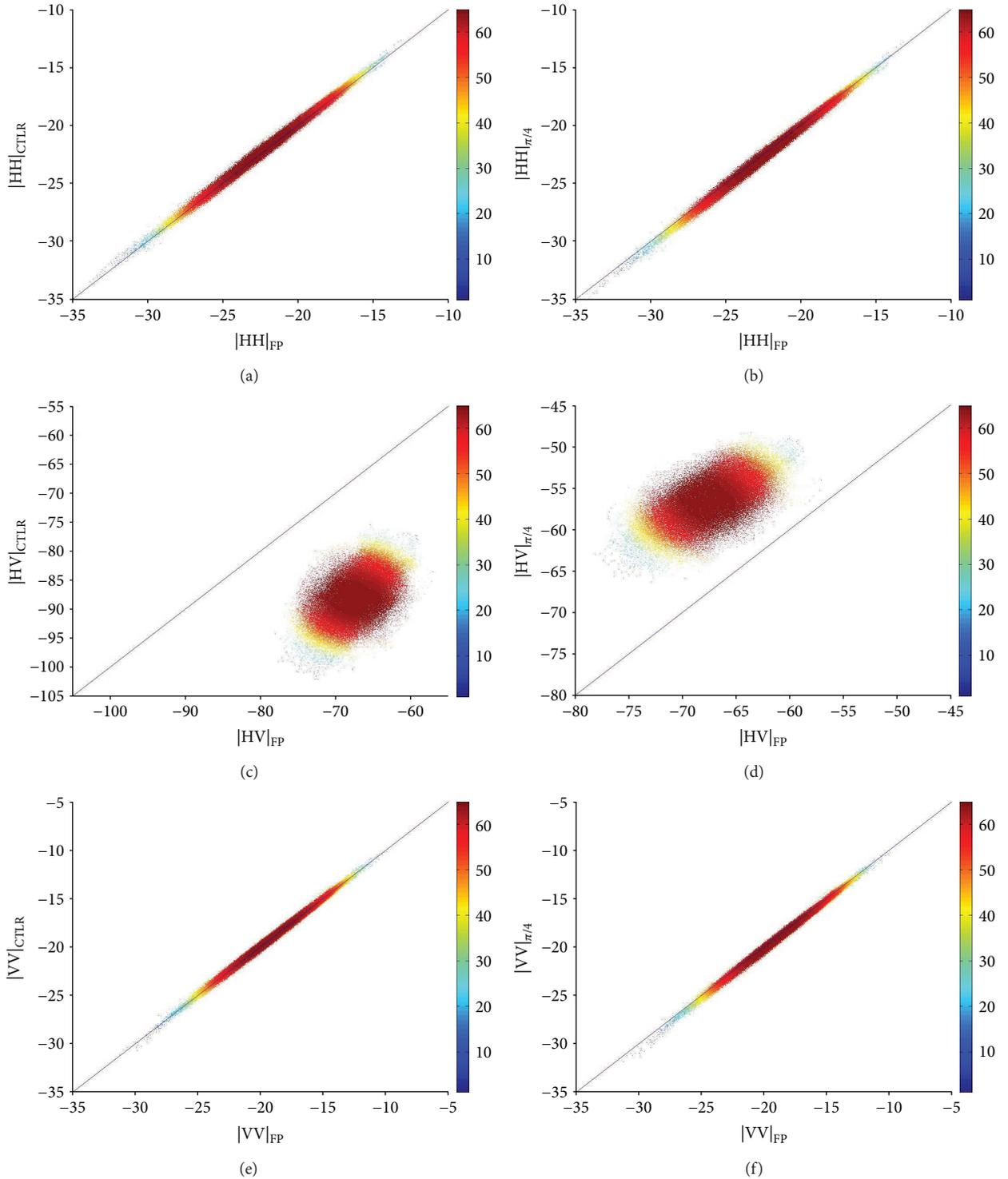
FIGURE 3: Reconstruction performance for HH, HV, and VV channels: scatter plots of FP and CP. (a), (c), and (e) correspond to the CTLR mode; (b), (d), and (f) correspond to the $\pi/4$ mode.

TABLE 4: Reconstruction accuracy of different polarized parameters for CTLR and $\pi/4$ modes.

Mode	σ_{hh}	σ_{hv}	σ_{vv}	ρ	l_1	l_2	l_3
CTLR	-0.0118	0.8658	-0.0050	0.0184	0.0014	-0.3394	0.8521
$\pi/4$	0.0205	-2.0466	0.0183	0.0202	0.0289	-0.3300	-0.3577

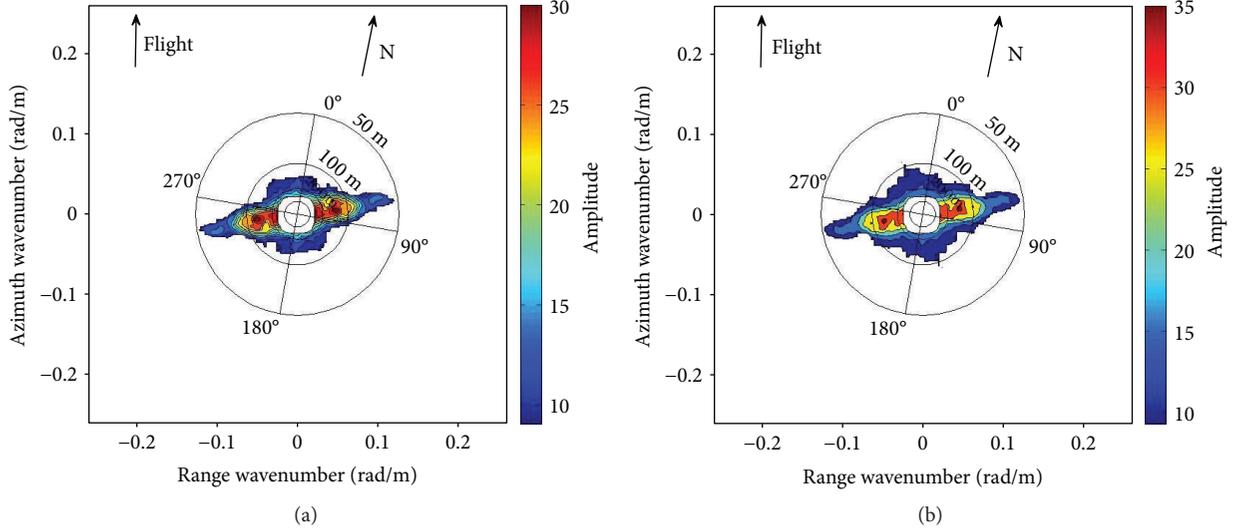


FIGURE 4: Ocean wave slope spectra derived from (a) FP data and (b) CP (CTLR mode) data (b). The color bar represents the wave energy amplitude.

FIGURE 5: The positions of subimages used for ocean information retrieval in scene 2, 1–4 represent four 512×512 pixel size subimage.

by $|\mathbf{S}_{hh} - \mathbf{S}_{vv}|^2 / |\mathbf{S}_{hv}|^2$, the value of which was calculated from FP data. We found that the convergence of Figure 2(a) was much inferior to Figure 2(b); the correlation coefficients, r , were equal to 0.53 and 0.93, respectively. To determine the optimum algorithm, we also calculated statistics for the non-convergence of pixels during the reconstruction process. The results for this are shown in Table 3. The number of nonconvergent pixels found when using the two subimages of the ocean surface was much lower for Nord's algorithm than for Souyris' algorithm. Nord's algorithm is, therefore, more suitable for reconstructing pseudo quad-polarized data of the ocean surface.

4.2. Evaluation of CP Reconstruction Accuracy. CP SAR has three main modes: CTLR mode, DCP mode, and $\pi/4$ mode.

Each CP mode has different orientation and ellipticity angles, which describe the transmitted and received wave status. Because of the linear relationship between the DCP and CTLR modes, these two modes have similar polarization information; therefore, we discuss only the $\pi/4$ mode and the CTLR mode in this paper. Figure 3 shows scatter plots detailing how well the derived pseudo quad-pol results fit the original quad-pol ones. The figure shows that the values of σ_{hh} and σ_{vv} between FP and CP are concentrated along the one-to-one line, which means that for σ_{hh} and σ_{vv} , the reconstruction of the pseudo quad-polarized data was good. However, the reconstruction of σ_{hv} shows a slight underestimation in the CTLR mode and an overestimation in the $\pi/4$ mode, which means that the reconstruction of σ_{hv} was not as good as for σ_{hh} and σ_{vv} . The reason for the low reconstruction accuracy for σ_{hv} may be related to the weak volume scattering at the ocean surface. It has been previously documented that the CP reconstruction algorithms are generally applied to fields and forests, where the total scattering contains a strong volume scattering component. However, the cross-polarization information is not directly used in our paper. Therefore, the inferior reconstruction performance of σ_{hv} will not make any influence on retrieval results.

To examine the reconstruction performance in more detail, we also selected 7 types of polarization parameters, including copolarization backscattering coefficient σ_{hh} , σ_{vv} , cross-polarization backscattering coefficient σ_{hv} , copolarization correlation coefficient ρ , and coherence matrix eigenvalues l_1 , l_2 , and l_3 , for each parameter, calculating the reconstruction accuracy for the CTLR and $\pi/4$ modes. The

TABLE 5: Comparison of ocean parameters derived from FP quad-pol and CP CTLR mode pseudo quad-pol data with NDBC buoy measurements.

Parameters	Mode	Zone	Retrieved value	Buoy	Error	Average error
Dominant wave height (m)	Fp	1	1.99	2.18	0.19	0.14
		2	2.35		0.17	
		3	2.34		0.16	
		4	2.14		0.04	
	Cp	1	2.37		0.19	0.36
		2	2.69		0.51	
		3	2.62		0.44	
		4	2.38		0.20	
Wave direction (°)	Fp	1	259.1	246	19.25	17.12
		2	264.57		19.38	
		3	265.06		18.02	
		4	263.73		15.19	
	Cp	1	265.25		13.1	17.96
		2	265.38		18.57	
		3	264.02		19.06	
		4	261.19		17.73	
Wave period (s)	Fp	1	14.41	14.29	0.12	0.61
		2	13.74		0.54	
		3	12.72		1.56	
		4	14.52		0.23	
	Cp	1	14.01		0.27	0.88
		2	13.31		0.97	
		3	12.58		1.70	
		4	14.88		0.59	

reconstruction accuracy $(\text{Pol}_{\text{cp}} - \text{Pol}_{\text{fp}})/\text{Pol}_{\text{fp}}$ was used to describe the amount of deviation in the CP reconstruction. The reconstruction accuracies for σ_{hh} , σ_{hv} , σ_{vv} , ρ , l_1 , l_2 , and l_3 are listed in Table 4. We used the descriptions “high,” “medium,” and “low” to describe the deviations, where “high” was defined as less than 10%, “medium” corresponded to 10%–30%, and “low” was defined as larger than 30%. Table 4 shows that, for both the CTLR and $\pi/4$ modes, the reconstruction accuracy for the parameters σ_{hh} , σ_{vv} , ρ , and l_1 was high, with the values for the CTLR mode being superior to those for the $\pi/4$ mode. We, therefore, decided to use the CRTL mode for the subsequent ocean surface parameter reconstruction and ocean wave slope spectrum retrieval.

4.3. Retrieval of Ocean Wave Slope Spectra. In this section, we describe how Nord’s algorithm was used to reconstruct the pseudo quad-pol data in the CTLR mode. The retrieved ocean wave slope spectra based on the reconstructed pseudo quad-polarized data are shown in Figure 4. Compared with the ocean wave slope spectrum derived from the FP data, the results for the CP CTLR mode obtained in this study are satisfied. It is also clear that the wave slope spectrum derived from FP presents more details compared to that derived from the CP data. However, the two wave slope spectra are similar on the whole.

In order to make a quantitative assessment of the wave retrieval results, the main wave parameters of dominant wave height, wave direction, and wave period were also extracted from the wave slope spectra and compared with the NDBC buoy data. The expressions of main wave parameters are

$$\begin{aligned}
 H_d &= \tan(S_{\text{rms}}) \frac{\lambda_d}{2}, \\
 \text{WD}_d &= \tan\left(\frac{\lambda_{d\text{-az}}}{\lambda_{d\text{-r}}}\right), \\
 \text{WP}_d &= \frac{2\pi}{\sqrt{k_d g \tan h(k_d h)}},
 \end{aligned} \tag{24}$$

where S_{rms} represents the root mean square of wave slope, λ_d represents the dominated wave length, $\lambda_{d\text{-az}}$ and $\lambda_{d\text{-r}}$, respectively, represent the components of λ_d in azimuth and range direction, k_d is the dominated wavenumber, g is acceleration of gravity, and h is the water depth.

Figure 5 shows four subimages’ position which is performed to retrieve result statistics, and the 180° ambiguity of wave slope spectrum has been removed according to the measured wind direction of buoy. The results shown in Table 5 show that the average errors in the dominant wave height, wave direction, and wave period derived from the FP data are 0.14, 17.12, and 0.61, while those derived from

the CP data are 0.36, 17.96, and 0.88, respectively. It should be noted that, although the average errors in the dominant wave height, wave direction, and wave period derived from the FP data were much larger than for the CP, the retrieved CP SAR results are also satisfactory for ocean wave measurements. It is also true that the retrieval accuracy obtained using CP SAR in this study was limited by the reconstruction of the CP pseudo quad-pol data—this was especially due to the loss in precision for σ_{hv} . In future studies, the retrieval accuracy obtained using CP SAR will need to be improved.

5. Conclusion

In this paper, we investigated the capacity and potential of CP for ocean wave measurement. A Radarsat-2 quad-polarized image was used to reconstruct CP pseudo quad-polarized data and validate the potential of this for ocean slope spectrum retrieval. The main conclusions of this study are as follows.

For ocean surface waves, Nord's reconstruction algorithm is much more effective than Souyris'. Because of its convergence capacity and iteration time, we recommend adopting Nord's algorithm for the reconstruction of CP pseudo quad-polarized data.

For both the CTLR and $\pi/4$ modes, the reconstruction accuracy for the parameters σ_{hh} , σ_{vv} , ρ , and l_1 was high whereas for σ_{hv} , l_2 , and l_3 it was low. In terms of the retrieval of ocean wave slope spectra using the CP mode, the CTLR mode is superior to the $\pi/4$ mode and, therefore, for the retrieval of ocean wave slope spectra, we recommend using the CTLR mode.

The potential of CP for ocean wave slope spectrum retrieval has been demonstrated in this paper. The values for the main wave parameters were also validated using in situ NDBC buoy data. A comparison of the CP SAR-based wave field information with the buoy outputs showed good agreement in terms of the dominant wave height, wave direction, and wave period, with biases of 0.36 m, 17.96°, and 0.88 s, respectively. Overall, the results for the retrieval of ocean wave slope spectra and ocean wave parameters using CP SAR were satisfactory. Moreover, considering only single SAR imagery is performed to wave retrieval algorithm study, the universality validation of ocean wave information retrieval from CP SAR will still need more results to support. In our future study, we will collect more quad-pol SAR to match up with buoy for plotting the wave parameter difference between FP and CP SAR precisely.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors gratefully acknowledge the financial support provided by the National Key Research and Development Program of China (2016YFB0502504, 2016YFB0502500) and National Science Foundation of China under Grant

nos. 41301500, 41431174, 61471358, and 41401427. This article is partly sponsored by the scholarship funding from the Chinese Academy of Sciences. The authors would also like to acknowledge Professor Zhang from the First Institute of Oceanography, State Oceanic Administration, for supporting research on CP theory. The authors would also like to acknowledge Dr. Bian and Dr. Liu from Institute of Remote Sensing and Digital Earth for polishing the original manuscript.

References

- [1] J. Chen, Z. Lee, C. Hu, and J. Wei, "Improving satellite data products for open oceans with a scheme to correct the residual errors in remote sensing reflectance," *Journal of Geophysical Research Oceans*, vol. 121, no. 6, pp. 3866–3886, 2016.
- [2] J. Chen, X. He, B. Zhou, and D. Pan, "Deriving colored dissolved organic matter absorption coefficient from ocean color with a neural quasi-analytical algorithm," *Journal of Geophysical Research Oceans*, vol. 122, no. 11, pp. 8543–8556, 2017.
- [3] S. K. Chaturvedi, P. Shanmugam, C. S. Yang, and U. Guven, "Detection of ocean wave parameters using synthetic aperture radar (SAR) data," *Journal of Navigation*, vol. 66, no. 2, pp. 283–293, 2013.
- [4] D. Kim, "Wind retrieval from X-band SAR image using numerical ocean scattering model," *Korean Journal of Remote Sensing*, vol. 25, no. 3, pp. 243–253, 2009.
- [5] M. Sugimoto, N. Shiroto, and K. Ouchi, "Estimation of ocean wave height using polarization ratio of synthetic aperture radar data," in *2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2821–2824, Vancouver, BC, Canada, July 2011.
- [6] P. W. Vachon, H. E. Krogstad, and J. S. Paterson, "Airborne and spaceborne synthetic aperture radar observations of ocean waves," *Atmosphere-Ocean*, vol. 32, no. 1, pp. 83–112, 1994.
- [7] K. Hasselmann and S. Hasselmann, "On the nonlinear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion," *Journal of Geophysical Research Oceans*, vol. 96, no. C6, pp. 10713–10729, 1991.
- [8] S. Hasselmann and K. Hasselmann, "Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. Part I: a new method for efficient computations of the exact nonlinear transfer integral," *Journal of Physical Oceanography*, vol. 15, no. 11, pp. 1369–1377, 1985.
- [9] C. Mastenbroek and C. F. de Valk, "A semiparametric algorithm to retrieve ocean wave spectra from synthetic aperture radar," *Journal of Geophysical Research Oceans*, vol. 105, no. C2, pp. 3497–3516, 2000.
- [10] H. Wang, J. Zhu, J. Yang, and C. Shi, "A semiempirical algorithm for SAR wave height retrieval and its validation using Envisat ASAR wave mode data," *Acta Oceanologica Sinica*, vol. 31, no. 3, pp. 59–66, 2012.
- [11] W. Shao, Z. Zhang, X. Li, and H. Li, "Ocean wave parameters retrieval from Sentinel-1 SAR imagery," *Remote Sensing*, vol. 8, no. 9, 2016.
- [12] B. Zhang, X. Li, W. Perrie, and Y. He, "Synergistic measurements of ocean winds and waves from SAR," *Journal of Geophysical Research Oceans*, vol. 120, no. 9, pp. 6164–6184, 2015.
- [13] D. L. SCHULER and J. S. LEE, "A microwave technique to improve the measurement of directional ocean wave spectra,"

- International Journal of Remote Sensing*, vol. 16, no. 2, pp. 199–215, 1995.
- [14] D. L. Schuler, J. S. Lee, D. Kasilingam, and E. Pottier, “Measurement of ocean surface slopes and wave spectra using polarimetric SAR image data,” *Remote Sensing of Environment*, vol. 91, no. 2, pp. 198–211, 2004.
- [15] Y. He, H. Shen, and W. Perrie, “Remote sensing of ocean waves by polarimetric SAR,” *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 12, pp. 1768–1773, 2006.
- [16] Y. He, W. Perrie, T. Xie, and Q. Zou, “Ocean wave spectra from a linear polarimetric SAR,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 11, pp. 2623–2631, 2004.
- [17] B. Zhang, W. Perrie, and Y. He, “Validation of RADARSAT-2 fully polarimetric SAR measurements of ocean surface waves,” *Journal of Geophysical Research Oceans*, vol. 115, no. C6, pp. 302–315, 2010.
- [18] Y. He, B. Zhang, and W. Perrie, “Validation of RADARSAT-2 polarimetric SAR measurements of ocean waves,” in *2009 IEEE International Geoscience and Remote Sensing Symposium*, pp. III-168–III-171, Cape Town, South Africa, July 2009.
- [19] R. L. Paes, F. Nunziata, and M. Migliaccio, “On the capability of hybrid-polarity features to observe metallic targets at sea,” *IEEE Journal of Oceanic Engineering*, vol. 41, no. 2, pp. 346–361, 2016.
- [20] R. K. Raney, “A perspective on compact polarimetry,” *IEEE Geoscience and Remote Sensing Letters*, vol. 160, pp. 12–18, 2011.
- [21] J.-C. Souyris, P. Imbo, R. Fjortoft, S. Mingot, and J.-S. Lee, “Compact polarimetry based on symmetry properties of geophysical media: the $\pi/4$ mode,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 634–646, 2005.
- [22] F. Nunziata, M. Migliaccio, and X. Li, “Sea oil slick observation using hybrid-polarity SAR architecture,” *IEEE Journal of Oceanic Engineering*, vol. 40, no. 2, pp. 426–440, 2015.
- [23] J. Yin, J. Yang, Z. S. Zhou, and J. Song, “The extended Bragg scattering model-based method for ship and oil-spill observation using compact polarimetric SAR,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 8, pp. 3760–3772, 2015.
- [24] T. Geldsetzer, F. Charbonneau, M. Arkett, and T. Zagon, “Ocean wind study using simulated RCM compact-polarimetry SAR,” *Canadian Journal of Remote Sensing*, vol. 41, no. 5, pp. 418–430, 2015.
- [25] H. Li, J. Wu, W. Perrie, and Y. He, “Wind speed retrieval from hybrid-pol compact polarization synthetic aperture radar images,” *IEEE Journal of Oceanic Engineering*, no. 99, pp. 1–12, 2017.
- [26] R. Shirvany, M. Chabert, and J. Y. Tourneret, “Ship and oil-spill detection using the degree of polarization in linear and hybrid/compact dual-pol SAR,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 3, pp. 885–892, 2012.
- [27] R. Touzi and F. Charbonneau, “Requirements on the calibration of hybrid-compact SAR,” in *2014 IEEE Geoscience and Remote Sensing Symposium*, pp. 1109–1112, Quebec City, QC, Canada, July 2014.
- [28] S. Boularbah, M. Ouarzeddine, and A. Belhadj-Aissa, “Investigation of the capability of the compact polarimetry mode to reconstruct full polarimetry mode using RADARSAT2 data,” *Advanced Electromagnetics*, vol. 1, no. 1, p. 19, 2012.
- [29] X. Zhang, J. Zhang, M. Liu, and J. Meng, “Assessment of C-band compact polarimetry SAR for sea ice classification,” *Acta Oceanologica Sinica*, vol. 35, no. 5, pp. 79–88, 2016.
- [30] P. C. Dubois-Fernandez, J. C. Souyris, S. Angelliaume, and F. Garestier, “The compact polarimetry alternative for spaceborne SAR at low frequency,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 10, pp. 3208–3222, 2008.
- [31] J. J. van Zyl, H. A. Zebker, and C. Elachi, “Imaging radar polarization signatures: theory and observation,” *Radio Science*, vol. 22, no. 4, pp. 529–543, 1987.
- [32] J.-S. Lee, D. L. Schuler, and T. L. Ainsworth, “Polarimetric SAR data compensation for terrain azimuth slope variation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 5, pp. 2153–2163, 2000.
- [33] J.-S. Lee, D. L. Schuler, T. L. Ainsworth, E. Krogager, D. Kasilingam, and W.-M. Boerner, “On the estimation of radar polarization orientation shifts induced by terrain slopes,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 1, pp. 30–41, 2002.
- [34] M. E. Nord, T. L. Ainsworth, J. S. Lee, and N. J. S. Stacy, “Comparison of compact polarimetric synthetic aperture radar modes,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 174–188, 2009.

Research Article

A DLM-LSTM Framework for North-South Land Deformation Trend Analysis from Low-Cost GPS Sensor Time Series

Fangling Pu ^{1,2}, Zhaozhuo Xu,³ Hongyu Chen,¹ Xin Xu,¹ and Nengcheng Chen ⁴

¹School of Electronic Information, Wuhan University, Wuhan, Hubei 430072, China

²Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan, Hubei 430079, China

³Electrical Engineering Department, Stanford University, Palo Alto, CA 94305, USA

⁴State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan, Hubei 430079, China

Correspondence should be addressed to Fangling Pu; flpu@whu.edu.cn

Received 18 February 2018; Revised 30 March 2018; Accepted 13 May 2018; Published 6 June 2018

Academic Editor: Biswajeet Pradhan

Copyright © 2018 Fangling Pu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Landslides endanger regular industrial production and human safety. Displacement trend analysis gives us an explicit way to observe and forecast landslides. Although satellite-borne remote sensing methods such as synthetic aperture radar have gradually replaced manual measurement in detecting deformation trends, they fail to observe displacement in a north-south direction. Wireless low-cost GPS sensors have been developed to assist remote sensing methods in north-south deformation monitoring because of their high temporal resolution and wide usage. In our paper, a DLM-LSTM framework is developed to extract and predict north-south land deformation trends from meter accuracy GPS receivers. A dynamic linear model is introduced to model the relation between measurement and the state vector, including the trend, periodic variation, and autoregressive factors in a discontinuous low-cost latitude time series. The deformation trend with submeter-level accuracy is extracted by a Kalman filter and smoother. With validated input as in previous work, the power of an LSTM network is also shown in its ability to predict deformation trends in submeter-level accuracy. A submeter-level deformation trend is detected from wireless low-cost GPS sensors with meter-level navigation error. The framework will have broad application prospects in geological disaster monitoring.

1. Introduction

Due to long-term mining, landslides threaten the safety of coal mine workers and nearby residents [1]. Continuous deformation on the surface of a coal mine is the main cause of landslides. The surface deforms slowly or relatively steadily most of the time [2]. Once the displacement accumulates and exceeds the maximum tolerance of the land surface, large slope failure becomes unstoppable [3, 4]. Constant deformation monitoring captures necessary information for landslide research. After summarizing the laws of long time land surface motion, early warnings can be communicated to local people [5].

Coal mining in the Fushun Western Open-Pit Coal Mine (FWOCM), which is located at $41^{\circ}50'38.33''$ north latitude and $123^{\circ}53'21.77''$ east longitude, started in the early 20th century. Production of coal used to be a pillar of the Fushun economy. Decades of mining has caused frequent surface deformation and landslides that threaten the city, Fushun. An open pit with 6600 m east-to-west length and 2200 m north-to-south width stands among the residential districts [6, 7]. Therefore, a landslide is a serious threat to miners and local people. According to our field investigation, deformation in the north-south direction is much worse than in the east-west direction. The velocity of displacement on the north and south sides is less than 3 m/year. A fault zone

was spotted between the coal mine's conveyor belts and residential buildings. To protect the buildings and people nearby, an effective approach is required to detect north-south land deformation.

For large-scale slopes, synthetic aperture radar (SAR) and differential interferometric synthetic aperture radar (DInSAR) are typical area surveying techniques [8, 9] with submeter- or centimeter-level accuracy. SAR systems provide full-time and full-weather observations for land motion. However, effective mine surface deformation monitoring requires high spatial resolution remote sensing images, which are expensive and hard to obtain. Direct temporal prediction of landslides is, however, scarce and constrained [10, 11]. Moreover, if the slope does not correspond with the ascending or descending passes of SAR satellites, it is concentrated in a few pixels of the satellite image. Therefore, traditional SAR or DInSAR can detect submeter or centimeter displacements in up-down and east-west directions (Figure 4 in [12, 13]), but have difficulty in observing north-south ground motion. Although a number of studies have engaged in modifying conventional SAR image processing algorithms [14] to detect north-south deformation, SAR and DInSAR solutions are still too sophisticated for practice uses. Therefore, north-south deformation time series should be observed directly.

Time series constructed from global positioning system (GPS) observation provide indispensable information about north-south land surface deformation. Compared with large-scale observation, such as SAR, GPS receivers have the advantage of high time resolution and north-south displacement observation. In past cases, GPS latitude time series of one location were obtained by reading the value periodically. These GPS series supplement remote sensing approaches in north-south deformation detection. With the rapid development of information and communication technology (ICT), interconnected GPS sensors form wireless sensor networks that can substitute manual measurement in long-term, high temporal resolution land deformation monitoring. Spatially distributed autonomous sensor systems generate massive data streams while saving plenty of labor costs.

However, time series obtained from wireless GPS sensors suffer in terms of both accuracy and continuity. Current wireless GPS sensor networks use low-cost GPS receivers that are widely found in smartphones, vehicles, and unmanned aerial vehicles (UAV). They receive satellite C/A (coarse acquisition) codes on the L1 frequency (1.57542 GHz). Under standard position services (SPS), the accuracy of a stand-alone C/A code receiver that uses only satellites is within 10 meters at the metric of 95% all-in-view horizontal, which exhibits significant errors in positioning. On the other hand, because most wireless GPS sensors are deployed in rural areas, unstable power supply becomes inevitable. Limited energy leads to GPS sensor dysfunction, which generates discontinuous points in latitude time series. At the current stage, problems in both sensor materials and power supply cannot be solved immediately. Therefore, an advanced processing paradigm is required to exploit a low-cost, discontinuous GPS latitude stream.

Submeter displacement monitoring is required at FWOCM. Real-time differential GPS techniques, such as real-time kinematic (RTK) [15], improve positional accuracy through sending corrections from a base station to an RTK terminal in real time. Network RTK (NRTK) uses several widely spaced permanent stations, named continuously operating reference stations (CORS), to provide correction information to the RTK user terminals. NRTK improves the accuracy, reliability, and availability of RTK. Centimeter or millimeter accuracy of displacement measurements can be obtained with NRTK. However, the cost of establishing RTK and NRTK and their operation is high, which hinders the application of RTK and NRTK. Combining only low-cost GPS receivers has demonstrated another solution [16, 17]. Averaging, linear filters and maximum likelihood estimation have been used to process the homogeneous GPS receivers with the aim of improving accuracy. The paper [18] suggests that using a dense array of low-cost GPS receivers may achieve high-precision displacement measurements. It also suggests that removing the fluctuation may improve the displacement measurement of low-cost GPS receivers.

Like most time series, the 21-month-long time series of the low-cost GPS sensors deployed at FWOCM can be assumed to consist of three components: trend, seasonal variation, and unsystematic irregular fluctuations. For noisy fragmented time series, the trend expresses the dynamic movement explicitly, and thus, we put trend analysis as one of our overriding tasks in land deformation monitoring using low-cost GPS sensors. Recently, methods proposed to trend analysis can be summarized as two-step procedures, which are trend extraction and prediction. The unprocessed, raw time series is firstly smoothed to obtain validated trends. Then the extracted trends are used as training samples to estimate future trends.

There is a wealth of literature concerning the trend extraction and prediction of low-cost GPS time series [18]. However, few comprehensive trend analysis frameworks have been developed. For trend extraction, researchers firstly utilized Kalman filters [19]. Although the Kalman filter has good performance in denoising Gaussian noise, the simplicity of Kalman state space function limits its power in filtering out complex periodical variations. To address this issue, empirical mode decomposition (EMD) is introduced in Gao et al.'s paper [20, 21] to enhance the Kalman filter in nonlinear and non-Gaussian fluctuation removal. However, the unexplained intrinsic mode functions (IMF) generated during EMD sifting makes it less reliable. In the mean time, experiments in early warning of abrupt displacement change at the Yemaomian Landslide of the Three Gorge Region, China [22], shows that displacement fluctuations can be detected by an autoregressive model (AR), which is seldom considered in previous Kalman or EMD approaches. Thus, a unified model that combines spline smoothing, periodical variation detection, and AR process estimations should be proposed.

For trends prediction, recurrent neural networks (RNNs), particularly with long short-term memory (LSTM) hidden units, are widely acknowledged to be powerful state-of-art

prediction models for sensor time series [23]. However, LSTM has rarely been shown in recent GPS land deformation trend analysis papers. Moreover, most LSTM networks for land deformation trend prediction received unprocessed noisy raw data as input, which confuses the network with wrong prior knowledge. Since the Kalman filter has been proved to enhance LSTM networks' performance [24], it is natural for us to think about applying LSTM networks with a unified trend extraction model as a preprocessed step.

This paper studies the feasibility and practical implementation of a unified comprehensive framework that can satisfy both land deformation trend extraction and prediction from low-cost GPS sensors. We model slowly moving displacement, which means annual movement of less than 2 meters [2], as variations of mean level in daily measurements, and treat collected time series as samples from the distribution of each measurement. In this case, all of the discontinuity within a day can be avoided. Then, by refining the dynamic linear model (DLM) proposed in [25] and combining it with LSTM networks, we develop an end-to-end DLM-LSTM framework to solve trend analysis in a discontinuous low-cost GPS sensor time series. The major contribution of our framework lies as follows:

- (1) The introduced DLM model expands both the observation matrix and the evolution matrix of a state space function to model periodical variation and autocorrelation in residuals. These expansions allow DLM to extract displacement trends with higher precision from noisy, discontinuous, periodically variable, and autoregressive low-cost GPS time series.
- (2) For sequence prediction, the introduced LSTM networks receive training-preprocessed samples. Those 5 training samples provide validated prior knowledge of land deformation trends, which enhances LSTM networks' power in both linear and nonlinear time series prediction. LSTM networks will receive a lower mean square error, which means that prediction sequences will be closer to real land deformation trends.
- (3) The fusion of DLM and LSTM provides an end-to-end north-south land deformation trend analysis workflow, which can undertake both trend extraction and prediction jobs. Our framework is able to model discontinuous low-cost GPS time series without other operations. The raw sensor data received by wireless sensor networks can be used directly as input.

To test our DLM-LSTM framework, a spatial-temporal distributed low-cost latitude time series (LCLTS) dataset is constructed using deployed wireless GPS sensors on both the moving southern slope of the Fushun Western Open-Pit Coal Mine (FWOCM) and the static comparative field in Baoxie, Wuhan, China. All GPS sensors that provide latitude time series are of the same type. Experiments have shown that our framework extracts and predicts submeter-level north-south deformation trends in FWOCM from

February to March in 2016, which coincides with manual measurement over the same period. However, the trend extracted in the Baoxie field from November to December 2016 is static at the submeter level. The two results validate our framework and prove its time and space independence. Like [26], our sensor network helps SAR in high-resolution north-south deformation detection. Steady observations reveal local features of 2-dimensional mine surface motion, which are valuable resources for researchers and authorities.

The paper is organized as follows. In Section 2, the DLM-LSTM framework is presented. The processing results for the LCLTS dataset are shown in Section 3. The conclusion is given in Section 4.

2. DLM-LSTM Framework for Deformation Trend Extraction and Forecast

2.1. Description of DLM-LSTM Framework. By comparing the characteristics of GPS latitude time series and field observations, the land deformation shows a stepwise character, which means the average daily latitude changed over adjacent days, but the latitudes within each day remain stationary. Long-term measurement implies that land displacement contains periodic components, which triggers fluctuations that mask the monotonic trend. Based on these observations, we propose a framework to filter out these disturbances and obtain monotonic deformation trends. Then the land deformation can be predicted. Figure 1 shows the flow chart of our DLM-LSTM framework. Firstly, the latitudes received every day are modeled as samples from Gaussian distributions. Mean level and standard deviation are regarded as the input variables to state space functions of the dynamic linear model. Then, we use a recursive Kalman filter to get the optimal states in the DLM model equations. Finally, the land deformation trend is extracted by a recursive Kalman smoother. We apply a simulation smoother in the last step of DLM to describe the full joint distribution of all states and the characteristic of extracted trends. After DLM, the deformation trend is transferred to LSTM networks for training. When network parameters are estimated, LSTM networks are able to predict the displacement using prior trend knowledge. In the end, extracted and predicted deformation trends are compared with observed ground truth displacement to evaluate the performance of our framework.

2.2. Dynamic Linear Model. The dynamic linear model was first defined and introduced in analyzing trends in stratosphere ozone time series [25]. From the definition, "dynamic" was explained as the evolution of the regression coefficient over time and "linear" means that the state space equations are all linear. From our understanding, DLM is the general form of Kalman filter functions. By adding additional operators to the observation matrix and the evolution matrix, simple Kalman Filters can denoise complex time-varying time series to obtain validated trends. In the following subsections, we will introduce the state space function of DLM designed for low-cost GPS time series and a Kalman recursive method for parameter estimation.

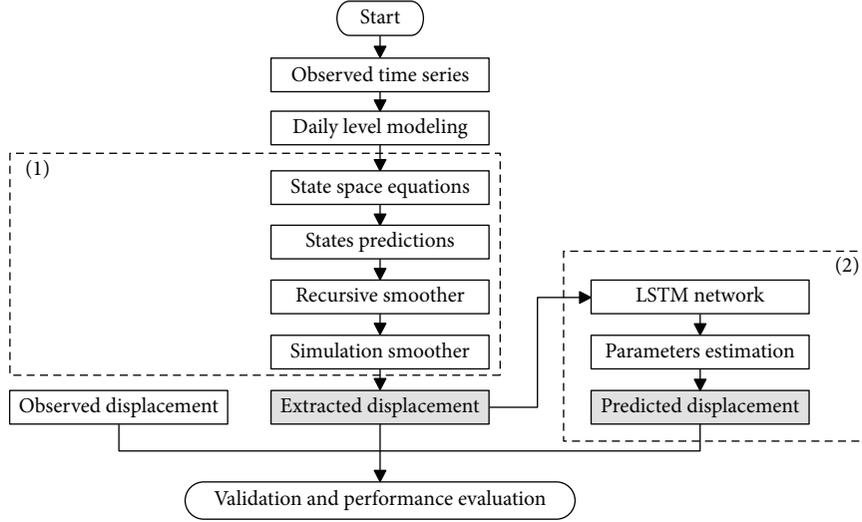


FIGURE 1: Flow chart of the proposed dynamic linear model (DLM)-long short-term memory (LSTM) framework. (1) Dynamic linear model using recursive Kalman formula for parameter estimation. (2) Recurrent LSTM networks.

The land deformation trend will be extracted after the parameters are fully estimated.

2.2.1. State Space Functions. According to [27], the general Kalman state space model is written as

$$\mathbf{y}_t = \mathbf{F}_t \mathbf{x}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathbf{N}(0, \mathbf{V}_t), \quad (1)$$

$$\mathbf{x}_t = \mathbf{G}_t \mathbf{x}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathbf{N}(0, \mathbf{W}_t), \quad (2)$$

where \mathbf{y}_t are the observations at time t , $t = 1, \dots, n$. The unobserved states, in our case the deformation trends, are defined as vector \mathbf{x}_t , which evolves in time with noise \mathbf{w}_t according to evolution matrix \mathbf{G}_t in (2). The relationship between observed and unobserved states is described in observation (1), where the result of \mathbf{x}_t transformed by observation matrix \mathbf{F}_t combining noise \mathbf{v}_t leads to \mathbf{y}_t . Both noises \mathbf{w}_t and \mathbf{v}_t lie in the Gaussian distribution with covariance \mathbf{W}_t and \mathbf{V}_t .

The DLM model first expands vectors in (1) and (2) in the following form:

$$\begin{aligned} \mathbf{x}_t &= [u_t, \alpha_t]^T, \\ \mathbf{G}_{\text{trend}} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \\ \mathbf{F}_{\text{trend}} &= [1, 0], \\ \mathbf{W}_{\text{trend}} &= \begin{bmatrix} \sigma_{\text{level}}^2 & 0 \\ 0 & \sigma_{\text{trend}}^2 \end{bmatrix}, \\ \mathbf{V}_t &= [\sigma_{\text{obs}(t)}], \end{aligned} \quad (3)$$

where u_t is the mean level of latitude each day and α_t is the change in the level from day t to day $t+1$. Correspondingly, \mathbf{W}_t is decomposed to two components: the Gaussian noise distribution with standard deviation σ_{level} and σ_{trend} . This form of DLM model is able to detect

smooth variation in the mean level and infer changes that happened in the land surface, but it does not consider periodic fluctuations.

To detect periodic fluctuations, the DLM models these fluctuations using harmonic functions. The corresponding operators in the observation matrix is defined as

$$\begin{aligned} \mathbf{G}_{\text{peri}(k)} &= \begin{bmatrix} \cos\left(k2\frac{\pi}{30}\right) & \sin\left(k2\frac{\pi}{30}\right) \\ -\sin\left(k2\frac{\pi}{30}\right) & \cos\left(k2\frac{\pi}{30}\right) \end{bmatrix}, \\ \mathbf{F}_{\text{peri}} &= [1, 0], \\ \mathbf{W}_{\text{peri}} &= \begin{bmatrix} \sigma_{\text{peri}(k)}^2 & 0 \\ 0 & \sigma_{\text{peri}(k)}^2 \end{bmatrix}, \end{aligned} \quad (4)$$

where k means the k th harmonic detected in daily level series within 30 days. We adjust number k to reach the best performance during the experiment.

Because paper [22] reveals an autoregressive characteristic in the time series of slowly moving surface displacement, we use a first-order autoregressive model for residual components, $\eta_t = \rho\eta_{t-1} + \epsilon_{\text{AR}}$, with $\epsilon_{\text{AR}} \sim N(0, \sigma_{\text{AR}})$. In DLM form, we have

$$\begin{aligned} \mathbf{G}_{\text{AR}} &= [\rho], \\ \mathbf{F}_{\text{AR}} &= [1], \\ \mathbf{W}_{\text{AR}} &= [\sigma_{\text{AR}}^2]. \end{aligned} \quad (5)$$

To extract land deformation trends from daily latitude time series, the DLM model must combine all the operators above, which leads to a larger-scale evolution and observation matrix as shown below:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{\text{trend}} & 0 & 0 & 0 \\ 0 & \mathbf{G}_{\text{peri}(5)} & 0 & 0 \\ 0 & 0 & \mathbf{G}_{\text{peri}(10)} & 0 \\ 0 & 0 & 0 & \mathbf{G}_{\text{AR}} \end{bmatrix}, \quad (6)$$

$$\mathbf{F}_t = \begin{bmatrix} \mathbf{F}_{\text{trend}} & \mathbf{F}_{\text{peri}(5)} & \mathbf{F}_{\text{peri}(10)} & \mathbf{F}_{\text{AR}} \end{bmatrix},$$

where $\mathbf{G}_{\text{peri}(5)}$ and $\mathbf{G}_{\text{peri}(10)}$ are the 5th and 10th harmonics of 30-day circulation. These two harmonics are found to outperform the denoising abilities of other combinations in periodic variations. Other operators remain the same as above. Respectively, we have a state vector \mathbf{x}_t as shown below:

$$\mathbf{x}_t = [\mu_t \quad \alpha_t \quad u_{\text{peri}(5)} \quad u_{\text{peri}(5)}^* \quad u_{\text{peri}(10)} \quad u_{\text{peri}(10)}^* \quad \eta_t]^T. \quad (7)$$

These DLM state space equations contain unknown parameters, including covariance matrix \mathbf{V}_t , AR coefficient ρ , and so on. In the following step, these parameters will be fully estimated to be able to apply DLM in real situations.

2.2.2. Model Parameter Estimation. In DLM, the unknowns are divided into two major categories: model state \mathbf{x}_t and auxiliary parameters defined in state space equations. In this paper, we select the Kalman formula method proposed in paper [25] to obtain the unknown trends and parameters. The essence of the Kalman formula is the recursive prediction and comparison. By continuously updating the covariance matrix in the state space equations, the Kalman formula constructs the connections between measurement and hidden real states. Firstly, we perform Kalman filter forward recursion to predict states \mathbf{x}_t , and we assume the conditional probability of predicted states \mathbf{x}_{t+1} follow normal distributions as below:

$$p(\mathbf{x}_{t+1} | \mathbf{x}_t, \mathbf{y}_{1:t}, \theta) = N(\bar{\mathbf{x}}_{t+1}, \hat{\mathbf{C}}_{t+1}), \quad (8)$$

where θ represents all the unknown parameters in vector form. Based on this assumption, the daily latitude mean level and standard deviation can be applied as prior input knowledge. Then, the Kalman filter is used as shown below:

$$\mathbf{v}_t = \mathbf{y}_t - \mathbf{F}_t \mathbf{G}_t \bar{\mathbf{x}}_{t-1}, \quad (9)$$

$$\hat{\mathbf{C}}_t = \mathbf{G}_t \bar{\mathbf{C}}_{t-1} \mathbf{G}_t^T + \mathbf{W}_t, \quad (10)$$

$$\mathbf{C}_t^y = \mathbf{F}_t \hat{\mathbf{C}}_t \mathbf{F}_t^T + \mathbf{V}_t, \quad (11)$$

$$\mathbf{K}_t = \hat{\mathbf{C}}_t \mathbf{F}_t^T (\mathbf{C}_t^y)^{-1}, \quad (12)$$

$$\bar{\mathbf{x}}_t = \mathbf{G}_t \bar{\mathbf{x}}_{t-1} + \mathbf{K}_t \mathbf{v}_t, \quad (13)$$

$$\bar{\mathbf{C}}_t = \hat{\mathbf{C}}_t - \mathbf{K}_t \mathbf{F}_t \hat{\mathbf{C}}_t. \quad (14)$$

In (9), we first calculate the residual between the predicted value and our prediction. Then, we obtain the Kalman gain in 12 via prior and posterior covariance calculated in 10 and 11. With the Kalman gain, we are able to get the next state's prior mean and covariance of \mathbf{x} . Then, we apply

the Kalman smoother backward to get the smoothed land deformation trend

$$\begin{aligned} \mathbf{L}_t &= \mathbf{G}_t - \mathbf{G}_t \mathbf{K}_t \mathbf{F}_t^T, \\ \mathbf{r}_t &= \mathbf{F}_t^T (\mathbf{C}_t^y)^{-1} \mathbf{v}_t + \mathbf{L}_t^T \mathbf{r}_{t+1}, \\ \mathbf{N}_t &= \mathbf{F}_t^T (\mathbf{C}_t^y)^{-1} \mathbf{F}_t + \mathbf{L}_t^T \mathbf{N}_{t+1} \mathbf{L}_t, \\ \tilde{\mathbf{x}}_t &= \mathbf{G}_t \bar{\mathbf{x}}_{t-1} + \hat{\mathbf{C}}_t \mathbf{r}_t, \\ \tilde{\mathbf{C}}_t &= \hat{\mathbf{C}}_t - \hat{\mathbf{C}}_t \mathbf{N}_t \hat{\mathbf{C}}_t, \end{aligned} \quad (15)$$

where \mathbf{r}_t is the covariance of the observation noise. $\tilde{\mathbf{x}}_t$ is the smoothed latitude trend, and $\tilde{\mathbf{C}}_t$ is the smoothed state covariance.

2.3. LSTM Recurrent Networks. Time series prediction is about taking some previous input terms, putting them through some hidden units, and predicting the next term. In the past, computer scientists have favored the hidden Markov model (HMM). As the sequence grows larger and larger, the bits of information HMM needs to remember increase largely. Therefore, recurrent neural networks (RNN) have been developed to substitute HMM by storing more prior knowledge in a distributed way. With several different inputs remembered at once, RNN is able to predict more complicated dynamics.

In our paper, we developed an LSTM RNN for land deformation trend prediction. LSTM, so-called long-short time memory units, is a special type of RNN neural. As Figure 2 shows, in each LSTM unit, three gates are designed. The gates serve to help LSTM interact between the memory cell itself and its environment. The input gate can allow an incoming signal to update the state of the memory cell or block it. At the other end, the output gate can allow the state of the memory cell to have an effect on other neurons or prevent it. Finally, the forget gate can modulate the memory cell's self-recurrent connection, allowing the cell to remember or forget its previous state, as needed. The whole network contains four LSTM memory cells. After the time sequence passes these units, the outputs will be pooled by mean pooling units to get the predicted value.

3. Case Studies

With the aim at illustrating the effectiveness of the proposed DLM-LSTM framework, one land surface deformation case and one static case are studied in the following subsections.

3.1. Case 1: Fushun Western Open-Pit Coal Mine

3.1.1. Dataset. We deployed a wireless sensor network (WSN) on the south slope near the fault zone. The yellow mark in Figure 3 indicates the position of our WSN, which is $41^\circ 49' 54.36''$ north latitude and $123^\circ 52' 00.31''$ east longitude. An inexpensive system including hardware and software was designed for monitoring the southern slope surface displacement of the FWOCM. The system consists of two parts, a WSN and a remote server. The sensor nodes of the WSN transmit the data to the remote server as soon

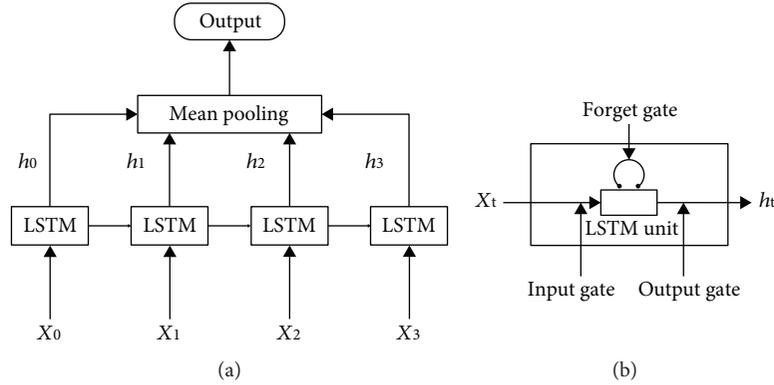


FIGURE 2: Applied LSTM recurrent network: (a) network structure and (b) LSTM unit structure.



FIGURE 3: The wireless sensor network (WSN) position on the southern slope of Fushun Western Open-Pit Coal Mine (FWOCM).

as the sensing data are received. The received data in the server are stored, processed, and published on the web. No processing is running on the sensor nodes due to the lack of electricity in the field.

The WSN with a star topology was composed of two kinds of nodes, sensor nodes and a coordinator. All the sensor nodes send their data to the coordinator, and the coordinator forwards the data to the server through a connected General Packet Radio Service (GPRS) module. Both the sensor node and the coordinator were developed using the CC2530 ZigBee development board. For sensor nodes, the sensor was connected to the board through peripherals, such as GPIO or UART. The sensors included a 3-dimensional digital compass (DCM308), GPS (Ublox NEO-6M), and low-cost inertial measurement unit (IMU), soil humidity, soil temperature, atmospheric pressure, atmospheric temperature, atmospheric humidity, and a rain gauge. The communication protocol for data transmission between multiple sensor nodes and coordinators was developed based on Z-stack (TI's API for low-cost ZigBee communication). The protocol adapts to the solar-powered network. The sensor node can find the coordinator and establish a communication link automatically. Figure 3 shows the position of the WSN. The WSN has been working well since December 1, 2014.

Figure 4 shows the star topology on the left and the layout of three GPS+IMU sensor nodes on the right. The three nodes were located near $N41^{\circ}50'$, $E123^{\circ}52'$. GPS sensors were positioned in a line on the north side, while another one was placed on the south side. The horizontal position accuracy of a stand-alone GPS receiver (L1 frequency, C/A code) is 2.5 m. As the sensor data accumulate, long-term land motion

calculations based on current and historical data become practical and feasible. In our paper, we selected the north-side GPS sensor as our data source and collected latitude data from February 1 to March 2, 2016. The latitudes are numbered according to the interval between the date of the latitude and February 1, which means the latitudes on March 1 are numbered as day 30.

Figure 5 shows the latitude samples collected each day. As shown in the figure, there is a significant discontinuity in the latitude time series due to the solar power. The latitudes from February 1 to February 6 are all missing, while latitudes from February 6 to February 20 are lacking in quantity. Under the assumption that the latitude changes daily, the time series were composed of daily mean latitude. Our DLM-LSTM framework will be proved effective to extract land deformation trends in submeter accuracy from the time series.

3.1.2. Results. We first calculated the mean level and standard deviation each day to produce a mean latitude level time series. Then, we chose time series from day 6 to day 31 as the input to our DLM model. The mean and variance of latitude within a day are used as input of DLM, which tackles the discontinuous latitude measurements in a day. For days that measurements are totally unavailable, we use linear interpolation of near days as input. Figure 6 shows the result of the DLM model, seasonal-trend decomposition procedure based on Loess (STL), and traditional Kalman filter in land deformation trend extraction. As shown in Figure 6, the DLM (full line) extracts a linearized trend with lesser oscillation than STL, while overfitting may occur in the Kalman filter. In addition, we constructed an estimated trend from the daily mean latitude points, and the trend of latitudes falls in the 95% confidential interval. The 95% probability envelope contains all the mean latitude-level points in the figure, which shows that DLM has a good understanding of the changes in latitude by deformation. Figure 7 shows the result of diagnostic analysis on the residuals by plotting an estimated autocorrelation function and normal probability. In the autocorrelation function (ACF) line, a dashed horizontal line represents the approximative region where the coefficient does not significantly deviate from zero. Figures 6 and 7 validate the extracted trend via estimating autocorrelation functions of DLM residuals. After latitude degrees are

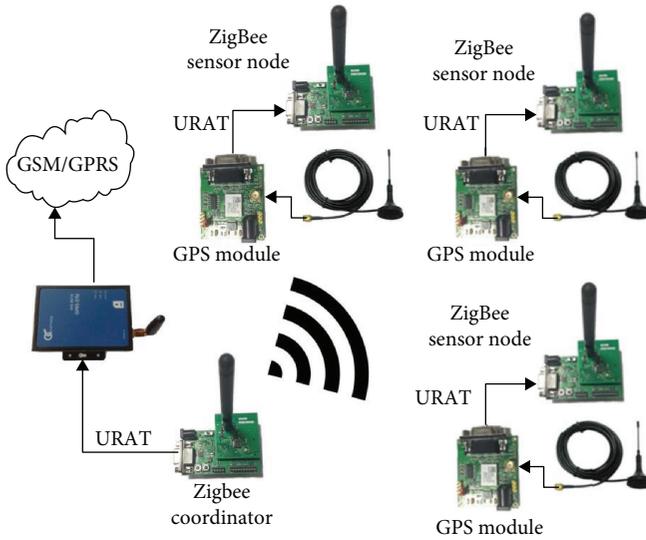


FIGURE 4: Layout of GPS + inertial measurement unit (IMU) sensor nodes in FWOCM.

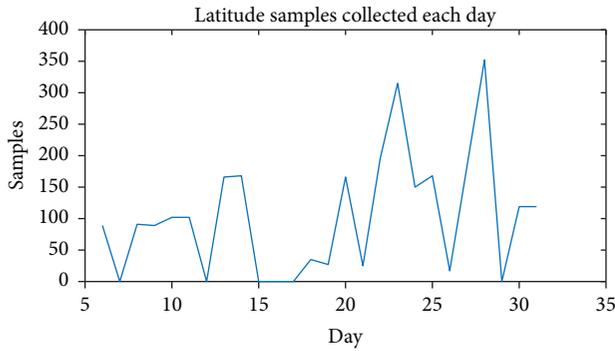


FIGURE 5: Latitude samples between February 1 and March 2, 2016.

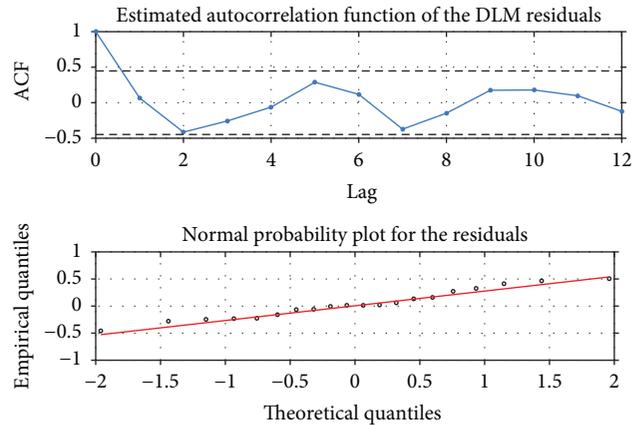


FIGURE 7: Analysis of the DLM residuals.

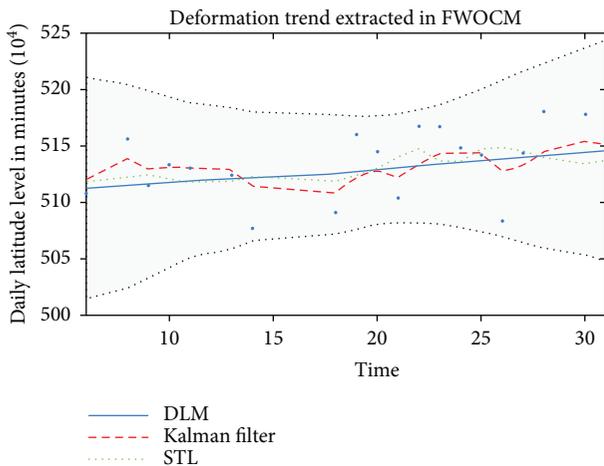


FIGURE 6: The result of deformation trend extraction. The dotted line represents the result of Kalman filter, the full blue line is the DLM results, and the dotted green line is the STL results.

transferred to submeters, we are able to compare the extracted displacement with ground truth, which is the professional measurement validated by local authorities.

The difference between the ground truth and the trends extracted by DLM, STL, and Kalman filter, respectively, is given in Table 1 by RMSE, MAPE, and R^2 . RMSE and R^2 are widely used in statistics. MAPE is the most common measure of forecast error. MAPE functions best when there are no extremes to the data (including zeros). According to the table, the computation results show consistent results with Figure 6 that DLM has better performance in characterizing land deformation trends, especially with lower error rates. Figure 8 shows the comparison of the extracted deformation trend with the ground truth. The ground truth comes from daily professional manual measurement of the displacement in the same period at the south slope of FWOCM. From Figure 8, we can see that our extracted trend shows a consistency with the manual measurement that was recognized by local authorities.

The results in Table 1 clearly indicate that the extraction performance of DLM is encouraging. The error between extracted trend and ground truth remains at the submeter level. As our GPS latitude sensors have a 2.5 m horizontal

TABLE 1: Performance of trend extraction methods.

Method	RMSE (mm)	MAPE	R^2
DLM	17.2129	0.0930	0.9903
Kalman	35.5304	0.3113	0.6994
STL	135.2132	0.4219	0.3907

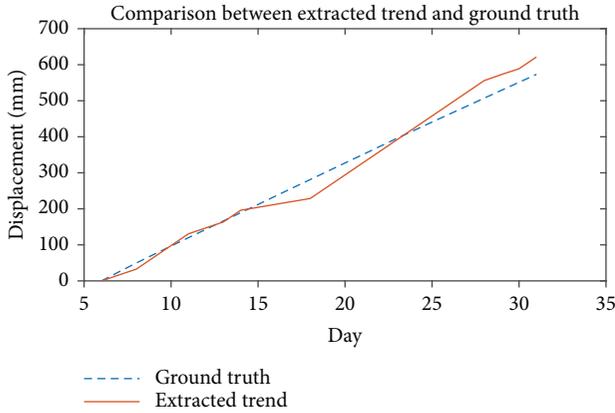


FIGURE 8: Comparison of the extracted deformation trend with the manual measurement displacement.

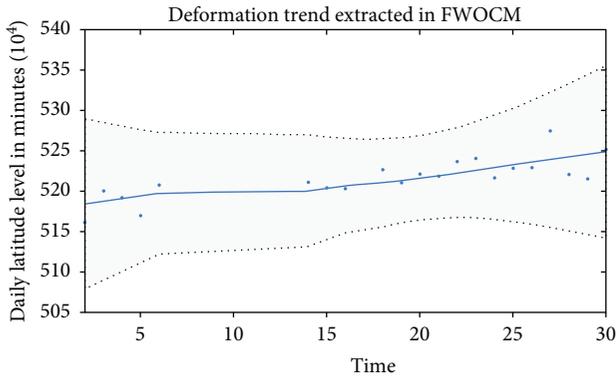


FIGURE 9: The result of deformation trend extraction. The dotted line represents the result of original Kalman filter, the full blue line is the DLM results.

TABLE 2: Parameters of LSTM networks.

Loss function	Optimizer	Batch size	Epoch
Mean square error	SGD	3	1500

navigation error, the result of DLM is a big step in accuracy improvement. Figure 9 shows the trend results extracted by the DLM model from April 2 to 30 2016. From Figures 6–9, DLM is proved to be effective for trend extraction from noisy and discontinuous time series of low-cost GPS sensors.

Because DLM’s performance has been validated, the trend can be used as input for our four-unit LSTM networks. We selected the trend from day 1 to day 26 as our training sample and tried to predict the following 13 days’ trend by the LSTM network. The start point of the break period was predicted by LSTM, and then this prediction was used as

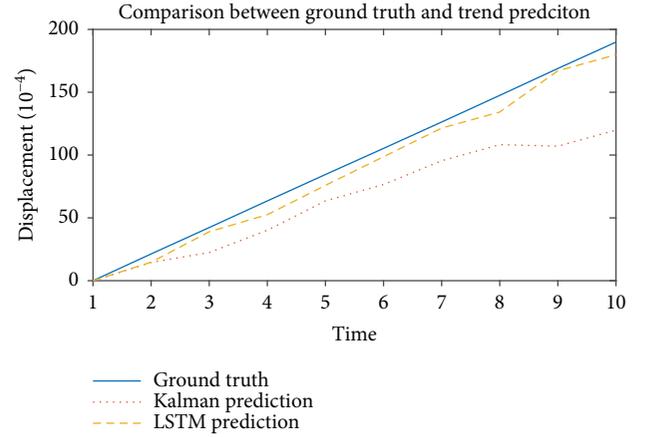


FIGURE 10: Displacement prediction by LSTM and Kalman filter.

TABLE 3: Performance of LSTM networks compared to Kalman filter.

Method	RMSE (mm)	MAPE	R^2
LSTM	7.7194	0.0917	0.9664
Kalman	36.7102	0.2920	0.7098

input until the validated input was available. The LSTM network parameters are shown in Table 2. The length of 1D mean pooling is 4 according to Figure 2. We use stochastic gradient descent with mini batches to optimize the cost function of the LSTM network. Mini-batch gradient descent is a variation of the gradient descent algorithm that splits the training dataset into small batches that are used to calculate model loss and update model coefficients. In our paper, we set the three training samples in a batch. In addition, we use the Kalman filter to get the dynamic function of the previous 26 days and to predict the next 13 days by this function. Figure 10 shows the prediction result of LSTM and the Kalman filter compared to ground truth measurement. Our LSTM network predicts 13 days’ deformation trend with 7.7194 mm RMSE, as shown in Table 3. Compared to Kalman filter’s 36.7102 mm, this proves that the LSTM network has an accuracy improvement effect in trend prediction. Future works may focus on improving the LSTM network’s performance.

3.2. Case 2: Baoxie Sensor Web Experimental Field

3.2.1. Dataset. To prove that our DLM has a locational invariant propriety, we need to find another place to conduct the validation experiment. The Baoxie sensor web experiment field, located in the east of the central Wuhan area, is the main Sensor Web research and experiment center for Wuhan University’s State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS). It contains four stations: Baoxie Landslide Monitoring Station, Baoxie Meteorological Experimental Station, Baoxie Edaphic and Meteorological Monitoring Station, and Baoxie Soil Temperature and Moisture Monitoring Station. These stations, in total, are about 800 m² with more



FIGURE 11: Baoxie Meteorological Experimental Station.

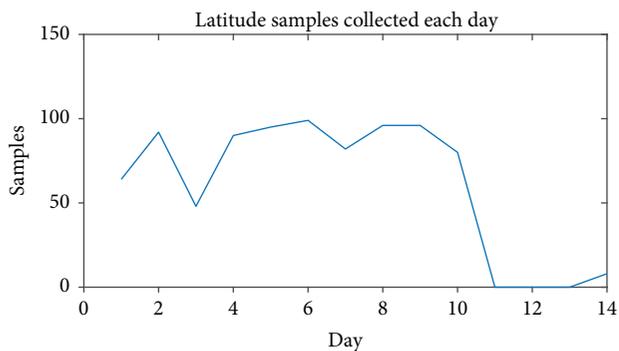


FIGURE 12: Latitude samples at Baoxie, Wuhan, China.

than 70 sensors. All of the sensors are powered by photovoltaic solar energy. Previously, LIESMARS successfully built a cyber-physical geographical information paradigm for in situ land surface monitoring [28], which provided a solid basis for our experiment. As Figure 11 implies, we selected a GPS sensor in a wireless sensor node deployed since 2013 as our experimental data source. The position of this GPS sensor is $N30^{\circ}28^{\prime}$, $E114^{\circ}31^{\prime}$. The horizontal position accuracy of the GPS receiver (L1 frequency, C/A code) is the same as the GPS receiver at FWOCM, which is 2.5 m.

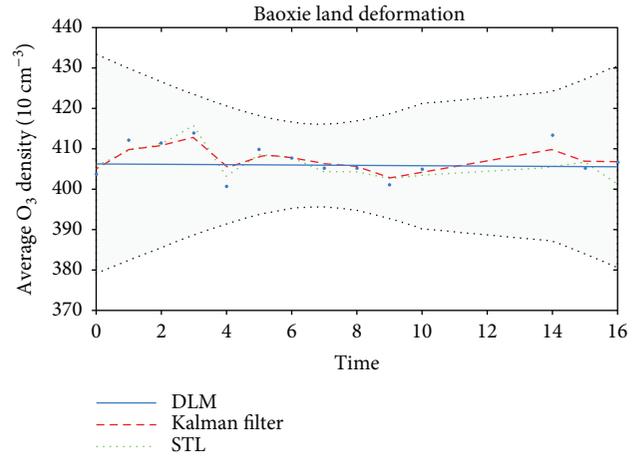


FIGURE 13: Trend extraction by DLM at Baoxie. The dotted line represents the result of the Kalman filter, the full blue line is the DLM results, and the dotted green line is the STL results.

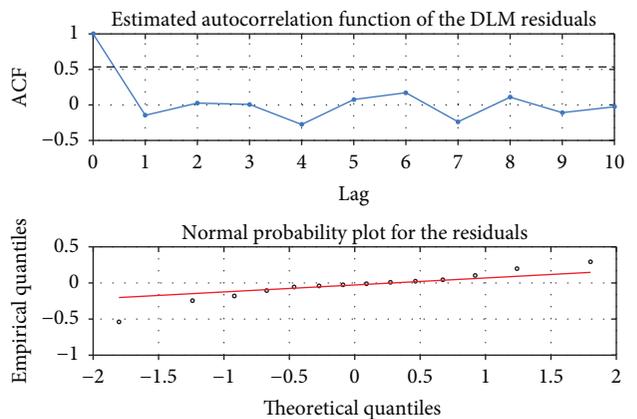


FIGURE 14: Residual analysis by DLM at Baoxie.

TABLE 4: Performance of trend extraction at static site.

Method	RMSE (mm)	MAPE	R^2
DLM	2.3542	0.0771	0.9891
Kalman	7.7823	0.4522	0.7709
STL	544.9739	0.9690	0.5004

We collected 14 days of latitude data from December 1 to December 14. Figure 12 shows the latitude samples received each day. It is clear that the GPS sensor in Baoxie is more stable than FWOCM, which has larger deviation in sample numbers each day. However, the Baoxie sensor suffers discontinuity, as does the FWOCM. Also, in situ station measurement shows that the location $N30^{\circ}28^{\prime}$, $E114^{\circ}31^{\prime}$ remains static with no submeter-level deformation in the same period, which provides the ground truth.

3.2.2. Results. Figures 13 and 14 show the extracted trend and residual analysis of GPS sensor from the Baoxie sensor web experimental field. From Figures 13 and 14, we can see that an obvious static trend is obtained by DLM. For further statistical tests, because the ground truth is recognized as zero,

only RMSE can be calculated in Baoxie. The RMSE shown in Table 4 is 2.3542 mm, which obviously proves DLM's capability in extracting a trend from low-cost GPS sensors.

4. Conclusions

A DLM-LSTM framework is proposed and applied for north-south deformation trend analysis from low-cost GPS time series. A dynamic linear model is introduced to model trend, noise, periodic variation, and autoregressive components hidden in a latitude sequence. Then, the land deformation trend is extracted using a recursive Kalman filter and smoother. LSTM recurrent neural networks are introduced to predict the next states in deformation trend. By analyzing two land deformation cases, the performance of the proposed method is evaluated quantitatively. Calculated displacement in the extracted trend from DLM is excellently consistent with manual measurement from the latitude series of FWOCM. The submeter-level accuracy can be obtained in a north-south deformation trend, which is difficult to obtain by satellite remote sensing approaches. Meanwhile, submeter-level accuracy is also achieved in trend prediction by LSTM networks. The results of this work indicate that the DLM-LSTM framework is a powerful and accurate end-to-end method to analyze land surface deformation trends.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key Research and Development Program of China under Grant no. 2016YFB0502601. It was also supported by the Fundamental Research Funds for the Central Universities under Grant no. 2042017kf0211.

References

- [1] D. Petley, "Global patterns of loss of life from landslides," *Geology*, vol. 40, no. 10, pp. 927–930, 2012.
- [2] B. Thiebes, *Landslide Analysis and Early Warning Systems: Local and Regional Case Study in the Swabian Alb, Germany*, [Ph.D. thesis], Springer Science & Business Media, 2012.
- [3] D. Zheng, J. D. Frost, R. Q. Huang, and F. Z. Liu, "Failure process and modes of rockfall induced by underground mining: a case study of kaiyang phosphorite mine rockfalls," *Engineering Geology*, vol. 197, pp. 145–157, 2015.
- [4] G. E. Hilley, R. Burgmann, A. Ferretti, F. Novali, and F. Rocca, "Dynamics of slow-moving landslides from permanent scatterer analysis," *Science*, vol. 304, no. 5679, pp. 1952–1955, 2004.
- [5] R. Macciotta, M. Hendry, and C. D. Martin, "Developing an early warning system for a very slow landslide based on displacement monitoring," *Natural Hazards*, vol. 81, no. 2, pp. 887–907, 2016.
- [6] L. Nie, Z. Li, M. Zhang, and L. Xu, "Deformation characteristics and mechanism of the landslide in west open-pit mine, Fushun, China," *Arabian Journal of Geosciences*, vol. 8, no. 7, pp. 4457–4468, 2015.
- [7] E. A. Johnson, "Geology of the Fushun coalfield, Liaoning province, people's republic of China," *International Journal of Coal Geology*, vol. 14, no. 3, pp. 217–236, 1990.
- [8] T. Strozzi, P. Farina, A. Corsini et al., "Survey and monitoring of landslide displacements by means of l-band satellite SAR interferometry," *Landslides*, vol. 2, no. 3, pp. 193–201, 2005.
- [9] M. Akbarimehr, M. Motagh, and M. Haghshenas-Haghighi, "Slope stability assessment of the Sarcheshmeh landslide, Northeast Iran, investigated using InSAR and GPS observations," *Remote Sensing*, vol. 5, no. 8, pp. 3681–3700, 2013.
- [10] S. C. Oliveira, J. L. Zêzere, J. Catalão, and G. Nico, "The contribution of PSInSAR interferometry to landslide hazard in weak rock-dominated areas," *Landslides*, vol. 12, no. 4, pp. 703–719, 2015.
- [11] G. Nico, L. Borrelli, A. Di Pasquale, L. Antronico, and G. Gulla, "Monitoring of an ancient landslide phenomenon by GBSAR technique in the Maierato town (Calabria, Italy)," in *Engineering Geology for Society and Territory-Volume 2*, pp. 129–133, Springer, 2015.
- [12] V. Tofani, F. Raspini, F. Catani, and N. Casagli, "Persistent scatterer interferometry (psi) technique for landslide characterization and monitoring," *Remote Sensing*, vol. 5, no. 3, pp. 1045–1065, 2013.
- [13] F. Raspini, F. Cigna, and S. Moretti, "Multi-temporal mapping of land subsidence at basin scale exploiting persistent scatterer interferometry: case study of gioia tauro plain (Italy)," *Journal of Maps*, vol. 8, no. 4, pp. 514–524, 2012.
- [14] L. He, L. Wu, S. Liu, Z. Wang, C. Su, and S.-N. Liu, "Mapping two-dimensional deformation field time-series of large slope by coupling DInSAR-SBAS with MAI-SBAS," *Remote Sensing*, vol. 7, no. 9, pp. 12440–12458, 2015.
- [15] T.-S. Bae, D. Grejner-Brzezinska, G. Mader, and M. Dennis, "Robust analysis of network-based real-time kinematic for GNSS-derived heights," *Sensors*, vol. 15, no. 10, pp. 27215–27229, 2015.
- [16] E. Trinklein and G. Parker, "Combining multiple GPS receivers to enhance relative distance measurements," in *2013 IEEE Sensors Applications Symposium Proceedings*, pp. 33–37, Galveston, TX, USA, 2013.
- [17] D. DeVon, T. Holzer, and S. Sarkani, "Minimizing uncertainty and improving accuracy when fusing multiple stationary GPS receivers," in *2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pp. 83–88, San Diego, CA, USA, 2015.
- [18] H. Jo, S.-H. Sim, A. Tatkowski, B. F. Spencer Jr, and M. E. Nelson, "Feasibility of displacement monitoring using low-cost GPS receivers," *Structural Control and Health Monitoring*, vol. 20, no. 9, pp. 1240–1254, 2013.
- [19] J. Gomez-Gil, R. Ruiz-Gonzalez, S. Alonso-Garcia, and F. Gomez-Gil, "A Kalman filter implementation for precision improvement in low-cost GPS positioning of tractors," *Sensors*, vol. 13, no. 11, pp. 15307–15323, 2013.
- [20] J.-x. Gao, H. Hu, F. Liu, Z.-k. Li, and Y.-f. Yao, "Signal extraction for GPS deformation monitoring in mining survey,"

- Transactions of Nonferrous Metals Society of China*, vol. 24, no. 12, pp. 3949–3954, 2014.
- [21] J. Zhang, W. Peng, F. Liu, H. Zhang, and Z. Li, “Monitoring rock failure processes using the Hilbert–Huang transform of acoustic emission signals,” *Rock Mechanics and Rock Engineering*, vol. 49, no. 2, pp. 427–442, 2016.
- [22] F. Pu, J. Ma, D. Zeng, X. Xu, and N. Chen, “Early warning of abrupt displacement change at the Yemaomian landslide of the Three Gorge Region, China,” *Natural Hazards Review*, vol. 16, no. 4, 2015.
- [23] Z. C. Lipton, D. C. Kale, C. Elkan, and R. Wetzell, “Learning to diagnose with LSTM recurrent neural networks,” 2015, <http://arxiv.org/abs/1511.03677>.
- [24] J. A. Pérez-Ortiz, F. A. Gers, D. Eck, and J. Schmidhuber, “Kalman filters improve LSTM network performance in problems unsolvable by traditional recurrent nets,” *Neural Networks*, vol. 16, no. 2, pp. 241–250, 2003.
- [25] M. Laine, N. Latva-Pukkila, and E. Kyrölä, “Analysing time-varying trends in stratospheric ozone time series using the state space approach,” *Atmospheric Chemistry and Physics*, vol. 14, no. 18, pp. 9707–9725, 2014.
- [26] X. Li, Q. Liu, R. Yang et al., “The combination of ground-sensing network and satellite remote sensing in Huailai county,” *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3819–3826, 2016.
- [27] R. H. Shumway and D. S. Stoffer, “State-space models,” in *Time series analysis and its applications*, pp. 319–404, Springer, 2011.
- [28] N. Chen, C. Xiao, F. Pu et al., “Cyber-physical geographical information service-enabled control of diverse in-situ sensors,” *Sensors*, vol. 15, no. 2, pp. 2565–2592, 2015.

Research Article

Wetland Change Detection Using Cross-Fused-Based and Normalized Difference Index Analysis on Multitemporal Landsat 8 OLI

Yan Gao , Zeyu Liang , Biao Wang , Yanlan Wu , and Penghai Wu 

School of Resources and Environmental Engineering, Anhui University, Hefei, Anhui 230601, China

Correspondence should be addressed to Biao Wang; wangbiao-rs@ahu.edu.cn

Received 30 November 2017; Revised 8 April 2018; Accepted 30 April 2018; Published 21 May 2018

Academic Editor: Biswajeet Pradhan

Copyright © 2018 Yan Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wetlands are one of the most important ecosystems on the Earth and play a critical role in regulating regional climate, preventing floods, and reducing flood severity. However, it is difficult to detect wetland changes in multitemporal Landsat 8 OLI satellite images due to the mixed composition of vegetation, soil, and water. The main objective of this study is to quantify change to wetland cover by an image-to-image comparison change detection method based on the image fusion of multitemporal images. Spectral distortion is regarded as candidate change information, which is generated by the spectral and spatial differences between multitemporal images during the process of image cross-fusion. Meanwhile, the normalized difference vegetation index (NDVI) and normalized difference water index (NDWI) were extracted from the cross-fused image as a normalized index image to enhance and increase the information about vegetation and water. Then, the modified iteratively reweighted multivariate alteration detection (IR-MAD) is applied to the generally fused images and normalized difference index images, providing a good evaluation of spectral distortion. The experimental results show that the proposed method performed better to reduce the detection errors due to the complicated areas under different ground types, especially in cultivated areas and forests. Moreover, the proposed method was tested and quantitatively assessed and achieved an overall accuracy of 96.67% and 93.06% for the interannual and seasonal datasets, respectively. Our method can be a tool to monitor changes in wetlands and provide effective technical support for wetland conservation.

1. Introduction

Wetlands are a unique ecosystem formed by the interaction between water and land, and they cover 6% of the Earth's surface [1]. Due to seasonal changes, the characteristics of wetlands vary among water, soil, and vegetation. This makes the wetland landscape more complex, and it becomes more difficult to extract information about changes in these regions. In addition, the combination of reflectance spectra of the underlying soil, the hydrologic regime, and atmospheric vapor makes optical classification more difficult, and these factors could introduce a reduction in spectral reflectance. Therefore, it is often difficult to achieve the expected results using a single method to extract information about wetland change [2].

Postclassification comparison (PCC), in which two multitemporal images are independently classified and then compared [3], is one of the methods used for wetland change detection. First, it is applied for detecting the trajectories of corresponding wetland cover types. More specifically, it includes many classification methods, such as the regression tree algorithm or the maximum likelihood classification [4, 5]. However, in some of these methods, high-accuracy classification and ground truth information are required [3–6].

The image-to-image (or direct) comparison change detection method, which is another method for wetland change detection, is used to obtain a difference image of spectral changes through the analysis and calculation of the spectral characteristics from multitemporal images, and then a binary image is generated in which the change areas are dis-

tinguished from unchanged areas [7]. The advantages of this method are first that it provides a faster comparison of images and second that it demands no ground truth information; however, it could not display the change trajectories of wetland cover types [8]. Change detection methods such as change vector analysis (CVA) [9], principal component analysis (PCA) [10], Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) [11], and multivariate alteration detection (MAD) [12] directly calculate multitemporal images.

However, the results of the change detection method based on the difference image largely depend on the spectral characteristics and may include some false positives. For this reason, a change detection method based on cross-fusion and spectral distortion is proposed to improve the accuracy of change detection in flood zones [13]. Successful change detection results have been achieved in coal-mining subsidence areas; nevertheless, there are still some false detections in wetland areas [14].

To mitigate the false positive results, we employ an image-to-image method through NDVI and NDWI extraction based on a cross-fusion image to detect the change information of wetlands. In the case of wetlands, vegetation, soil, and water coexist, and the application of the cross-fusion method is beneficial for improving spatial resolution and enhancing information about wetland change. In addition, based on the cross-fused image, NDVI and NDWI are extracted in order to enhance the information about vegetation and water. We can then derive the change information using the modified IR-MAD algorithm, which is a well-established change detection method for multitemporal multispectral images [12]. Finally, the change area of wetlands is obtained using the automated threshold method.

2. Study Area and Dataset

In this study, we collected three Landsat 8 OLI multitemporal images covering the Shengjin Lake Nature Reserve area. These images represent the seasonal (25 July 2016 and 16 December 2016) and interannual (6 November 2013 and 16 December 2016) changes in the types of land cover, as shown in Figures 1(a)–1(c). In the preprocessing, the relevant bands were selected, including the 30-meter-resolution MS (multispectral) bands 2–7 and the 15-meter-resolution PAN (panchromatic) band 8. Then, the Shengjin Lake-protected area vector data were used to cut the images, with the specific parameters shown in Table 1.

The Shengjin Lake National Nature Reserve (30°15′~30°30′N, 116°55′~117°15′E) is located in Chizhou City, Anhui Province. The protected region, with a total area of 333.40 km², consists of Shengjin Lake, cultivated areas, urban areas, forest, and bare land. The Shengjin Lake wetland ecological environment is well-preserved, with rich natural and cultural landscapes. It is one of the most complete areas in the wetland ecosystem of inland freshwater lakes in the lower reaches of the Yangtze River. It connects with the Yangtze River, and the water level of the lake is regulated by the Huangpen sluice. The location of the study area is shown in Figure 1(d). The water level of Shengjin Lake varies between

3.4 and 7.4 m due to the sluice. Water level changes lead to the largest lake area in the summer wet season and a smaller area in the winter dry season. During the dry season, two largest *Carex* meadows (“upper lake meadow” and “lower lake meadow”) provide suitable living environments and food sources for Greater White-fronted Geese and Bean Geese, and it has become a critical winter habitat for rare birds [15, 16].

3. Methodology

In this section, we detail the process of extracting wetland change information from bitemporal images using a modified IR-MAD. We consider the two datasets $F_{t_1}^H$ and $F_{t_2}^H$ consisting of high-resolution PAN and low-resolution MS images acquired in the same geographical area at different times t_1 and t_2 , respectively. The process flow is shown in Figure 2, and the details of each step are described below.

3.1. Cross-Fused Image Generation. Generally, the image fusion method fuses high-resolution PAN images and low-resolution MS images into a high-resolution MS image. In this paper, the PAN and MS images are generally fused using Gram-Schmidt adaptive (GSA) to produce high-resolution multispectral $F_{t_1}^H$ and $F_{t_2}^H$ images. The GSA algorithm is applied as a representative component substitution- (CS-) based fusion algorithm, and there is no limit to the number of bands of the fused image [17]. The major drawback of the CS-based fusion method is spectral distortion, also called color (or radiometric) distortion. This spectral distortion is caused by the mismatch between the spectral responses of the MS and PAN bands according to the different bandwidths [18]. In this study, the spectral distortion will be regarded as a candidate detection feature for wetland cover change [17]. To this end, we extract the high-resolution NIR image instead of high-resolution PAN image. The NIR band has a narrower bandwidth, and by using this band, the mismatch level of spectral response outside the NIR spectral range will increase, and the spectral distortion will become more pronounced [13]. At the same time, the NIR band is a very useful information source for detecting water or vegetation areas, because the water area appears dark since it has strong absorption characteristics, and the vegetation has the opposite characteristics and appears light. Thus, the NIR band is useful in extracting change information. Then, the cross-fused image CF_1^H is generated where the $F_{t_1}^H$ MS image is fused with the NIR band of $F_{t_2}^H$ by using the GSA image fusion algorithm. CF_2^H can be obtained in the same way as CF_1^H , and the formula is as follows:

$$CF_1^H = NIR_{F_{t_2}}^H + MS_{t_1}^L, \quad (1)$$

$$CF_2^H = NIR_{F_{t_1}}^H + MS_{t_2}^L. \quad (2)$$

3.2. NDVI and NDWI Extraction. The NDVI, which is a classic index used to monitor vegetation changes, is calculated from a normalized transform of the NIR and red reflectance ratio. Application of the NDVI strives to minimize the solar

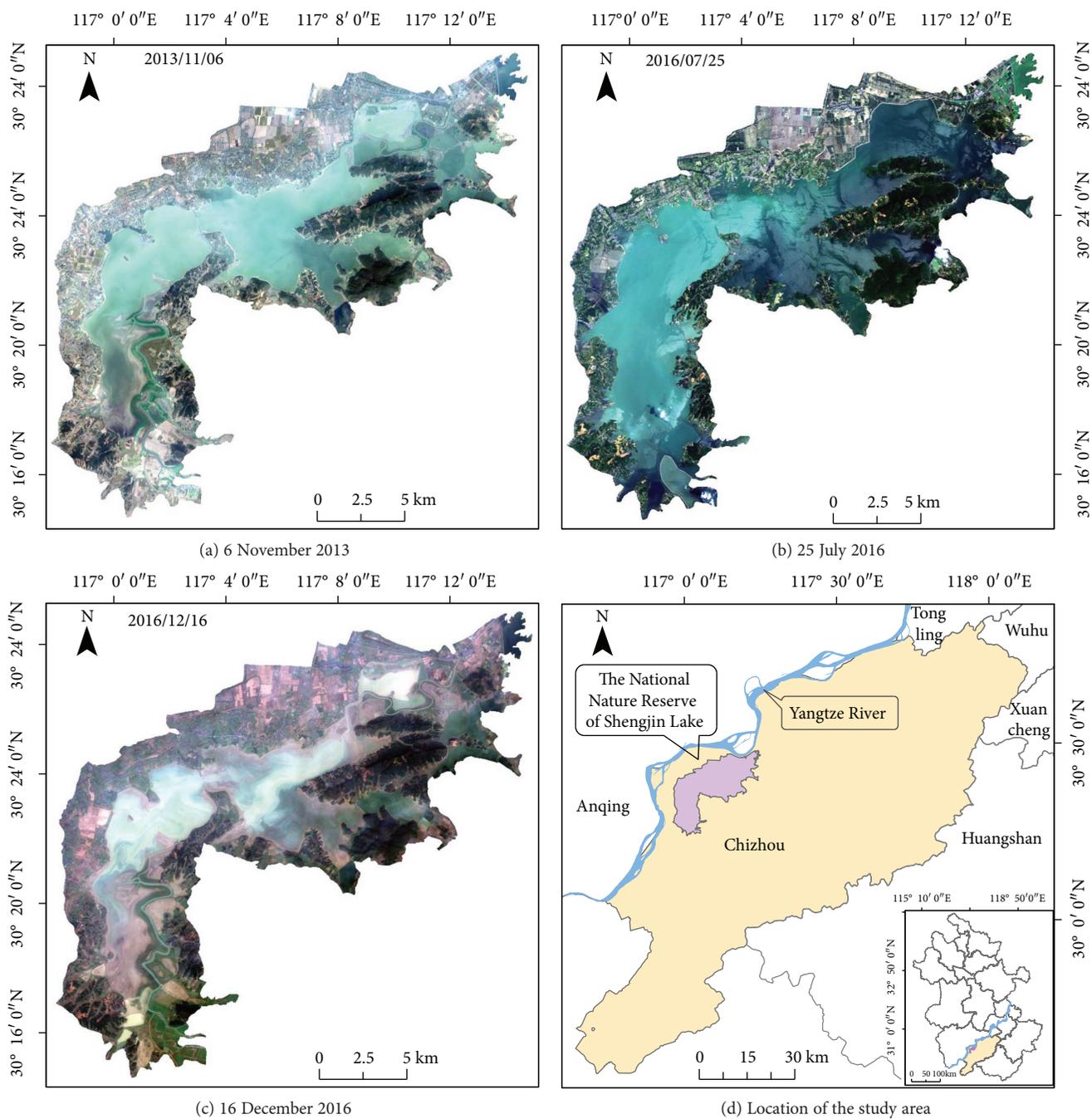


FIGURE 1: Multitemporal images of Shengjin Lake used in the experiment.

TABLE 1: Data specifications.

Date	Spatial resolution	Image size (pixels)	Wavelength (μm)
06/11/2013	MS: 30 m	MS: 3650 \times 3570	Band 2 blue: 0.450–0.515
25/07/2016			Band 3 green: 0.525–0.600
16/12/2016			Band 4 red: 0.630–0.680
	PAN: 15 m	PAN: 1825 \times 1785	Band 5 NIR: 0.845–0.885
			Band 6 SWIR: 11.560–1.660
			Band 7 SWIR: 12.100–2.300
			Band 8 pan: 0.500–0.680

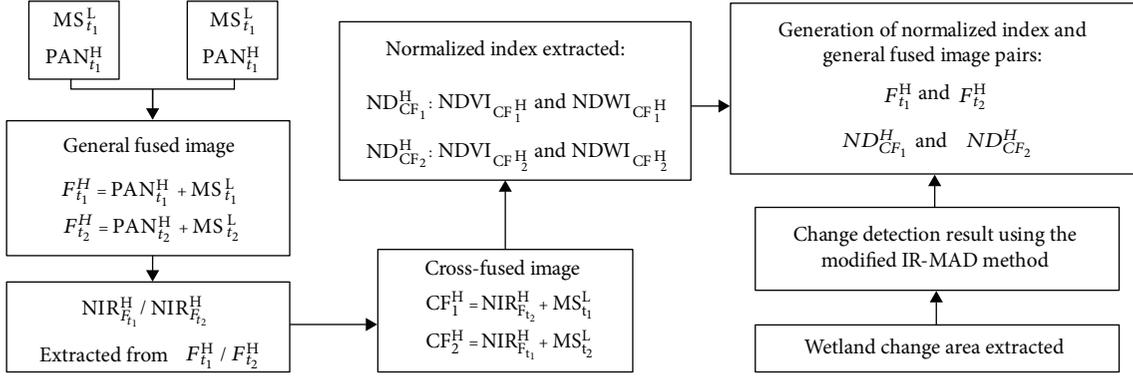


FIGURE 2: Workflow of the proposed methodology for wetland vegetation extraction using multitemporal satellite images. The superscript H and L represent high resolution and low resolution, respectively.

irradiance and soil background effects and improve the vegetation signal. Similarly, the NDWI is a commonly used remote sensing water monitoring index [19, 20].

The cross-fused image has high spatial and temporal resolution, and the vegetation and water details of the bitemporal images are preserved. By using the NDVI and the NDWI, the change information from water and vegetation is further optimized, and the distinction among water bodies, wet soil, and vegetation is improved. The spectral difference is enhanced twofold, and the detection sensitivity of the vegetation and water increases. The equation is as follows:

$$\text{NDVI}_F = \frac{\text{NIR}_F - R_F}{\text{NIR}_F + R_F}, \quad (3)$$

$$\text{NDWI}_F = \frac{\text{green}_F - \text{NIR}_F}{\text{green}_F + \text{NIR}_F}, \quad (4)$$

where NIR_F , R_F , and green_F are the NIR, red, and green bands of the cross-fused image, respectively. According to (3) and (4), the $\text{NDVI}_{CF_1}^H$ and $\text{NDVI}_{CF_2}^H$ images are generated from the cross-fused images CF_1^H and CF_2^H and include the NDVI and NDWI bands.

3.3. Wetland Change Area Extracted. The IR-MAD algorithm, which is based on canonical correlation analysis, considers two K band multispectral images R and T of the same area, but acquired at two different times, and it has an important application in the field of multitemporal multispectral image change detection. The random variables U and V , generated by any linear combination of the intensities of the spectral bands using coefficient matrices \mathbf{a} and \mathbf{b} , are defined. The equation is as follows [21].

$$\begin{aligned} U &= \mathbf{a}^T R, \\ V &= \mathbf{b}^T T, \end{aligned} \quad (5)$$

where the superscript T is the transpose of each matrix. The task is to find suitable vectors \mathbf{a} and \mathbf{b} by maximizing the variance of $U - V$. This leads to solving two generalized eigenvalue problems for \mathbf{a} and \mathbf{b} from a canonical matrix analysis.

The MAD variate M_k , which is generated by taking the paired difference between U and V , represents the changed information [22], where the equation is as follows (6).

$$\begin{aligned} M_k &= U_{K-k+1} - V_{K-k+1}, \\ k &= \mathbf{a}_{K-k+1}^T R - \mathbf{b}_{K-k+1}^T T, \\ K &= 1, 2, \dots, K. \end{aligned} \quad (6)$$

In this study, corresponding to the two multitemporal normalized difference index images ($\text{NDVI}_{CF_1}^H$ and $\text{NDVI}_{CF_2}^H$) and two general fused images ($F_{t_1}^H$ and $F_{t_2}^H$), MAD variate M_p is generated by using optimal coefficients \mathbf{a} and \mathbf{b} through (7) and (8):

$$M_1 = \mathbf{a}_{(1)}^T F_{t_1}^H - \mathbf{b}_{(1)}^T F_{t_2}^H, \quad (7)$$

$$M_2 = \mathbf{a}_{(2)}^T \text{ND}_{CF_1}^H - \mathbf{b}_{(2)}^T \text{ND}_{CF_2}^H, \quad (8)$$

where M_1 is the MAD variate of the generally fused image with the same temporal data. Meanwhile, two normalized index images ($\text{NDVI}_{CF_1}^H$ and $\text{NDVI}_{CF_2}^H$) in M_2 are extracted from the cross-fusion image.

The probability of the changed information for pixel j , which is calculated by the sum of the squares of the standardized MAD variate, is defined in (9).

$$Z_j = \sum_{k=1}^K \left(\frac{M_{kj}}{\sigma_{M_k}} \right)^2, \quad (9)$$

where variable Z_j represents a weight for the probability of changes in each pixel to identify a greater chi-square value, M_{kj} is the MAD variate of the k th band for pixel j , and σ_{M_k} is the variance of the no-change distribution. The above results can be regarded as the weights of the observations. The iteration process would continue for a number of fixed iterations or until there is no significant change in the canonical correlations. The latter is used in this study [22]. Then, the optimal matrices \mathbf{a} and \mathbf{b} are recalculated by the weight factor.

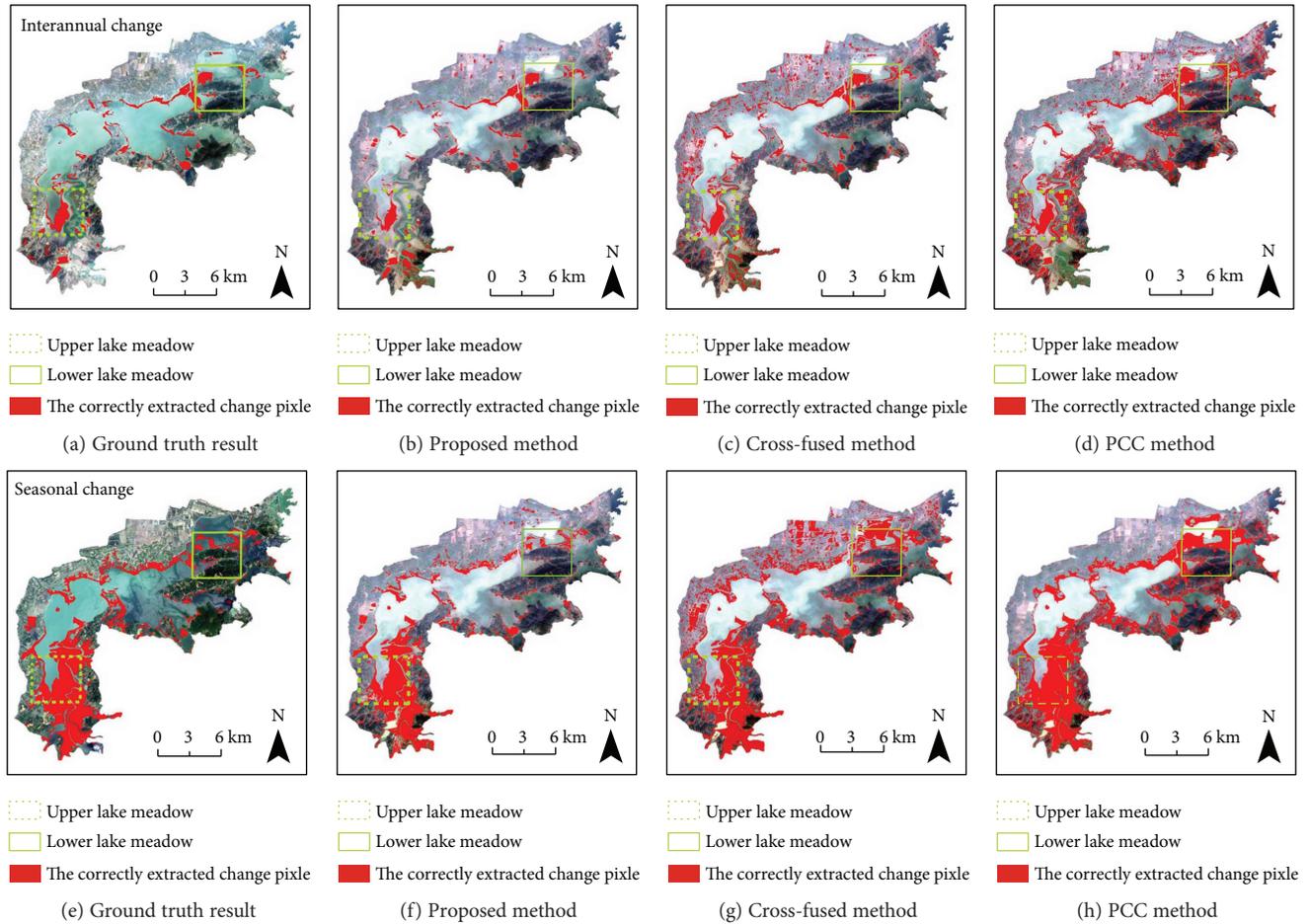


FIGURE 3: Results of wetland change area by using the tested methods: (a)–(d) interannual change detection result; (e)–(h) seasonal change detection result.

In this study, based on the combination of M_1 and M_2 , the final change detection index Z'_j is calculated.

$$Z'_j = \sum_{p=1}^2 \sum_{k=1}^K \left(\frac{M_{(p)kj}}{\sigma_{(p)M_k}} \right)^2. \quad (10)$$

$M_{(p)kj}$ is the k th band MAD variate of the p th pair of fused images for pixel j . This method can effectively reduce the falsely detected changes by considering Z'_j values twice in (10).

This modified IR-MAD algorithm can not only alleviate the problem of spectral distortion that caused massive false change alarms in the process of using bitemporal images to generate the cross-fused images but also reduce the interaction between bands of multispectral images [14, 21]. Therefore, this algorithm can yield better change detection results in multitemporal images. Finally, the Otsu thresholding algorithm, which is based on histogram image segmentation [23], has effective performance and easy application, and it was applied to the modified IR-MAD image to obtain binary data of the changed and the unchanged area [23].

4. Experimental Result and Discussion

To evaluate the effectiveness of this method, we analyze and discuss the seasonal and interannual variations in the study area. We use the cross-fused and PCC change detection methods to compare with our result. In the cross-fused change detection method, original IR-MAD is applied between CF_1^H and CF_2^H images to extract the changed area. In the process of PCC, three general fused images are classified into 7 classes (water, bare land, meadow, cultivated area, city, forest, and mudflats) through the maximum likelihood classification.

To quantitatively compare the performance of these methods, ground truths of seasonal and interannual variation images were generated from GSA-fused images ($F_{t_1}^H$ and $F_{t_2}^H$) by manually digitizing the changed areas of Shengjin Lake Nature Reserve as shown in red in Figures 3(a) and 3(e). The results are overlain on the multispectral images of 6 November 2013 and 25 July 2016, respectively. In the quantitative analysis process, the confusion matrix method was applied to evaluate the statistical accuracy of the tested methodologies, and some indices such as overall accuracy (OA), kappa coefficient (KC), commission error (CE), omission error (OE), and false alarm rate (FAR) were calculated

TABLE 2: Change detection accuracy results: overall accuracy (OA), kappa coefficient (KC), commission error (CE), omission error (OE), and false alarm rate (FAR).

		OA (%)	KC	CE (%)	OE (%)	FAR
Interannual	Proposed method	96.67	0.72	27.60	25.02	0.02
	Cross-fused method	92.07	0.52	58.97	15.02	0.07
	PCC	86.26	0.39	70.22	11.13	0.14
Seasonal	Proposed method	93.06	0.76	14.88	24.00	0.03
	Cross-fused method	85.28	0.59	43.06	13.93	0.15
	PCC	84.45	0.57	44.56	16.36	0.15

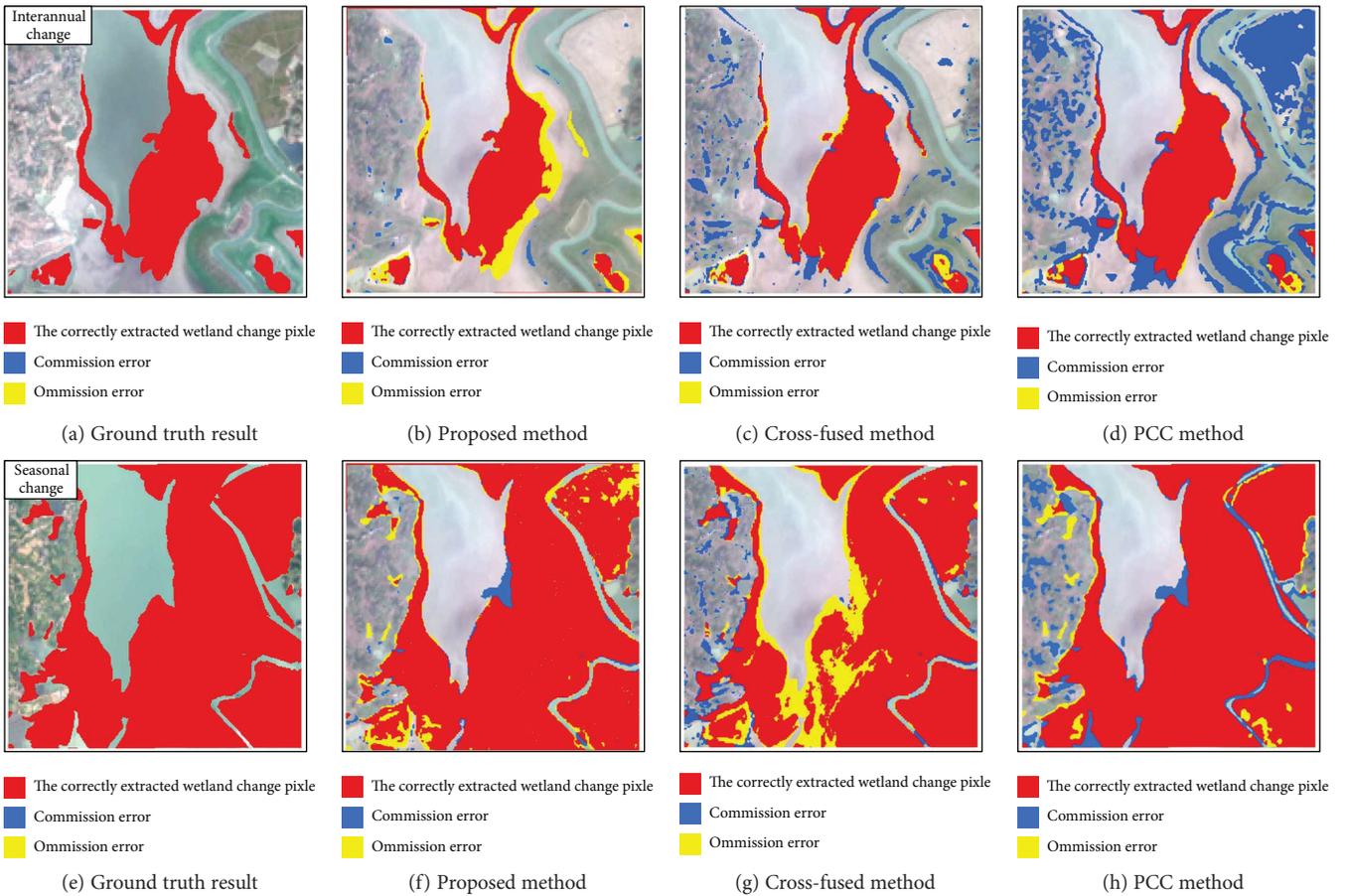


FIGURE 4: Detailed images from upper lake meadow regions of Figure 3: (a)–(d) interannual change detection result; (e)–(h) seasonal change detection result.

[24]. The detailed quantitative change detection accuracy assessment results for each method are shown in Figure 3 and Table 2. The red color indicates the change pixels extracted from the change detection results of the different methods, and the results are overlain on the multispectral image of 25 December 2016.

Through the observation and analysis of the results in Figure 3, in areas where different ground types coexist (water, bare land, meadow, cultivated area, city, forest, and mudflats), the proposed method, compared to PCC based on the general fused image and cross-fused method, can more accurately detect the wetland change information and

effectively reduce the change detection errors for interannual or seasonal wetland cover change. In the results of the PCC and cross-fused methods, some parts of the unchanged area are considered to be changed areas. As shown in Table 2, the CE value of PCC reaches 70%. In addition, the results of our study are more accurate than the PCC; the OA value reaches 90% and the FAR value reaches 0.02.

Figure 3 includes the whole study area, allowing for an initial visual assessment of the results of wetland change extent extraction. Figures 4 and 5 show the subimages extracted from “upper lake meadow” and “lower lake meadow” regions of Figure 3.

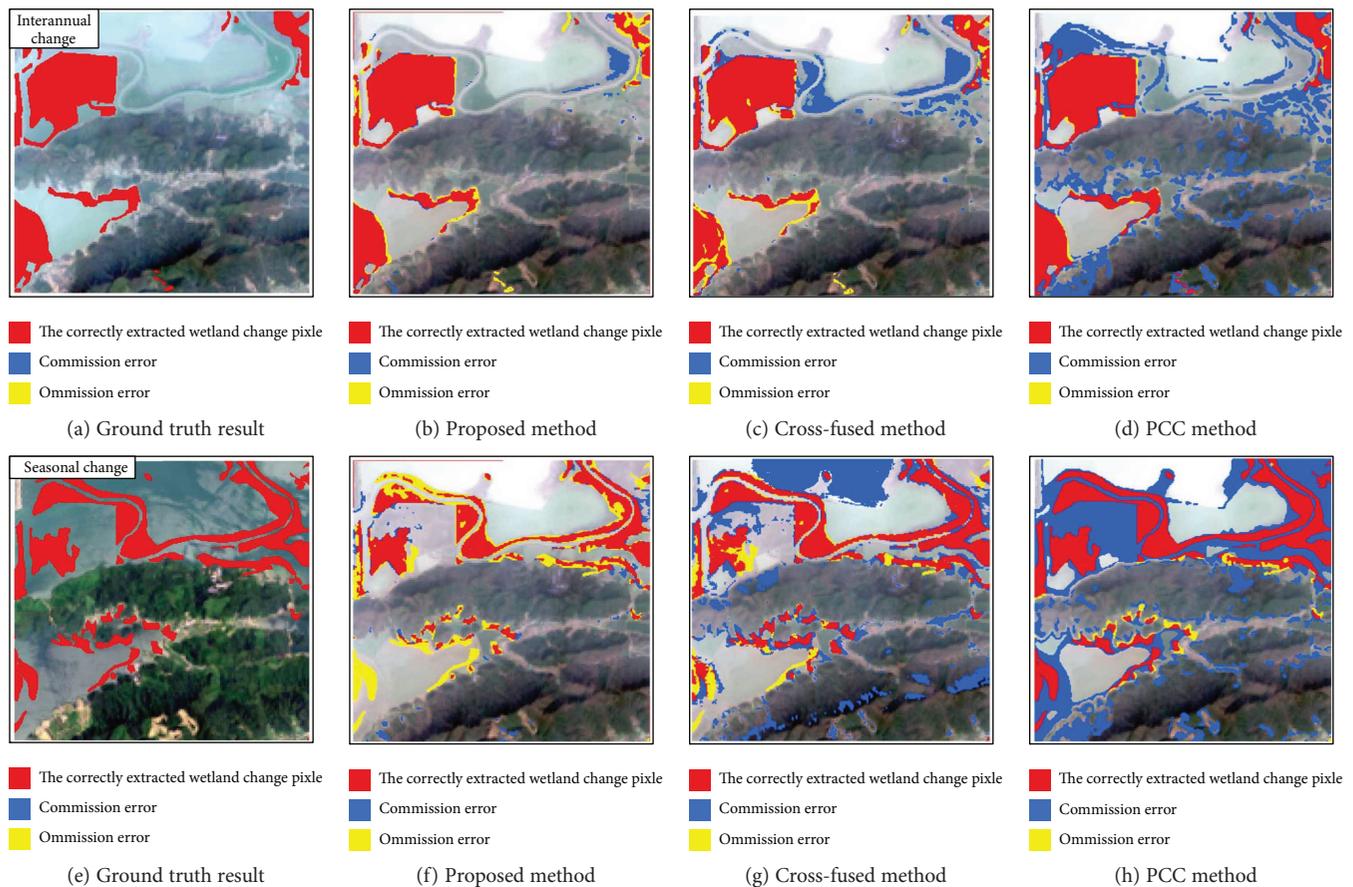


FIGURE 5: Detailed images from lower lake meadow regions of Figure 3: (a)–(d) interannual change detection result; (e)–(h) seasonal change detection result.

Figure 4 mainly consists of a complex area composed of water, meadow, cultivated area, city, and mudflats. In addition, the area shown in Figure 5 mainly consists of complex areas composed of water, meadow, forest, and mudflats. The blue and yellow colors indicate the commission and omission error, respectively, and the results are overlain on the multispectral image of 25 December 2016.

As shown in Figures 4 and 5, the PCC and cross-fused method are not sensitive to artificially cultivated areas and forest areas affected by seasonal changes, and they detect too many false positives caused by similar spectral characteristics.

The proposed method efficiently detects the changed area that corresponds to the complex area with similar spectral characteristics, improves the accuracy of wetland change detection, and minimizes the impacts of seasonality and artificiality. The proposed method also has good performance in meadow areas. However, it produces false positives in some mudflat edge regions, such as yellow areas in Figures 4(b) and 5(f). On the one hand, this is because spatial inconsistency occurred due to the different look angles of bitemporal imagery. On the other hand, the proposed method is based only on the NDWI, which is sensitive to water. As a result, some omission errors can occur in areas where the water level falls and the mudflats are exposed because the mudflats still contain a certain amount of water.

5. Conclusions

In this paper, we proposed an image-to-image change detection method using multitemporal images to quantify wetland cover changes; the method is based on a combination of a cross-fusion image and normalized difference index image. For multitemporal Landsat 8 OLI images, the GSA fusion method is used to generate cross-fusion images, and then NDVI and NDWI are extracted. The optimal change information was calculated through the modified IR-MAD, which used pairs of normalized difference index images and general fused images. The experimental results showed that the proposed method increases the accuracy of change detection and minimizes the error detection in the complex areas under different ground types. Especially in the cultivated area affected by manmade alterations, change information can be identified more accurately, and a lower FAR can be achieved. This allows us to help wetland managers implement effective management plans. Further, our method is of guiding importance in the monitoring of wetland health and wetland conservation.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was supported in part by the Natural Science Foundation of Anhui Province under Grant no. 1608085MD83, the Science and Technology project of the Department of Land and Resources of Anhui Province: Study on the Data Integration Technology for the Real Estate Registration (no. 2016-K-12), the Educational Commission of Anhui Province of China (no. KJ2018A0007), and the Department of Human Resources and Social Security of Anhui: Innovation Project Foundation for Selected Overseas Chinese Scholar.

References

- [1] G. Liu, L. Zhang, Q. Zhang, Z. Musyimi, and Q. Jiang, "Spatio-temporal dynamics of wetland landscape patterns based on remote sensing in Yellow River Delta, China," *Wetlands*, vol. 34, no. 4, pp. 787–801, 2014.
- [2] E. Adam, O. Mutanga, and D. Rugege, "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review," *Wetlands Ecology and Management*, vol. 18, no. 3, pp. 281–296, 2010.
- [3] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989–1003, 1989.
- [4] L. Yang, C. Homer, J. Brock, and J. Fry, "An efficient method for change detection of soil, vegetation and water in the Northern Gulf of Mexico wetland ecosystem," *International Journal of Remote Sensing*, vol. 34, no. 18, pp. 6321–6336, 2013.
- [5] C. Munyati, "Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset," *International Journal of Remote Sensing*, vol. 21, no. 9, pp. 1787–1806, 2000.
- [6] F. Bovolo, S. Marchesi, and L. Bruzzone, "A framework for automatic and unsupervised detection of multiple changes in multitemporal images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 6, pp. 2196–2212, 2012.
- [7] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *International Journal of Remote Sensing*, vol. 25, no. 12, pp. 2365–2401, 2004.
- [8] D. Renza, E. Martinez, I. Molina, and D. M. Ballesteros L., "Unsupervised change detection in a particular vegetation land cover type using spectral angle mapper," *Advances in Space Research*, vol. 59, no. 8, pp. 2019–2031, 2017.
- [9] J. Chen, X. Chen, X. Cui, and J. Chen, "Change vector analysis in posterior probability space: a new method for land cover change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 2, pp. 317–321, 2011.
- [10] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: from pixel-based to object-based approaches," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 80, pp. 91–106, 2013.
- [11] D. Renza, E. Martinez, and A. Arquero, "A new approach to change detection in multispectral images by means of ERGAS index," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 1, pp. 76–80, 2013.
- [12] A. A. Nielsen, "The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 463–478, 2007.
- [13] Y. Byun, Y. Han, and T. Chae, "Image fusion-based change detection for flood extent extraction using bi-temporal very high-resolution satellite images," *Remote Sensing*, vol. 7, no. 8, pp. 10347–10363, 2015.
- [14] B. Wang, J. Choi, S. Choi, S. Lee, P. Wu, and Y. Gao, "Image fusion-based land cover change detection using multi-temporal high-resolution satellite images," *Remote Sensing*, vol. 9, no. 8, p. 804, 2017.
- [15] C. Li, G. Beauchamp, Z. Wang, and P. Cui, "Collective vigilance in the wintering hooded crane: the role of flock size and anthropogenic disturbances in a human-dominated landscape," *Ethology*, vol. 122, no. 12, pp. 999–1008, 2016.
- [16] M. Barter, L. Cao, L. Chen, and G. Lei, "Results of a survey for waterbirds in the lower Yangtze floodplain, China, in January-February 2004," *Forktail*, vol. 21, pp. 1–7, 2005.
- [17] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS +Pan data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3230–3239, 2007.
- [18] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 5, pp. 1301–1312, 2008.
- [19] R. B. Myneni, F. G. Hall, P. J. Sellers, and A. L. Marshak, "The interpretation of spectral vegetation indexes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 481–486, 1995.
- [20] A. K. Bhandari, A. Kumar, and G. K. Singh, "Improved feature extraction scheme for satellite images using NDVI and NDWI technique based on DWT and SVD," *Arabian Journal of Geosciences*, vol. 8, no. 9, pp. 6949–6966, 2015.
- [21] B. Wang, S.-K. Choi, Y.-K. Han, S.-K. Lee, and J.-W. Choi, "Application of IR-MAD using synthetically fused images for change detection in hyperspectral data," *Remote Sensing Letters*, vol. 6, no. 8, pp. 578–586, 2015.
- [22] P. R. Marpu, P. Gamba, and M. J. Canty, "Improving change detection results of IR-MAD by eliminating strong changes," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 799–803, 2011.
- [23] A. S. Abutaleb, "Automatic thresholding of gray-level pictures using two-dimensional entropy," *Computer Vision Graphics and Image Processing*, vol. 47, no. 1, pp. 22–32, 1989.
- [24] R. G. Congalton, "A review of assessing the accuracy of classification of remotely sensed data," *Remote Sensing of Environment*, vol. 37, no. 1, pp. 35–46, 1991.

Research Article

Development of Data Fusion Method Based on Topological Relationships Using IndoorGML Core Module

Junho Park , Dasol Ahn , and Jiyeong Lee 

Department of Geoinformatics, University of Seoul, 163 Seoulsiripdearo, Dongdaemun-gu, Seoul, Republic of Korea

Correspondence should be addressed to Jiyeong Lee; jlee@uos.ac.kr

Received 2 March 2018; Accepted 18 April 2018; Published 13 May 2018

Academic Editor: Biswajeet Pradhan

Copyright © 2018 Junho Park et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Geospatial datasets are currently constructed, managed, and utilized individually according to the spatial scale of the real world, such as the ground/surface/underground or indoor/outdoor, as well the particular purpose of the geospatial data used for location-based services. In addition, LBS applications use an optimal data model and data format according to their particular purpose, and thus, various datasets exist to represent the same spatial features. Such duplicated geospatial datasets and geographical feature-based GIS data cause serious problems in the financial area, compatibility issues among LBS systems, and data integration problems among the various geospatial datasets generated independently for different systems. We propose a geospatial data fusion model called the topological relation-based data fusion model (TRDFM) using topological relations among spatial objects in order to integrate different geospatial datasets and different data formats. The proposed model is a geospatial data fusion model implemented in a spatial information application and is used to directly provide spatial information-based services without data conversion or exchange of geometric data generated by different data models. The proposed method was developed based on an extension of the AnchorNode concept of IndoorGML. The topological relationships among spatial objects are defined and described based upon the basic concept of IndoorGML. This paper describes the concept of the proposed TRDFM and shows an experimental implementation of the proposed data fusion model using commercial 3D GIS software. Finally, the limitations of this study and areas of future research are summarized.

1. Introduction

Geoinformation has always been a challenge owing to a variety of data models, data formats, spatial resolutions, and methods of geometric and topological representations. In general, the real world is a huge feature connected by spaces made up of a combination of geographical objects. However, a geospatial dataset has been constructed, managed, and utilized individually according to the spatial scale of the real world, such as ground/surface/underground or indoor/outdoor, as well as the purpose of the geospatial data used for location-based services. In addition, LBS applications use an optimal data model and data format according to their particular purpose, and thus, various datasets exist to represent the same spatial features. Such duplicated geospatial datasets and geographical feature-based GIS data have caused serious problems in the financial area, compatibility issues among LBS systems, and problems with data

integration among the various geospatial datasets generated independently for different systems. In other words, the spatial data representing the real world have fundamental problems in a variety of data formats, including spatial relationships among objects, data integration between 2D and 3D data, data compatibility among various types of spatial data, spatial resolution, and consistent representations in geometric and topological data [1].

To provide seamless LBSs between indoor and outdoor spaces, one of the most important requirements is to connect indoor with outdoor spaces. In general, although geographical features in an outdoor space have been represented as 2D objects, spatial entities in a micro built environment such as buildings have been represented as 3D objects in the building information model (BIM) or CityGML datasets. These 2D and 3D spatial data are generated based on various data models and formats according to the generation method applied. In this way, according to the purpose of the spatial

information service, indoor spatial data and outdoor spatial data are constructed, managed, and utilized, where certain problems are caused with regard to the interoperability among spatial information systems, data compatibility, and utilization of data linkages [2].

In addition, the convergence or fusion of spatial data is required in integrating geospatial data generated based on geographical features individually. In other words, spatial data on transportation are constructed and used individually based on geographical features such as roads, railways, sidewalks, and subways. However, the need for linkage among individual traffic-related data has been revealed through the complex transportation system based on various types of transportation infrastructure. The system consists of various transportation subsystems such as passenger cars, buses, trains, and sidewalks. However, because most transportation services utilize multiple transport modes, it is difficult to independently utilize individual transportation systems. Therefore, research on interdependent and interactive multi-modal transport systems is required. To provide such multi-modal transport services, the fusion of various transportation data must first be studied [3]. This demand is related to urban facility management issues, as seen in the California power outages in early 2001. The California power outages have had a major impact on many urban-based service sectors such as oil, natural gas production in California, and water transport for crop irrigation, which has led to interest in the understanding and analysis of connectivity between major types of urban infrastructure, including water, sewer, electric and gas systems, among others, which are considered critical infrastructure [4, 5].

To solve these problems, research has been conducted to fuse geospatial data generated by various construction methods and data models [6]. As the developed methods for improving the interoperability and compatibility of data by combining spatial data with other data of different formats, there are three approaches: a data fusion method using geometric data conversion, a data fusion method based on visualization, and recently proposed geospatial data fusion methods based on topological relations of spatial objects.

The purpose of this study is to propose the development of a spatial data fusion model called topological relation-based data fusion model (TRDFM) using topological relations among spatial objects in order to utilize different geospatial datasets and different data formats. TRDFM is a data fusion model implemented in LBS applications and is used to directly provide spatial information-based services without a data conversion or exchange of geometric data generated by different data models. The proposed method was developed through an extension of the AnchorNode concept of IndoorGML adopted by the Open Geospatial Consortium (OGC) in 2015. The topology relations between objects are defined and described based on the basic concept of IndoorGML. In the following section, we describe the existing spatial data fusion methods. In the third section, we present the proposed TRDFM using topological relations between spatial objects based on IndoorGML. In the next section, an experiment on the data fusion method based on the proposed model is described, and the final section

summarizes the limitations of this study and areas of further research.

2. Previous Studies

As mentioned in the previous section, existing spatial data representing the real world have limitations in implementing seamless LBS systems, such as various data formats, lacks of spatial relations among objects, and a lack of compatibility among geospatial data. To overcome these limitations, research on the convergence of spatial data has been carried out. The developed methods are classified into three groups: a data fusion method using a geometric data transformation, a data fusion method with a visualization aspect, and a data fusion method based on the topological relations of spatial entities. A review of previous studies is given to examine the direction and considerations of a more efficient approach to combining geospatial data.

The first spatial data fusion method is data fusion through a geometric data transformation [7] and can be grouped into two types. The first method in converting the data is in accordance with the format of the target data. The second method is to define one standardized format when there is a plurality of data to be combined and to convert the other data into a standardized format and combine them.

A data fusion method through a geometric data transformation is being actively studied for exchanging BIM data with the CityGML datasets [8, 9]. CityGML is an international standard data model adopted by OGC to represent urban spaces in 3D GIS, called a 3D city model. Industry Foundation Classes (IFC) is a representative standard data model of building information model (BIM) data. Research on converting datasets between two standards has been ongoing in the development of data fusion methods for the management and utilization of indoor spatial data [8]. BIM has more detailed information on buildings than CityGML. Therefore, research on converting BIM data into CityGML has been mainly conducted [9]. Data conversion from IFC to CityGML requires the transformation of attribute and geometric data. It is necessary to define the property data of the object to be converted from CityGML to IFC, and there is a need for an algorithm for simplifying the object data of the IFC geometry model based on rules for mapping between the two CityGML and IFC models [10]. In IFC and CityGML, the developed method has only been used to categorize the objects related to buildings and map the corresponding relations, defining the mapping rules that apply to each level of detail (LoD) of CityGML [7].

The data fusion method using a geometric data transformation requires analyzing how to map objects to objects of the target data, and mapping rules and geometric simplification algorithms should therefore be defined. As in the case of converting IFC to CityGML, there is a need to create additional objects because they are difficult to map perfectly owing to differences in the amounts of data.

Second, a data fusion method between spatial data based on visualization has been developed. This method visually represents different datasets in one application system.

Typically, V-World [11], which is implemented as an open space information platform, visually integrates various topographical 2D data, 3D building data, and 3D indoor space modeling data using digital elevation models (DEM) and orthographic images. Google Earth [12] also provides a 3D visualization system displaying 2D geospatial data through 3D topographical modeling and texturing. Currently, because of problems related to coordinate systems and scale issues, the linkage between 3D outdoor spatial data and indoor spatial data in a 3D visualization application cannot be expressed in a single viewer window or a smooth screen transition cannot be performed. To provide seamless navigation service or route guidance between indoor and outdoor spaces, a topological model expressing the spatial relations of indoor and outdoor spaces is needed. Therefore, a spatial data convergence method in terms of visualization has certain limitations in providing seamless indoor and outdoor services.

The third approach is a data fusion method based on topological relationships among spatial entities. This method merges data by connecting topological relationship data models of different datasets. The topological relation method (TRM) has been proposed using topological relationships between spatial objects to directly apply various geometric data generated by different data models for indoor spatial location services [2]. The ConnectEdge feature class defined in the 3D navigable data model (3D NDM), which abstracts the connectivity relationships of a building's internal structure, defines the connectivity relationships of spatial entities such as rooms and corridors, which are the movement paths of people [13]. Data fusion between IFC and CityGML, which are different data formats, was applied using topological representations through a node-link graph [14]. The same types of network-based topological representations utilizing node-link graphs are generated from the two geometry models by defining the topology data generation procedure and necessary attribute information for each geometry model. IndoorGML proposed an additional element for connecting indoor and outdoor spaces, which is presented as an anchor node concept. In all indoor spaces, there is an entrance for outdoor use, which is used as an anchor node to connect the indoor and outdoor areas [15]. Therefore, the TRM was developed not at the data level but at the application service level [16].

In this way, the developed methods for fusing spatial data are grouped into fused composite methods using a geometric data transformation, data fusion methods with a visualization aspect, and data fusion methods based on topological models. However, a data fusion method using a geometric data transformation requires the transformation of existing data, as well as corresponding relations of the mapping data, mapping rules, and simplification algorithms for various application services. A method based on visualization has difficulty in identifying the topological relationships among spatial entities represented in different datasets because the datasets are not merged into one geometric dataset with topological consistency even though they use the same coordinate system. The data fusion method based on topological models for connecting indoor and outdoor areas using

AnchorNode proposed by IndoorGML was presented at the conceptual level, the model of which has limitations in implementing a practical integrated system to provide seamless LBSs in the real world. In this paper, we propose a generic spatial data fusion model by extending the concept of AnchorNode of IndoorGML. The proposed spatial data fusion model can be used to directly provide seamless LBSs from the urban scale (macro space) to the human scale (micro space) without converting or exchanging the geometric data of space-based spatial data and geographical feature-based spatial data.

3. Geospatial Data Fusion Model Based on IndoorGML Core Module

This section describes the IndoorGML core module used to develop a data fusion model based on IndoorGML. We also describe the proposed TRDFM, which extends the concept of an AnchorNode class defined in IndoorGML, and detail the procedure used for generating the proposed data fusion model.

3.1. IndoorGML Core Module. In this study, the proposed data fusion model is developed based on the basic concept of IndoorGML, which presents the topological relationships among the spatial entities in a graphical structure. IndoorGML was established as the OGC international standard and is an open data model for indoor navigation applications and XML-based formats to represent and exchange indoor spatial data for indoor navigation [15]. Because the IndoorGML core module is a topological data model used to represent spatial relationships among indoor spatial entities, the module does not provide a geometric representation of spatial entities but has interface classes associated with a space and spatial boundary feature classes defined in the existing feature data models such as the CityGML and IFC models [17].

IndoorGML uses a node-relation structure (NRS) to express the connectivity and adjacency relationships among spatial objects. The NRS utilizes Poincaré duality to represent the topological relationships using dual graphs. Through a duality transformation, solid objects (3D) such as rooms within a building in a primal space are transformed into vertices (0D) in a dual space. The common surface (2D) shared by two solid objects is mapped to an edge (1D) linking two vertices in a dual space. Thus, the edges of a dual graph represent topological relationships among 3D objects in a primal space, such as doors, windows, or hatches, between rooms in a building. Similar to node-edge graphs, which use a dual graph to represent space-activity interactions [18], the NRS was developed to represent topological relationships such as adjacency relationships as $G = (V(G), E(G))$ and connectivity relationships as $H = (V(H), E(H))$ among spatial entities in the real world. In IndoorGML utilizing a network model, the nodes represent rooms, staircases, elevators, doors, building exits, and hallways. The edges in IndoorGML represent the topological relationships among spatial entities, which indicate the paths of pedestrian movement between nodes within a building.

To analyze human activity using multimodal transportation systems in urban areas, the network-based topological model (called IndoorGML) needs to be integrated with a 2D network of the ground transportation system, such as a road network. In a road network represented by a node-edge graph, the nodes represent intersections of the roads, and the edges connected by two nodes represent road segments. To connect indoor and outdoor spaces, IndoorGML provides a concept for defining additional topology elements between indoor and outdoor spaces, called *anchor nodes*. The anchor node represents the entrance of a building as a special node in the topological graph as a mediator of an indoor and outdoor space connection. As shown in Figure 1, two-way access between nodes is possible when referring to an outdoor network through an anchor node, and the topological network of a building can be obtained from an external node when a vehicle enters the building. In addition, the geospatial data of an underground water pipe and ground water pipe, which are generated separately, can be connected through an anchor node. For smooth spatial connections, the anchor nodes have attributes such as node reference data of the external reference network and parameters for supporting coordinate system transformation of indoor and outdoor spaces [18].

3.2. Topological Relation-Based Data Fusion Model. As mentioned in the previous section, IndoorGML introduced an *anchor node* to connect indoor and outdoor spaces at the conceptual level, which means that the proposed data fusion method in IndoorGML has limitations in implementing a practical integrated system to provide seamless LBSs in an urban environment. In this paper, we propose the generic spatial data fusion model, TRDFM, by extending the concept of an anchor node of IndoorGML, as shown in Figure 2. The proposed spatial data fusion model can be used to directly provide seamless LBSs from the urban scale (macro space) to the human scale (micro space).

Figure 2 shows a UML diagram of the proposed TRDFM. The orange-colored classes in the UML are GML geometry objects, and the gray-colored classes represent the classes of the core module of IndoorGML. The pink-colored classes are those derived from the navigation module of IndoorGML used to determine the topology relations from the geometric data generated based on different geometric data models. The *NavigableBoundary* and *NonNavigableBoundary* feature classes represent the surfaces of 3D spatial entities such as rooms, which are the geometric representation elements of 3D surface-oriented data models. The *NavigableBoundary* and *NonNavigableBoundary* feature classes are mapped to the *Transition* feature, which is realized as an edge in the topological model. The *NavigableSpace* and *NonNavigableSpace* feature classes are associated with the spatial entities (*CellSpace* feature) and mapped to the *State* feature, which is realized as a node in IndoorGML. The blue-colored classes are feature classes defined to integrate datasets that represent spatial entities in different spaces such as the ground, surface, or underground.

To analyze human activity using multimodal transportation systems in urban areas, the building's network generated

by IndoorGML needs to be integrated with a street network. The first step of the integration process is to define the connectivity relations between the building and ground streets. In other words, the connectivity relationships can be defined between the building network generated by IndoorGML and the street network representing road centerlines. The connectivity relations are represented by an edge, called *AnchorEdge* [19]. One node of *AnchorEdge* represents the entrance halls of the buildings, and the other node of *AnchorEdge* is on the street network, which is defined by the projection of the node onto the edge of the street network. The former node is called *AnchorNode* and represents an entrance hall of a building, whereas the latter is also *AnchorNode* representing the corresponding node of a street network generated in the manner described above. *AnchorEdge* is an edge that connects two *AnchorNodes*. *AnchorNode* has attributes such as absolute coordinates and variables for transforming the coordinate system and links the geometry data with the topology data. As shown in Figure 3, *AnchorNode* and *AnchorEdge* are defined as additional topology elements to connect between the building network and street networks in the proposed TRDFM.

AnchorNode and *AnchorEdge* are presented in the UML diagram of the proposed TRDFM, as shown in Figure 3. In Figure 4, the *State* class defined in the core module of IndoorGML represents a node in a dual graph of the geometric model of spatial entities in the real world. The spatial entities within a building can be associated with a room, corridor, door, or other elements. The *Transition* class in the UML diagram is an edge that represents the topological relationships among spatial entities. The *Transition* connects two *States* in the topological model of IndoorGML. The *State* class has two subclasses: *GeneralState* and *SpecialState*, indicated by an inheritance relationship in the UML object diagram. *GeneralState* is a node in the topological model of IndoorGML representing a space similar to a room in the real world, whereas *SpecialState* is a node introduced in the proposed TRDFM, called *AnchorNode*. The *Transition* class has two subclasses: *GeneralTransition* and *SpecialTransition*, indicated by an inheritance relationship in the UML object diagram. *GeneralTransition* is an edge in the topological model of IndoorGML representing the topological relationships among spatial entities, like rooms in the real world, and *SpecialTransition* is an edge as described above, called *AnchorEdge*.

To formalize the topology-based data fusion model, TRDFM, including *SpecialState* and *SpecialTransition*, the schemas of the objects are shown in Tables 1 and 2. The primal classes of the TRDFM are *GeneralState*, *SpecialState*, *GeneralTransition*, and *SpecialTransition*, which are inherited from *State* and *Transition* classes defined in IndoorGML. A *State* consists of an identifier and position data in 3D (x-, y-, and z-coordinates), and *Transition* consists of an identifier and two *States*. Each *GeneralState* in the database has an identifier (Id), space type (Type), a space ID (RelatedSpaceId) represented by the corresponding space, and an edge ID (ConnectedEdgeId) of an edge connected to a node. *SpecialState* has additional attributes, as shown in Table 1, which are reference node data of the street network and conversion parameters used to support the coordinate system conversion

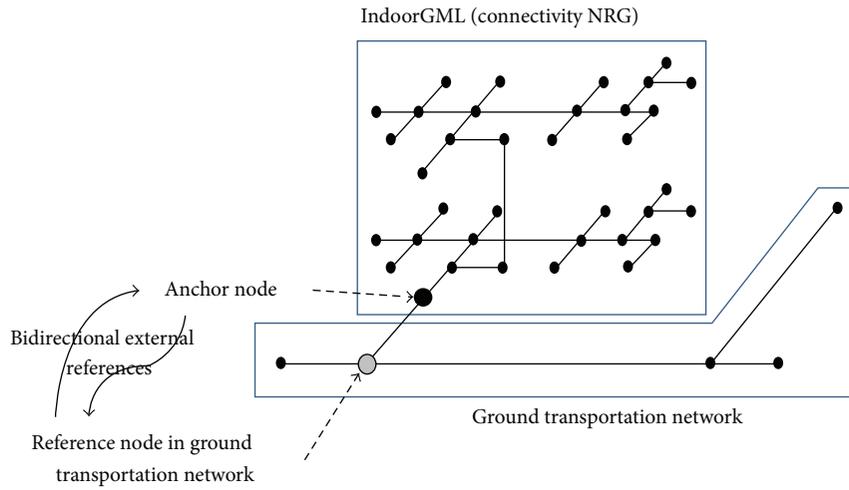


FIGURE 1: Anchor node connecting indoor and outdoor spatial networks (OGC [15]).

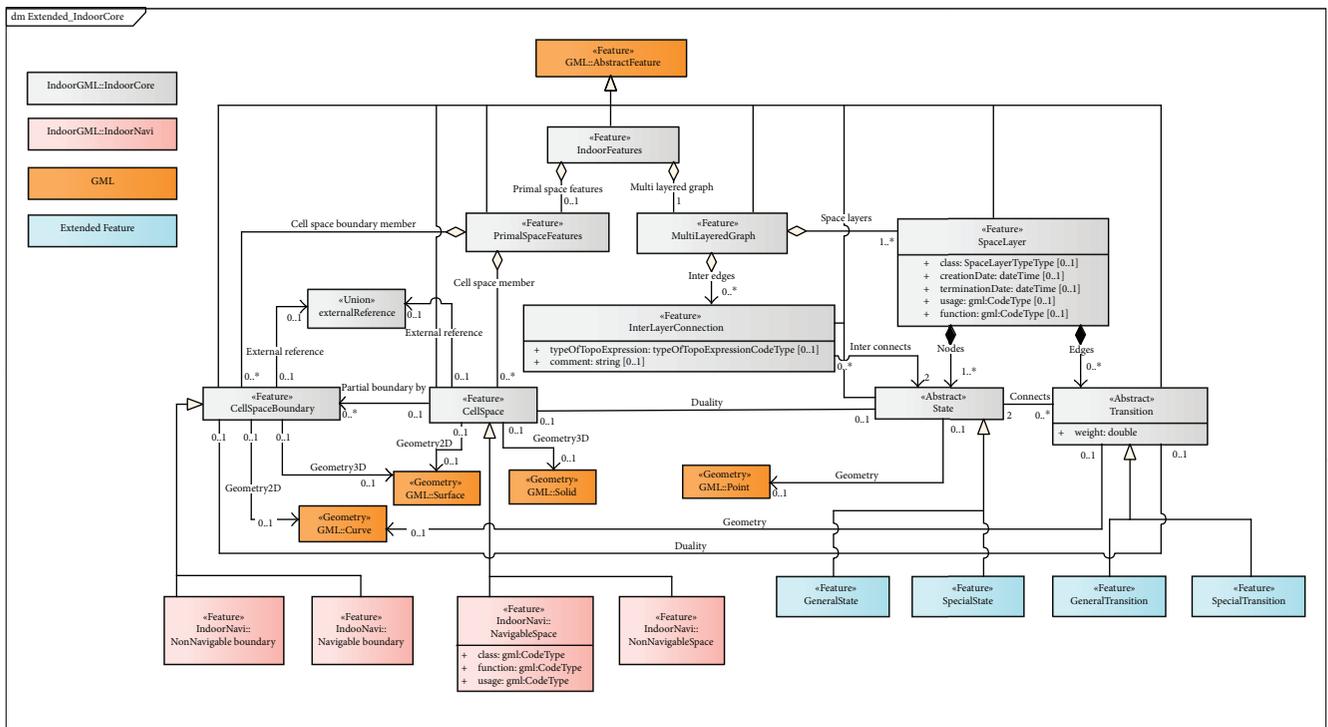


FIGURE 2: UML diagram of TRDFM based on IndoorGML core module.

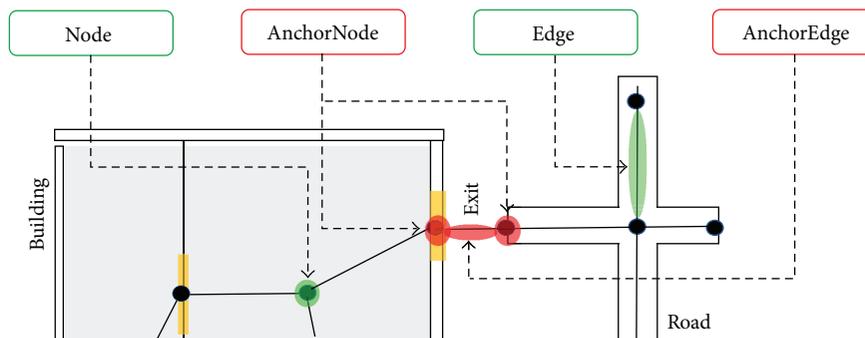


FIGURE 3: Spatial data fusion method based on topological relationships.

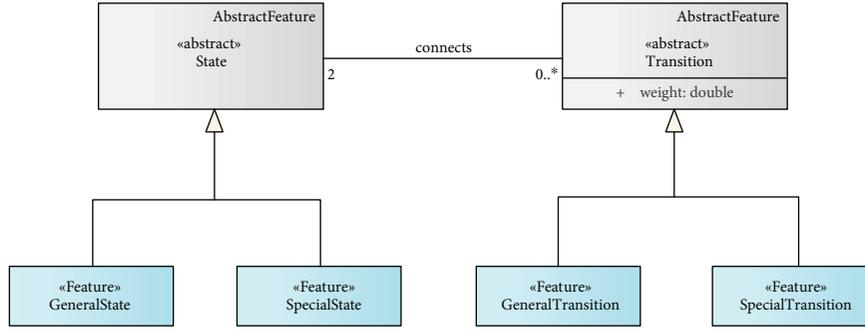


FIGURE 4: Topology data model: anchor portion.

TABLE 1: Required attribute information for *SpecialState* class.

Name	Type	M/O	Description
Id	String	M	Identifier
Type	NodeType	M	GeneralSpace/TransferSpace/AnchorSpace
RelatedSpaceId	String	M	Identifier of presented CellSpace or CellSpaceBoundary
ConnectedEdgeId	Set of edge	M	A set of edge to adjacent node
OriginPoint	Set of coordinates	O	Coordinates of node
RotationAngles	Set of value	O	CRS-converting parameter
RescalingFactor	Set of value	O	CRS-converting parameter
TranslationVector	Set of value	O	CRS-converting parameter

TABLE 2: Required attribute information for *SpecialTransition* class.

Name	Type	M/O	Description
Id	String	M	Identifier
L_Node	String	M	Identifier of left-side connected node
R_Node	String	M	Identifier of right-side connected node
Type	EdgeType	M	Connectivity/adjacency/anchor
RelatedSpaceId	String	M	Identifier of CellSpace or CellSpaceBoundary presented by connected node
LinkType	LinkType	O	Surface-ground/surface-underground/ground-underground
Activate	Boolean	O	Activate condition

and link the indoor spatial data, such as a building network, into outdoor spatial data, such as a street network. It consists of OriginPoint (x_0 , y_0 , and z_0), RotationAngles (α , β , γ , x , y , and z), RescalingFactor (s_x , s_y , and s_z), and TranslationVector (t_x , t_y , and t_z).

Each *GeneralTransition* in the database has an edge ID (Id), a left node connected to the edge (L_Node), a right node connected to the edge (R_Node), an edge type representing a topology relation, and a space ID (RelatedSpaceId). *SpecialTransition* has additional attributes, as shown in Table 2, which are LinkType and Activate_State data, in order to indicate the special type of connected network data, and thus, it can be utilized in various application services such as the opening and closing of entrances and exits.

3.3. Procedure Used to Generate Topological Relation-Based Data Fusion Model. Figure 5 illustrates the data that follow from the generation process of the proposed TRDFM. The

input files are two GIS datasets including network datasets representing the centerlines of streets in an outdoor space and 3D solid GIS datasets representing a building and describing subunits within the building. Each solid representing a spatial unit has a label, which is a node abstracted from a solid in the TRDFM generated by the 3D Poincaré duality. All nodes representing spatial entities such as rooms, hallways, or entrances of a building are connected to each other based on their topological relations. The combined network graph is generated through the above processes as a logical node-relation graph (NRG) and geometric NRG [15]. The geometric NRG is the navigable data model of a building. The logical NRG is a pure graph representing the adjacency and connectivity relationships among the internal units of a building, and does not represent the geometric properties such as the distances among the units. However, the geometric NRG accurately represents their geometric properties so as to implement a network-based analysis, such as pathfinding.

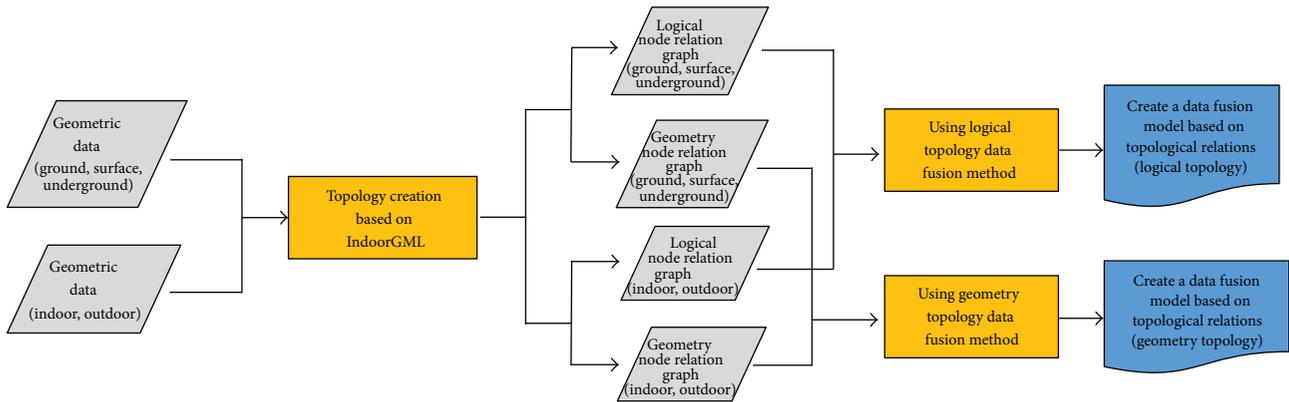


FIGURE 5: Flowchart of the proposed method.

As mentioned previously, geoinformation has always been a challenge because of the variety of data models, data formats, spatial resolutions, and methods of geometric representation. Despite the real world being a huge object, a geospatial dataset has been constructed, managed, and utilized individually according to the spatial scale of the real world, such as the ground/surface/underground or indoor/outdoor, as well as the purpose of LBS applications. In addition, the spatial dataset used for LBS applications is generated based on an optimal data model and data format according to their particular purpose. The 3D geometric modeling data formats for expressing an indoor space include 3D B-reps, CSG, IFC, and CityGML. These are largely divided into volume-oriented geometric modeling methods and surface-oriented geometric modeling methods. IFC is a representative volume-based data model. Therefore, the interior space geometry modeling of IFC is represented as a volumetric hexahedral geometry object. CityGML is representative surface-based geometric model. The geometry model of CityGML is a geometric element in a plane form, where each wall is represented, and each room is expressed as surrounded by the generated plane.

Figures 6 and 7 illustrate the method for generating topological data from 3D geometric data representing an indoor space generated by volume-oriented geometric models and surface-oriented geometric models. In the case of the geometric data of a surface-oriented data model, the spatial objects are divided into *NavigableBoundary* and *NavigableSpace* objects defined in IndoorGML. In the thin door model, a room is mapped to *NavigableSpace*, and a door is mapped to *NavigableBoundary*, which is mapped to the *State* object of the NRG represented by a node. A node of *NavigableBoundary* connects to the nearest *NavigableBoundary* node and connects to the adjacent *NavigableSpace* node. The edges represent the connectivity relationship of each door and the connectivity relationships of a door and a room. In the case of geometric data based on volume-oriented data models, which is called a thick door model, rooms are mapped to the *NavigableSpace* objects of the NRG realized by nodes in the TRDFM. In addition, doors are mapped to the *NavigableSpace* objects because the doors are represented as solid geometric objects in the geometric modeling data. Next, each of the nodes of *NavigableSpace* and the neighboring nodes are

connected to the edge to show the connectivity relations of the spatial entities.

For road network data, as shown in Figure 8, the link generated along the centerline of the road corresponds to the *Transition* object, and the node generated at the road intersection point and the broken point corresponds to the *State* object and is utilized as the topological data.

The fusion of indoor spatial data and outdoor spatial data is achieved using *SpecialState* and *SpecialTransition* objects, as shown in Figure 9. *SpecialState* is represented as an *AnchorNode* object, and *SpecialTransition* is realized as an *AnchorEdge* object. *AnchorNode* is an anchor node located at the end of the road network that connects to the entrance of the building and represents the entrance of the building. *AnchorEdge* connecting the two *AnchorNodes* is integrated with the two datasets.

4. Experimental Implementation of the TRDFM

To evaluate the potential benefit of a topology-based data fusion model for developing a seamless service application through indoor and outdoor spaces, we conducted an experimental implementation of a system based on the TRDFM described in the last section. The dataset used for our implementation was drawn from a comprehensive GIS database of the National Geographic Information Institute in Korea, which is located in the study area of the University of Seoul. In this section, topological data were generated based on the TRDFM to integrate the building and street data. To investigate the convergence of different spatial datasets based on the generated NRGs, a shortest path search from an underground building space to a ground building space and a network analysis for a specific event occurrence scenario were carried out.

4.1. Experimental Spatial Datasets. The geometric datasets used in this study are shown in Figure 10. Figure 10(A) shows the 3D geometric data of 21C building (called building A) at the University of Seoul, Korea, generated from a building layer of the 1/5000 digital topographic map. Figure 10(B) shows the 3D geometric data of one of the underground buildings (called building B), also located at the University of Seoul (UoS), Korea. The 3D geometric data of 21C

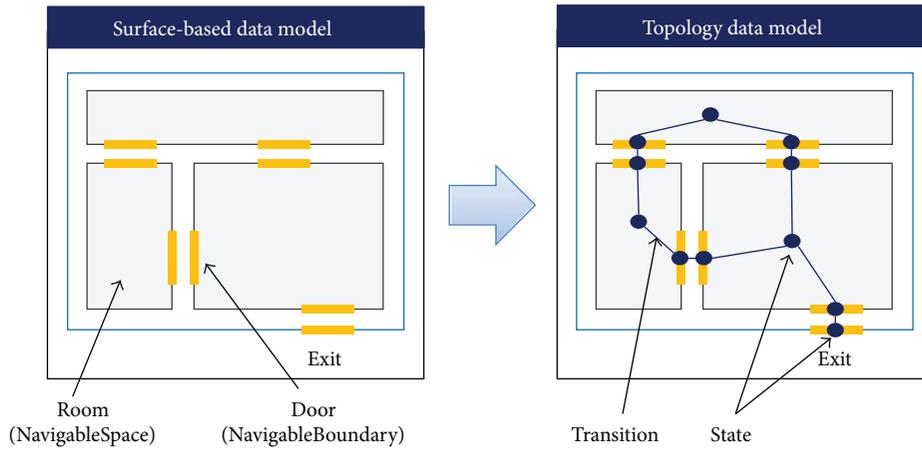


FIGURE 6: How to generate topology data using a surface-based data model.

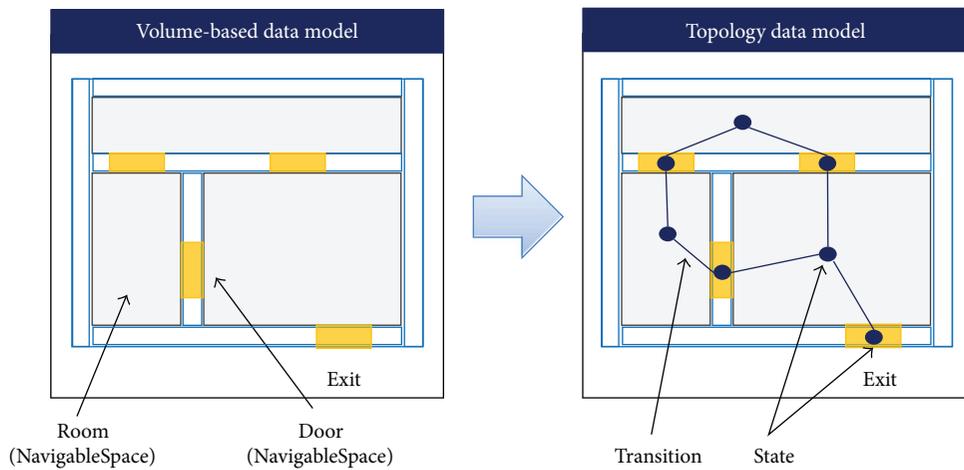


FIGURE 7: How to generate topology data using a volume-based data model.

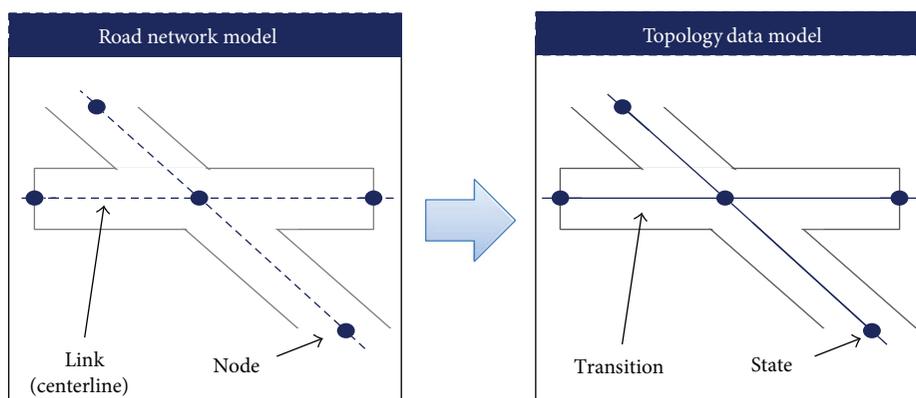


FIGURE 8: How to use topology data using the road network model.

building is formatted using the IFC geometric model, and the 3D geometry data of the underground building is generated in the CityGML data format.

The topological data of the study area were generated according to the procedure described in the previous section. In the case of 21C building, nodes in the topological data

were constructed based on the thick door model, one of the volume-oriented data models used for each floor, from the first to the seventh, and the nodes were connected to construct an edge based on the connectivity relationships among the spatial entities. When the spatial relations of an indoor space are expressed using a node-edge structure, if the long

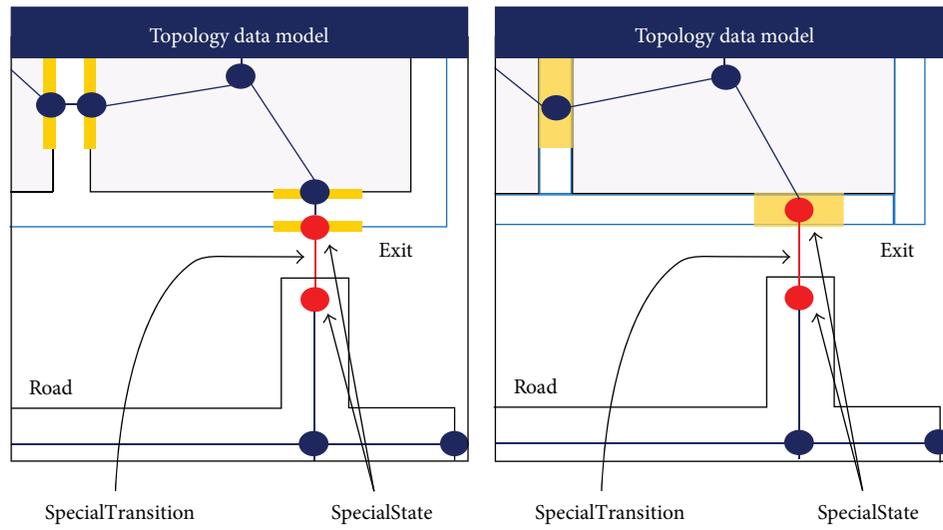
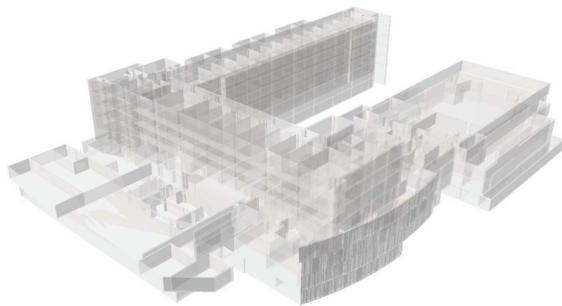
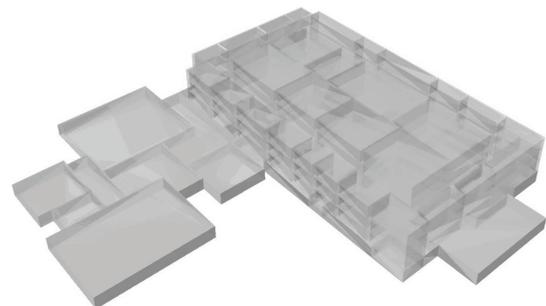


FIGURE 9: How to create *SpecialTransition* and *SpecialState*.



(A) The 21st century building 3D geometry data



(B) Underground building 3D geometry data

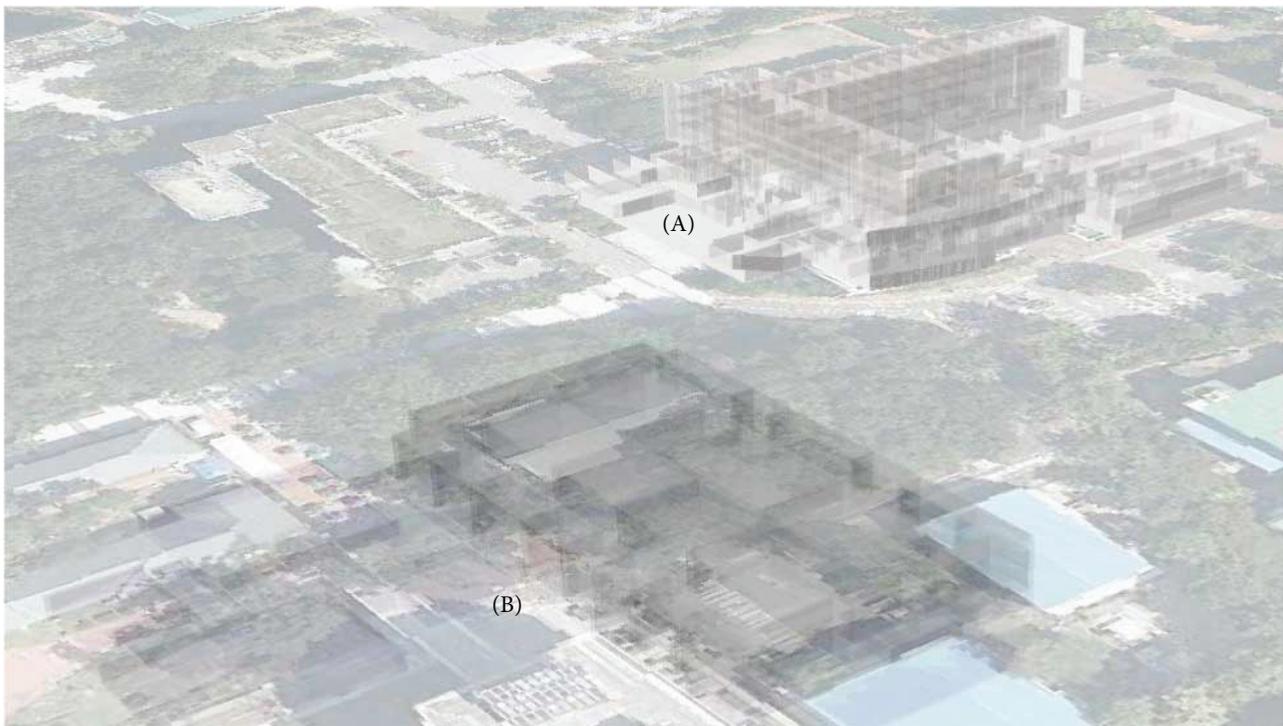


FIGURE 10: Geometric datasets of two buildings in the study area.

corridor space connected to the various rooms is expressed as a single node, the connection between rooms is defined through only one hallway node. In this case, the distance between the nodes is calculated incorrectly for network-based analyses. An effective spatial partitioning is needed to take into account the geometric aspects of the actual space and to derive more effective analysis results. In this study, we used a medial axis algorithm to generate effective topology data for the corridors. In the case of a long corridor, we used the medial axis algorithm to divide one space into several rooms, as shown in Figure 11. The medial axis was divided based on the door installed in each room and the built corridor nodes in front of the door. Figure 11(b) shows the results of the above process, called a geometric NRG in IndoorGML.

For the 3D vertical connectivity of each floor in the building, the nodes that have height information of each floor in the vertical movement space are constructed, and the edge connecting the constructed nodes representing a vertical movement space is constructed. The street network of the roads in the study area was constructed according to the method of topologic data generation based on the road network model. The topological data for the underground building are constructed based on the method described above, which are a logical NRG and a geometric NRG.

Figure 12 illustrates the combined network graph in a 2D viewer, which is integrated with the geometric NRG of buildings A and B, and the street network in the study area, using *AnchorEdges* and *AnchorNodes*. The combined network representing the connectivity relations among spatial entities can be used to provide seamless service applications through indoor and outdoor spaces. The experiment for network-based analysis will be described in the next section. In this experiment, nodes and edges are constructed with minimal attributes for connectivity analysis.

4.2. Connectivity Analysis Using TRDFM. In this study, we developed a demonstration program to visualize and analyze 3D network data by loading the generated topological data based on TRDFM. The demo program visualizes all of the loaded topological data in a 3D viewer, enables a data display through 3D rotation and movement, visualizes the results of the network analysis conducted by the user's selected node, and outputs the results. The topological data constructed for each layer are visualized in a 2D space, and the list of nodes and edges in the loaded topological data is provided and can be selected by the user. This allows users to conduct network analyses by selecting a specific space (node). The attribute data of the selected node or edge are also displayed.

Spatial data representing different spaces and data in different formats are expressed in the same topological model of the node-edge structure. All constructed topological network data are integrated through *AnchorEdge* objects. As a result, it is possible to analyze not only the network in each existing space but also the network connected to the other space. In this experiment, we implemented 3D spatial analyses by applying Dijkstra's algorithm to the combined 3D topological network data in the study area. A network analysis according to the shortest path search and the specific space

event occurrence scenario was conducted to confirm the data fusion of different spatial datasets based on the topological relations among spatial entities.

Dijkstra's algorithm is a shortest path search algorithm that finds the path that minimizes the sum of the costs from one node to another in a network [19]. This can be used to search for paths with the smallest cost (distance, etc.) from one specific space to another. According to the value of the edge having the cost data ranging from node to node, it is possible to search for a path based on the occurrence of an event such as the blocking or detouring of a specific space. The operation of the shortest path utilized by Dijkstra's algorithm is shown in Algorithm 1. The algorithm is implemented as follows: (1) find all edges connected to the starting node, (2) compare the connected edges and find the edge with the least cost, (3) make sure that the end node connected to the edge with the least cost has another path, and (4) compare the cost of the path and allocate the minimum cost. The algorithm searches for the shortest path by repeating the process until reaching the destination node.

In this experiment, the connection of the topological node is the connectivity relationship in a 3D continuous space. The shortest path search is performed by computing the minimum cost of the path of the 3D space using Dijkstra's algorithm based on the distance between nodes in the 3D space, that is, the distance property data of the edge, as described in Algorithm 1. The results of the shortest path search implemented in the network visualization and analysis program developed in this study are shown in Figure 13.

To verify the connectivity relationships of different spatial datasets based on the topological relations, we created the following setup: node 40021 (type, room) of the first basement floor of a virtual underground building is set as the start node and node 50043 (type, room) in 21C building as the destination node. As shown in Figure 13, node 40021 as the starting node is connected to the entry point of the underground building (building B), which is the anchor node 40053 (type = VerticalWay/anchor). It is connected to a node of the nearest road segment, which is anchor node 50076 (type = road/anchor). Next, it is also possible to use node 50082 (type = road/anchor) in the street network to connect to building A. The anchor node is connected to the final destination, node 50043, through the nodes on the first floor of the 21C building. The distance between the selected start and destination nodes is 138.2m, and the minimum cost of the route is calculated. The derived shortest path is visualized in the 3D viewer of the demonstration program, as shown in Figure 13.

As a result of the shortest path search, the anchor node (40053) of building B and the anchor node (50076) of an underground building are connected to each other through an anchor edge (4005350076). The anchor node (50082) of the road and the anchor node (50045) indicating the entrance of 21C building (building A) were connected through the anchor edge (5004550082). Based on this result, the connectivity through the same topological models of different spatial data is confirmed.

The second network analysis in this experiment for data fusion based on the topological relations was conducted to

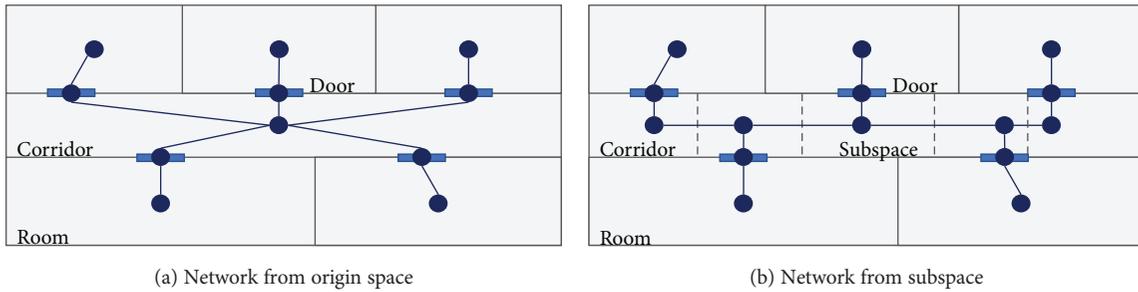


FIGURE 11: Example of topology data construction through spatial division.

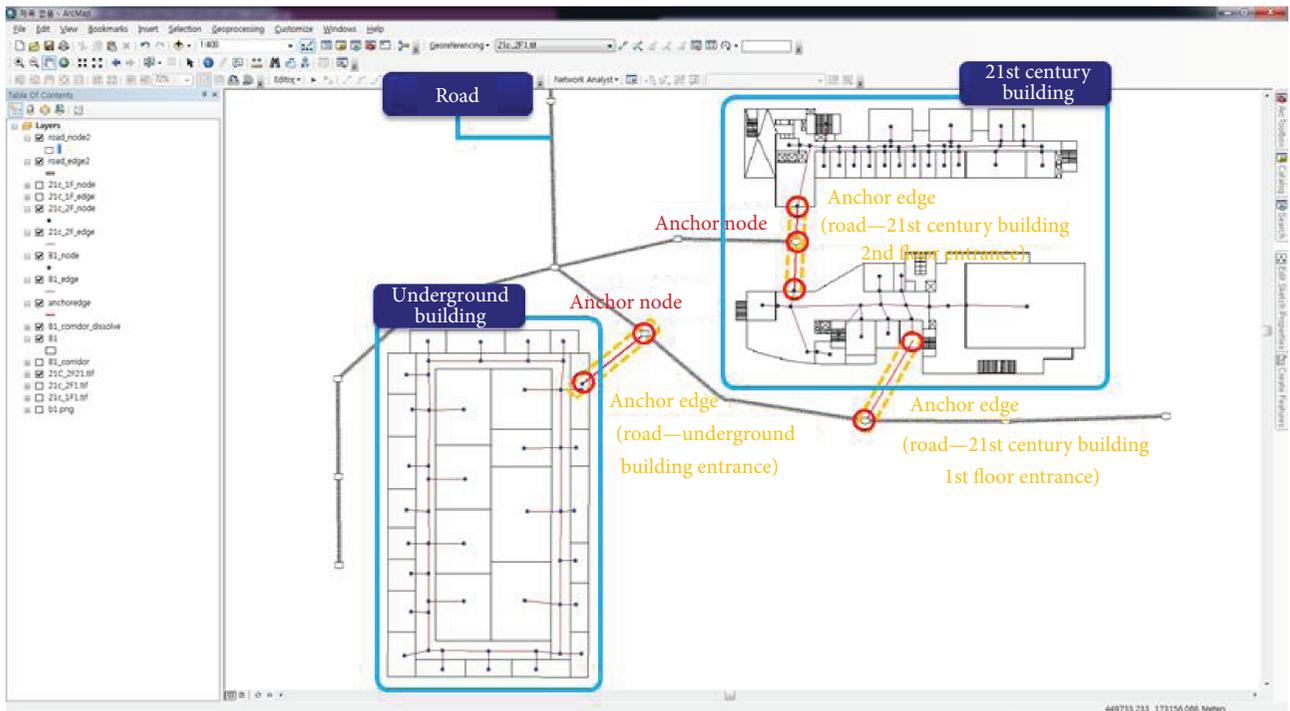


FIGURE 12: Topological data based on the TRDFM in the study area.

$(i \neq j \neq k \neq l)$

(a) Compare the costs of the edges connected to the start node(N_i)
 (b) Add node(N_j) connected to the edge(E_i) with the least cost to the path
 (c) The cost(C_i) of the edge(E_i) is allocated to the added node(N_j)
 (d) Performs (a) operation starting from node(N_j)
 (e) Add node(N_k) connected to the edge(E_j) with the least cost to the path
 (f) The cost of the route ($C_k=C_i+C_j$) is allocated to the added node(N_k)
 (g) Identify different paths between node(N_i) and node(N_k)
 (g1) Compared with the cost of another route(C_l) and the cost of allocated route(C_k)
 (g2) Reallocate the smaller cost to the node(N_k) with the minimum cost ($C_k:=C_l$ or C_k)
 (h) Repeating (a) ~ (g) operations up to the destination node, calculating the minimum cost of the destination node

ALGORITHM 1: Dijkstra’s algorithm operation.

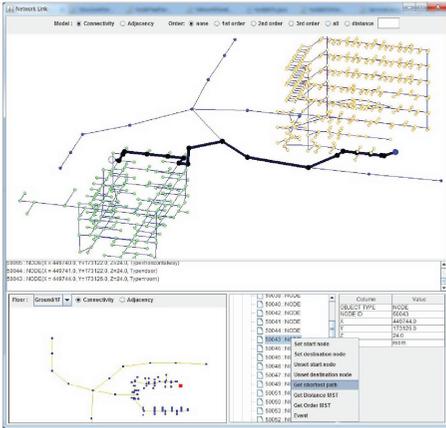


FIGURE 13: Shortest path search result screen.

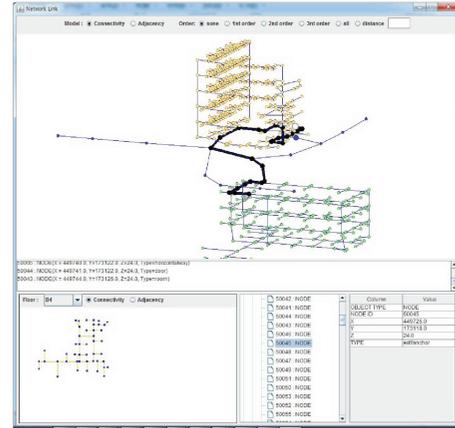


FIGURE 14: Shortest path search result screen showing event occurrence.

```

SetEvent (Graph, s)
in      Graph G = (Node N, Edge E)
        set node  $n_s(x,y,z) \in N$ 
        distance of edge
Out
 $d(n_p, n_j)$  : distance of edge  $(n_p, n_j) \in E$ 
s, i, j : integer
Begin
  Choose a Node  $n_i \in N$ 
  FOR j = 1 to n
    IF  $(n_p, n_j) \in E$  then
       $d(n_p, n_j) < -\infty$ 
    End
  End
End

```

PSEUDOCODE 1: Pseudocode for event generation algorithm.

determine the shortest path search according to the event occurrence within a specific area. A network analysis is used to search for the shortest path as determined based on the assumed scenario of the event, such as fire in a specific space, using Dijkstra's algorithm. With Dijkstra's algorithm, which calculates the minimum cost, the cost of all edges connected to the event space (node) at the time of the event occurrence is increased to the maximum value. The pseudocode of the event generation algorithm used to implement this is shown in Pseudocode 1.

As shown in Figure 14, the network analysis at the event occurrence has the following settings used to test the different spatial data connectivity implementations based on the topological relationships. The network analysis was conducted by generating an event at the entrance anchor node of 21C building where the ground building and road data are connected. The event scenario assumes that a fire occurred at the south entrance (node 50045) on the first floor of 21C building and that the second-floor doorway should be used (nodes 60076 and 60077). The search is for the shortest path from a specific room (node 40021) on the first basement floor of building B to a specific room (node 50043) on the first floor of building A.

As a result, the shortest path to the destination node 50043 was searched through the anchor node (60077), which is the entrance to the second floor of the western part of building A, without passing through the south entrance of the first floor. Likewise, the connectivity through the same topological relationships of different spatial data has been confirmed through an implementation of the network analysis, ranging from underground to the surface and to the ground. By setting the property values of the anchor edge connecting different spatial data, it is possible to utilize service applications such as a fire occurrence or opening/closing of the entrance/exit.

The result of the shortest path search using the network analysis based on the topology relation conducted in this experiment can be visualized along with the existing independent 3D geometric data, as shown in Figure 15. Independent from the geometric data, but based on the spatial relationships, different data can be fused to conduct such network analyses. Visualization is possible by displaying the results together with geometry datasets.

In this experiment, different geometric data are expressed in the same topological model according to the proposed convergence method. Through the experimental implementation, it is possible to provide a network analysis and application service linked to the existing independent data, and the analysis results can be visualized as geometric data to help the user's understanding.

5. Conclusion

Although the real world is a huge object, geospatial datasets have been constructed, managed, and utilized individually according to the spatial scale of the real world, such as the ground/surface/underground or indoor/outdoor, and the purpose of LBS applications. In addition, the spatial datasets used for LBS applications are generated based on an optimal data model and data format in terms of their particular purpose. Such duplicated geospatial datasets and geographical feature-based GIS data cause serious problems in terms of financial issues, compatibility among LBS systems, and data integration among various geospatial datasets independently

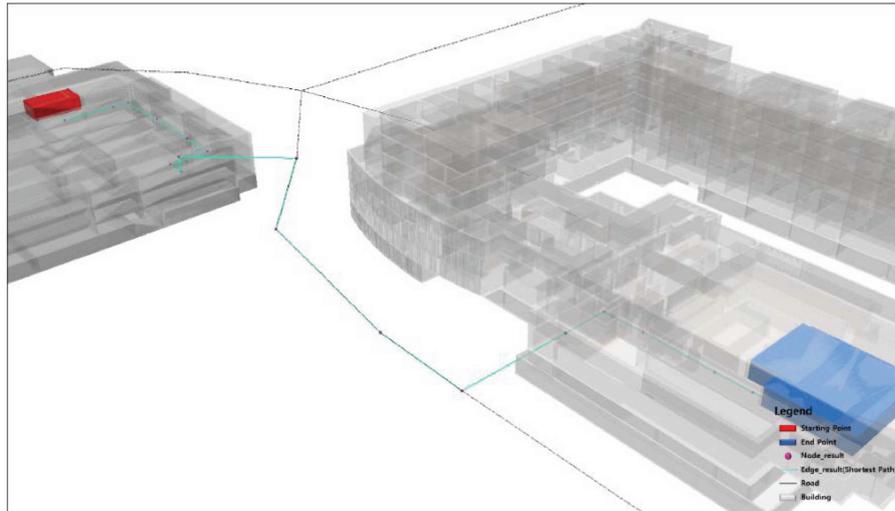


FIGURE 15: Visualization of geometry data and results of network analysis.

generated for different systems. These problems need to be addressed to better understand the interdependencies of major urban infrastructure that affect a wide range of modern urban societies and to analyze the flows and connectivity. To solve this problem, research has been conducted to fuse information generated using various methods and data models. This study proposed the development of a spatial data fusion model called the topological relation-based data fusion model (TRDFM) using topological relations among spatial objects in order to utilize different geospatial datasets and different data formats.

To realize the connectivity of different spatial data based on the topological relations using the proposed convergence method, we implemented a demo program to visualize topological datasets and conduct a 3D network analysis using Dijkstra's algorithm. In this manner, we implemented data connectivity based on topological relations. It is possible to query various 3D spatial data through connected network data through the fusion of different formatted geometric datasets based on topological relations. It is also possible to link indoor pedestrian navigation used in a car navigation system and indoor spaces in an outdoor space. In addition, a network analysis and application services are possible regardless of the scale of the space, such as route guidance from one subway station to another subway station, to ground transportation, or to the interior space of a building.

The proposed TRDFM contributes to the literature in significant ways because current geographical information systems still have huge problems in using differently formatted data in the same application. The most common method remains data fusion through a geometric data conversion. However, this study has several limitations that need to be addressed. First, efficient data utilization and various application services should be provided by expanding various 3D spatial query functions based on a topology-based data fusion. In addition, for the user's understanding and convenience, it will be necessary to improve the visualization functions so that the network analysis results based on the

performed topological relations can be visualized along with the geometric data.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by a grant (17NSIP-B135746-01) from the National Spatial Information Research Program (NSIP) funded by the Ministry of Land, Infrastructure and Transport of the Korean Government.

References

- [1] M. Ouyang, "Review on modeling and simulation of interdependent critical infrastructure systems," *Reliability Engineering & System Safety*, vol. 121, pp. 43–60, 2014.
- [2] J. Lee, H. Y. Kang, and Y. J. Kim, "Developing data fusion method for indoor space modeling based on IndoorGML core module," *Journal of Korea Spatial Information Society*, vol. 22, no. 2, pp. 31–44, 2014.
- [3] B. Si, X. Yan, H. Sun, X. Yang, and Z. Gao, "Travel demand-based assignment model for multimodal and multiuser transportation system," *Journal of Applied Mathematics*, vol. 2012, Article ID 592104, 22 pages, 2012.
- [4] H. Anderson, *Increased Threat of Outages in California*, UPI, Washington, DC, 2001.
- [5] P. Behr and W. Booth, *Hot, Dark Summer Ahead for California: Drought Worsens Power Crunch, Senators Told*, The Washington Post, Washington, DC, 2001.
- [6] M. Gunduz, U. Isikdag, and M. Basaraner, "A review of recent research in indoor modelling & mapping," in *The International*

- Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B4, pp. 289–294, Copernicus GmbH, Göttingen, Germany, 2016.
- [7] J. R. Hwang, T. W. Kang, and C. H. Hong, “A study on the correlation analysis between IFC and CityGML for efficient utilization of construction data and GIS data,” *Journal of Korea Spatial Information Society*, vol. 20, no. 5, pp. 49–56, 2012.
- [8] OGC (Open Geospatial Consortium), “CityGML v.2.0,” <https://www.citygml.org/>.
- [9] S. J. Tang, Q. Zhu, W. W. Wang, and Y. T. Zhang, “Automatic topology derivation from IFC building model for in-door intelligent navigation,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-4/W5, no. 4, pp. 7–11, 2015.
- [10] U. Isikdag and S. Zlatanova, “Towards defining a framework for automatic generation of buildings in CityGML using building information models,” in *3D Geo-Information Sciences*, pp. 79–96, Springer, Berlin, Heidelberg, 2009.
- [11] “Korea National Spatial Data Infrastructure portal, V-world,” <http://map.vworld.kr/map/maps.do>.
- [12] “Google, Google Earth,” <https://www.google.co.kr/intl/ko/earth/>.
- [13] J. Lee, “A Spatial Access Oriented Implementation of a Topological Data Model for 3D Urban Entities,” *Geoinformatica*, vol. 8, no. 3, pp. 235–262, 2004.
- [14] I. Hijazi, M. Ehlers, S. Zlatanova, T. Becker, and L. van Berlo, “Initial investigations for modeling interior utilities within 3D geo context: transforming IFC-interior utility to CityGML/UtilityNetworkADE,” in *Advances in 3D Geo-Information Sciences*, pp. 95–113, Springer, Berlin, Heidelberg, 2011.
- [15] OGC (Open Geospatial Consortium), “IndoorGML v.1.0,” <http://docs.opengeospatial.org/is/14-005r3/14-005r3.html>.
- [16] H. Y. Kang, J. R. Hwang, and J. Y. Lee, “A study on the development of indoor spatial data model using CityGML ADE,” *Journal of Korea Spatial Information Society*, vol. 21, no. 2, pp. 11–21, 2013.
- [17] K. J. Li and J. Y. Lee, “Basic concepts of indoor spatial information candidate standard IndoorGML and its applications,” *Journal of Korea Spatial Information Society*, vol. 21, no. 3, pp. 1–10, 2013.
- [18] J. Lee, “A three-dimensional navigable data model to support emergency response in microspatial built-environments,” *Annals of the Association of American Geographers*, vol. 97, no. 3, pp. 512–529, 2007.
- [19] E. W. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

Research Article

Research on Real-Time Supervisory System for Compaction Quality in Face Rockfill Dam Engineering

Shengxiang Huang,^{1,2} Wen Zhang ,^{1,2} and Gen Wu ³

¹School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

²Collaborative Innovation Center for Geospatial Technology, Wuhan 430079, China

³Xianning Vocational Technical College, Xianning 437100, China

Correspondence should be addressed to Wen Zhang; wzhang@whu.edu.cn

Received 1 December 2017; Revised 18 April 2018; Accepted 30 April 2018; Published 10 May 2018

Academic Editor: Ghassan Beydoun

Copyright © 2018 Shengxiang Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Compaction quality control in filling construction is of great significance to the stability and durability of the face rockfill dam. The conventional method of quality control mainly relies on manual process control and inspection for a limited number of test holes, which cannot meet the high requirements of modern mechanized construction and schedule anymore, with increasing of scale of face rockfill dams. There is an urgent need to propose a new quality control method of face rockfill dams during the entire compaction process. In this paper, a supervisory system based on GNSS (Global Navigation Satellite System) technology, wireless data communication technology, Internet of things technology, and computer technology is developed to supervise the real-time roller compaction parameters of the working surface including rolling track, rolling times, rolling speed, thickness, and smoothness. The system obtains continuous and high-precision spatial position information of roller compaction machines through GNSS technology and then calculates the roller compaction parameter information. The compaction quality control for the face rockfill dam is achieved through the supervision of roller compaction parameters. The feasibility and robustness of the developed supervisory system are validated by a case study in the face rockfill dam of Shuibuya project in China. The practice shows that the system provides a new and effective method of process control for the construction quality of the roller compaction in dam engineering and realizes real-time, precision, and automatic supervising of roller compaction parameters and ensures better construction quality.

1. Introduction

In the construction of the face rockfill dam, the compaction quality of the filling material is very important to the stability and durability of the dam. Therefore, the quality control in the process of filling construction is the key point to ensure the quality of dam construction. Insufficient compaction and quality control result in decreased strength and bearing capacity and increased settlement, volume change, and permeability. And it will directly affect the operation safety of the dam [1]. How to supervise the rolling construction process and make it meet the corresponding requirements is an important research question for scholars and quality managers.

The quality control methods of “dual controls” are mainly adopted in the quality management of filling construction of the face rockfill dam [2], of which one is manually controlling the roller compaction parameters including rolling times and driving speed of compaction machines, thickness of filling layer, and smoothness of storehouse surface, and the other is inspecting the test holes sampled manually in the working surface. However, the conventional method is difficult to ensure construction quality, because it is hard to accurately control these roller compaction parameters, interfered by human factors and extensive managements. Moreover, with the increase of scale of face rockfill dams, it puts forward higher requirements on filling construction quality control. In a word, the conventional

manual quality management mechanism cannot meet the high requirements of modern mechanized construction and schedule anymore. Considering the shortages of conventional quality control method, it is necessary to propose a real-time and automatic compaction quality supervising method, which can realize timely and quickly supervising and feedback control impersonally.

Recently, many scientific research institutions and university science and technology workers paid attention to the automatic quality control method of compaction supervising. Different supervising systems for compaction quality have been developed and applied in engineering practice. In the middle of the 1980s, GEODYN company developed a compaction documentation system (CDS) [3, 4]. The CDS is a system that conceptually controls the process of pressure implementation. The data needs to be entered manually by the driver during the compaction process. Moreover, the CDS does not use sensors to orient and move the mobile roller compactor, and the driver must follow a predetermined route. In 1996, Froumentin and Peyret developed a prototype system called MACC, whose basic goal is to define a man-machine interface for assisting the driver to complete the compaction construction of asphalt pavement [5]. In the early development of the MACC, a laser positioning system is used. However, due to the laser positioning system requires expensive infrastructure, the cost is too high to be accepted by the construction unit. Then the MACC uses the GPS (Global Position System) as its positioning system. On the basis of the MACC prototype system, the CIRC (computer integrated road construction) project supported by the European Union Brite/EuRam plan, for roller compaction machine, defined and developed a system called CIRCUM aided by driver construction [6, 7]. The user of MACC and CIRCUM is the driver, so only one roller compaction machine is supervised. In 1996, Krishnamurthy and other scholars developed a system called AutoPave that realized the roller compaction machine walking path design and real-time navigation function by using GPS technology, communication technology, and computer technology to reach the requirements of asphalt pavement compaction [8, 9]. The AutoPave system warns and guides the driver to drive according to the design path in the form of graphics and sound, which can achieve uniform compaction of the road surface. In 1996, Oloufa and other scholars developed the Compactor Tracking System (CTS) with the help of GPS technology. Through continuous improvement, CTS-II and CTS-III have been developed successively [10, 11]. Both CTS and CTS-II can only control one roller compaction machine. CTS-III has improved the composition of CTS-II from the point of view of the reality of the compaction of several compacting machines on the road surface and the reduction of equipment costs. Adding a remote processing unit for processing positioning data and using the data transmission wireless communication unit, four roller compaction machines can be tracked in real time. In 2003, the German company BMW, the world's most famous road roller equipment manufacturer, introduced a GPS applied to compacting compaction quality control system called BVC [12]. However, the BVC system is only developed for

the driver of the roller compaction machine, and the construction management unit (such as the supervisor) cannot control it in real time, and the complete system is expensive. The system is only suitable for the use of the equipment of its own company. In summary, the research of compaction quality supervising system has undergone the transformation and development from manual to automation, single to multiple machines, from the driver to the construction management unit, and the laser positioning technology to the GPS technology. The researches mentioned above are mainly focused on the road construction. Compared with road construction, the dam construction has different roller compaction parameters and construction techniques; thus, the systems are not very suitable.

Based on the existing research and our previous research achievements and according to the quality control needs of filling construction of the face rockfill dam, this paper mainly carries on proposing a real-time compaction quality supervising method by the application of GNSS technology [13], wireless data communication technology, Internet of things technology, computer technology, and data processing and analysis technology. A real-time supervisory system for compaction quality is developed, which is suitable for the quality management needs of roller compaction construction in filling engineering. Several benefits are conferred by the proposed method. First of all, the real-time supervisory system for compaction quality automates the acquisition of real-time roller compaction data by using sensor-based technologies and effectively reduces the interference between data collection activities and other construction activities. Then the proposed method can obtain the real-time compaction quality of the working surface, making timely feedback possible for construction control and quality improvement. This helps to ensure high-quality construction while reducing the chance of costly rework. Lastly, the system has comprehensive functions of being real time, continuous, automatic, high precision, and so on. It can be applied to real-time supervising compaction quality of filling construction for dam, highway, airport, and so on and has become an effective assistant to ensure the construction quality.

The rest of this paper is organized as follows: Section 2 presents the research methodology of this study. Section 3 delineates the principles and composition of the real-time compaction quality supervising system. Section 4 introduces the roller compaction parameter calculation method. Section 5 illustrates the application of the real-time compaction quality supervising system in the face rockfill dam of Shuibuya project in China. Finally, Section 6 summarizes the article with a discussion of conclusion and future research directions.

2. Research Methodology

The current research is devoted to the design and development of a software system. The research paradigm that suits this inquiry is design science research (DSR) [14, 15] in which a feasible software system addressing a relevant solution to an unsolved problem is developed and evaluated.

We conduct three phases of a typical DSR tailored for the purpose of this research, as delineated in the following:

Phase 1. Problem identification has been already described in Section 1, that is, how to supervise the construction process in face rockfill dam engineering and make it meet the corresponding requirements. Our research objective is set up as “developing of a new compaction quality supervising system of face rockfill dams during the entire compaction process, which has characteristics of real time, continuous, automatic, and high precision.”

Phase 2. Design and develop through which, the system, constituting four core components, was developed. These components will be described in detail in Section 3.3.

Phase 3. Validate appraises the efficacy of the system resulting from phase 2 through a case study in the face rockfill dam of Shuibuya project in China, providing real-time roller compaction parameters to ensure the construction quality.

3. Principles and Composition of the Real-Time Supervisory System

3.1. Principles of the System. The quality control of roller compaction is mainly the control of roller compaction parameters. According to China Electricity Council [16], the roller compaction parameters, including rolling track, rolling times, rolling speed, thickness, and smoothness, must be supervised and compared with the quality standards during the whole process of roller compaction. These parameters have a common physical quantity: the spatial position. Thickness of filling layer is the elevation difference between two space surfaces. Smoothness of storehouse surface is the concave and convex of the space surface. Driving speed is the distance between two spaces in a unit time. Rolling times are the number of rolling machines through the same space position. Therefore, a real-time, continuous, automatic, high precision spatial positioning system can be used to achieve the quality control of roller compaction.

In the past few decades, automatic real-time positioning technology has developed rapidly. At present, there are mainly two kinds of automatic real-time positioning technology, one is laser positioning and the other is GNSS. GNSS is the standard generic term for satellite navigation systems that provide autonomous geospatial positioning with global coverage. This term includes, for example, the GPS, GLONASS, Galileo, BDS (BeiDou Navigation Satellite System), and other regional systems. Laser positioning accuracy is quite high, but the laser positioning needs a large number of beacon stations, which are expensive and therefore limits its application [5]. GNSS is cheap and the technology is very mature. As a new technology of modern spatial satellite navigation and positioning system, GNSS has gradually replaced the normal optical and electrical surveying equipment in many fields [17–20], and it has been more and more widely used in engineering field. The combination of GNSS technology and modern data communication technology and computer technology makes it possible for applying GNSS to high

precision, real-time, continuous, automatic, and all-weather compaction quality control.

The real-time supervisory system for compaction quality is to place GNSS receivers on the main machines of roller compaction and to control rolling track, rolling times, driving speed, thickness, and smoothness by supervising the change of the spatial position of the roller compaction machines. The quality control method based on GNSS can record the construction process information automatically and in real time and avoid the disadvantages of traditional manual recording, which has the factual basis and strong reliability.

3.2. Precision of the System. The real-time supervisory system for compaction quality uses the carrier phase differential positioning technique, which is also called real-time kinematic (RTK). With the GLONASS system to realize the 24 satellite full deployment, China’s BDS system has the ability to provide positioning services in China and the surrounding areas [21, 22]; compared with the original GPS single system RTK in the past, the three-system (GPS + BDS + GLONASS) RTK greatly increases the number of visible satellites, effectively enhancing the graphic intensity of the observed satellite and improving the accuracy and reliability of the positioning result [23–25]. At present, the nominal accuracy of RTK plane positioning is 1 cm+ 1 ppm. As long as the construction area is in the service range of GNSS reference station, the plane position precision of the system fully meets calculation precision needs of rolling track, rolling speed, and rolling times, which is less than 10 cm. The positioning accuracy of elevation is 2~3 times lower than the plane precision. Considering the data acquisition method and the reasonable elevation fitting model, the elevation precision can achieve $\pm(1\sim 2)$ cm, which can fully meet the quality control needs of thickness of filling layer and smoothness of storehouse surface.

3.3. Composition of the System. The system hardware unit mainly includes the following three parts:

- (1) GNSS satellite signals receiving equipment
- (2) Wireless data communication equipment
- (3) Computers

According to the requirements of compaction quality for filling engineering and construction quality management to system, the system is composed of supervisory center, data center, GNSS reference station, and mobile terminals. Figure 1 is the schematic diagram of the GNSS real-time construction quality supervisory system for filling construction project.

3.3.1. Supervisory Center. Supervisory center is the core of the system, receiving the real-time position information from the mobile stations through wireless data communication and automated processing. From the electronic display screen equipped in the supervisory center, the accurate position and speed of the roller compaction machines in construction and the state of compaction quality of filling engineering can

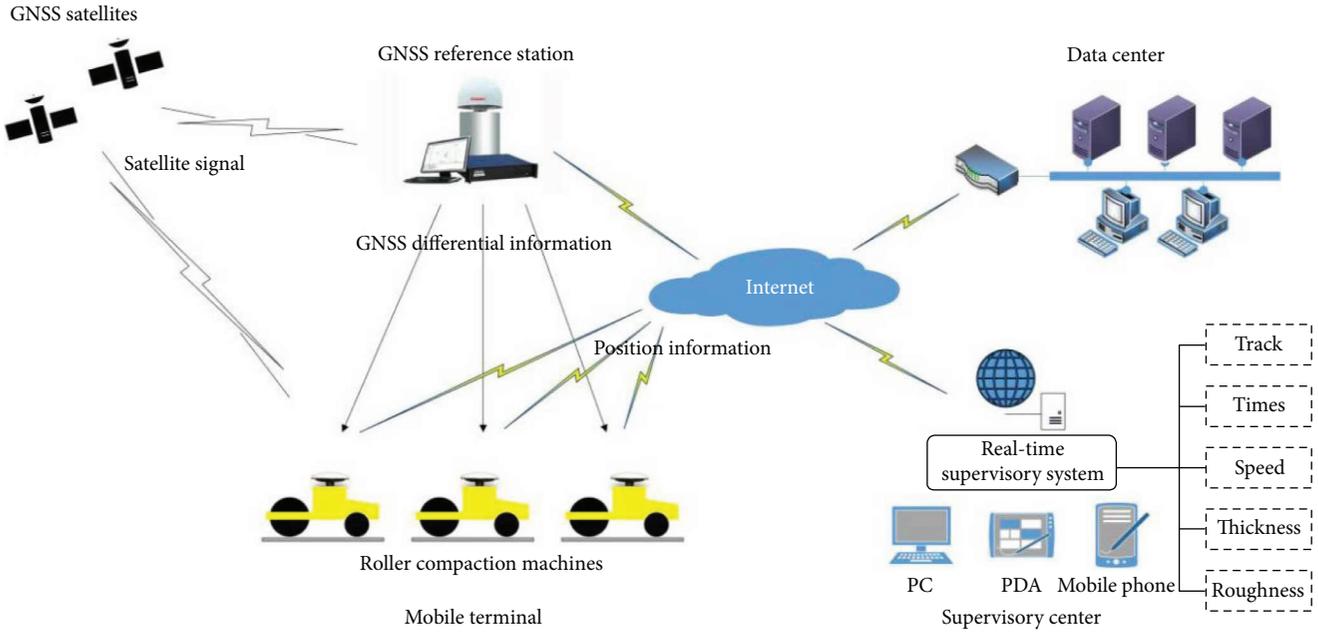


FIGURE 1: Schematic diagram of the real-time supervisory system.

be viewed in real time. The supervisory center is responsible for the data processing, analysis, and other work of the system and runs the information platform for all parties to visit through the Internet to view the real-time construction quality information.

3.3.2. Data Center. The data center is responsible for the storage and management of all data in the system to ensure data security and stability. The supervisory center accesses data center for data access operations through the Internet.

3.3.3. GNSS Reference Station. In order to improve the positioning accuracy, the GNSS reference station is established at a point with known coordinates to provide the carrier phase differential information. The real-time differential data of GNSS reference station is continuously sent to the GNSS mobile stations. GNSS reference station needs to be built in a place that can meet close to the construction area, has a wide-open sky (with few things blocking it like trees and buildings), has a solid foundation, and has a stable network and power supply.

3.3.4. Mobile Terminal. A mobile terminal is mainly composed of an integrated system unit, a GNSS receiver antenna, and a wireless communication antenna. The system unit integrates GNSS receiver, communication module, power module, and so on (Figure 2). Mobile terminal is installed in a roller compaction machine as a GNSS moving station and is responsible for collecting data. With the real-time positioning data, the roller compaction parameters for quality control can be calculated.

3.4. Work Flow of the System. The work flow of the real-time supervisory system for compaction quality is shown in Figure 3.

4. Supervising of Roller Compaction Parameters

4.1. Rolling Track. Rolling track, in fact, is roller wheel's track. As the GNSS antenna is placed on the top of the cab of roller compaction machine, it is necessary to calculate the axis midpoint and the left and right end coordinates of the roller wheel from the coordinates of the GNSS antenna according to the relative position relationship between the GNSS antenna and the roller wheel. After confirming the position of wheel axle of roller compaction machine, the rolling track in the passing sampling time is described by constructing the quadrangle continuously by connecting the roller wheel axis in chronological order. As is shown in Figure 4, it is a schematic diagram of rolling track within a continuous period of time. In the diagram, there are four black thick solid lines of L_1R_1 , L_2R_2 , L_3R_3 , and L_4R_4 , and each of which is the roller wheel axis of the roller compaction machine at four sampling times t_1 , t_2 , t_3 , and t_4 . And T_1 , T_2 , T_3 , and T_4 are the points of the GNSS antenna at each moment, P_1 , P_2 , P_3 , and P_4 are the center points of the rolling wheel axis at each moment, and L_1 , R_1 , L_2 , R_2 , L_3 , R_3 , L_4 , and R_4 are the left and right end points of the roller wheel axis at each moment.

Figure 4 shows that the real rolling track is with a width. In the specific rolling track when painting, in order to express the graph simply and clearly and recognize the mechanical path clearly and legibly, a line is generally used to constantly connect the midpoint of drum axle line to describe. As shown in Figure 4, these lines made by connecting P_1 , P_2 , P_3 , and P_4 and marked with arrows indicate more sections.

4.2. Rolling Times. Rolling times, which are an important parameter of construction quality, refer to the number of roller compaction machines through the same space position. If the number of rolling times is too small and packing



FIGURE 2: Schematic diagram of mobile terminal.

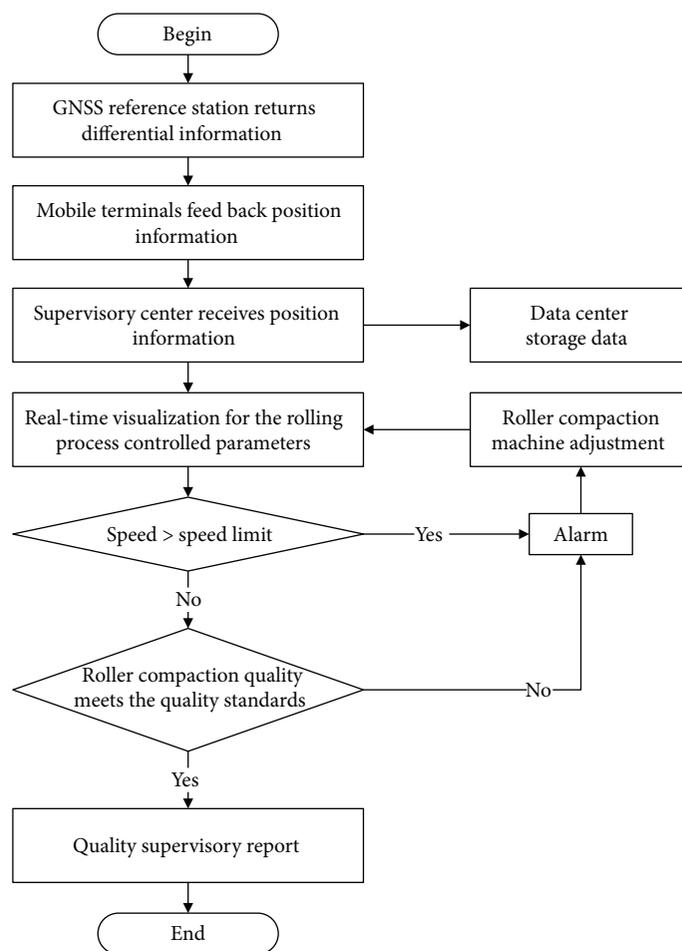


FIGURE 3: Flow chart of the real-time supervisory system for compaction quality.

compression deformation does not reach the standard, it will lead to the consequence that the final compaction, dry density, and other quality evaluation indicators fail to meet the

requirements. Therefore, how to get the crush number of the current area quickly and accurately is the focus of the roller compaction quality monitoring system. The GNSS

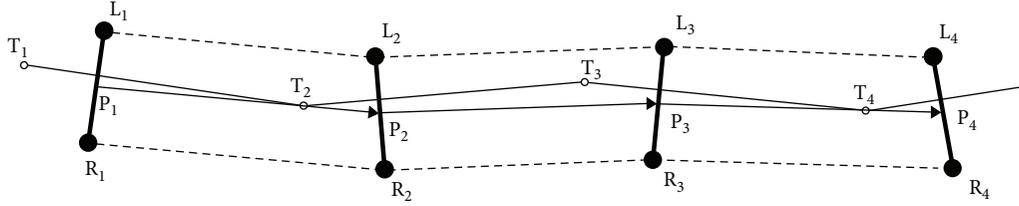


FIGURE 4: Real rolling track in continuous time.

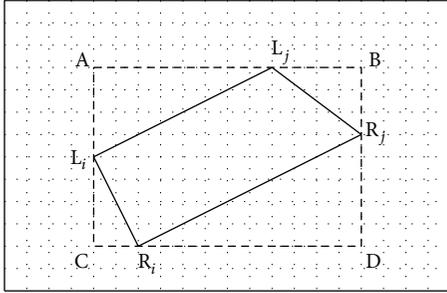


FIGURE 5: Schematic diagram of rolling times calculation by grid method.

real-time supervisory system adopts the grid method to calculate the rolling times.

Grid method is to grid the entire rolling area into an array of regular grid, as shown in Figure 5. In real time, we determine the real rolling track in the time of t_i to t_j is illustrated as quadrangle $L_iR_iR_jL_j$ in Figure 5. Whether or not the grids' centers are in $L_iR_iR_jL_j$ should be identified. If the center of grid is in $L_iR_iR_jL_j$, the rolling times of this grid should be added one more time. The detailed calculation steps are as follows:

- (1) According to the rolling range, the rolling area is divided into lots of grids, and the rolling times of the grid unit are initialized to 0.
- (2) The grid analysis area ABCD is determined according to $L_iR_iR_jL_j$.
- (3) The ray method is used to determine whether the central point of the grid unit in the ABCD is within the quadrangle $L_iR_iR_jL_j$.
- (4) According to the rolling times of grid units, the corresponding color is filled and the rolling time graph (Figure 6) is obtained.

4.3. *Rolling Speed.* Assuming that two consecutive sampling times t_1 and t_2 , and $P_{t_1}(x_{t_1}, y_{t_1}, h_{t_1})$ $P_{t_2}(x_{t_2}, y_{t_2}, h_{t_2})$ are the center points of the rolling wheel axis at each moment, the driving speed of roller compaction machine can be calculated as follows:

$$v = \sqrt{\frac{(x_{t_2} - x_{t_1})^2 + (y_{t_2} - y_{t_1})^2 + (h_{t_2} - h_{t_1})^2}{(t_2 - t_1)}} \quad (1)$$

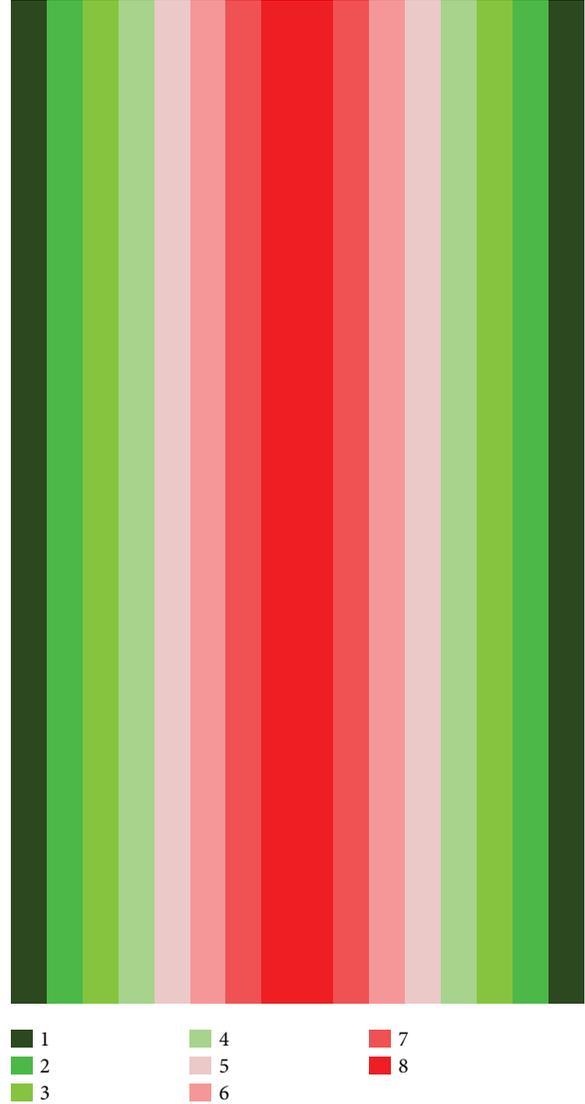


FIGURE 6: An example of the rolling times graph.

In this paper, the speed measurement method is called average speed method, which is long as the selected speed sampling period Δt and the two positioning data P_{t_1} and P_{t_2} , and does not require other new observations. Average speed method is very effective for low speed (usually not more than 4 km/h). With time as the abscissa axis, the velocity as the vertical axis, it can be plotted in a rolling speed diagram (Figure 7).

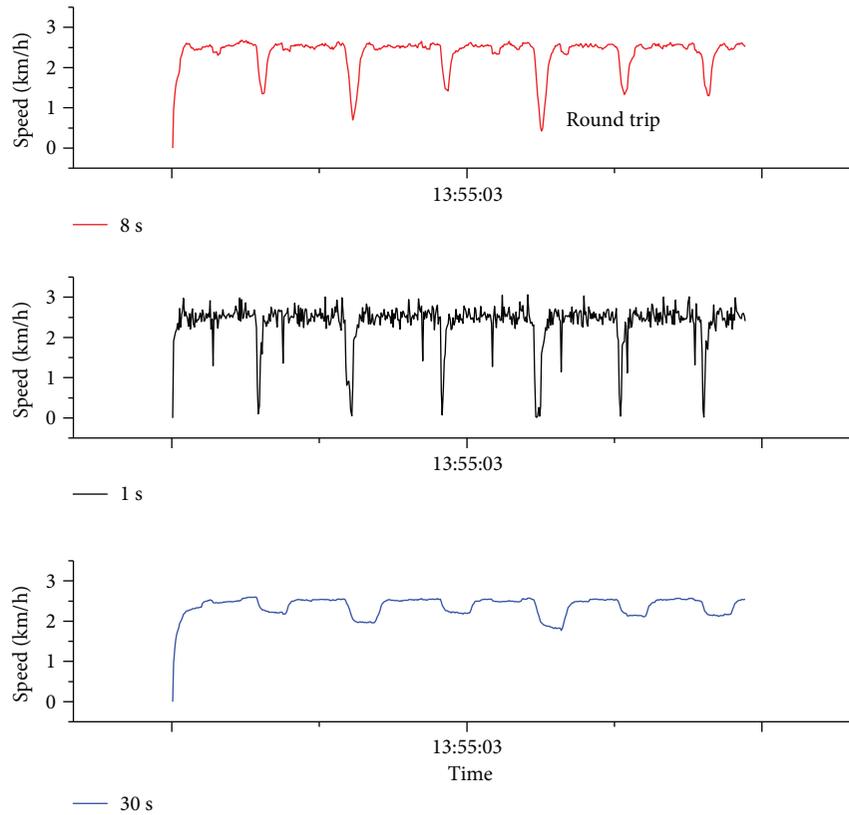


FIGURE 7: Rolling speed diagram.

When using the average speed method, it should be noted that the sampling period Δt should be suitable. If the period is too long or too short, it will not accurately represent the actual speed. In the actual calculation of rolling speed, the value of Δt was tested. When $\Delta t = 1$ s, rolling speed is abrupt, and it is shown in the velocity diagram with a large sawtooth, which is not consistent with the actual speed. When $\Delta t = 30$ s, rolling speed is almost invariable, and it is a straight line in the velocity diagram. After repeated experiments, it is concluded that Δt is suitable for 6~15 s. As shown in Figure 7, it is a typical rolling speed diagram when $\Delta t = 8$ s.

4.4. Thickness of Filling Layer. Thickness of filling layer is the elevation difference between the elevation of the surface after the filling and the elevation of the surface before the filling. The average elevation of a storehouse surface is H_1 , and the average elevation of the previous storehouse surface is H_0 , then the thickness of the filling layer shall be $H_1 - H_0$. As shown in Figure 8, the thickness of the upper filling layer can be obtained at 60 cm.

4.5. Smoothness of Storehouse Surface. Smoothness of storehouse surface is a very important quality control index. The smoothness reflects the concave and convex condition of rolling storehouse surface, which further reflects the distribution of the compaction of the storehouse surface. Through GNSS positioning, it can continuously collect the coordinate information of roller compaction machine. But if the points are discrete and irregular, it is necessary to use

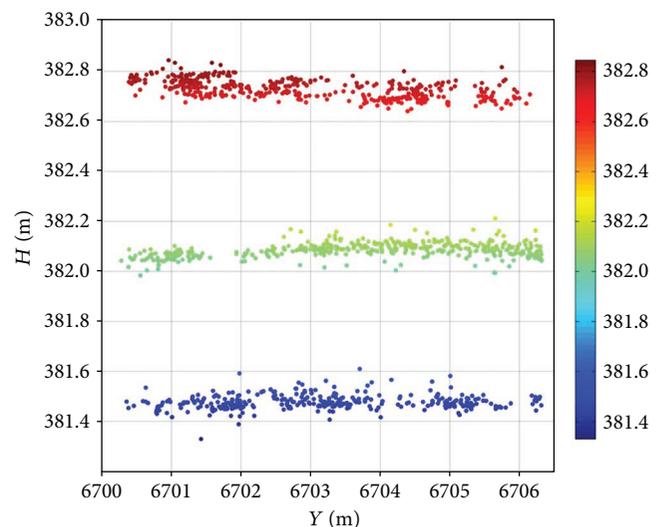


FIGURE 8: An example of thickness calculation.

some interpolation methods to generate regular grid digital elevation model (DEM). And then DEM visualization can graphically describe the smoothness of storehouse surface. In this paper, moving least square surface fitting algorithm is adopted. This method has high fitting precision and smooth surface, which can reflect the smoothness of storehouse surface directly and accurately (Figure 9).

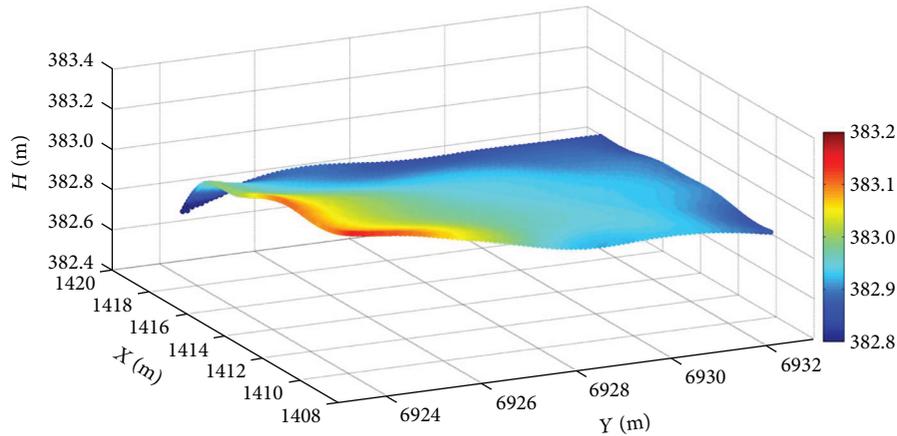


FIGURE 9: An example of storehouse surface after fitting.

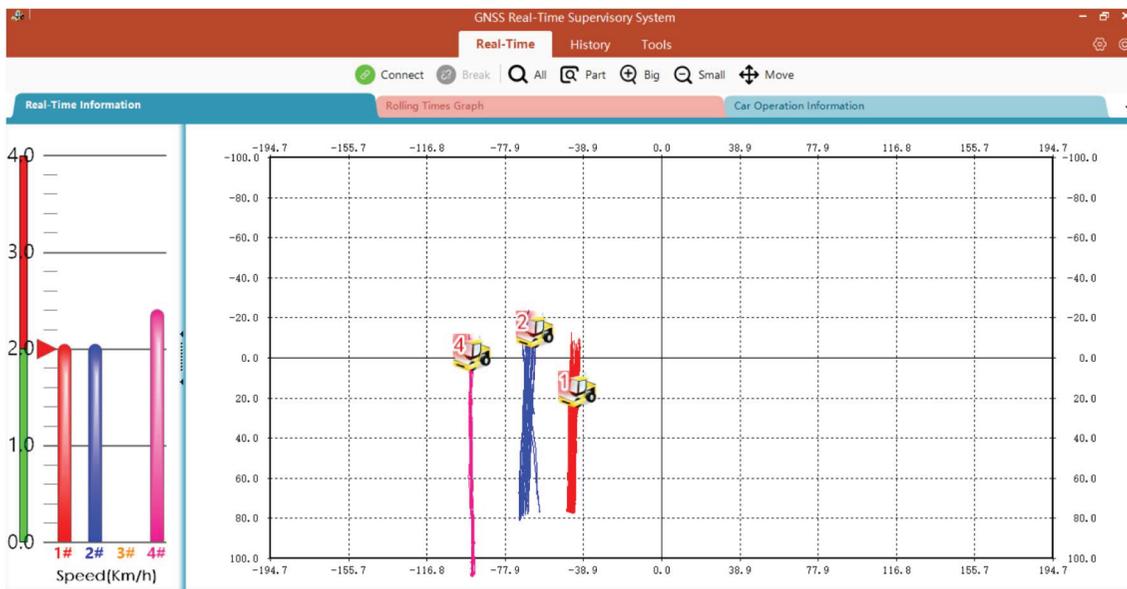


FIGURE 10: Main interface of the real-time supervisory system.

5. Case Study

The face rockfill dam of Shuibuya project in Qingjiang River of Hubei Province, China, ranking first in dam height of the same type in the world, is as high as 233 m. The project has a total filling volume of $1.6 \times 10^7 \text{ m}^3$ and more than $6 \times 10^5 \text{ m}^3$ for peak filling volume of a single month. So the filling construction of the dam needs all-weathered 24 hours without break, and many roller compaction machines work simultaneously. Therefore, the supervisory system has a great meaning for improving the construction quality of Shuibuya project.

The real-time compaction quality supervising system (described in Section 3) is applied in the actual construction process of the Shuibuya dam. The whole operation is in good condition. The main interface of the system is shown in Figure 10, which illustrates the supervising of running speed of roller compaction machines and rolling times at any

position. With the use of this system, the 4 roller compaction machines served for the project are all under supervising real timely. The field supervisors can conduct roller compaction machines to make complementary compaction if poor quality is found in some places. When the compaction process is finished, the charts of rolling tracks (Figure 11), rolling times (Figure 12), and rolling speed will be reported as the supporting documents of construction quality inspection.

6. Conclusions and Future Work

With the development of dam construction technology, face rockfill dams are playing an increasingly significant role in hydraulic and hydroelectric projects. The quality control during the filling construction process is of great significance to the stability and durability of the face rockfill dam. The conventional method obtaining the compaction quality mainly based on some test holes cannot reflect the quality

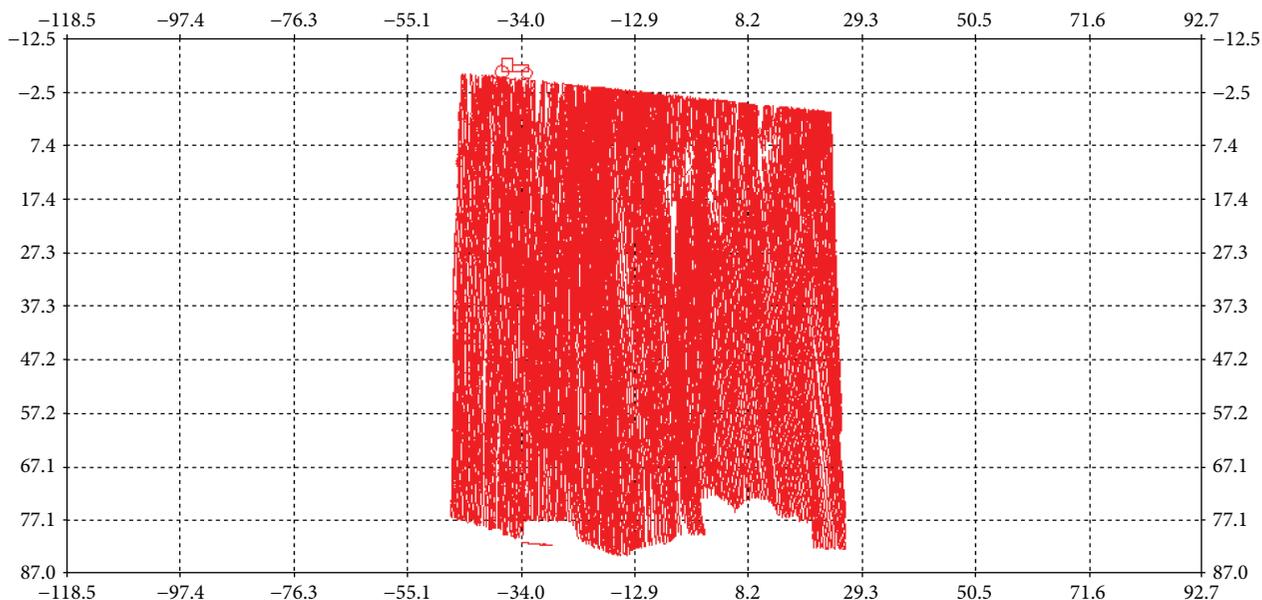


FIGURE 11: Rolling tracks of storehouse surface.

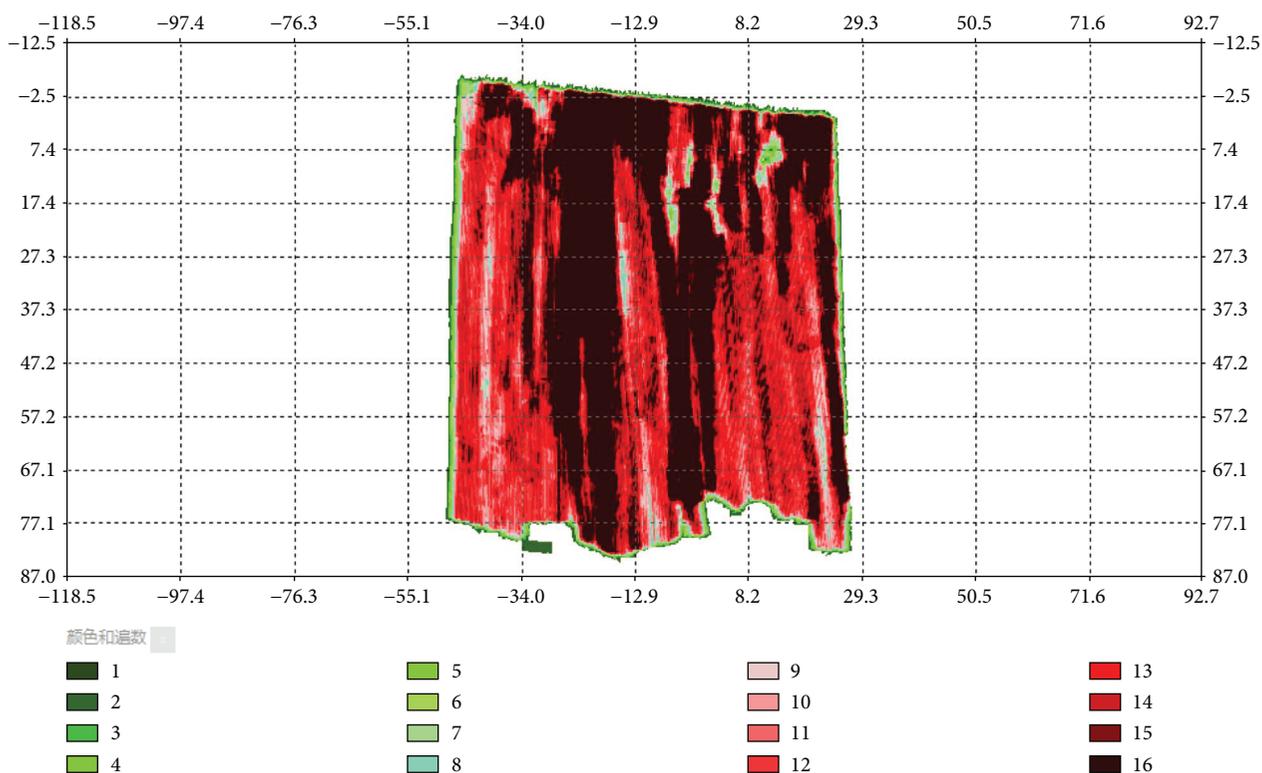


FIGURE 12: Rolling times of storehouse surface.

of an entire construction area. At the same time, the conventional manual quality management mechanism cannot meet the high requirements of modern mechanized construction and schedule anymore. In this study, a real-time compaction quality supervising method with the application of GNSS technology, wireless data communication technology, Internet of things technology, and computer

technology was proposed, which involves data collection, data transmission, data integration, and data analysis. The method solves the real-time roller compaction parameter information by supervising the change of the spatial position of the roller compaction machines to realize monitoring of the rolling construction process. And the method overcomes the shortcomings of traditional methods, realizes the on-site

quality control and process control in all aspects, and has a factual basis in the construction quality monitoring through data recording and storage. A real-time construction quality supervising system for face rockfill dams is developed, and the construction process of the Shuibuya dam is taken as a case study. The results show that the system is effective for construction quality control of face rockfill dams.

However, the compaction quality of face rockfill dams is affected by many factors, including not only the roller compaction parameters listed in this paper but also the excitation force of compaction machines and the water content of the material. Therefore, using more sensors to obtain additional construction information is a clear area for further research. Meanwhile, in order to apply to various compaction machines and benefit supervision of construction quality, the system needs to be further improved and perfected in the aspect of unit bodily form and module integration of mobile terminal.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Shengxiang Huang devised the research, and Wen Zhang and Gen Wu designed the methods and developed the program. All authors wrote the paper. All authors have read and approved the final manuscript.

Acknowledgments

This paper was supported by the National Natural Science Foundation of China (Grant no. 41274020).

References

- [1] R. L. Peurifoy, C. Schexnayder, and A. Shapira, *Construction Planning, Equipment and Methods*, McGraw-Hill, New York, NY, USA, 8th edition, 2010.
- [2] H. Shengxiang, L. Jingnan, and Z. Huai'en, "GPS real-time supervisory system and application in the construction of face rockfill dam," *Geo-spatial Information Science*, vol. 8, no. 4, pp. 240–246, 2005.
- [3] *The Geodynamic Compaction Documentation System*, GEO-DYN, Inc., Stockholm, Sweden, 1985.
- [4] H. F. Thurner and A. Sandstrom, "Compaction meter and compaction documentation system," *Journal of the Acoustical Society of America*, vol. 106, no. 5, pp. 2900–2912, 1989.
- [5] M. Froumentin and F. Peyret, "An operator aiding system for compactors," in *1996 Proceedings of the 13th International Symposium on Automation and Robotics in Construction*, pp. 359–368, Tokyo, Japan, June 1996.
- [6] L. H. Pampagnin, F. Peyret, and G. Garcia, "Architecture of a GPS-based guiding system for road compaction," in *1998 Proceedings. 1998 IEEE International Conference on Robotics and Automation (Cat. No.98CH36146)*, pp. 2422–2427, Leuven, Belgium, May 1998.
- [7] F. Peyret, "The computer integrated road construction project," in *Proceedings of the 16th IAARC/IFAC/IEEE International Symposium on Automation and Robotics in Construction*, pp. 593–598, Madrid, Spain, September 1999.
- [8] B. K. Krishnamurthy, H.-P. Tserng, R. L. Schmitt, J. S. Russell, H. U. Bahia, and A. S. Hanna, "AutoPave: towards an automated paving system for asphalt pavement compaction operations," *Automation in Construction*, vol. 8, no. 2, pp. 165–180, 1998.
- [9] H.-P. Tserng, J. S. Russell, B. K. Krishnamurthy, and R. L. Schmitt, "An operations planning system for asphalt pavement compaction," in *1996 Proceedings of the 13th International Symposium on Automation and Robotics in Construction*, pp. 349–358, Tokyo, Japan, June 1996.
- [10] A. A. Oloufa, W. Do, and H. R. Thomas, "An automated system for quality control of compaction operations: receiver tests & algorithms," in *1999 Proceedings of the 16th International Symposium on Automation and Robotics in Construction*, pp. 67–71, Madrid, Spain, September 1999.
- [11] A. A. Oloufa, "Quality control of asphalt compaction using GPS-based system architecture," *IEEE Robotics and Automation Magazine*, vol. 9, no. 1, pp. 29–35, 2002.
- [12] J. Briaud and J. Seo, *Intelligent Compaction: Overview and Research Needs*, Federal Highway Administration, Washington, DC, USA, 2003.
- [13] R. Zekavat and R. Buehrer, *Overview of Global Navigation Satellite Systems*, John Wiley & Sons, New York, NY, USA, 2011.
- [14] R. H. Von Alan, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004.
- [15] M. Fahmideh and G. Beydoun, "Reusing empirical knowledge during cloud computing adoption," *Journal of Systems and Software*, vol. 138, pp. 124–157, 2018.
- [16] China Electricity Council (CEC), *D/T5129-2013-Specifications for Rolled Earth-Rockfill Dam Construction*, China Electric Power Press, Beijing, China, 2013.
- [17] J. M. Dow, R. E. Neilan, and C. Rizos, "The International GNSS Service in a changing landscape of Global Navigation Satellite Systems," *Journal of Geodesy*, vol. 83, no. 3–4, pp. 191–198, 2009.
- [18] A. Fellner and H. Jafernig, "Implementation of satellite techniques in the air transport," *Reports on Geodesy and Geoinformatics*, vol. 100, no. 1, pp. 39–53, 2016.
- [19] S. Huang, C. Li, and L. Luo, "Landslide monitoring network establishment within unified datum and stability analysis in the Three Gorges Reservoir Area," *Journal of Sensors*, vol. 2017, Article ID 9819214, 13 pages, 2017.
- [20] R. Sabatini, L. Rodriguez, A. Kaharkar, C. Bartel, and T. Shaid, "Carrier-phase GNSS attitude determination and control system for unmanned aerial vehicle applications," *ARNP Journal of Aeronautics & Aerospace Engineering*, vol. 2, no. 11, pp. 297–322, 2012.
- [21] Y. X. Yang, "Progress, contribution and challenges of compass/Beidou satellite navigation system," *Acta Geodaetica et Cartographica Sinica*, vol. 39, no. 1, pp. 1009–1018, 2016.
- [22] China Satellite Navigation Project Center, *Compass/Beidou Navigation Satellite System Development*, CSNPC, Beijing, China, 2009.
- [23] L. Wanninger, "Carrier-phase inter-frequency biases of GLONASS receivers," *Journal of Geodesy*, vol. 86, no. 2, pp. 139–148, 2012.

- [24] Z. A. Ye and J. G. Miao, "Research on accuracy and reliability of GNSS-RTK combination positioning," *Journal of Navigation and Positioning*, vol. 4, no. 3, pp. 77–81, 2016.
- [25] Y. B. Yao, M. X. Hu, and C. Q. Xu, "Positioning accuracy analysis of GPS/BDS/GLONASS network RTK based on DREAMNET," *Acta Geodaetica et Cartographica Sinica*, vol. 45, no. 9, pp. 1009–1018, 2016.

Research Article

Performance Analysis of Network-RTK Techniques for Drone Navigation considering Ionospheric Conditions

Tae-Suk Bae  and Minho Kim

Department Geoinformation Engineering, Sejong University, Seoul 05006, Republic of Korea

Correspondence should be addressed to Tae-Suk Bae; baezae@sejong.ac.kr

Received 26 February 2018; Accepted 15 April 2018; Published 10 May 2018

Academic Editor: Hyung-Sup Jung

Copyright © 2018 Tae-Suk Bae and Minho Kim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, an accurate positioning has become the kernel of autonomous navigation with the rapid growth of drones including mapping purpose. The Network-based Real-time Kinematic (NRTK) system was predominantly used for precision positioning in many fields such as surveying and agriculture, mostly in static mode or low-speed operation. The NRTK positioning, in general, shows much better performance with the fixed integer ambiguities. However, the success rate of the ambiguity resolution is highly dependent on the ionospheric condition and the surrounding environment of Global Navigation Satellite System (GNSS) positioning, which particularly corresponds to the low-cost GNSS receivers. We analyzed the effects of the ionospheric conditions on the GNSS NRTK, as well as the possibility of applying the mobile NRTK to drone navigation for mapping. Two NRTK systems in operation were analyzed during a period of high ionospheric conditions, and the accuracy and the performance were compared for several operational cases. The test results show that a submeter accuracy is available even with float ambiguity under a favorable condition (i.e., visibility of the satellites as well as stable ionosphere). We still need to consider how to deal with ionospheric disturbances which may prevent NRTK positioning.

1. Introduction

The Global Navigation Satellite System (GNSS) is widely used for accurate positioning in various civil applications. The stand-alone positioning using C/A code pseudoranges is the most popular approach for aviation and/or ground vehicle navigation. However, a submeter or better accuracy is required for Mobile Mapping System (MMS) or smart cars for identifying lanes, especially, the Lane Departure Warning System (LDWS) or Advanced Driver Assistance Systems (ADAS). Contrary to the ground vehicle, there is no designated paths for drones, which make it difficult to provide accurate control points. Rather than code-based GNSS positioning in either stand-alone or differential mode, the carrier phase observables are commonly used for precision geodetic and surveying applications where an accurate positioning is the most important consideration. A Real-time Kinematic (RTK) approach can provide a centimeter-

level accuracy in a favorable environment. A more accurate and stable positioning technique was first proposed in mid-1990 by rigorously modeling all error components based on a network of known references, called Network-based RTK (NRTK) technique [1–3]. Since then, several methods were developed and applied in practice such as Virtual Reference Station (VRS), Flächen Korrektur Parameter (FKP), and Master-Auxiliary Concept (MAC), which are basically similar approaches [4–7].

With a widespread use of drones in the field of geoinformation, we need to understand the accuracy requirements for the topographic mapping purpose in an open sky. For an accurate position of the camera exposure, it is necessary to achieve decimeter accuracy of the drones in 3-dimensional. Even with the carrier phase observations, it is challenging to reach such an accuracy without proper integer ambiguity resolution. However, it is well known that fixing integer ambiguity of the carrier phase

signals is difficult, or sometimes exaggerated errors may occur on the resolved ambiguities depending on the geometric configuration and/or surrounding environment.

Therefore, we analyzed the stability and the accuracy of NRTK systems under various conditions to verify the possibility of applying this to drone navigation. We used the currently available NRTK systems in Korea for the test. The National Geographic Information Institute (NGII) of Korea is operating two NRTK services, that is, VRS and FKP [3, 8]. The reference stations of Korea are densely located ranging from 30 to 50 km for most of the baselines, which should be a worthy infrastructure for NRTK positioning.

2. NRTK System

2.1. Overview of NRTK System. The accuracy and stability of NRTK are considerably affected by the ambiguity resolution of the carrier phases [9–11]. However, it is difficult to accomplish accurate resolution due to many factors such as insufficient number of observations, the distance to the reference stations, the signal blockage near the moving rover, and the erratic ionospheric condition [12]. Therefore, it is necessary to understand the relationship with an external information (e.g., ionospheric conditions) to apply for drone navigation, especially for topographic mapping purposes in an open sky. These analyses were conducted in two ways, that is, (1) internal factors of NRTK technique and GNSS receiver specification and (2) the satellite configuration, the behavior of the rover, and the space radio signal disturbance as external factors.

The NRTK correction information is generally transmitted to the user (mostly the rover) via Radio Technical Commission for Maritime Services (RTCM) format, although the proprietary format such as Compact Measurement Record (CMR) or CMR+ is also available [12–14]. The RTCM Special Committee (SC) 104 on Differential GNSS (DGNSS) provides the standards for differential GPS and RTK operations. The RTCM was mainly used for the correction information in this study.

2.2. NRTK Reference Network of Korea. The actual NRTK surveying was conducted for kinematic positioning of ground vehicles across the country, instead of a real drone like a multicopter, due to the limitation of payload of the receivers. In addition, the static NRTK positioning was also carried out using zero-baseline configuration. The reference network of Korea at the time of the experiment is shown in Figure 1.

Although two NRTK systems were operational by the same institute (NGII), the reference network for the two systems is not the same due to the different licensing policy. In addition, the spatial density of the NRTK reference network is not identical to both systems. While VRS uses all Continuously Operating Reference Stations (CORS) for modeling of the correction information, only 30 CORS are involved in the network of FKP. The performance and the stability of the two systems can change depending on the internal and/or external environment of the receiving system (mostly the rover). The status of the rover can be either static

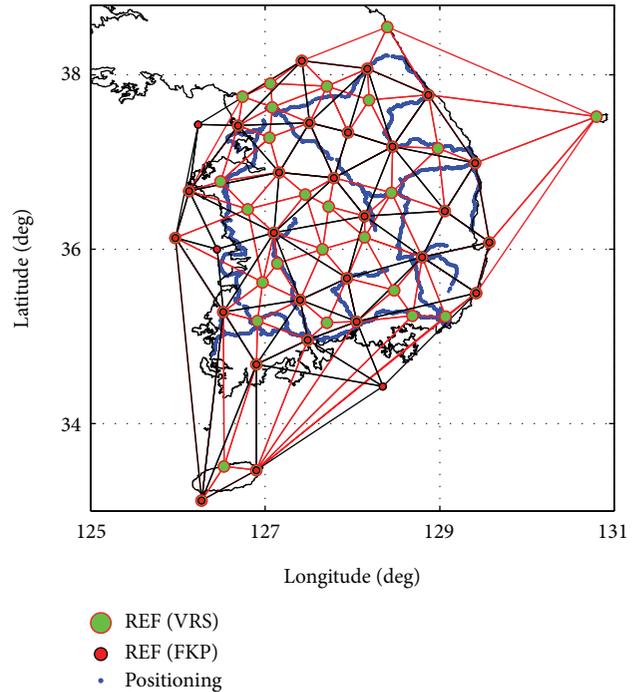


FIGURE 1: Configuration of NRTK network of Korea [8].

(stationary) or kinematic (moving), and the same condition should be guaranteed during the experiment except the NRTK techniques.

2.3. System Configuration of the Study. We set up a system to test NRTK performance and stability in three cases. Figure 2 shows the basic system configuration where the GNSS antenna (Leica GRX1200) was installed in an open sky rooftop by splitting the signal to ensure the identical condition of the two systems. The correction information was sent to the receiver via serial port, and the NRTK solutions were transferred back to the main computer through the same ports. Each receiver was set up to conduct NRTK positioning with the corresponding correction information of VRS or FKP.

The two other cases of the study have basically the same system configuration, but some modification was applied to the GNSS receivers, antenna, and/or the mode of operation (static or kinematic).

3. Experiments

It is generally known that the integer ambiguity resolution is essential for high-accuracy NRTK solutions based on the carrier phase observations. We used a dual-frequency GNSS receiver as well as a single-frequency low-cost receiver for static and/or kinematic positioning. Three experiments were conducted to evaluate the performance of NRTK systems in different environments and the effect of positioning modes (see Table 1). First, since the NRTK performance is generally affected by the atmosphere [15, 16], especially ionospheric disturbances, we analyzed two types of NRTK systems under restricted experimental conditions (case 1). For this

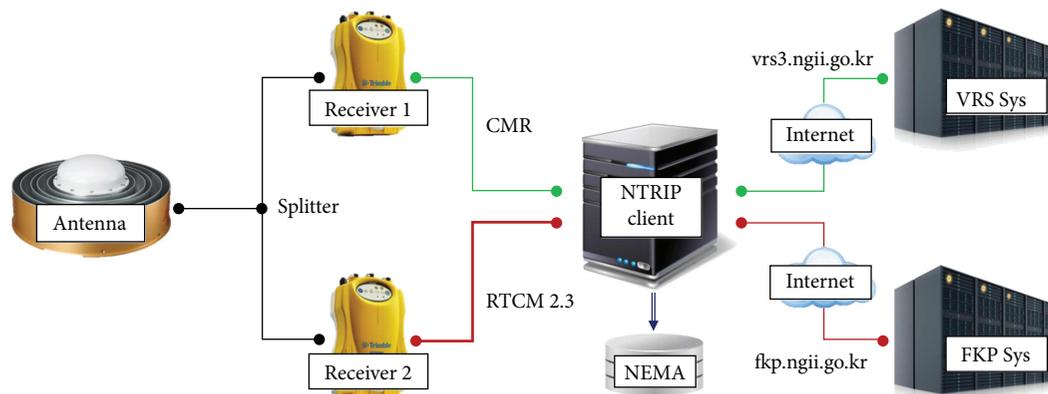


FIGURE 2: Systematic configuration for the comparison of static NRTK solutions (VRS versus FKP) [3].

TABLE 1: Overview of the experiment for NRTK positioning.

Component	Case 1		Case 2		Case 3	
	VRS	FKP	VRS	FKP	VRS	VRS
Receiver	Trimble 5700		Trimble R6		Trimble R6	
Frequency	L1/L2		L1/L2		L1/L2	
Sampling	1 Hz		1 Hz		1 Hz	
Period	1 month		5 days		1.5 h (ublox) 11 days (Trimble R6)	
GNSS	GPS		GPS + GLO		GPS + GLO	
Antenna	Leica AT504GG		Internal		Internal	
Protocol	CMR	RTCM 3.1	CMR	RTCM 3.1	RTCM 3.1	
Common epochs	165,719		7516		3462	
Fixed epochs	116,023		6989		3251	
Software	Internal	Internal	Internal	Internal	Internal	RTKLIB

experiment, the same type of GNSS receivers was connected to a single antenna on the rooftop of a building by splitting the signals. The main objective is to evaluate the effect of ionospheric conditions on the NRTK performance.

The second test (case 2) was carried out in a static mode where both receivers were connected to a single antenna, similarly to case 1. In this case, however, we analyzed the performance of two NRTK techniques with identical receivers, as well as different grades of GNSS receivers. The last experiment (case 3) is a kinematic NRTK positioning with a low-cost receiver with respect to a geodetic grade receiver. Every single epoch of NRTK solutions was analyzed for 24 hours. In kinematic positioning, the success rate of the ambiguity resolution was mainly evaluated along with relative baselines with respect to the reference solution. All experiments used the RTCM message as a principal data format, and CMR/CMR+ was also used for some VRS operations.

3.1. Effect of Ionospheric Condition (Case 1). As already known in GNSS community, the double-differenced observables can eliminate the satellite and receiver clock biases, orbit errors, and interchannel biases [17]. However, atmospheric delays, such as ionosphere and troposphere, still

remain in the observables; thus, these should be eliminated for precision positioning and stability of the system. A new approach was proposed to deal with these atmospheric delays in RTK [1]. It is difficult to eliminate the ionospheric residual errors even with NRTK when the atmospheric disturbance occurs due to a solar flare. Therefore, it is still a critical issue to minimize the effects of space radio signal disturbances on NRTK. The Korean Space Weather Center (KSWC) forecasts the disturbance grades for every event. As can be seen in Table 2, several geomagnetic disturbances and radio black-outs happened during the entire period of the experiment from September 26 to October 31, 2013. We compared the effects of solar flares, radio signal blockages, and geomagnetic disturbance on NRTK and analyzed the degradation of the system performance by these disturbances.

There was no significant degradation of the NRTK accuracy and stability in the case of the disturbance stage 1 (see Figure 3). However, the FKP system seems not to work properly for stage 2 or 3. Particularly, two warnings in stage 3 on October 25, 2013, prohibited the users from accessing the FKP server, and there were cases that the performance was highly correlated with the disturbances. Therefore, it is necessary to consider the signal disturbances to apply NRTK for the navigation of the drones.

TABLE 2: Disturbances and radio signal interruptions during the experiment [8].

Type	Local time (start)	Grade	Local time (end)	Max. value
Geomagnetic disturbances	2013-10-09 06:00	G1	2013-10-09 15:00	Kp = 5
Geomagnetic disturbances	2013-10-09 18:00	G1	2013-10-09 21:00	Kp = 5
Radio blackout	2013-10-24 09:26	R2	2013-10-24 09:36	M9.3
Radio blackout	2013-10-25 16:59	R3	2013-10-25 17:09	X1.7
Radio blackout	2013-10-25 23:58	R3	2013-10-26 00:12	X2.1
Radio blackout	2013-10-28 10:58	R2	—	M9.5
Radio blackout	2013-10-28 11:02	R3	2013-10-28 11:13	X1.0
Radio blackout	2013-10-28 13:41	R2	2013-10-28 13:46	M5.1
Radio blackout	2013-10-30 06:48	R2	—	M9.4
Radio blackout	2013-10-30 06:50	R3	2013-10-30 07:01	X2.3

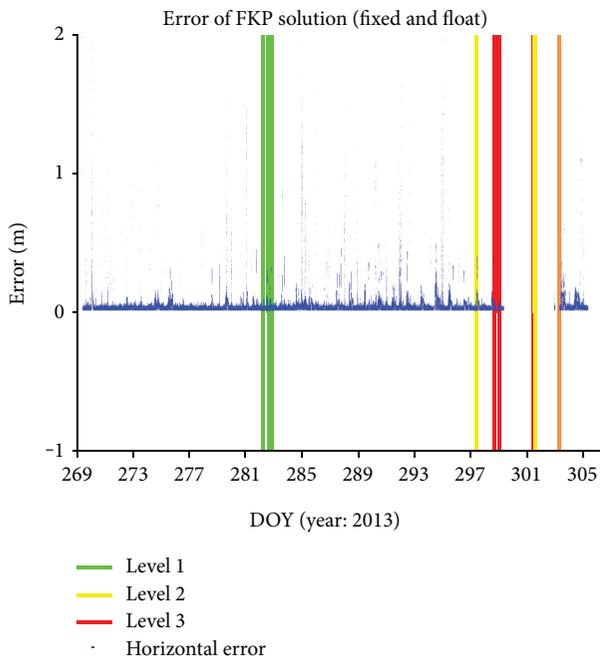


FIGURE 3: NRTK positioning errors (fixed and float solutions) with respect to the ionospheric disturbances [3].

3.2. *Characteristics of NRTK Systems (Case 2)*. For the analysis of the NRTK characteristics, we conducted a long-term experiment with two currently operating NRTK systems of Korea. Two identical GPS receivers were connected to a single antenna on the open sky rooftop. The resulting output, GGA message (fix information) of National Marine Electronics Association (NMEA) sentences, was saved at 1 Hz rate for 5 days, and the CMR (VRS) and RTCM 2.3 (FKP) were used as a communication protocol to compare different NRTK techniques (see Table 3).

Based on 5 days of NRTK survey in a static mode, the integer ambiguities were resolved to be 99.6% and 88.6% on average for VRS and FKP, respectively. The horizontal error of VRS solutions was less than 5 cm for all epochs with fixed ambiguities, while almost 94% of FKP solutions were ensured within 5 cm. Therefore, it can be concluded

TABLE 3: Summary of two NRTK surveying [18].

UTC	2012-12-15-19 (5 days)
Receiver/antenna	Trimble 5700/Leica AT504GG
NRTK	VRS/FKP
Protocol	CMR (VRS), RTCM2.3 (FKP)
GNSS	GPS

TABLE 4: Accuracy of NRTK solutions [18].

Component	Error (m)	Fixed solution (%)		Ambiguity fixing rate (%)	
		VRS	FKP	VRS	FKP
North	≤0.03	99.7	80.8	99.6%	88.6%
East		100	93.1		
North	≤0.05	100	93.4		
East		100	95.5		

that VRS is relatively more stable than FKP in a normal condition (see Table 4).

On the other hand, the daily solution shows that ambiguity fixing rate for VRS was consistently stable for the entire period. However, there were some variations in the case of FKP, ranging from 82% to 91%, particularly almost a 17% difference on December 17, 2012 (see Figure 4).

To analyze the reason of this gap, we classified the data into two groups, that is, daytime (UTC 0–9, which corresponds to 9 am to 6 pm in local time) and nighttime (UTC 9–24). It shows that the ambiguity fixing rate of FKP during daytime is clearly worse than nighttime, while there is no significant difference for the VRS system (see Table 5). Based on the results, it can be concluded that the FKP system is more sensitive to ionospheric disturbances than VRS. The overall ambiguity fixing rate reached 97.92% when adding GLONASS in the long run [19]. In addition, FKP shows similar statistics as VRS in positioning accuracy when the integer ambiguities are fixed (3.7 ± 2.1 cm). Therefore, GLONASS needs to be included when the FKP system is used, which will certainly increase the stability of the NRTK positioning results.

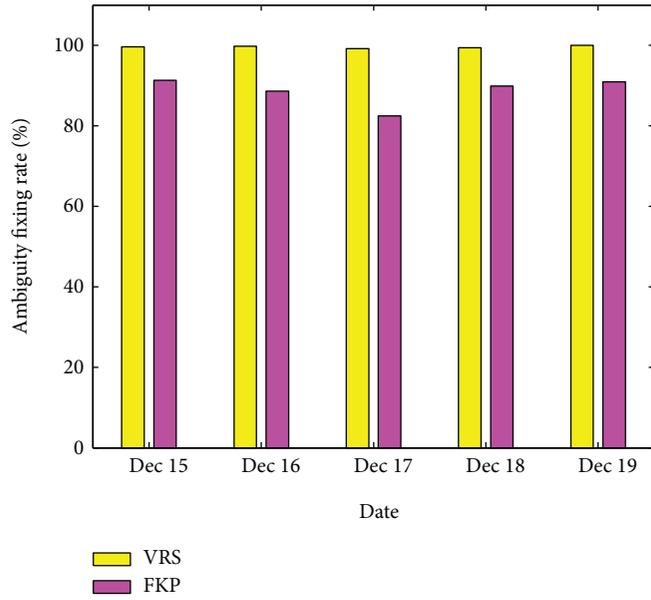


FIGURE 4: Daily ambiguity fixing rates.

TABLE 5: Ambiguity fixing rate by the time of day [18].

Time	Ambiguity fixing rate (%)	
	VRS	FKP
Daytime (UTC 00–09)	99.2	70.9
Nighttime (UTC 09–24)	99.2	94.2

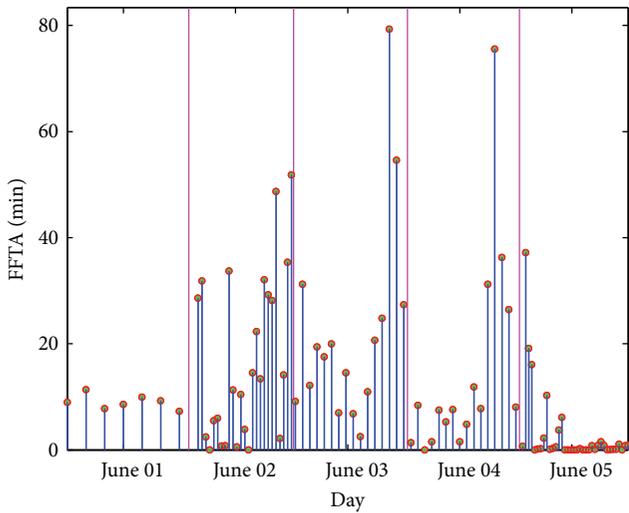


FIGURE 5: Time to first ambiguity fixing (TTFA) by dates.

3.3. *Mobile Application of NRTK (Case 3)*. The NRTK correction information was applied to kinematic positioning with the same system configuration as the static case. The geodetic class receiver equipped with an internal antenna (Trimble R6) was used at one hand, where NRTK positioning was accomplished within the receiver. As a rover, the low-cost single-frequency GNSS receiver (ublox EVK-6T) was connected to the same antenna for checking receiver

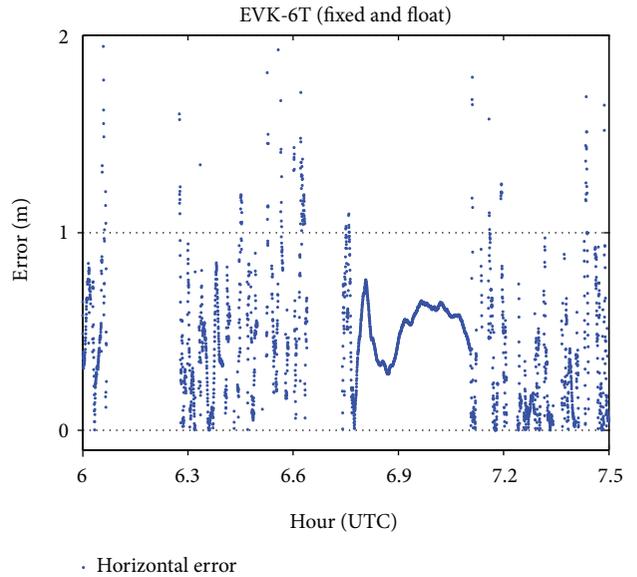


FIGURE 6: Comparison of NRTK solutions in kinematic mode.

performance, and the NRTK positioning was done by RTKLIB open source software.

Since the ambiguity resolution is the key to NRTK positioning, we analyzed the time to first ambiguity fixing (TTFA) based on 100 times of practices for the low-cost GNSS sensors. The interesting thing is that even the low-cost receiver shows a high performance due to the antenna of GNSS CORS, as was seen in the static case. The results showed unstable duration ranging from several seconds up to 80 minutes, and the integer ambiguities were fixed within several minutes for more than half of epochs. Figure 5 shows the TTFA result that behaves differently depending on the date of observation. The best performance in ambiguity fixing time was accomplished on June 5, 2015.

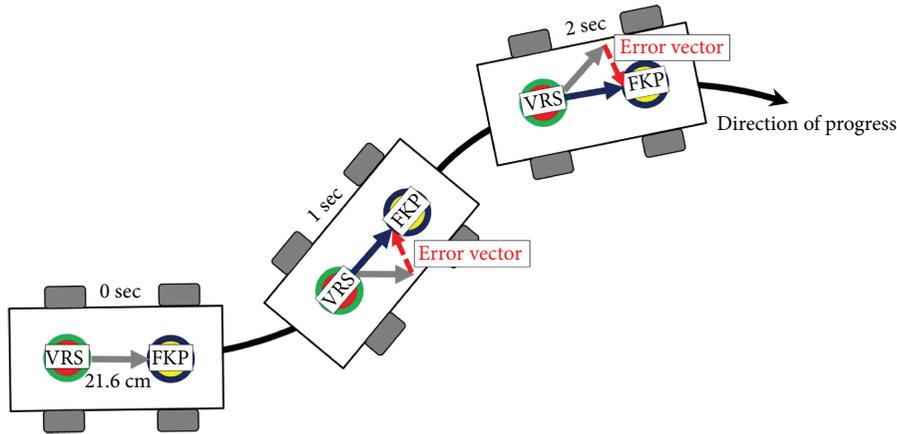


FIGURE 7: System configuration of two NRTK systems.

It is understood that the initialization was not performed properly when reestablishing GNSS receivers to measure TTFA, and the ionospheric condition may be considerably related with the performance of the ambiguity resolution. Therefore, it is concluded that the initial ambiguity fixing can be accomplished within several minutes on average. Since the dual-frequency GNSS receivers usually fix the integer ambiguity in a few seconds, the ambiguity fixing rates can be significantly dropped down for moving single-frequency receivers.

As an additional experiment, we analyzed the performance of NRTK based on a low-cost GNSS receiver in moving environment but with a patch-type antenna. As mentioned before, the baseline length between detachable GNSS antennas was estimated beforehand. Overall 1043 epochs of measurements were used to estimate the horizontal baseline lengths, resulting in about 1.52 m with a standard deviation of 0.58 m. Therefore, the bias can affect the NRTK results at a level of about 1 meter. Figure 6 shows the result of the EVK-6T solution with respect to the reference solution by Trimble R6. As expected, it is difficult to resolve the integer ambiguities with low-cost GNS receivers in a moving condition (only 8% of epochs were fixed to integer). Even the solution with fixed ambiguities shows a similar accuracy as float solution. Since the low-cost receiver can hardly fix the integer ambiguities and may produce outliers in kinematic mode, we set up criteria to eliminate these outliers to calculate the reasonable horizontal baseline lengths. If we assume the outliers to be over 1.5 m, there is no significant difference in the stability between a fixed solution (1.14%) and a float solution (2.5%), and the outliers are at a level of a few meters.

Lastly, two geodetic grades of the NRTK receiver with the same type (Trimble R6) were firmly installed on top of a roof of a car with the baseline length of 21.6 cm (see Figure 7).

Each R6 receiver is connected to a VRS and a FKP server to perform NRTK positioning, and the horizontal baseline vector was estimated to evaluate the performance of both systems. The calibrated baseline length is 21.6 cm which is considered to be the true value. Assuming the direction

TABLE 6: The combined statistics of NRTK for the horizontal component [3].

	Fixed	Float
Mean (m)	0.03	2.90
RMSE (m)	0.07	10.36

remains constant during the consecutive epochs, we calculated the horizontal component of the vector for the analysis of NRTK navigation performance. A total of 11 days of NRTK tests were conducted across the country (see the blue lines in Figure 1). Table 6 shows the statistics of relative positioning of two NRTK solutions. The solution can provide a result within several centimeters of precision for a fixed ambiguity, but several meters of uncertainty with about 1.5 m of bias is expected for the float solution. The integer ambiguity fixing rates are 78.42% and 68.29% for VRS and FKP, respectively. When compared with the static NRTK positioning, the ambiguity resolution rates were decreased about 20%, but still VRS system shows slightly better performance than FKP.

4. Summary and Conclusions

We analyzed the effect of the atmospheric disturbances on NRTK positioning as well as the performance of the system in static and/or kinematic mode. A decimeter level accuracy is preferable for the high-accuracy drone navigation for topographic mapping in an open sky. Three cases of experiments were conducted for two currently operating NRTK systems (VRS and FKP) in this study. Since the ambiguity resolution is the most important factor for precision NRTK positioning, we analyzed the performance and the stability of the NRTK system in several aspects: the (1) NRTK techniques, (2) GNSS receiver performance, and (3) status of the rover. In a static mode, there is a difference in the NRTK performance depending on the technique, that is, VRS showed a higher stability than FKP. FKP seems to be influenced by the ionospheric condition when the results are classified into time of observation.

However, if GLONASS is combined, the stability can be considerably improved as pointed out in the previous section. Even the NRTK solutions by a low-cost receiver do not show an apparent difference in terms of accuracy and stability if connected to a high-quality GNSS antenna. However, the geodetic class antenna is not feasible for the drone navigation; the patch-type antenna was tested as well.

Kinematic NRTK positioning also shows a different characteristic depending on NRTK techniques, and VRS generally performs better than FKP in terms of stability and accuracy. Time of observation does not affect the NRTK performance significantly, but overall the NRTK positioning during daytime shows a better accuracy by several centimeters. The low-cost receiver cannot resolve integer ambiguities in kinematic mode; thus, the performance cannot be guaranteed. Nevertheless, the NRTK positioning of a low-cost receiver provides a submeter level accuracy even under float ambiguities in an open sky, which is a promising result for drone navigation. It should be considered that even a drone navigation sometimes encounters a GNSS signal blockage when modeling buildings in 3-dimensional with detail or taking images under the bridge. Therefore, we need to consider integrating other sensors such as inertial measurement unit (IMU) for the seamless navigation.

The conventional NRTK provides the combined total error of the GNSS observables thus is difficult to model stochastic properties to each specific error source. Another representation of NRTK errors, called the state space representation (SSR), models all physical errors on GNSS observations separately, resulting in effective and flexible correction due to the spatiotemporal knowledge of the parameters [20]. Since a single-frequency operation is possible in SSR, it should be considered for the navigation of drones in the further research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by a grant (18RDRP-B076564-05) from Regional Development Research Program funded by the Ministry of Land, Infrastructure and Transport of the Korean government.

References

- [1] G. Wubbena, A. Bagge, and M. Schmitz, "Network-based techniques for RTK applications," in *Presented at the GPS Symposium, GPS JIN 2001, GPS Society, Japan Institute of Navigation, 2001*, pp. 14–16, Tokyo, Japan, 2001.
- [2] L. Wanninger, "Introduction to network RTK," in *International Association of Geodesy, Working Group 4.5.1: Network RTK*, Dresden, Germany, 2008 <http://www.wasoft.de/e/iagwg451/wegener/communication.html>.
- [3] M. H. Kim and T. S. Bae, "Preliminary analysis of network-RTK for navigation," *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, vol. 33, no. 5, pp. 343–351, 2015.
- [4] B. Park, *A Study on Reducing Temporal and Spatial Decorrelation Effect in GNSS Augmentation System [Ph.D. thesis]*, Seoul National University, Seoul, Korea, 2008.
- [5] F. Takac and O. Zelzer, "The relationship between network RTK solutions MAC, VRS, PRS, FKP and i-MAX," in *Proceedings of ION GNSS 2008*, pp. 348–355, Savannah, GA, USA, 2008.
- [6] V. Janssen, "A comparison of the VRS and MAC principles for network RTK," in *International Global Navigation Satellite Systems Society, IGNS Symposium 2009*, pp. 1–13, Surfers Paradise, QLD, Australia, 2009.
- [7] X. Zou, W. M. Tang, M. R. Ge, J. N. Liu, and H. Cai, "New network RTK based on transparent reference selection in absolute positioning mode," *Journal of Surveying Engineering*, vol. 139, no. 1, pp. 11–18, 2013.
- [8] M. Kim, *Algorithm Development of Integrated Navigation System for Ground Vehicle Using Network-RTK [M.S. thesis]*, Sejong University, Seoul, Korea, 2015.
- [9] X. Zou, M. Ge, W. Tang, C. Shi, and J. Liu, "URTK: undifferenced network RTK positioning," *GPS Solutions*, vol. 17, no. 3, pp. 283–293, 2013.
- [10] B. Li, Y. Shen, Y. Feng, W. Gao, and L. Yang, "GNSS ambiguity resolution with controllable failure rate for long baseline network RTK," *Journal of Geodesy*, vol. 88, no. 2, pp. 99–112, 2014.
- [11] T. S. Bae, D. Grejner-Brzezinska, G. Mader, and M. Dennis, "Robust analysis of network-based real-time kinematic for GNSS-derived heights," *Sensors*, vol. 15, no. 10, pp. 27215–27229, 2015.
- [12] H. J. Euler and B. E. Zebhauser, "The use of standardized network RTK messages in rover applications for surveying," in *Proceedings of the 2003 National Technical Meeting of The Institute of Navigation*, vol. 1, pp. 377–384, Anaheim, CA, USA, 2003.
- [13] N. Talbot, G. Lu, T. Allison, and U. Vollath, "Broadcast network RTK-transmission standards and results," in *Proceedings of the 15th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2002)*, pp. 2379–2387, Portland, OR, USA, 2002.
- [14] W. Tang, X. Meng, C. Shi, and J. Liu, "Algorithms for sparse network-based RTK GPS positioning and performance assessment," *The Journal of Navigation*, vol. 66, no. 3, pp. 335–348, 2013.
- [15] C. Wang, Y. Feng, M. Higgins, and B. Cowie, "Assessment of commercial network RTK user positioning performance over long inter-station distances," *Journal of Global Positioning Systems*, vol. 9, no. 1, pp. 78–89, 2010.
- [16] R. Odolinski, "Temporal correlation for network RTK positioning," *GPS Solutions*, vol. 16, no. 2, pp. 147–155, 2012.
- [17] T. S. Bae, "Parametric analysis of the solar radiation pressure model for precision GPS orbit determination," *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, vol. 35, no. 1, pp. 55–62, 2017.

- [18] M. H. Kim and T. S. Bae, "Stability assessment of FKP system by NGII using long-term analysis of NTRIP correction signal," *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, vol. 31, no. 4, pp. 321–329, 2013.
- [19] J. S. Park, *Applicability Analysis of FKP for the Cadastral Resurveying and Public Survey through Comparative Analysis of VRS and FKP Position Accuracy [M.S. thesis]*, University of Seoul, Seoul, Korea, 2014.
- [20] G. Wübbena, M. Schmitz, and A. Bagge, "PPP-RTK: precise point positioning using state-space representation in RTK networks," in *ION GNSS-05 18th International Technical Meeting*, pp. 2584–2594, Long Beach, CA, USA, 2005.

Research Article

Inversion Model of GPR Imaging Characteristics of Point Objects and Fracture Detection of Heritage Building

Gao Lv ¹, Ning Li ^{1,2}, Jie Yang^{1,3}, Xianchun Yao ², Dexiu Hu^{1,3} and Rixuan Pang⁴

¹Institute of Water Resources and Hydro-Electric Engineering, Xi'an University of Technology, Xi'an 710048, China

²Institute of Geotechnical Engineering, Xi'an University of Technology, Xi'an 710048, China

³State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi'an University of Technology, Xi'an, Shaanxi 710048, China

⁴Geochemical Exploration Team of Shaanxi Geological Mineral Survey Group, Xi'an 710023, China

Correspondence should be addressed to Ning Li; ningli@xaut.edu.cn

Received 29 December 2017; Accepted 20 February 2018; Published 7 May 2018

Academic Editor: Hyung-Sup Jung

Copyright © 2018 Gao Lv et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are often many hidden structural defects in heritage buildings. As a convenient and effective nondestructive detecting method, ground-penetrating radar (GPR) has a technical advantage in detecting and protecting heritage buildings depending on the advanced image interpretation. The analytic relationship between buried depth and radius of point object and long and short axis of hyperbolic equation was established according to derivations of formulas. The image characteristics of hyperbolic curves with different depth and radius were studied by finite-difference time-domain method (FDTD). And then, inversion models of buried depth and radius of point object were established. The buried depth and radius can be accurately deduced by long and short axis of hyperbolic image. This result was applied in the detection of pedestal defects of the heritage building, and the depth and distribution range of hidden fracture can be accurately interpreted. It provides an effective and fast method to detect hidden defects in civil engineering.

1. Introduction

There have been many point defects in the hidden areas of civil engineering and environmental investigations, such as buried utility mapping, concrete and pavement inspection, and tunnel [1, 2]. Nondestructive testing (NDT) is an effective, flexible, and accurate method for detection of point object. Nowadays, nondestructive techniques which can detect the positions and sizes of point objects are rebound method, ultrasonic method, electrical method, and electromagnetic wave method [3]. Because of high resolution and high flexibility, ground-penetrating radar (GPR) is nowadays considered as one of the most effective and powerful NDT methods [4, 5]. The information of hidden objects, such as the location and geometry, can be deduced by detecting imaging research of point object. It provides a reliable guarantee and basis for the quality inspection and accuracy evaluation.

The detection principle of GPR is as follows. When emitted electromagnetic waves (frequency range is 10–2200 MHz) transmit in the underground, the electromagnetic waves in inhomogeneous relative permittivity of interface will produce the phenomenon of reflection and refraction. And then, the reflected waves, also called echo, can reflect the characteristics of underground medium [6, 7]. However, the relative permittivity of the geological background was complex and unknown in real measurement. This would lead to solution multiplicity in inversion of radius and depth of object [8]. So it is necessary to establish an effective and accurate inversion method.

The most common inversion methods of GPR data are linear inversion (the steepest descent method, the conjugate gradient method, the Gauss-Newton method, the gradient regularization method, etc.) and nonlinear algorithm (simulated annealing algorithm, genetic algorithm, ant colony algorithm, particle swarm optimization, fish

swarm algorithm, etc.) [9, 10]. Because the inversion calculation of nonlinear algorithm needs a lot of time and extensive data, there was less research applied in practical engineering [11, 12]. Linear inversion method exhibits fast convergence speed and excellent numerical stability. But the computed results are multiple local minima points because extremal function tends to be convex.

Herein, a one-to-one relationship between object parameters and inversion parameters was established. We employed a forward modeling method to accurately analyze the relationship between depth and radius of object and long and short axis of hyperbolic imaging. Combined with theoretical derivation and field verification, the imaging features and regularities with different depths and radius were studied. And their multiple regression models with long and short axis of hyperbolic curves were established to achieve the purpose of interpreting the geometry of object. Furthermore, a nondestructive testing of heritage building was carried out to achieve fast interpretation of hidden fracture distribution range and the depth.

2. Mathematical Model of GPR Imaging of Point Object

The antenna of GPR emits electromagnetic waves in the underground. When electromagnetic wave arrives at object, the reflection of signals will occur, and echo signals will be received by receiving antenna [13, 14]. The echo signals contain direct waves, ground reflections, object echoes, and interference signals. Two-dimensional echo data can be obtained by continuous one-dimensional echoes formed by movement of GPR equipment. The scheme is shown in Figure 1. h represents the distance from x to the object. h_0 represents the distance from x_0 to the object. r represents the radius of the object. When geological radar reaches x , the scanning area is a fan-shaped surface. So the object vertically located under x_0 point can be detected. In the time-domain recording, the reflection feature can only be recorded at the x point. And when the GPR reaches x_0 , the detection of the object belongs to the vertical detection.

According to the geometric relationship, we can get the equilibrium

$$(h+r)^2 - (x-x_0)^2 = (h_0+r)^2, \quad (1)$$

$$h = \frac{v \cdot t}{2}, \quad (2)$$

$$h_0 = \frac{v \cdot t_0}{2}, \quad (3)$$

where v is the propagation speed of the electromagnetic wave in the medium. t is the two-way travel time of electromagnetic wave from x to O . t_0 is the two-way travel time of electromagnetic wave from x_0 to O .

Formula (2) is plugged into (1), and the analytic formula of radar imaging of point object is obtained.

$$\left(\frac{t + (2r/v)}{t_0 + (2r/v)} \right)^2 - \left(\frac{x - x_0}{(v/2)t_0 + r} \right)^2 = 1. \quad (4)$$

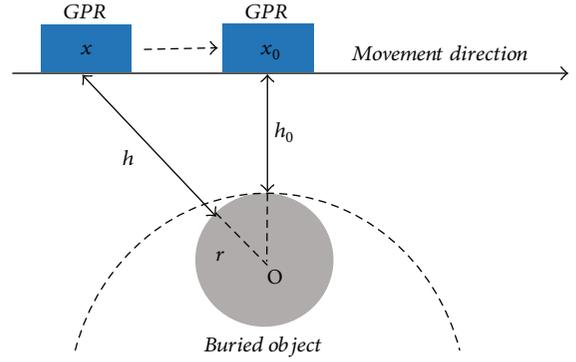


FIGURE 1: Basic principle of point object imaging of GPR.

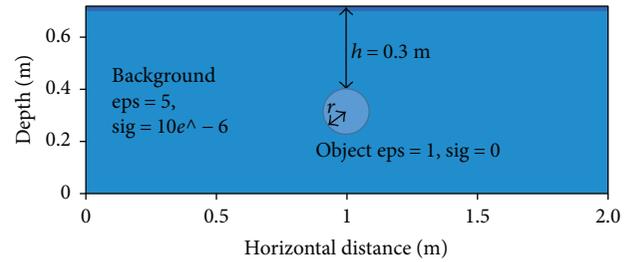


FIGURE 2: Geoelectric model with different radius ($r=0.005$ m, 0.025 m, 0.05 m, and 0.1 m).

Formula (4) conforms to the basic characteristics of hyperbolic equation, and the general formula of hyperbolic equation is as follows:

$$\frac{(y-y_0)^2}{a^2} - \frac{(x-x_0)^2}{b^2} = 1. \quad (5)$$

According to comparative coefficient method, we can get

$$a = t_0 + \frac{2r}{v}, \quad (6)$$

$$b = \frac{v}{2} \left(t_0 + \frac{2r}{v} \right), \quad (7)$$

where a is the long axis of hyperbolic equation. b is the short axis of hyperbolic equation. After transforming (6) and (7), (8) and (9) can be obtained.

$$r = \frac{b(a-t_0)}{a}, \quad (8)$$

$$h_0 = \frac{vt_0}{2} = \frac{bt_0}{a}. \quad (9)$$

Therefore, the characteristics of point object (O) with different depths and radius can be expressed by the long and short axis of the hyperbolic equation. At the same time, the imaging asymptotic lines, eccentricity, and other parameters of the GPR can also be obtained by the ratio of the long axis to short axis.

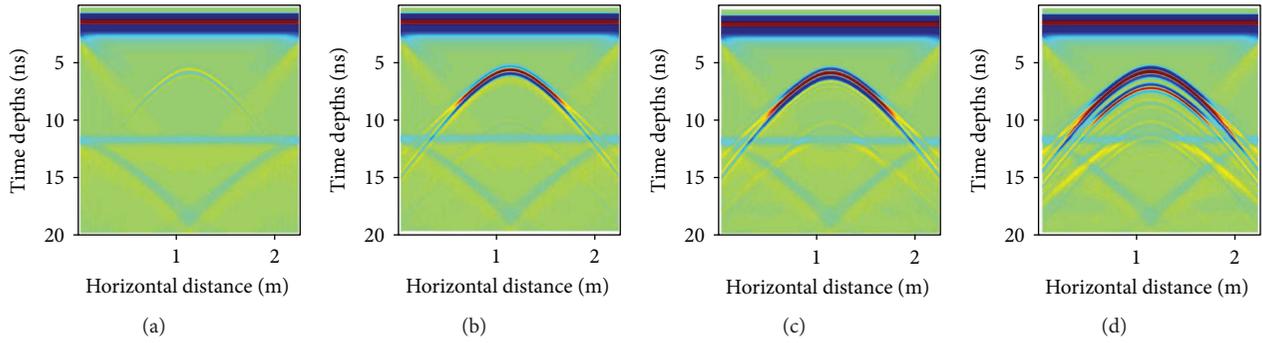


FIGURE 3: (a–d) GPR imaging of different radius ($r = 0.005$ m, 0.025 m, 0.05 m, and 0.1 m).

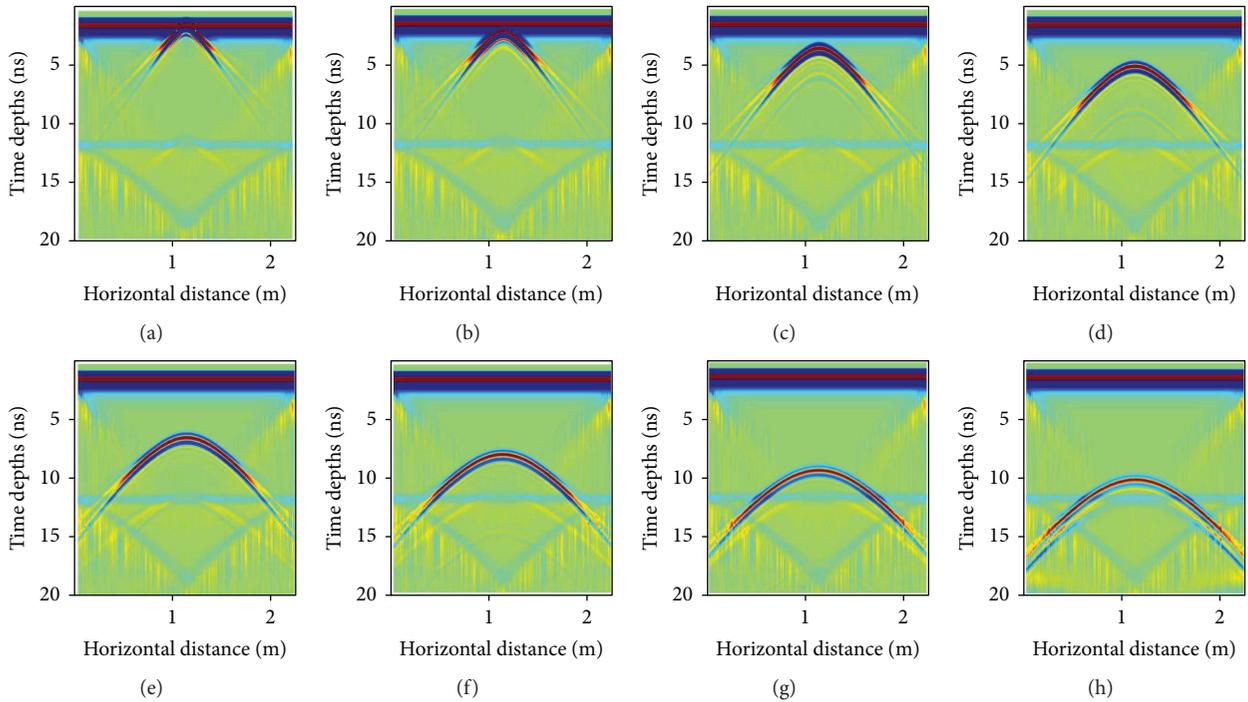


FIGURE 4: (a–h) GPR imaging of different depth ($h = 0.05$ m, 0.10 m, 0.20 m, 0.30 m, 0.40 m, 0.50 m, and 0.60 m).

3. FDTD Forward Modeling of Point Object

In this study, GPR imaging characteristics in different radius and depth of the object were simulated by finite-difference time-domain (FDTD) method, and we employed GprMax software. GprMax [15, 16] is an open source software that simulates electromagnetic wave propagation for the numerical modelling of GPR and has fast computational speed and high computing precision.

3.1. Influence of Different Radius to GPR Imaging. The influence of radius of point object on the GPR imaging is studied by FDTD method. The model design of point object is as follows. The geological background is loess, and the point object is air. The geological model scheme is shown in Figure 2.

The horizontal distance of this model is 2.0 m. The depth of this model is 0.7 m. The cell size is 0.0025 m by

0.0025 m. The time depth is 20 ns. The relative permittivity of geological background is 5. The electrical conductivity of geological background is 0.000001 S/m. The relative permittivity of point object is 1. The electrical conductivity of point object is 0.00 S/m. The depth of point object is 0.4 m. The radius of the point body is 0.005 m, 0.025 m, 0.050 m, and 0.100 m. Dominant frequency is set as 900 MHz. The excitation source is Ricker wavelet. In the numerical simulation, there are 180 step calculations, and each step calculation contains 3391 times. The imaging features of point object with different radius under the same depth are shown in Figure 3.

In Figure 3, the reflectors of point object with different radius appear at about 5 ns. The intensity of the reflected curves is obviously increased, with the increasing of the radius from 0.005 m to 0.100 m. In Figure 4(a), there are only very weak multiple curves. However, there are obviously multiple curves in Figure 4(d). The arc reflected at 5 ns

represents the upper edge of the point object. And the reflection at 6.5 ns represents the lower edge. Moreover, there is no obvious change of opening degree of hyperbola with the increase of the radius of the point object. So the point objects with different radius are not easy to be identified by the opening degree of hyperbola in real detection.

3.2. Influence of Different Depth to GPR Imaging. The influence of depth of point object on the GPR imaging is studied by FDTD method. The depth of object can generally be converted by the time depth and the relative permittivity of the geological background. As the depth of the object changed, the characteristics of radar imaging are bound to be changed. In this section, we established the geological model of objects with different depth, which is as shown in Figure 5.

The basic parameters of this model are as follows. The horizontal distance of this model is 2.0 m. The depth is 0.7 m. The cell size is 0.0025 m by 0.0025. The time depth is 20 ns. The relative permittivity of the geological background is 5, and the conductivity is 0.000001 S/m. The radius of the object is 0.05 m. The relative permittivity of the object is 1. The electrical conductivity of the object is 0 S/m. The depth of the object is set as 5 cm, 10 cm, 20 cm, 30 cm, 40 cm, 50 cm, 60 cm, and 65 cm. The main frequency of wavelet is 900 MHz, and the excitation source is Ricker wavelet. There are a total of 180 calculation steps, and each one is 3392 times. The imaging features of this model are shown in Figure 4.

In Figure 4, the reflection intensity of the hyperbolic image is not obviously changed with the increase of object depth. It demonstrates that reflection intensity is almost not affected by object depth in the same geological background. In addition, the depths of peaks of hyperbolic curves are increasing with the object depths increased. So the key points of hyperbolic curves can be extracted and transformed into time depths to deduce the object depths. Furthermore, opening degree of hyperbola is gradually increased as the object depth increases. According to (8), it can be deduced that object radius is simultaneously determined by propagation time, long axis, and short axis of hyperbolic curves. Therefore, long axis and short axis of hyperbolic curves can be analyzed by opening degree of hyperbola, and then object radius can be deduced.

4. Inversion Model of Geometric Parameters of Object

4.1. Imaging Characteristic Analysis of Objects with Different Radius. The imaging characteristics of point objects are mainly hyperbolic curves. The coordinates of key points of hyperbolic curves can be extracted, and the fitting formulas of hyperbolic curves can be established for regression analysis. In this study, the coordinates of key points were extracted by GetData software (Figure S1). The coordinate values of time depths of different radius were obtained after manual sampling (Table S1). Furthermore, the coordinate values of apparent depths of different radius were obtained

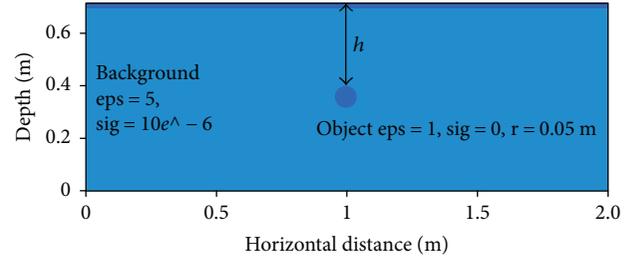


FIGURE 5: Geoelectric model of different depth ($h = 0.05$ m, 0.10 m, 0.20 m, 0.30 m, 0.40 m, 0.50 m, and 0.60 m).

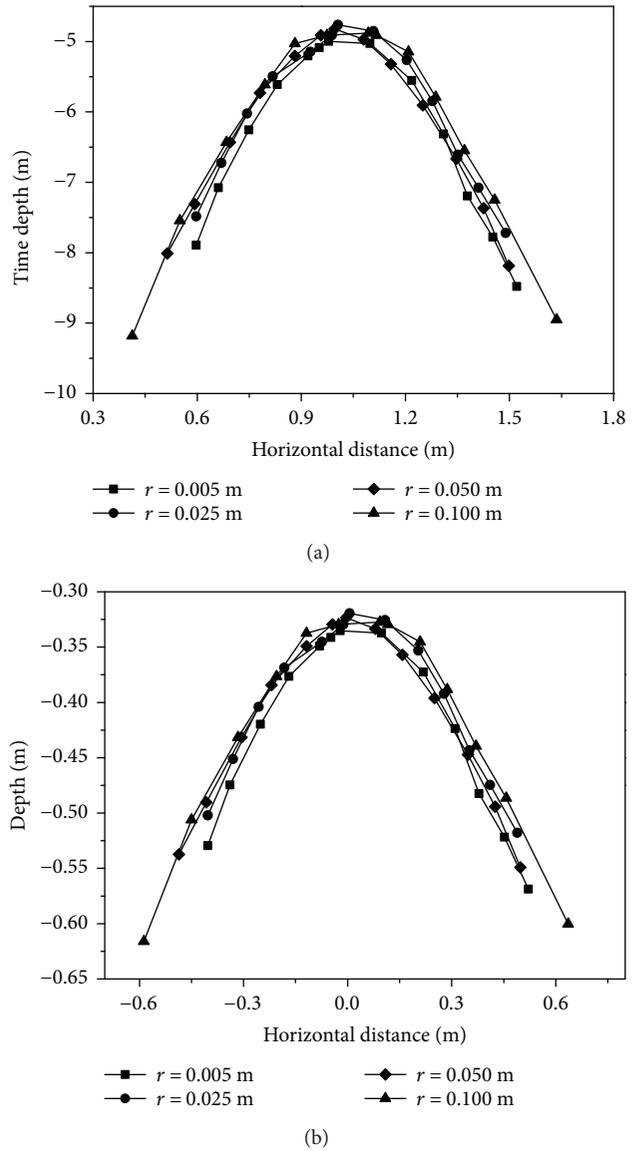


FIGURE 6: (a) Data conversion of time depth and distance in different object radius. (b) Data conversion of depth and distance in different object radius.

through the relationship between time depth and apparent depths (Formula S1). And then, the abscissa values of apparent depths were translated through horizontal axis

TABLE 1: Imaging feature of the object in different radius.

r/m	h/m	Long axis (a_r)	Short axis (b_r)	Hyperbolic formula
0.005	0.3	0.347	0.434	$\frac{y^2}{0.347^2} - \frac{x^2}{0.434^2} = 1$
0.025	0.3	0.333	0.388	$\frac{y^2}{0.333^2} - \frac{x^2}{0.388^2} = 1$
0.05	0.3	0.330	0.379	$\frac{y^2}{0.330^2} - \frac{x^2}{0.379^2} = 1$
0.100	0.3	0.342	0.408	$\frac{y^2}{0.342^2} - \frac{x^2}{0.408^2} = 1$

TABLE 2: Imaging feature of the object with different depth.

r/m	h/m	Long axis (a)	Short axis (b)	Hyperbolic formula
0.05	0.05	0.089	0.104	$\frac{y^2}{0.089^2} - \frac{x^2}{0.104^2} = 1$
0.05	0.10	0.133	0.159	$\frac{y^2}{0.133^2} - \frac{x^2}{0.159^2} = 1$
0.05	0.20	0.231	0.228	$\frac{y^2}{0.231^2} - \frac{x^2}{0.228^2} = 1$
0.05	0.30	0.330	0.379	$\frac{y^2}{0.330^2} - \frac{x^2}{0.379^2} = 1$
0.05	0.40	0.430	0.487	$\frac{y^2}{0.430^2} - \frac{x^2}{0.487^2} = 1$
0.05	0.50	0.531	0.598	$\frac{y^2}{0.531^2} - \frac{x^2}{0.598^2} = 1$
0.05	0.60	0.632	0.713	$\frac{y^2}{0.632^2} - \frac{x^2}{0.713^2} = 1$

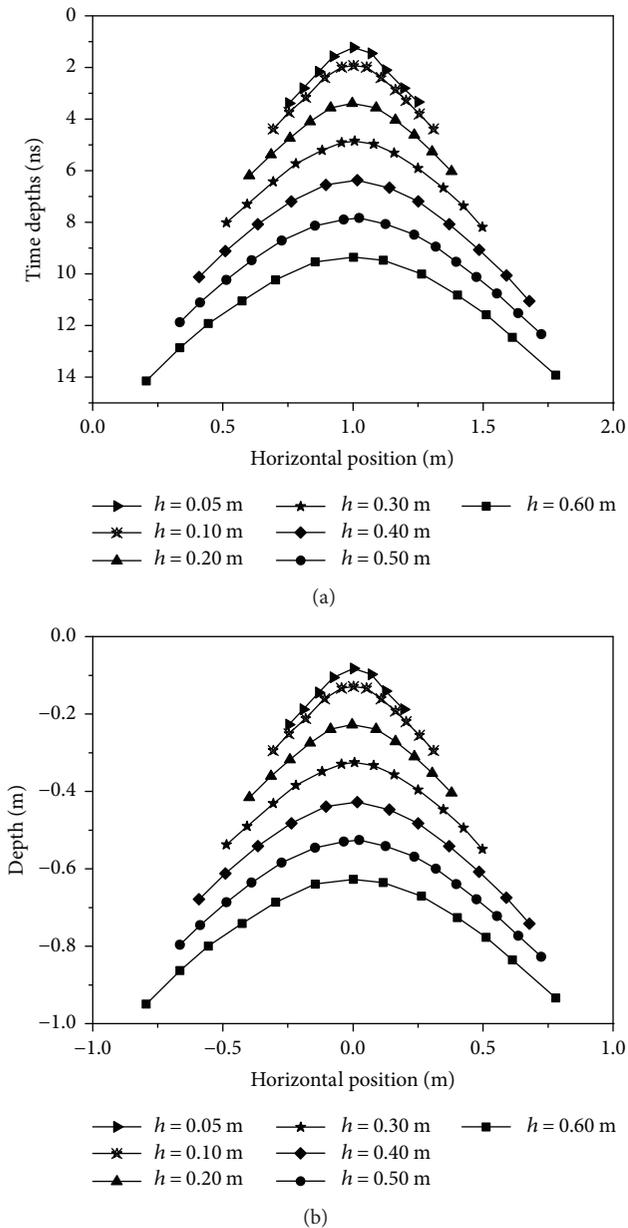


FIGURE 7: Data conversion of time depth and depth in different object depth.

in order that the abscissa values of peaks of hyperbolic curves reached zero (Table S2). So it is easy to fit the hyperbolic formulas. The imaging characteristics of time depths and apparent depths of different radius are shown in Figure 6.

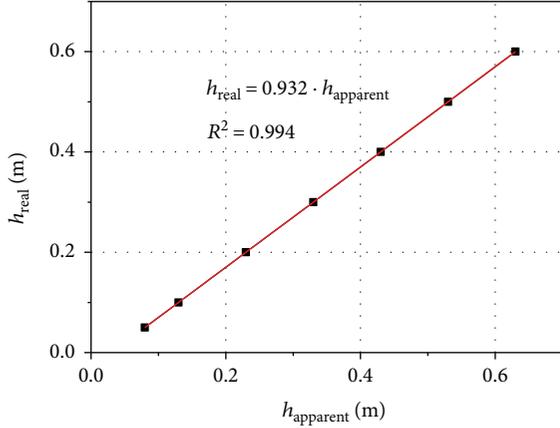
In Figure 6, the hyperbolic curves of depths of different radius are very close, and even some points are overlapped. It demonstrates that objects with different radius ($r=0.005\sim 0.100$ m) cannot be distinguished by imaging characteristics. In addition, the real depth of objects with different radius in the model is 0.30 m. However, the apparent depth of different radius is about 0.33 m. There is an error between apparent depth and real depth for the same radius, and it has been modified in Section 4.3.1.

The data of key points of hyperbolic curves are fitted by conjugate gradient method, and the hyperbolic formulas are obtained. The values of long and short axis are obtained by comparing hyperbolic formulas with (5). The detailed data are listed in Table 1. It can be seen that the values of long axis are changed from 0.330 to 0.347, when the object radius changed from 0.005 m to 0.100 m. And the values of short axis are changed from 0.379 to 0.434. In other words, the values of long and short axis are changed very small even though the radius changed 20 times. So the imaging characteristics do not show clear changes when the radius range is from 0.005 m to 0.100 m.

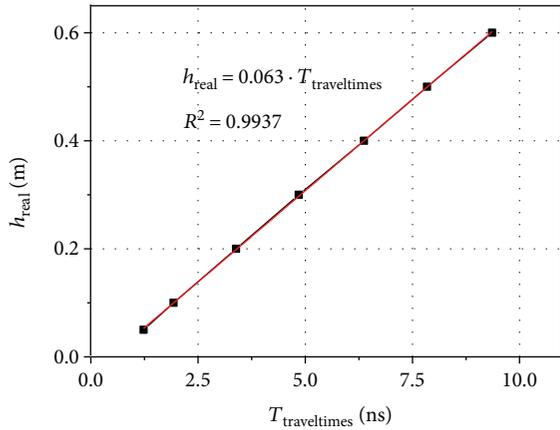
4.2. Imaging Characteristic Analysis of Objects with Different Depth. The method for extracting data is the same with that in Section 4.1. The coordinate values of time depths of different buried depth were obtained after manual sampling (Table S3). The coordinate values of apparent depths of different buried depth were obtained after depth transformation and coordinate translation (Table S4). The imaging characteristics of time depths and apparent depths of different buried depth are shown in Figure 7, when h is equal to 0.05 m, 0.10 m, 0.20 m, 0.30 m, 0.40 m, 0.50 m, and 0.60 m. In Figure 7, the hyperbolic curves of different buried depth are obviously changed. With the

TABLE 3: Comparison of the apparent depth and real depth with different depth.

Contents	Data						
Real depth h_{real}/m	0.05	0.10	0.20	0.30	0.40	0.50	0.60
Apparent depth $h_{\text{apparent}}/\text{m}$	0.08	0.13	0.23	0.33	0.43	0.53	0.63
Time depth $T_{\text{time}}/\text{ns}$	1.23	1.93	3.39	4.85	6.37	7.84	9.36



(a)



(b)

FIGURE 8: (a) Correction factor of real depths and apparent depths. (b) Correction factor of real depths and time depths.

increase of buried depths of objects, the depths of peaks of hyperbolic curves are increased. So the buried depths can be reflected by the depths of peaks of hyperbolic curves. In addition, the opening degree of hyperbola is increased with the buried depth increased.

The conjugate gradient method is also used to realize hyperbolic formula fitting. And the values of long and short axis hyperbolic formula are obtained by comparing hyperbolic formulas with (5). The detailed data are listed in Table 2. When the buried depths of objects changed from 0.05 m to 0.60 m, the values of long axis are changed from 0.089 to 0.632 (about 7 times). And the values of short axis are changed from 0.104 to 0.713 (about 7 times). In other words, both of the values of long and short axis are almost changed 7 times when the buried depths changed 0.55 m.

TABLE 4: Parameters of depth and radius of object, the long axis (a) and short axis (b).

t/ns	h/m	a_c^*	b_c^*	b_c/a_c	$(b_c/a_c)^*t$
1.23	0.05	0.08	0.10	1.24	0.76
1.93	0.10	0.12	0.15	1.25	1.21
3.39	0.20	0.21	0.26	1.22	2.07
4.85	0.30	0.31	0.37	1.22	2.96
6.37	0.40	0.40	0.48	1.21	3.85
7.84	0.50	0.49	0.60	1.21	4.74
9.36	0.60	0.59	0.71	1.20	5.64

* a_c and b_c represent the correction of long axis (a) and short axis (b) of hyperbolic curves, respectively.

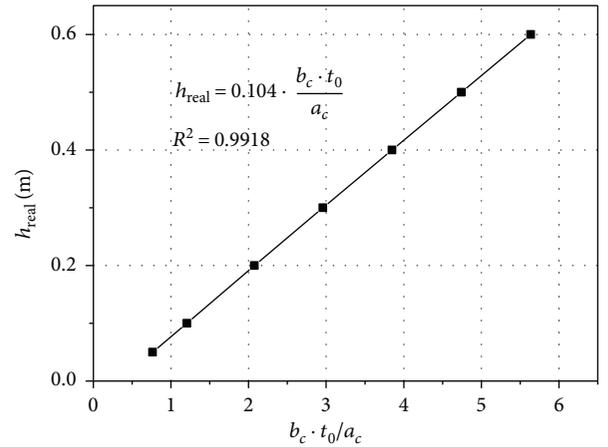


FIGURE 9: Correction factor of real depths and coefficient of the short and long axis.

So there are clear changes of imaging characteristics when the buried depths range is from 0.05 m to 0.60 m.

4.3. Inversion Model of Geometric Parameters and Long and Short Axis

4.3.1. Modified Model of Real Depth. In the forward imaging for FDTD method, there is an error between apparent depth and real depth. The main reason is that there are truncation errors between forward and backward differences of FDTD method. The value of truncation error is equal to the square of time-domain lattice (Δt). The real depths, apparent depths, and time depths of objects with different buried depths are listed in Table 3.

It can be seen that there are errors between real depths and apparent depths. Interestingly, the errors with different buried depths are the same, and are equal to 0.03 m. So

TABLE 5: Comparison of the real radius and inversion radius in different depth.

Contents	Data						
Real depth h_{real}/m	0.05	0.10	0.20	0.30	0.40	0.50	0.60
Real radius r_{real}/m	0.050	0.050	0.050	0.050	0.050	0.050	0.050
Inversion radius $r_{\text{inversion}}/\text{ns}$	0.089	0.065	0.054	0.051	0.049	0.049	0.047
Error e	78%	30%	8%	2%	2%	2%	6%

the errors can be modified and accurately real depth can be obtained.

$$h_{\text{real}} = 0.932 \cdot h_{\text{apparent}} \quad (10)$$

or

$$h_{\text{real}} = \frac{c}{\sqrt{\epsilon_r}} \cdot \frac{T}{2} = 0.063 \cdot T_{\text{traveltimes}} \quad (11)$$

Figure 8(a) is the fitting curve of real depth and apparent depth, and (10) is the modified relation between real depth and apparent depth. The depth error can be reduced using the modified relation. Figure 8(b) is the fitting curve of real depth and time depth, and (11) is the modified relation between real depth and time depth. The time depth of imaging can be quickly transformed into real depth using (11), and it is more valuable in application.

4.3.2. Inversion Model of Geometric Parameters and Long and Short Axis. The values of the long and short axis of hyperbolic imaging can be calculated by (8), (9), and the basic parameters of the forward model. Because the long and short axis of hyperbolic curves is not obviously affected by the change of radius, the imaging characteristics of objects are mainly affected by the change of buried depth. The long axis (a) and short axis (b) corresponding with real depth can be calculated by (11) and listed in Table 4.

According to (7) and (11), the relationship among real depth and long and short axis (a_c , b_c) of hyperbolic curves is fitted by binary fitting method. The correlation coefficient between real depth and axis is 0.104 from Figure 9; so the empirical (12) can be obtained.

$$h_{\text{real}} = 0.104 \cdot \frac{b_c \cdot t_0}{a_c} \quad (12)$$

The empirical (12) is a modified relation among real depth and long and short axis of hyperbolic curves. It can be seen that there is a difference between (12) and (7). And the difference is just reflected by the correction coefficient, which is equal to 0.104. Therefore, the accurate values of real depth can be quickly obtained by imaging characteristics of hyperbolic curves in practical application.

According to 12 and the data of Tables 3 and 4, the inversion model of object radius (13) can be obtained by nonlinear algorithm. The data of inversion radius and real radius of objects are listed in Table 5.

$$r_{\text{real}} = \frac{0.23}{h} \cdot \frac{b_c(a - (0.104/2) \cdot t_0)}{a_c} \quad (13)$$

TABLE 6: The antenna parameters of GPR.

Frequency	400 MHz
Emissivity	100 KHz
Ranges (ns)	40
Scanning speed (scan/s)	50
Gain	5

In Table 5, the errors of real radius and inversion radius are less than 8%, and inversion radius calculated by (13) is accurate when the depth range is from 0.20 m to 0.60 m. The corresponding errors of real depth of 0.05 m and 0.10 m are more than 30%. The reason is that there are big errors of fitting formulas of hyperbolic curves when the real depths are between 0.05 m and 0.10 m.

5. Inversion of the Measured GPR Data of the Heritage Building

5.1. Overview of the Heritage Building and Arrangement of Measuring Line. The heritage building built in 1384 during the early Ming Dynasty is a symbol of the city of Xi'an and one of the grandest of its kind in China. It covers an area of 1377 square meters and has 36 meters high and consisted of a foundation, a pedestal, and a tower. The structure of the tower is mainly made up of brick and wood and was built on the square base compacted by soil. At present, the pedestal shows some apparent fractures. It may be due to artificial disturbance and natural erosion, such as earthquake and subway.

GPR method is used to detect the internal deformations and distribution area in the pedestal. A distinct advantage of GPR is the ability to provide high-resolution continuous profiling for nondestructive site investigations. In addition, GPR has become economically feasible and convenient method in the protection of the heritage. Combining with the actual situation, antenna of 400 MHz frequency is an optimal choice because of the proper precision and detection depth. The detailed parameters of this antenna are listed in Table 6. Figure 10 shows the location of the heritage buildings, the leakage in the northeast of the pedestal, and monitoring survey area and the line layout. The survey lines are located in the northeast of the pedestal. Each line is 25 m length, and the distance between each line is 0.5 m.

After data processing and image interpretation, the results of GPR images show that abnormal regions of each survey line are distributed in the same location. So we use the data fragments of the abnormal regions to analyze the



FIGURE 10: (a–b) Location of the heritage buildings. (c) Leakage in external wall of the pedestal. (d) Monitoring survey area and the line layout.

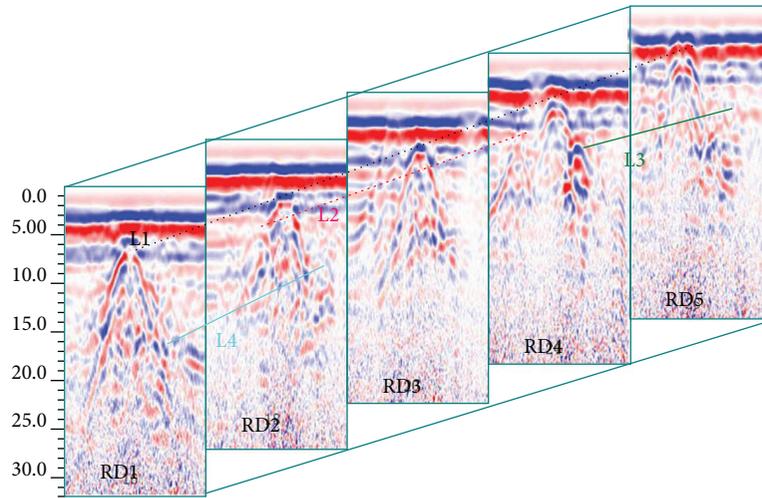


FIGURE 11: Geological section of anomaly region by GPR.

reasons of the complicated imaging. The length of each fragment is 1 m, and the slice figure consisted of the GPR images of each fragment in Figure 11.

In Figure 11, it can be obviously seen that there are typically hyperbolic reflection in the similar position of each line, and the depth of peaks of hyperbolic curves is about 0.4m. According to above results, the hyperbolic reflection can be deduced as point defects. The RD1 line is close to the main building, and the image of RD1 line

TABLE 7: Inversion parameters of depth and radius of the defects.

	<i>a</i>	<i>b</i>	<i>h/m</i>	<i>r/m</i>	<i>t/ns</i>
RD1 line	0.70	0.56	0.52	0.13	6.26
RD2 line	0.74	0.58	0.54	0.13	6.55
RD3 line	0.55	0.49	0.56	0.08	6.14
RD4 line	0.50	0.46	0.54	0.08	5.56
RD5 line	0.74	0.61	0.50	0.17	5.81

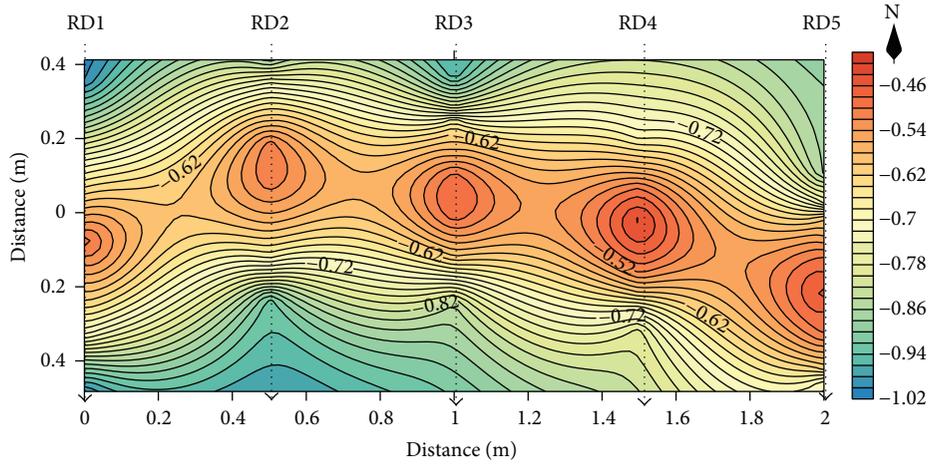


FIGURE 12: Inversion result of radius and depth of fracture.

has the strongest reflection. The reflection intensity of RD2, RD3, RD4, and RD5 lines is weakened change in order. This phenomenon reveals that the defects are consecutive and the geometry sizes of the defects are decreasing from RD1 to RD5 line. Through the contrast of the five GPR reflection images of the pedestal, it can deduce that the reflection characteristics of hyperbolic curves in the profile accord with penetrating fracture, which is marked as L1 line. Meanwhile, some filling materials (impurities or air) can be found nearby the main defects which are obtained from the GPR imaging, such as L2, L3, and L4 line.

5.2. Comprehensive Interpretation of Pedestal Defects of the Heritage Building. Feature points are extracted from the hyperbolic curves which have the strongest reflection intensity of each GPR images in Figure 11. And then the feature points are fitted to calculate the analytical equation of hyperbolic curves by conjugate gradient method. The depth and radius of objects in each GPR images could be estimated from the long axis and short axis of the hyperbolic curves. The detailed data are listed in Table 7.

In Table 7, the depth of the objects is 0.52 m, 0.54 m, 0.56 m, 0.54 m, and 0.50 m. The inversion depths are close to the seepage depths on the external wall (Figure 10(c)). The radius of the objects is 0.13 m, 0.13 m, 0.08 m, 0.08 m, and 0.08 m. The inversion radius is close to the width of the brick in the pedestal. So the falling bricks and the erosion of compacted soil are thought to be reasonable interpretation. On the basis of data interpolation and fitting, the initial values of depth and radius of L1 line are used to draw geological section map, which is shown in Figure 12.

In Figure 12, it can be visually observed that the five-point objects are connected to form a nonstraight penetrability fracture. The depth of the fracture is from 0.46 m to 0.60 m, and the width is about 0.15 m. The fracture reinforcement should be carried out, and the pedestal should be early protected with waterproof in case of monsoon.

6. Conclusions

In summary, we established the analytic relationship between depth and radius of object and long and short axis of hyperbolic equation, according to derivations of formulas.

$$\begin{aligned} r &= \frac{b(a-t_0)}{a}, \\ h_0 &= \frac{vt_0}{2} = \frac{bt_0}{a}. \end{aligned} \quad (14)$$

The image characteristics of hyperbolic curves with different depth and radius were studied by FDTD method. When $h = 0.3$ m, the intensity of reflected curves is obviously increased, with the increase of the radius from 0.005 m to 0.100 m but there is no obvious change of opening degree of hyperbolic curves. When $r = 0.05$ m, the depths of peaks of hyperbolic curves are increasing with the object depths increased from 0.05 m to 0.60 m, and the opening degrees of hyperbolic curves are gradually increased.

Inversion model of geometric parameters and long and short axis was established. The relationship among real depth and long and short axis (a_c , b_c) of hyperbolic curves was obtained.

$$h_{\text{real}} = 0.104 \cdot \frac{b_c \cdot t_0}{a_c}. \quad (15)$$

The inversion model of object radius was obtained by nonlinear algorithm.

$$r_{\text{real}} = \frac{0.23}{h} \cdot \frac{b_c(a - (0.104/2) \cdot t_0)}{a_c}. \quad (16)$$

The practical detection in heritage building was implemented according to the above research results, and its feasibility was verified. The detecting of pedestal defects of the heritage building is obviously effective by the inversion model. It is demonstrated that this inversion model is feasible, since the observed and calculated radius and depth of pedestal defects agree fairly well with each other.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors gratefully acknowledge the financial supports by China Postdoctoral Science Foundation (Grant no. 2017M613175), Shaanxi Province Science Foundation for youths (Grant no. 2018JQ5203) and Shaanxi Province Major Science Foundation (Grant no. 2018JZ5010). And the work was supported by the National Science Foundation of China (Grant no. 11572246, 51779207) and the Open Foundation of State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area (Grant no. 2016ZZKT-8).

Supplementary Materials

Formula S1: the equation of the object depth and two-way travel time. Figure S1: GETDATA obtains the key point coordinates by taking points continuously at the maximum amplitude of the reflection curve in GPR images. Table S1: data of the GPR imaging with different object radius. Table S2: transfer data of the GPR imaging with different object radius. Table S3: data of the GPR imaging with different object depth. Table S4: transfer data of the GPR imaging with different object depth. (*Supplementary Materials*)

References

- [1] J. L. Porsani, Y. B. Ruy, F. P. Ramos, and G. R. B. Yamanouth, "GPR applied to mapping utilities along the route of the line 4 (yellow) subway tunnel construction in São Paulo City, Brazil," *Journal of Applied Geophysics*, vol. 80, pp. 25–31, 2012.
- [2] S. Lahouar and I. L. Al-Qadi, "Automatic detection of multiple pavement layers from GPR data," *NDT & E International*, vol. 41, no. 2, pp. 69–81, 2008.
- [3] D. Breyse, "Nondestructive evaluation of concrete strength: an historical review and a new perspective by combining NDT methods," *Construction and Building Materials*, vol. 33, pp. 139–163, 2012.
- [4] S. K. U. Rehman, Z. Ibrahim, S. A. Memon, and M. Jameel, "Nondestructive test methods for concrete bridges: a review," *Construction and Building Materials*, vol. 107, pp. 58–86, 2016.
- [5] G. Lv, J. Yang, N. Li, D. Hu, Y. Zhang, and F. Zhao, "Dielectric characteristics of unsaturated loess and the safety detection of the road subgrade based on GPR," *Journal of Sensors*, vol. 2018, Article ID 5185460, 8 pages, 2018.
- [6] L. Lo Monte, D. Erricolo, F. Soldovieri, and M. C. Wicks, "Radio frequency tomography for tunnel detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1128–1137, 2010.
- [7] E. Slob, "Interferometry by deconvolution of multicomponent multioffset GPR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 3, pp. 828–838, 2009.
- [8] E. Slob, M. Sato, and G. Olhoeft, "Surface and borehole ground-penetrating-radar developments," *Geophysics*, vol. 75, no. 5, pp. 75A103–75A120, 2010.
- [9] G. A. Meles, J. Van der Kruk, S. A. Greenhalgh, J. R. Ernst, H. Maurer, and A. G. Green, "A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 9, pp. 3391–3407, 2010.
- [10] A. Di Matteo, E. Pettinelli, and E. Slob, "Early-time GPR signal attributes to estimate soil dielectric permittivity: a theoretical study," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 3, pp. 1643–1654, 2013.
- [11] A. P. Tran, M. R. M. Ardekani, and S. Lambot, "Coupling of dielectric mixing models with full-wave ground-penetrating radar signal inversion for sandy-soil-moisture estimation," *Geophysics*, vol. 77, no. 3, pp. H33–H44, 2012.
- [12] G. A. Meles, S. A. Greenhalgh, A. G. Green, H. Maurer, and J. van der Kruk, "GPR full-waveform sensitivity and resolution analysis using an FDTD adjoint method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 5, pp. 1881–1896, 2012.
- [13] H. Chen and A. G. Cohn, "Probabilistic robust hyperbola mixture model for interpreting ground penetrating radar data," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, Barcelona, Spain, July 2010.
- [14] W. Al-Nuaimy, Y. Huang, M. Nakhkash, M. T. C. Fang, V. T. Nguyen, and A. Eriksen, "Automatic detection of buried utilities and solid objects with GPR using neural networks and pattern recognition," *Journal of Applied Geophysics*, vol. 43, no. 2–4, pp. 157–165, 2000.
- [15] C. Warren, A. Giannopoulos, and I. Giannakis, "gprMax: open source software to simulate electromagnetic wave propagation for ground penetrating radar," *Computer Physics Communications*, vol. 209, pp. 163–170, 2016.
- [16] I. Giannakis, A. Giannopoulos, and C. Warren, "A realistic FDTD numerical modeling framework of ground penetrating radar for landmine detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 1, pp. 37–51, 2016.

Research Article

Evaluation and Analysis of Dam Operating Status Using One Clock-Synchronized Dual-Antenna Receiver

Yunlong Zhang ¹, Songlin Yang,¹ Jiankun Liu,¹ Dongwei Qiu ², Xiaoyan Luo,³
and Jianhong Fang⁴

¹School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China

²School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

³Beijing Key Laboratory of Urban Spatial Information Engineering, Beijing Institute of Surveying and Mapping, Beijing 100038, China

⁴Qinghai Research Institute of Transportation, Xining 810001, China

Correspondence should be addressed to Yunlong Zhang; 13115311@bjtu.edu.cn and Dongwei Qiu; qiudw@bucea.edu.cn

Received 18 September 2017; Revised 26 January 2018; Accepted 4 March 2018; Published 22 April 2018

Academic Editor: Yee-wen Yen

Copyright © 2018 Yunlong Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Higher monitoring cost is the main limiting factor in the application of geodetic receivers in dam deformation monitoring, as the number of receivers applied in one dam is always too large. In this paper, an improved GNSS monitoring scheme is proposed. Firstly, three-dimensional finite element model is built to simulate the main deformation part of the dam, and only one clock-synchronized dual-antenna receiver is used to set the reference and monitoring stations; then, the dam displacement warning standard is determined based on the relationship between the corresponding surface displacement and safety factor of the different mechanical parameter combination; finally, through BDS/GPS dam deformation monitoring system for Lijiaxia in China, the operating status of dam is evaluated. Through comparative analysis, it is found that the scheme is feasible.

1. Introduction

External deformation is an important monitoring part of reservoir dam, which can provide important data for dam safety operation and deformation mechanism analysis [1, 2]. Accuracy, cost, and automation are the important factors that need to be considered in dam deformation monitoring. The traditional monitoring methods, such as total station, level, and theodolite, did not allow the real-time, continuous and automatic monitoring of dam [3]; GNSS (Global Navigation Satellite System) has been used for the surface deformation monitoring of the engineering structure based on its high-precision and high-automation characteristics [4–6] but with higher cost of geodetic receivers severely limiting its application in the dam deformation monitoring.

Multiantenna method is proposed for reducing the monitoring cost. He et al. conducted a remote-controlled GPS monitoring system using GPS multiantenna technology and General Packet Radio Service (GPRS) wireless data

communication to monitor the dam deformation and steep slope movement for reducing cost in large projects [7–9]. However, the multiantenna method is not multiantenna clock-synchronized technology in the true sense. The monitoring needs continuous switch between antennas using microwave switches, causing phase cycle-slip error; moreover, each switch needs reinitialization, and single-epoch positioning is not available.

GNSS deformation monitoring integrated system can realize the continuous and automatic monitoring of the dam external deformation. Jiang et al. developed the GPS deformation monitoring system of reservoir dam and achieve the automated, high-precision dam deformation monitoring [10–13]. Unfortunately, the cost of the system is also too high, caused by the geodetic receivers.

Clock-synchronized dual-antenna receivers are totally diverse from the geodetic receivers, which are very expensive (>\$4600) and cumbersome (>6 kg). Meanwhile, not like the traditional GNSS multiantenna technology, the dual-

channel RF front-end and digital correlator share the same clock module, avoiding cycle slip error and realizing single-epoch positioning, owing to electronic component reduction of clock-synchronized dual-antenna receivers, the new-generation N71JN from Shanghai Huace Navigation Technology Ltd. in Figure 1, which has low cost (<\$1565) and light (<3.6 kg). Moreover, most of the dam GNSS surface monitoring schemes are unreasonable, and GNSS monitoring station overabundance results in tremendous waste of equipment resources [4, 6, 10, 12, 14].

In this paper, three-dimensional finite element model (TFEM) of Lijiaxia Dam is built in order to analyze the main deformation area and extract the critical point of Lijiaxia Dam. The stage early warning standard for dam deformation is determined based on the relationship between displacement and safety factors corresponding to different parameter combination. Through BDS (BeiDou Navigation Satellite System)/GPS (Global Positioning System) dam deformation monitoring system (DDMS) of Lijiaxia Hydropower Station and BJMonitor software using embedded robust weighting based on moving window (RWMW) developed independently, the operation status of the dam is evaluated based on the warning standard and the displacement measured with only one clock-synchronized dual-antenna receiver, to verify the effectiveness of the monitoring scheme.

2. Warning Criteria and Model

2.1. TFEM of Lijiaxia Dam. Lijiaxia Hydropower Station is a large hydropower project on the Yellow River upstream, and the dam is a double-curvature arch dam with three-centered arc. The dam height is 155 m, the top elevation of the dam is 2185 m, and the dam top width is 8 m.

In order to ensure the dam operating safety, the number of GNSS observation stations set on the crest is usually larger than three, resulting in a great waste of equipment resources. In theory, as long as the main deformation area of the dam is determined, the dam operation status can be evaluated based on displacement data measured with only one clock-synchronized dual-antenna receiver. To minimize the cost and make more reasonable use of GNSS equipment, this paper proposes that TFEM is built by using the ANSYS software to predict the deformation of the Lijiaxia Dam. The Drucker-Prager criterion is adopted for numerical simulation analysis; according to the predicting displacement and the stress condition, the primary deformation area is determined [15–18]. In order to simplify the calculation, the structural materials are mainly divided into three categories: good rock mass, weak rock mass, and concrete. Through the relevant literature, the physical and mechanical parameters are determined and shown in Table 1, the concrete temperature expansion coefficient is set to $1 \times 10^{-5}/^{\circ}\text{C}$ [19]. As shown in Figure 2, the geometric dimensions of TFEM are $480\text{ m} \times 700\text{ m} \times 305\text{ m}$ (along the river, to cross the river, and to height difference). The upper surface is unconstrained, and both sides of the dam and the dam bottom are normal constraint. The upstream water level is set to the normal water level 2180 m and the downstream water level 2050 m. The upstream and downstream

interception boundaries are treated with impervious boundaries, as well as the boundaries between the left and right banks are treated with known heads, which are extrapolated based on the long-term measured data of observation holes in groundwater [20].

2.2. Main Deformation Dam Section. Compared to the upstream, the water depth downstream is lower, generally at around 30 m and has less influence on the dam deformation, thus only the upstream water weight, water pressure, and dam deadweight are considered in the numerical calculation of each dam segment displacement. Combined with the field temperature and upstream water depth data which is considered in the numerical calculation of each dam segment displacement, TFEM is established to calculate the deformation of the dam. As shown in Figure 3, the dam is divided into 20 sections, and the displacement information of critical point of each dam section is extracted. According to the numerical simulation result in the right part of Figure 3, the main forced location of the dam is determined. Figure 4 shows that the numerical simulation deformation of each dam section is proportional to its mean water depth upstream in 2011, and the correlation coefficient reaches 0.99. Due to the maximum water depth, the relative displacement of the critical point at number 11 dam section to the number 1 dam section is maximal. The relative deformation is about 13 mm, that is, the critical point of the number 11 dam can obviously reflect the dam operation state as the unique GNSS monitoring point, to make real-time monitoring of dam operating status.

Compared with manual measurement methods, the surveying robot is widely applied in dam deformation monitoring field, and relevant scholars have verified the surveying robot has the high precision [3]. The local independent coordinate system was established in Figure 3; the TM30 total station was adopted to survey the critical point coordinate at dam section numbers 1 to 20; the equal interval coordinates per day in 2011 were selected to calculate the displacement of other dam sections relative to dam section number 1 in Figure 5; and the relative displacement at dam section 11 is the largest in most time, approximately up to 10 mm. Strictly speaking, the structural materials of Lijiaxia Dam are not only three categories, and for simple numerical calculation, the model is just a simplified one; meanwhile, due to the impact of external observation conditions, such as wind, temperature, and atmospheric refraction, the surveying robot measurement error is inevitable, which can reach $0.6 + 1 \times 10^{-6} \cdot D$ mm (D represents ranging length, the unit is km), so compared to the displacement measured by the surveying robot, the numerical simulation results have some deviation. However, the deviation is within acceptable limits, just about 3 mm; thus, it is verified that determining the main deformation section based on the finite element model is effective.

2.3. Deformation Warning Criteria. It is important to make warning criteria of surface displacement for evaluating the operating status of engineering structure. To determine the displacement warning standard, Zhang et al. [21] designed

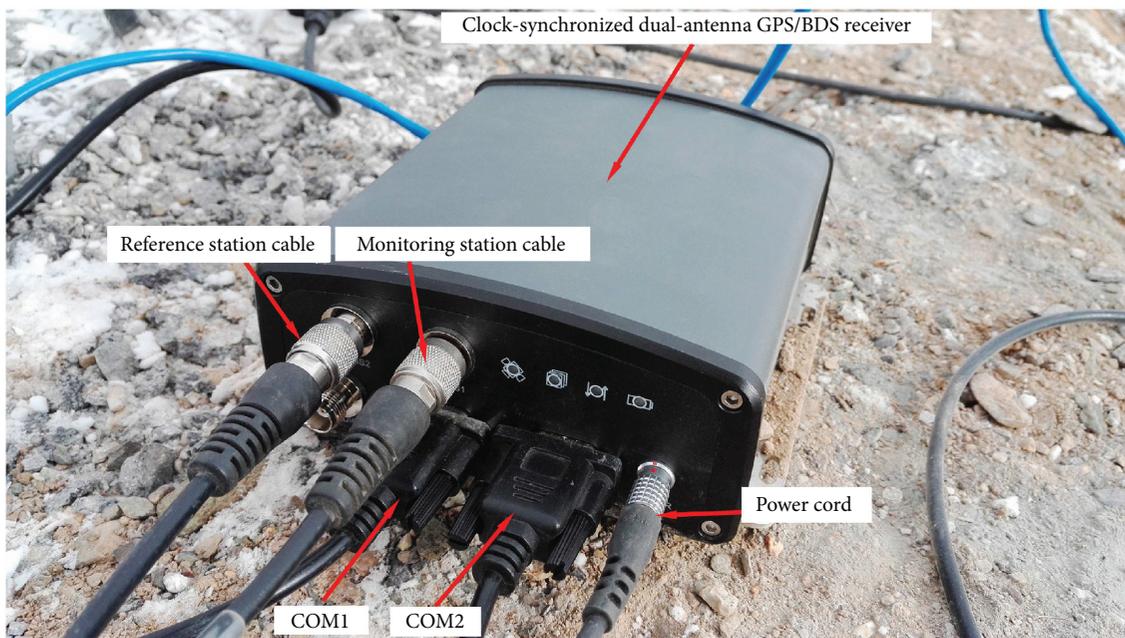


FIGURE 1: New generation of clock-synchronized dual-antenna receivers.

TABLE 1: Mechanical parameters of structural materials in Lijixia Dam.

Material	Elastic modulus (GPa)	Cohesion (MPa)	Position ratio	Appearance density (t/m ³)	Friction angle (deg)
Good rock mass	10	1.0	0.30	1×10^{-5}	32
Weak rock mass	8	0.9	0.30	1×10^{-5}	32
Concrete	30	2.5	0.18	2.4	54.9

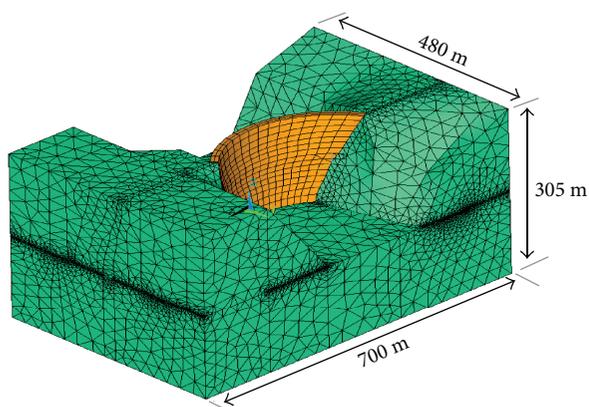


FIGURE 2: TFEM of Lijixia Dam.

the different combination in the value range of mechanical parameters to represent the impact on the slope stability by water level, temperature, and other factors and to calculate the critical point displacement and corresponding safety factor through the plastic analysis module of ANSYS. The method is also applicable to the Lijixia Dam, and the critical point at dam section 11 can also be chosen to determine warning criteria. Figure 6 shows the exponential function is used to fit the data of displacement and safety factor; after derivation calculus to the equation in the figure, the

displacement trend can be predicted. In the right graph, points 1–4 are selected for which displacements are larger, where the corresponding displacement of point 1 is 15 mm (as the yellow alarm value) and the safety factor is 3; the displacement and safety factor at point 2 are 20 mm (as the orange alarm value) and 2.5, which at points 3 and 4 corresponding to 40 mm (as the red alarm value), 2 and 120 mm, 1.8, respectively. It can be seen that as the station of GNSS02 in Figure 3 up to the displacement at point 1, the dam is in the initial instability state; after the displacement at point 2, the dam gets into the rapid instability state; after point 3, the dam is in the accelerated instability state, and to reach point 4, the dam is in the destruction of the collapse stage.

3. DDMS for Lijixia Dam

3.1. *System Composition and Function.* BJMonitor, PostgreSQL, and BDS/GPS data publishing management system are integrated in order to formulate DDMS. The communication system between host computer and receivers is held based on TCP/IP protocol, then data reception, processing, and storage are realized. Data publishing management system is connected with the database through .NET, including real-time data display, query, alarm, predictive analysis, and report generation function. The system architecture is shown in Figure 7.

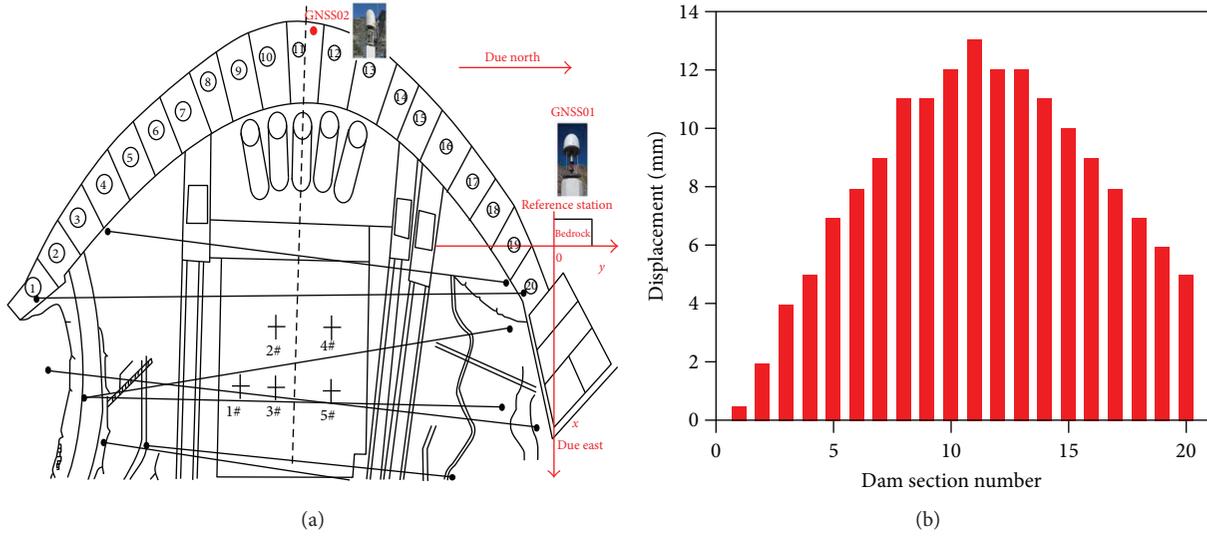


FIGURE 3: Numerical simulation deformation and monitoring point placement of the Lijixia Dam.

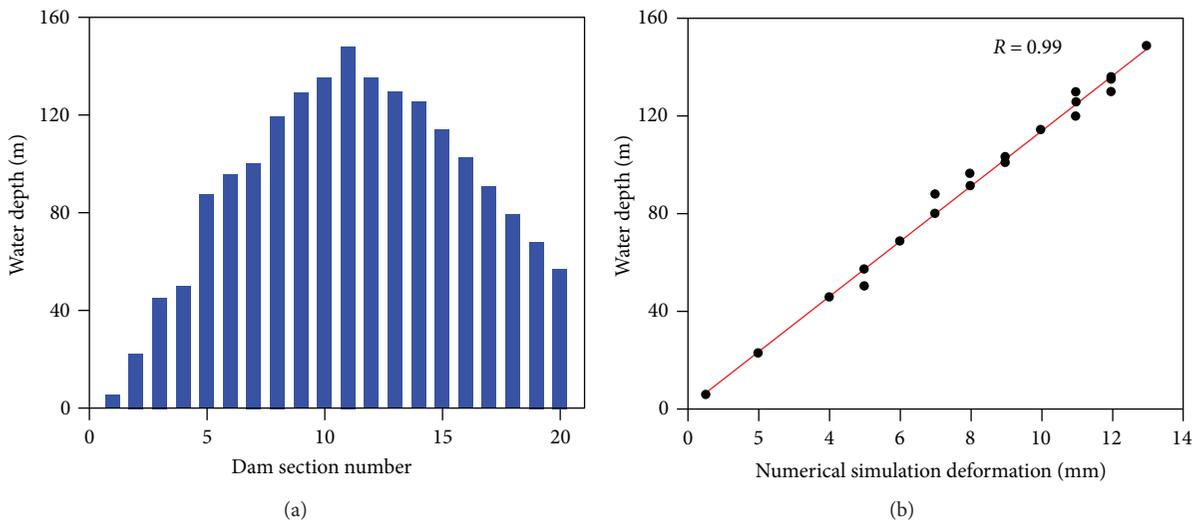


FIGURE 4: Mean water depth upstream in 2011 and relation between water depth and numerical simulation deformation.

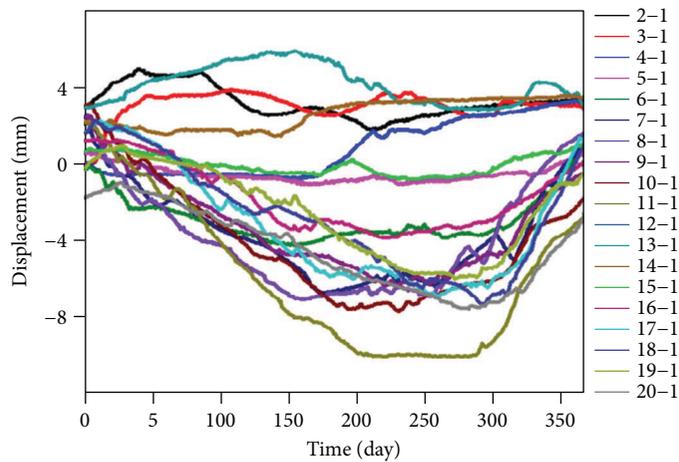


FIGURE 5: Displacement per day of the key point of each dam section relative to number 1 dam section in 2011.

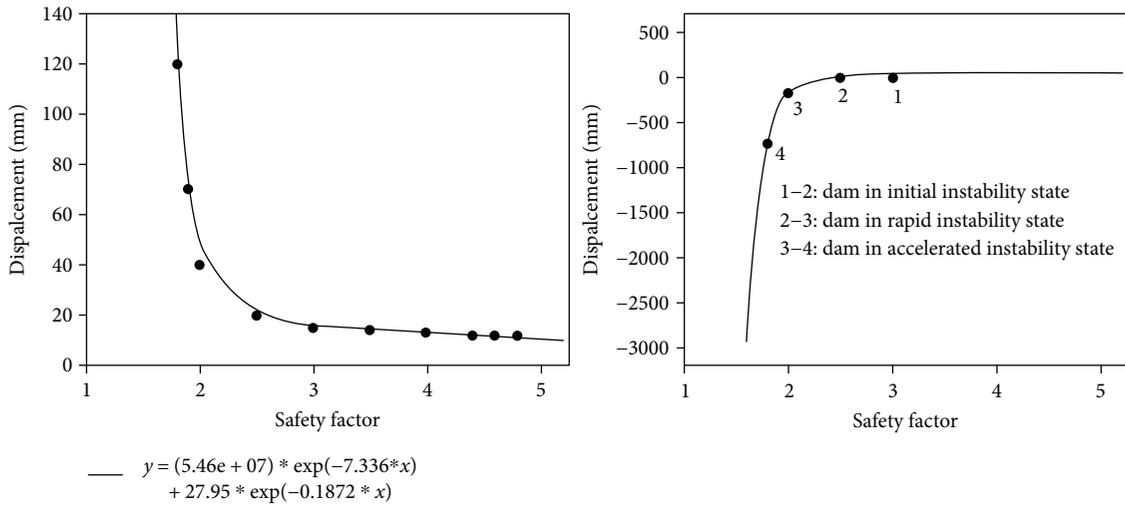


FIGURE 6: Relationship of the key point deformation and safety factor of the number 11 dam section.

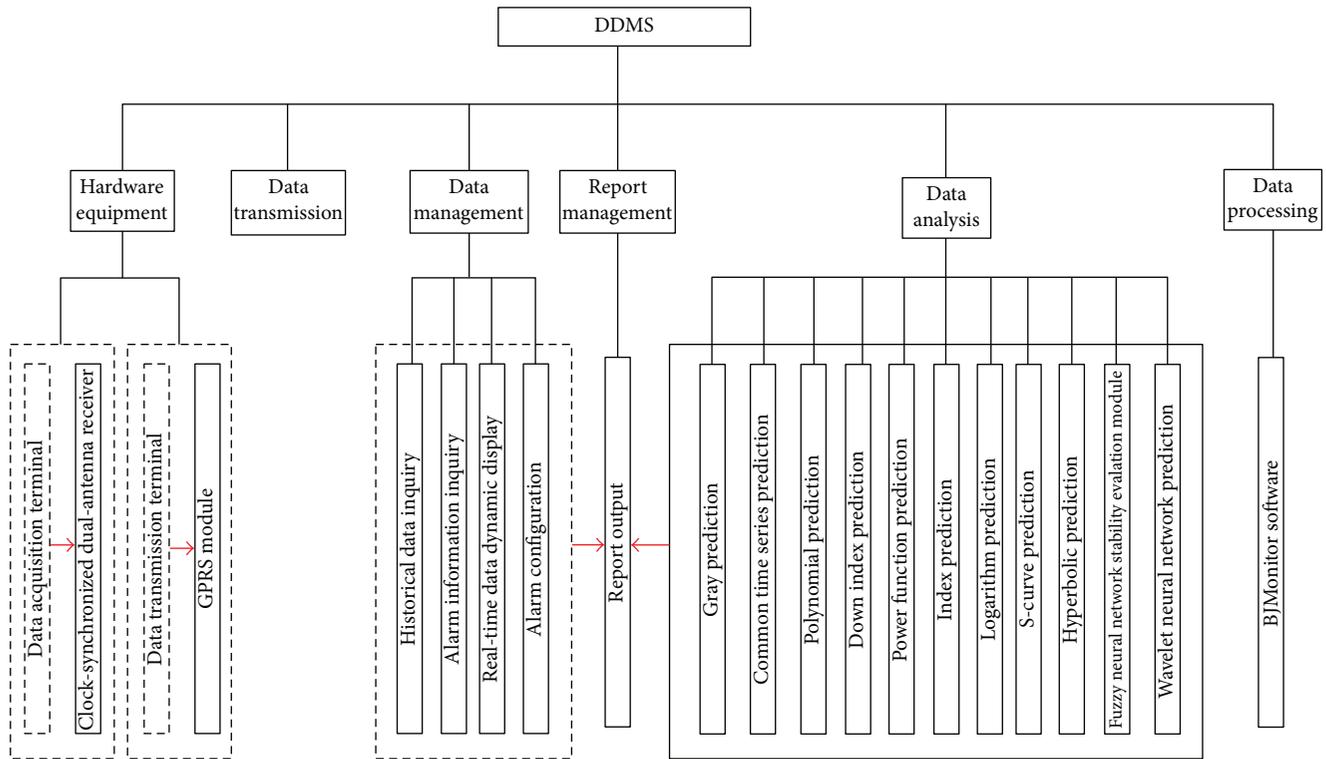


FIGURE 7: DDMS architecture.

3.2. *Hardware Equipment.* The hardware device mainly includes the data acquisition terminal, and the N71JN receiver system is adopted, consisting of the reference station antenna, the monitoring station antenna, and the clock-synchronized dual-antenna receiver. As shown in Figure 8, N71JN receiver composes of a dual-channel RF front-end, a dual-channel digital correlator, and a processor connected in order. The most prominent characteristic is that the clock signal is provided by the same clock module for the above components; moreover, the same processor is used to avoid

a variety of hardware channel inconsistencies, and the reference and monitoring station antenna system have the common crystal oscillator, so compared with traditional geodetic receivers, the pseudorange and phase observations are more accurately extracted from the dual-channel digital correlator.

3.3. *Data Transmission.* BDS/GPS observation data is transferred to the GPRS module through the receiver ports COM1 and COM2 in Figure 1, then sent to the data

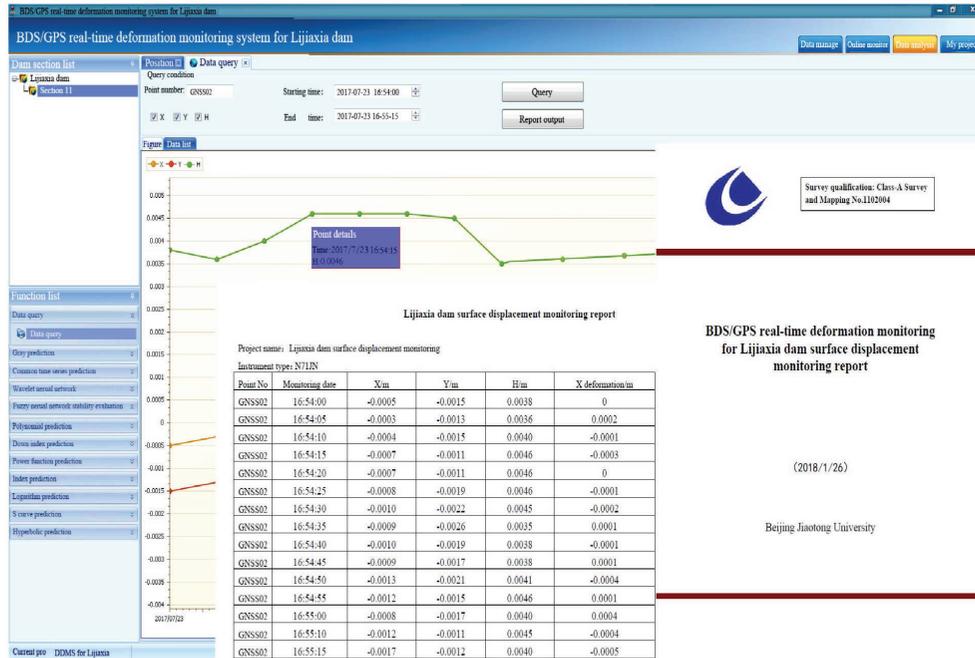


FIGURE 10: Data query and report generation.

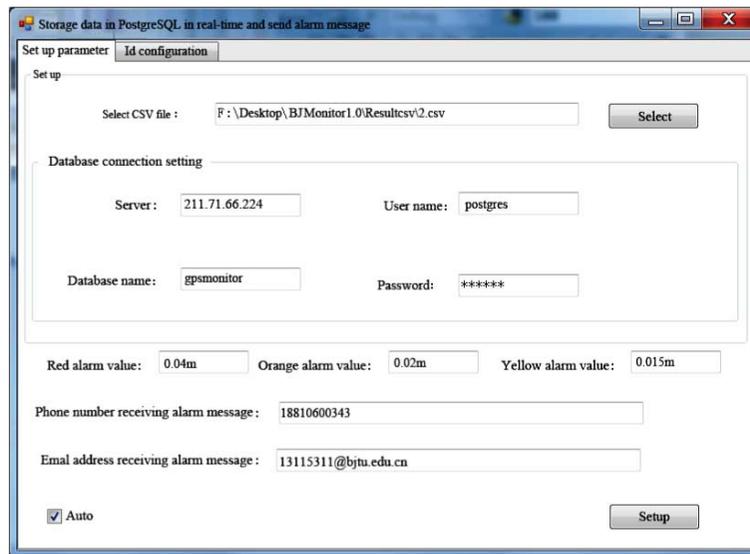


FIGURE 11: Data loader module in real time.

processing center by the GPRS module. The parameters need to be configured, including the server fixed IP address and port written into the GPRS module. Data transmission can be achieved between the GPRS module and server through the public network IP and port mapping program. For the central server has a fixed IP, the router can map the public network port to the corresponding ports in the local area network. GPRS is connected with the central public network IP port by the DNS server, and the wireless data communication between the GPRS module and server can be achieved [22].

3.4. Data Processing. BJMonitor is taken as data processing software, using embedded RWMW module. GNSS positioning accuracy is mainly affected by the error such as antenna, receiver and cable hardware delay, and multipath effects. In order to improve the estimation accuracy of observation noise variance, the moving window parameter can appropriately be increased properly, and the Helmert variance component estimated by introducing time-correlated forgetting factor is used to achieve the precise weight ratio between GPS and BDS observations; then, the history window data

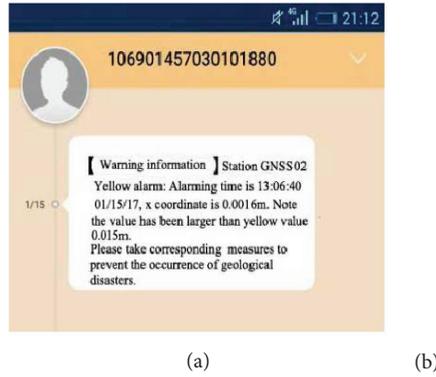


FIGURE 12: Short message and email alarm.

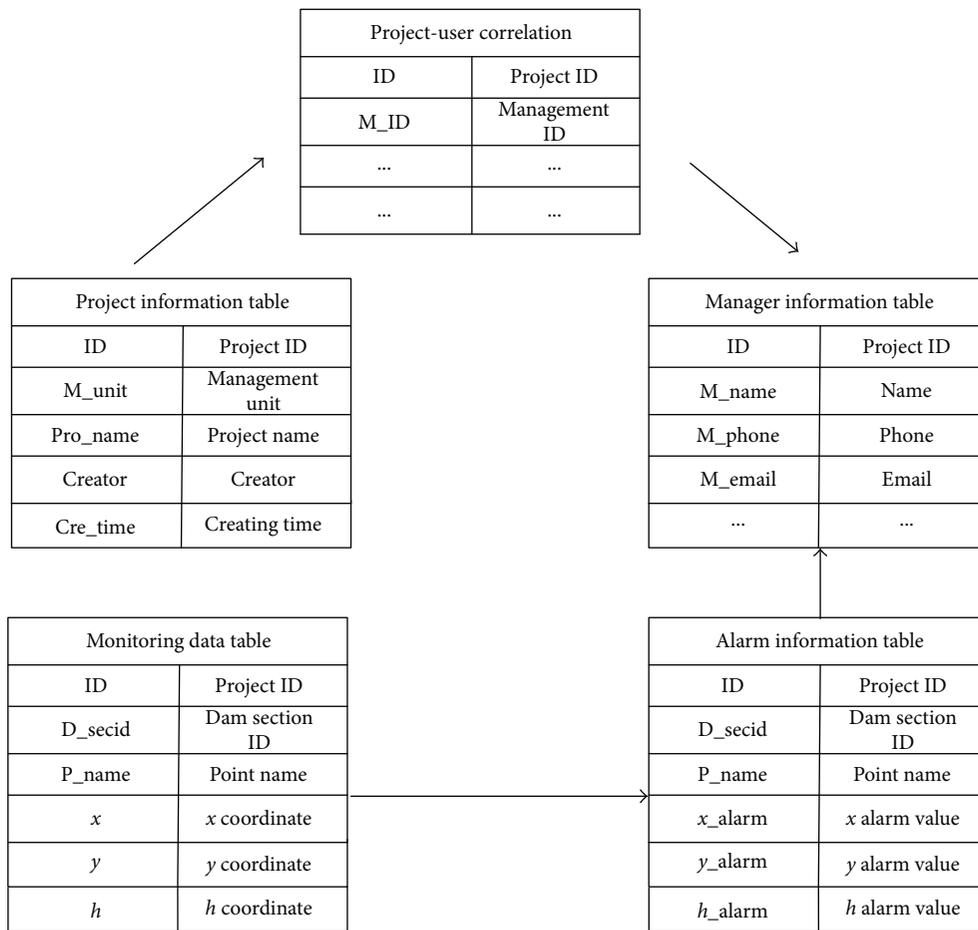


FIGURE 13: ER diagrams.

is used to improve the accuracy of observation variance at the current epoch by smoothing the noise; finally, the displacement detection with submillimeter accuracy can be achieved based on the KALMAN filter [23]. The specific calculation principle has been presented in literature 22, so here, it is not described in detail again.

3.5. *Data Management.* Data management module mainly consists of data storage, real-time display, query, alarm

information sending, and query and report generation functions and provides C/S platform for managers to manage the monitoring data. As shown in Figures 9–12, the module can be used to query, display three-dimensional coordinate data in real-time, generate the corresponding report, store the data, and send the alarm short message and mail.

PostgreSQL is an open-source database; moreover, interactive interface is good; therefore, the system takes PostgreSQL as a data management platform, to store and



FIGURE 14: Forecast analysis module of Lijiaxia Dam.

manage three-dimensional coordinate data, alarm information, and so on. The core table ER is shown in Figure 13 [24].

3.6. *Data Analysis.* In view of the forecast analysis of the monitoring data of Lijiaxia Dam, DDMS was developed comprising the gray scale prediction, common time series, polynomial regression, and fuzzy neural network stability evaluation modules in Figure 14, training the sample data of the same kind of dam, to get the corresponding model of physical mechanics parameters and safety factor, then input the parameters obtained from the laboratory experiment, to calculate the safety factor of Lijiaxia Dam, and to evaluate the stability of the dam to test the reliability of the displacement warning standard [25–27].

4. Lijiaxia Dam Stability Analysis

The deformation mechanism of the dam is explored to ensure safety operation of the generator sets, as shown in Figure 3; according to the finite element numerical analysis results of Lijiaxia Dam and the precision requirement, the joint monitoring system was established based on GNSS and GeoMos surveying robot. The reference station of GNSS01 (the origin of local independent coordinate system) was set up on the bedrock near the dam, and the due east, due north, and due height were taken as the x -, y -, and h -axis; meanwhile, one monitoring station of GNSS02 was set up at dam section number 11. The vertical axis concentric observation devices with eccentric prisms and two C220GR3D choke-ring antennas, connected with one clock-synchronized dual-antenna receiver, were installed at the two stations. The coordinates can be obtained through GNSS and total station TM30 at the same time, then the dam operating status is analyzed according to the displacement warning standard.

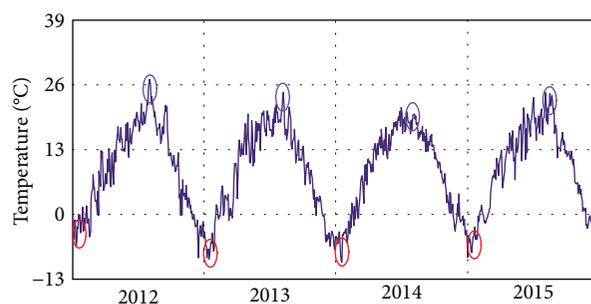


FIGURE 15: Temperature curve of Lijiaxia Dam during 2012–2015.

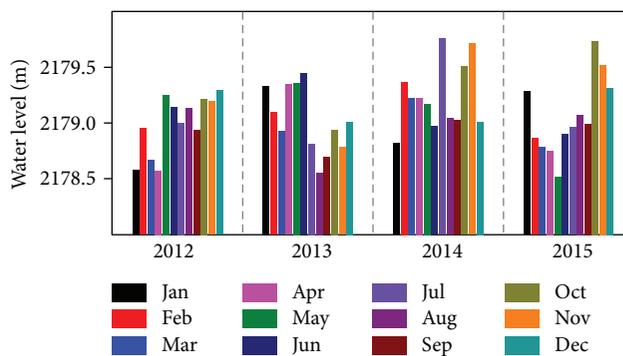


FIGURE 16: Statistics of the seasonal variations of the water level at dam section 11 during 2012–2015.

4.1. *Deformation Law Analysis.* As shown in Figure 3, the dam section number 11 is located in the middle of Lijiaxia arch dam, and it is the main position to undertake the pressures from landslides of both sides and water, which

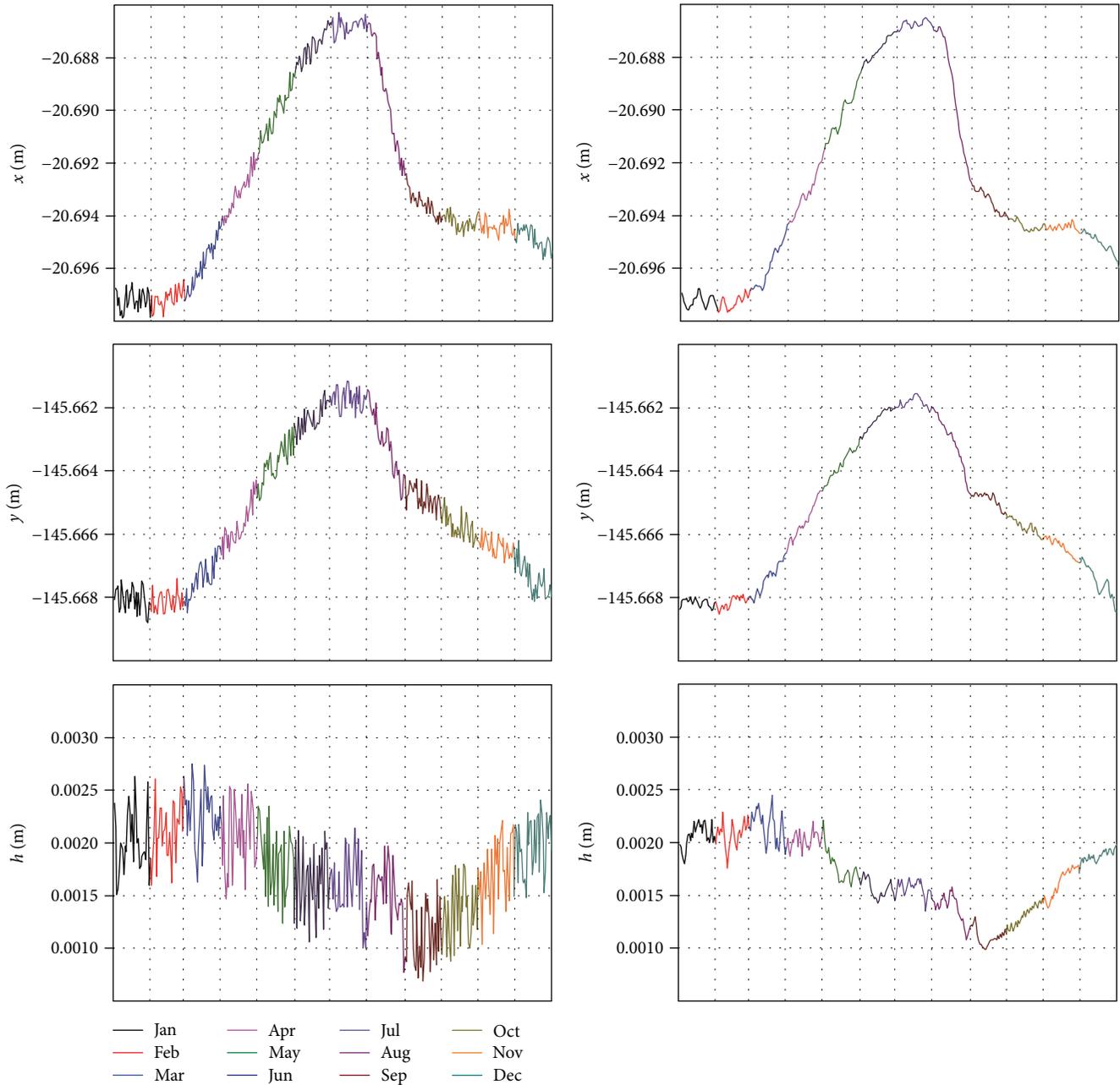


FIGURE 17: Coordinate series of station, GNSS02, at dam section 11, using the RWMW in 2015 (the left figures correspond to moving windows of 1 h; the right figures correspond to moving windows of 2 h).

is related to the safety operations of all the generators in the dam. And also, to analyze the influencing factors of Lijiaxia arch dam deformation and explore the availability of RWMW, the coordinate time series of station GNSS02 are emphatically analyzed.

To analyze the seasonal variation law of the data conveniently, the equal interval sampling (1 day) is adopted due to the large amount of data in real-time sampling (sampling interval is 5 s). Figure 15 shows that the marked red circle represents the minimum temperature, and the marked purple circle represents the maximum temperature. The maximum temperature difference can reach 36°C between January and July. In the comparative analysis of

Figures 15–18, the dam absolute deformation is mainly caused by the seasonal temperature change. Since the seasonal water level of number 11 section during 2012~2015 changes slightly, the deformation has no obvious correlation with water level. Thus, the temperature seasonal change is the main factor for causing the dam deformation.

To study the main direction of dam deformation, the observation data in 2015 is selected for analysis. As shown in the right part of Figure 17, the moving window parameter can be appropriately increased to two hours; time forgetting factor is introduced, and then the history window data is used to improve the accuracy of observation variance at the current epoch by smoothing the noises, to achieve high-

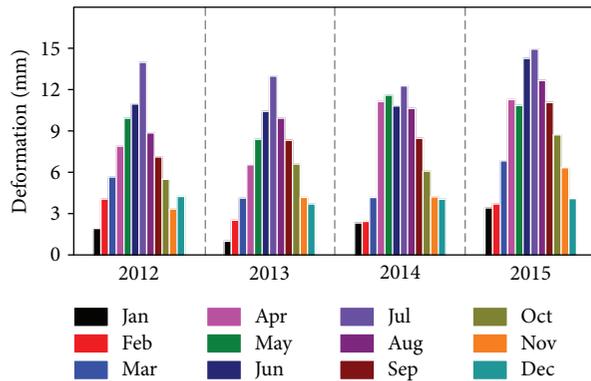


FIGURE 18: Seasonal deformation of station, GNSS02, at dam section 11, using the TM30, during 2012–2015.

precision positioning [23]. Due to both ends of dam arch that are pinned down by the mountains, which have very limited space of moving to the both banks and due height, and the arch dam expands on heating and contracts on cooling, Figure 17 shows it can only move upstream and downstream and can hardly move to both sides of the river. When the temperature reached the highest in July 2015, the arch dam moved upstream (due east), and the creep at station GNSS02 reached the maximum value of the coordinate, which was -20.6875 m. The dam body moved downstream (due west) in January 2015, and the creep reached the minimum value of the coordinates, which was -20.6975 m; its displacement was around 10 mm. However, the seasonal temperature variation had little influence on both sides of the dam body, and the creep in due north reached the maximum value of the coordinates in July 2015, which was -145.6625 m; meanwhile, the creep in due south direction reached the minimum value of -145.6678 m; its displacement was around 5.3 mm. Due height direction was barely influenced, and the coordinate was fluctuated within 1 mm. The total displacement of January and July was 11.3 mm, which was in accordance with the result obtained through TM30 presented in Figure 18; moreover, it is further explained that the RWMW algorithm is suitable for arch dam monitoring of Lijiaxia Hydropower Station with creep characteristics. According to the displacement warning standard, the displacement was less than 15 mm, and the dam was in stable state. The safety function of the dam was predicted by the fuzzy neural network model according to the physical and mechanical parameters obtained from the relative literature [25]. Figure 14 shows the safety factor was 3.8, larger than 3, and it is further proved that the displacement warning criteria was effective.

5. Conclusion

The results of this research show that the proposed monitoring scheme to evaluate and analyze the operating status of Lijiaxia Dam using one clock-synchronized dual-antenna receiver is feasible. There are six conclusions as follows from the completed work:

- (1) According to TFEM numerical simulation of Lijiaxia Dam, the main bearing part is considered as the GNSS

monitoring station, and one clock-synchronized dual-antenna receiver is used to evaluate the dam operation status, one antenna as the reference station and the other as the monitoring station, greatly reducing the monitoring cost.

- (2) TFEM is simplified, and the relationship between the safety factor of the dam and the critical point displacement is obtained based on different physical and mechanical parameter combination. The displacement warning criteria is established, and the validity of the standard is verified by the fuzzy neural network model.
- (3) BJMonitor, PostgreSQL, and BDS/GPS data publishing management system are integrated to formulate DDMS, consisting of real-time dynamic display, query data, report generation, forecasting, and early warning functions. Through the long-time application in Lijiaxia Dam, the stability of the DDMS system is verified; moreover, it can provide technical support and basis for the similar engineering structure.
- (4) Combined with the example of Lijiaxia Dam monitoring project, it is proved that the RWMW algorithm is suitable for the structure with creep characteristics, and the operation status of the dam is evaluated through the displacement warning criteria, and the effectiveness of the monitoring scheme is further verified.
- (5) Through the numerical simulation result and water depth data, the relative displacement to number 1 dam section is proportion to water pressure, and the water depth of the number 11 section is highest; thus, the relative displacement at the center section to number 1 dam section is the largest.
- (6) The water level of number 11 dam section almost maintains unchanged during 2012–2015, and the absolute deformation changes with temperature seasonal variations; thus, the temperature is the main influence on the absolute deformation of Lijiaxia Dam. Moreover, both ends of the dam arch are pinned down by the mountains, and the space of moving to both ends and due height is limited; the absolute deformation of number 11 dam section in the upstream and downstream direction is largest, both ends smaller, due height smallest.

This paper provides a new idea for dam monitoring of hydropower stations, greatly reducing the cost of equipment. Whether the monitoring scheme is applicable to other similar structures remains to be further validated.

Conflicts of Interest

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by the National Key R&D Program of China (no. 2017YFB0503700) and Beijing Key Laboratory of Urban Spatial Information Engineering under Grant no. 2016101.

References

- [1] J. Wang and C. C. Zhang, "Deformation monitoring of earth-rock dams based on three-dimensional laser scanning technology," *Chinese Journal of Geotechnical Engineering*, vol. 36, no. 12, pp. 2345–2350, 2014.
- [2] D. González-Aguilera, J. Gómez-Lahoz, and J. Sánchez, "A new approach for structural monitoring of large dams with a three-dimensional laser scanner," *Sensors*, vol. 8, no. 12, pp. 5866–5883, 2008.
- [3] C. O. Yigit, S. Alcay, and A. Ceylan, "Displacement response of a concrete arch dam to seasonal temperature fluctuations and reservoir level rise during the first filling period: evidence from geodetic data," *Geomatics, Natural Hazards and Risk*, vol. 7, no. 4, pp. 1489–1505, 2015.
- [4] D. Galán-Martín, M. Marchamalo-Sacristán, R. Martínez-Marín, and J. A. Sánchez-Sobrino, "Geomatics applied to dam safety DGPS real time monitoring," *International Journal of Civil Engineering*, vol. 11, no. 2, pp. 134–141, 2013.
- [5] Y. Kalkan, L. V. Potts, and S. Bilgi, "Assessment of vertical deformation of the ataturk dam using geodetic observations," *Journal of Surveying Engineering*, vol. 142, no. 2, article 04015011, 2016.
- [6] Y. Kalkan, "Geodetic deformation monitoring of Ataturk Dam in Turkey," *Arabian Journal of Geosciences*, vol. 7, no. 1, pp. 397–405, 2014.
- [7] X. He, D. Jia, and W. Sang, "Monitoring steep slope movement at Xiaowan Dam with GPS multi-antenna method," *Survey Review*, vol. 43, no. 323, pp. 462–471, 2013.
- [8] X. He, G. Yang, X. Ding, and Y. Chen, "Application and evaluation of a GPS multi-antenna system for dam deformation monitoring," *Earth, Planets and Space*, vol. 56, no. 11, pp. 1035–1039, 2004.
- [9] X. L. Ding, Y. Q. Chen, D. F. Zhu et al., "Slope monitoring using GPS," *GPS World*, vol. 11, no. 3, p. 52, 2000.
- [10] W. P. Jiang, H. F. Liu, W. K. Liu et al., "CORS development for Xilongchi Dam deformation monitoring," *Geomatics and Information Science of Wuhan University*, vol. 37, no. 8, pp. 949–952, 2012.
- [11] Z. H. Li and Z. Z. Liu, "Study on monitoring dam deformation with GPS positioning," *Journal of Wuhan University of Hydraulic and Electric Engineering*, vol. 29, no. 6, pp. 26–29, 1996.
- [12] C. D. Ince and M. Sahin, "Real-time deformation monitoring with GPS and Kalman filter," *Earth, Planets and Space*, vol. 52, no. 10, pp. 837–840, 2000.
- [13] X. F. He, W. G. Sang, and D. Z. Jia, "Deformation monitoring of steep rock slopes by means of GPS," *Journal of Hydraulic Engineering*, vol. 37, no. 6, pp. 746–750, 2006.
- [14] E. Gökalp and L. Taşçı, "Deformation monitoring by GPS at embankment dams and deformation analysis," *Survey Review*, vol. 41, no. 311, pp. 86–102, 2013.
- [15] A. Lokke and A. K. Chopra, "Direct finite element method for nonlinear analysis of semi-unbounded dam-water-foundation rock systems," *Earthquake Engineering & Structural Dynamics*, vol. 46, no. 8, pp. 1267–1285, 2017.
- [16] Y. R. Liu, Q. Yang, and L. Zhu, "Abutment stability analysis of arch dam based on 3D nonlinear finite element method," *Chinese Journal of Rock Mechanics and Engineering*, vol. 27, no. s1, pp. 3222–3228, 2008.
- [17] M. Zhou, B. Zhang, C. Peng, and W. Wu, "Three-dimensional numerical analysis of concrete-faced rockfill dam using dual-mortar finite element method with mixed tangential contact constraints," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 40, no. 15, pp. 2100–2122, 2016.
- [18] C. P. Zhang, D. L. Zhang, J. J. Luo, M. S. Wang, and J. P. Wu, "Remote monitoring system applied to the construction of metro station undercrossing existing metro tunnel," *Rock and Soil Mechanics*, vol. 30, no. 6, pp. 1861–1866, 2009.
- [19] J. Bai, P. Lin, P. Li, and W. Zhou, "Reinforcement effects of Lijixia arch dam on complicated rock foundations and its back analysis," *Chinese Journal of Rock Mechanics and Engineering*, vol. 27, no. 5, pp. 902–912, 2008.
- [20] T. F. Bao, C. S. Gu, and Z. R. Wu, "Analysis of uplift pressure anomaly of Lijixia Dam foundation," *Chinese Journal of Geotechnical Engineering*, vol. 30, no. 10, pp. 1460–1466, 2008.
- [21] Y. H. Zhang, H. X. Li, Q. Sheng, K. Wu, Z. Y. Li, and Z. P. Yue, "Study of highway landslide monitoring and early warning based on surface displacements," *Rock and Soil Mechanics*, vol. 31, no. 11, pp. 3671–3677, 2010.
- [22] K. Wu, Q. Sheng, Y. H. Zhang, Z. Y. Li, H. X. Li, and Z. P. Yue, "Development of real-time remote monitoring and forecasting system for geological disasters at subgrade slopes of mountainous highways and its application," *Rock and Soil Mechanics*, vol. 31, no. 11, pp. 3683–3687, 2010.
- [23] Y. L. Zhang, S. L. Yang, X. Y. Luo et al., "BeiDou/GPS displacement detection using robust precise weighting based on moving window," *Acta Geodaetica et Cartographica Sinica*, vol. 45, no. S2, pp. 72–81, 2016.
- [24] X. M. Song, X. J. Ye, X. Q. Zeng et al., "Study and improvement on equivalence classes of PostgreSQL query optimization," *Computer Engineering and Applications*, vol. 50, no. 14, pp. 31–38, 2014.
- [25] C. Y. Chen, S. J. Wang, and X. K. Shen, "Predicting models to estimate stability of rock slope based on artificial neural network," *Chinese Journal of Geotechnical Engineering*, vol. 23, no. 2, pp. 157–161, 2001.
- [26] J. Q. Sun and H. B. Xu, "Fuzzy-neural network based on T-S model," *Journal of Tsinghua University*, vol. 37, no. 3, pp. 76–80, 1997.
- [27] S. Y. Chen, Y. P. Zhu, and L. H. Huang, "Application of chaotic time series analysis to the prediction of tunnel surrounding rock displacement," *Modern Tunnelling Technology*, vol. 52, no. 3, pp. 75–81, 2015.

Research Article

Tile-Based Semisupervised Classification of Large-Scale VHR Remote Sensing Images

Haikel Alhichri , Essam Othman , Mansour Zuair , Nassim Ammour, and Yakoub Bazi 

Computer Engineering Department, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

Correspondence should be addressed to Haikel Alhichri; haikel.alisr@gmail.com

Received 14 September 2017; Revised 14 January 2018; Accepted 21 January 2018; Published 5 April 2018

Academic Editor: Biswajeet Pradhan

Copyright © 2018 Haikel Alhichri et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper deals with the problem of the classification of large-scale very high-resolution (VHR) remote sensing (RS) images in a semisupervised scenario, where we have a limited training set (less than ten training samples per class). Typical pixel-based classification methods are unfeasible for large-scale VHR images. Thus, as a practical and efficient solution, we propose to subdivide the large image into a grid of tiles and then classify the tiles instead of classifying pixels. Our proposed method uses the power of a pretrained convolutional neural network (CNN) to first extract descriptive features from each tile. Next, a neural network classifier (composed of 2 fully connected layers) is trained in a semisupervised fashion and used to classify all remaining tiles in the image. This basically presents a coarse classification of the image, which is sufficient for many RS application. The second contribution deals with the employment of the semisupervised learning to improve the classification accuracy. We present a novel semisupervised approach which exploits both the spectral and spatial relationships embedded in the remaining unlabelled tiles. In particular, we embed a spectral graph Laplacian in the hidden layer of the neural network. In addition, we apply regularization of the output labels using a spatial graph Laplacian and the random Walker algorithm. Experimental results obtained by testing the method on two large-scale images acquired by the IKONOS2 sensor reveal promising capabilities of this method in terms of classification accuracy even with less than ten training samples per class.

1. Introduction

The intent of the classification process is to categorize all pixels in a remote sensing (RS) image into one of several land-cover classes. Most classification solutions have used RS images with small sizes (typically less than 1000×1000 pixels). However, the latest generation of satellite-based imaging sensors (Pleiades, Sentinel, etc.) acquires big volumes of Earth's images with high spatial, spectral, and temporal resolution. This leads to images of very large sizes and creates new challenges for land-use classification algorithms. In this case, traditional pixel-based algorithms become unfeasible due to (1) increased computational and memory costs, (2) increased spectral variation within the same land-cover class, and (3) increased variability in the data distributions over this large-scale image, and (4) the need

to collect a large set of training samples to properly model the underlying distribution of the data in the image. The latter issue may be caused by the increased resolution which heightens the effect of lighting conditions, the depth dimension, and the sensor's noise. The same land-covers and even the same objects can be found in images belonging to different classes. Thus, there is a great need to develop efficient solutions to classify large-scale images.

One solution to this problem is to use object-based techniques. This can be accomplished first via segmenting the large image into regions and then performing object analysis on the regions. However, segmenting large images is, in itself, still a difficult problem [1]. Another solution is to divide the image into tiles and then perform coarse classification of the image by classifying the tiles instead of pixels as was introduced in [2–4].

The problem of tile classification becomes then similar to object classification or recognition in computer vision applications. In those applications, many local descriptors, like bag of words (BOW) [5], local binary patterns (LBP) [6], and histograms of oriented gradients (HOG) [7], with their invariance to geometric and photometric transformations, have been proven effective especially for object recognition. They can be extracted both in sparse and in dense ways. Recently, deep convolutional neural networks (CNN) [8–10] have shown outstanding results in many computer vision applications such as image classification [11], object recognition [12], face recognition [13], medical image analysis [14], speech recognition [15], and traffic flow prediction [16]. Several CNN architectures have been already proposed and tested in computer vision tasks, and most of them have been implemented and made available online [17].

CNNs use images directly as input, and they automatically learn a hierarchy of features instead of using hand-crafted features. This is accomplished by successively convolving the input image with learned filters to build up a hierarchy of feature maps. The hierarchical approach allows learning more and more complex and highly descriptive set of features. One disadvantages of CNN is the high computational costs during the training phase due to the high number of weights to be estimated (in the range of hundreds of thousands). Another stringent requirement is the need for a huge number of training images [18]. One way to solve this problem is to use CNNs that have been pretrained on huge auxiliary datasets and then transfer the knowledge embedded in them to help extract highly descriptive features in our new domain. These highly descriptive features are then used to train another classifier (such as a smaller neural network) for our particular recognition task.

In the remote sensing community, some works have introduced CNN as a solution to classification and object detection problems [19–23]. For the scene classification problem, the reader is referred to this work by Cheng et al. [24] for a recent review of the state-of-the-art review in the area. As a set of sample work, Castelluccio et al. proposed a solution for the classification of RS scenes by starting with a pretrained CNN and then fine-tuning it on the target data in order to improve the accuracy [19]. Zhang et al. proposed a method called gradient boosting random convolutional network (GBRCN), which trains and fuses the result of several CNNs for the purpose of RS scene classification [22]. Chen et al. proposed a method to detect small objects such as vehicles in satellite images using a CNN to extract and learn rich features from the training data [21]. Marmanis et al. extracted feature representations using a pretrained CNN and then fed them into another CNN, which is trained in a supervised manner [20]. Another more recent work proposes a novel new feature for scene classification called bag of convolutional features (BoCF) [23]. This is similar to bag of words concept, except that the codebook is composed of convolutional feature vectors collected from the 5th layer of the VGGNet-16 CNN when it is applied to all training and testing scenes.

Another important issue in RS classification is that the collection of a statistically significant and representative

amount of training (labelled) samples is a tedious task. Thus, the RS community has introduced semisupervised learning (SSL) methods to tackle this problem [25–29]. SSL methods attempt to exploit the wealth of information provided by unlabelled data, besides the few available labelled data, to improve the performance of the classifier. SSL is based on the assumption that points within the same structure (such as a cluster or a manifold) are likely to have the same label [29]. We can use the large amount of unlabelled data to uncover such structure. For example, in RS images, it is reasonable to assume that samples are likely to have similar labels, if they have close spectral information or if they are neighbours spatially. Many machine-learning researchers have found that unlabelled data, when used in conjunction with a small amount of labelled data, can produce considerable improvement in learning accuracy [30].

Semisupervised learning is natural for deep CNN architectures since the latter requires huge training data and in practice, either we never have enough labelled data or it is quite costly to get them. Nonlinear embedding algorithms are popular for use with shallow semisupervised learning techniques such as kernel methods, but they can be applied to deep multilayer architectures as well, as a regularizer either at the output layer or on the hidden layers of the architecture [31]. Graph-based techniques are one way to perform semisupervised learning. In this approach, one tries to preserve the structure present in the unlabelled data by building a graph where each sample data point is a node and the edges encode the similarity between data points. In the literature, graph Laplacian is embedded at the hidden layer and optimized in a joint way within the error function of the network [29]. Another approach is to use the graph Laplacian as a regularizer at the output layer such as in Markov random field (MRF).

In this work, we present an effective solution for the classification of large-scale VHR RS images, where the number of labelled samples is limited (no more than ten tiles per class). To deal with the huge size of the VHR images, we subdivide the large image into square tiles and classify them instead of classifying pixels. This is our first contribution in this paper. Our second contribution is presenting a deep architecture for semisupervised classification of large-scale images. This deep architecture is composed of a pretrained CNN, for feature extraction, followed by two fully connected layers for classification. The proposed solution accomplishes semisupervised learning via simultaneous embedding of a graph Laplacian in the hidden layer followed by a spatial graph Laplacian after the output layer for regularization. The experimental results, carried out on two large-scale VHR images, have shown that embedding a graph Laplacian at the feature level and at the output level simultaneously provides significant improvement in the classification result.

The rest of the paper is organized as follows: in Section 2, we present the details of proposed method. In Section 3, we present the experimental results and discuss them. Finally, we present our concluding remarks and future works in Section 4.

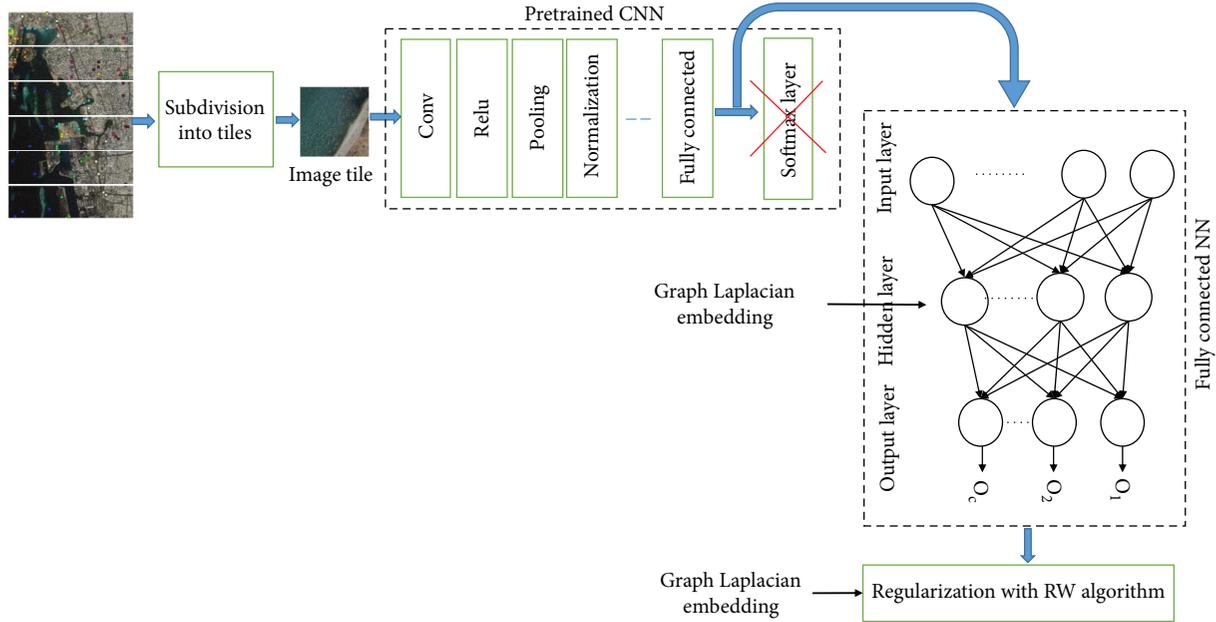


FIGURE 1: Overview of the proposed deep CNN solution for the semisupervised classification of large RS images.

2. Materials and Methods

The general overview of our proposed solution is shown in Figure 1. The first step in this solution is to divide the large image into a grid of tiles, and the goal is to classify these tiles instead of pixels to produce a coarse tile-based classification of the RS image. Let I be a large VHR RS image that is divided into a grid of N_t tiles $\{I_i\}_{i=1}^{N_t}$, where $I_i \in \mathcal{R}^{w \times w}$ represents tile i of size $w \times w$. The second step in the algorithm is feature extraction using one of the pretrained CNNs found in the literature. Let $\{\mathbf{x}_i\}_{i=1}^{N_t}$ be the set of features extracted from each one of the N_t tiles. Furthermore, let $\text{Tr} = \{\mathbf{x}_i, y_i\}_{i=1}^n$ be a training set of labelled sample tiles, where $y_i \in \{1, 2, \dots, C\}$ are the corresponding class labels.

The classification step is carried out using a fully connected neural network (NN), as shown in Figure 1, with one hidden layer and an output (softmax) layer. As mentioned previously, using this proposed architecture, we can implement semisupervised learning in different ways. We can embed the graph Laplacian at the hidden layer of the fully connected neural network (i.e., at the feature level). We can also embed a graph Laplacian at the output level, which implements what is called output regularization. Finally, we can implement a suitable combination thereof. Our final proposed solution shown in Figure 1 embeds a graph Laplacian (spectral or spatial) in the hidden layer of the NN and also adds an additional layer that performs regularization of the output labels. In the remainder of this section, we explain how to use a pretrained CNN for feature extraction. Then the next subsection shows how to perform classification using a fully connected neural network. Finally, we explain how to achieve semisupervised learning via embedding a graph Laplacian at the hidden layer of the neural network and also embedding another graph Laplacian to perform output regularization spatially.

2.1. Feature Extraction from Tiles. The second step in the proposed solution is to extract feature descriptors from the image tiles. Several types of handcrafted feature descriptors are presented in the literature for image representation, such as bag of words (BOW) [5], local binary patterns (LBP) [6], or histograms of oriented gradients (HOG) [7]. However, more recently, the new state-of-the-art approach is to use feature descriptors that can be learned instead of handcrafted features. CNNs play a major role in this regard as they have been proven effective in feature learning from images. However, as CNN needs a huge amount of training data, one can resort to pretrained CNN as a compromise solution. In this case, a pretrained CNN becomes like a feature extraction tool instead of a classifier as it is originally designed. This approach exploits the huge information embedded in pretrained CNNs to provide highly descriptive features from the tiles.

A CNN is composed of several layers that do different functions such as convolution, pooling, activation function, and classification as shown in Figure 1. The last layer is always a softmax layer which is used for classification. Given this deep nature, one can extract features at different layers of the CNN. In this work, we take the output of the hidden layer before the softmax layer as feature vector representation. Thus, given a CNN with L layers, we feed each tile image I_i , $i = 1, \dots, N_t$ as input to the CNN and generate a feature representation $\mathbf{x}_i \in \mathcal{R}^D$ at layer k , of the CNN, with $k = L - 1$, that is,

$$\mathbf{x}_i = f_k^{\text{CNN}} \left(\dots f_2^{\text{CNN}} \left(f_1^{\text{CNN}} (I_i) \right) \right), \quad i = 1, \dots, N_t, \quad (1)$$

where D is the dimensionality of the feature vector.

2.2. Classification via a Fully Connected Layer. Given a training set $\text{Tr} = \{\mathbf{x}_i, y_i\}_{i=1}^n$ composed of feature vectors and their corresponding labels, the hidden layer takes the

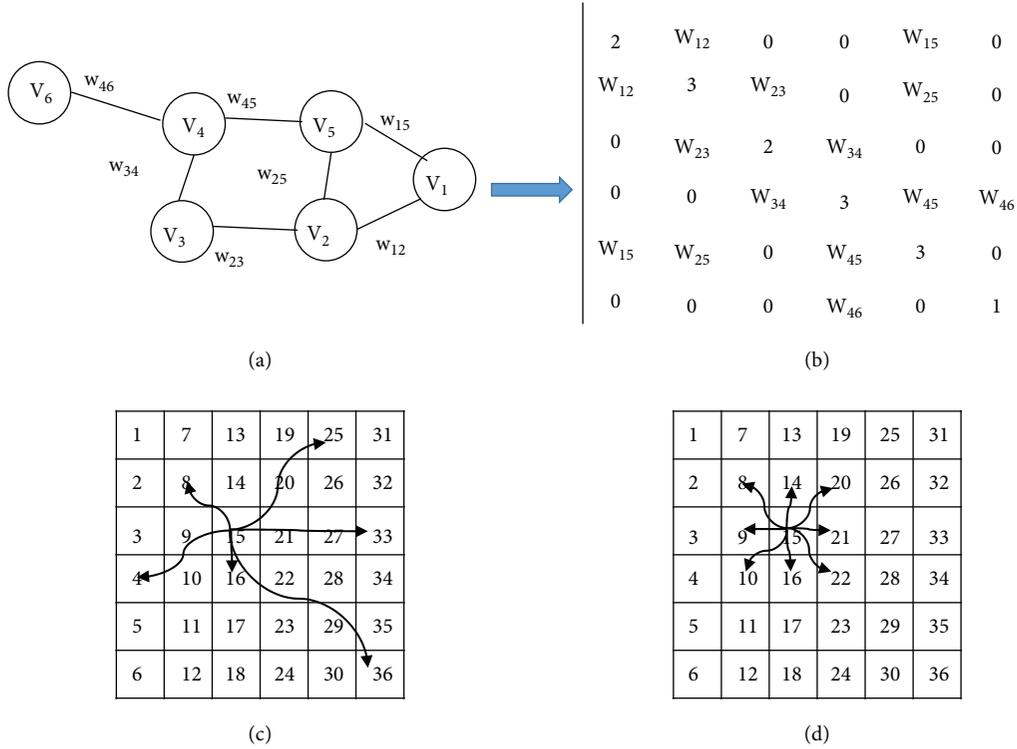


FIGURE 2: Illustration of graph Laplacian building. (a) Sample graph with 6 vertices and 7 edges. (b) Laplacian matrix. (c) Spectral graph Laplacian. Each tile (a node in the graph) is connected to the k -most similar tiles. (d) Spatial graph Laplacian, each tile (a node in the graph) is connected to its 8 neighbours.

input x_i and maps it to another representation $h_i^{(1)} \in \mathfrak{R}^{d^1}$ of dimension d^1 through the nonlinear activation function f as follows:

$$h_i^{(1)} = f\left(\mathbf{W}^{(1)}x_i\right), \quad (2)$$

where $\mathbf{W}^{(1)} \in \mathfrak{R}^{d^1 \times d}$ is the mapping weight matrix. A typical choice of the activation function is the sigmoid function, that is, $f(v) = 1/(1 + \exp(-v))$. For simplicity, we omit the bias vector in the expression as it can be incorporated as an additional column vector in the mapping matrix; then in that case, the feature vector should be appended by the value 1.

The *softmax* layer performs multiclass classification and takes as input the resulting hidden representation $h_i^{(1)}$. It produces, as a result, an estimate of the posterior probability for each class label $c = 1, 2, \dots, C$ as follows:

$$p(\hat{y}_i = c | x_i) = \frac{\exp\left(\left(w_c^{(2)}\right)^T h_i^{(1)}\right)}{\sum_{j=1}^C \exp\left(\left(w_j^{(2)}\right)^T h_i^{(1)}\right)}, \quad (3)$$

where $\mathbf{W}^{(2)} = [w_1^{(2)} w_2^{(2)} \dots w_C^{(2)}] \in \mathfrak{R}^{d \times C}$ are the weights of the *softmax* layer and the superscript $(\bullet)^T$ refers to the transpose operation. To learn the weights $\theta = \{\mathbf{W}^{(1)}, \mathbf{W}^{(2)}\}$ representing the complete network structure, we minimize the training error on the labelled training data. The cost function is then formulated as follows:

$$J(\theta, D^{(1)}) = \arg \min E_{\text{net}}(D^{(1)}), \quad (4)$$

where $D^{(1)}$ is the set of labelled data and $E_{\text{net}}(D^{(1)})$ is the cross-entropy loss, which measures the error between the actual network outputs and the desired outputs of the labelled source data. As the outputs of the network are probabilistic, we propose to maximize the log-posterior probability to learn the network weights, which is equivalent to minimizing the so-called cross-entropy error:

$$E_{\text{net}}(D^{(1)}) = -\frac{1}{n} \sum_{i=1}^{n_s} \sum_{k=1}^K 1(y_i = k) \ln \left(\frac{\exp\left(\left(w_k^{(2)}\right)^T h_i^{(1)}\right)}{\sum_{j=1}^K \exp\left(\left(w_j^{(2)}\right)^T h_i^{(1)}\right)} \right), \quad (5)$$

where $1(\bullet)$ is an indicator function that takes 1 if the statement is true otherwise it takes 0 and the superscript T refers to matrix transpose.

2.3. Graph-Based Semisupervised Learning. The idea of graph-based semisupervised learning is to build a graph that connects similar sample data points to each other, and then use this graph to propagate estimated labels among similar data points. This follows directly from the assumption that similar data points should have similar labels, which the graph tries to encode.

Figure 2 illustrates how a Laplacian matrix is built using a simple example. Let $G = (V, E)$ be a graph with vertices $v \in V$

and edges $e \in E$. Let e_{ij} denote an edge spanning two vertices \mathbf{v}_i and \mathbf{v}_j and ω_{ij} is the associated weight. The degree of a vertex \mathbf{v}_i is $d_i = \sum \omega_{ij}$ for all edges e_{ij} incident on \mathbf{v}_i . Then the tile-based Laplacian matrix indexed by vertices \mathbf{v}_i and \mathbf{v}_j is given by:

$$\mathbf{L}_{ij} = \begin{cases} d_{ij}, & \text{if } i = j, \\ -\omega_{ij}, & \text{if } i \text{ and } j \text{ are connected,} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

In our case, each tile image is associated with a vertex in the graph, while edges are defined in several ways. The k -nearest graph, for example, connects each vertex to the k -most similar vertices. The number k is usually left as a parameter to be determined. Another approach is to connect each vertex to all vertices whose similarity is below a certain threshold value. This is known as a spectral graph and is illustrated in Figure 2(d).

However, here, the tiles are also spatially dependent. Thus, we can also consider adjacency as a similarity, that is, we can make the assumption that neighbouring tiles are similar and should have similar labels. In this case, the number of connections per node is usually limited to two choices: 8 or 4 depending on whether we want to consider all 8 neighbouring tiles or only 4. This is known as a spatial graph and is illustrated in Figure 2(c).

Finally, a weight w_{ij} is associated with each edge e_{ij} in the graph. This weight is defined as the similarity between image tiles corresponding to vertices \mathbf{v}_i and \mathbf{v}_j . It can be calculated using many types of similarity measures or distance measures. A common choice for obtaining these weights is the Gaussian weighting function, that is, $\omega_{jk} = \exp(-\beta \|\mathbf{x}_j - \mathbf{x}_k\|^2)$ and β is a free parameter (usually set to 1).

2.3.1. Embedding Graph Laplacian in the Hidden Layer. The first approach to implement semisupervised learning is by embedding a spectral or spatial graph Laplacian in the hidden layer. The fully connected neural network is supplied with labelled and unlabelled data. Then to learn the weights $\theta = \{\mathbf{W}^{(1)}, \mathbf{W}^{(2)}\}$ representing the complete network structure, we propose to simultaneously minimize: (i) the training error on the labelled data and (ii) the energy of the graph Laplacian built on all unlabelled tiles. The proposed cost function is then formulated as follows:

$$J(\theta, D^{(1)}, D^{(u)}) = \arg \min E_{\text{net}}(D^{(l)}) + \lambda_1 \text{Lap}(D^{(u)}), \quad (7)$$

where λ_1 is a regularization parameter. The first term is the same as (4), while the second term is added for graph regularization. This basically minimizes the energy of the graph and can be written as follows:

$$\text{Lap}(D^{(u)}) = \frac{1}{2} \sum_{e_{jk} \in E} \omega_{jk} \left\| \mathbf{W}^{(1)} \mathbf{x}_j - \mathbf{W}^{(1)} \mathbf{x}_k \right\|_2^2. \quad (8)$$

It can be shown that (8) can be written in the following matrix form:

$$\text{Lap}(D^{(u)}) = \frac{1}{2} \text{Trace} \left(\left(\mathbf{W}^{(1)} \mathbf{X}^{(u)} \right) \mathcal{L} \left(\mathbf{W}^{(1)} \mathbf{X}^{(u)} \right)^T \right), \quad (9)$$

where $\mathbf{X}^{(u)} = [\mathbf{x}_1 \ \dots \ \mathbf{x}_{n_u}] \in \mathcal{R}^{n_u \times d^l}$ represents the unlabelled samples.

By substituting the terms in (5) and (9), the total cost function $J(\theta)$ is then given by

$$\begin{aligned} J(\theta, D^{(1)}, D^{(u)}) = \arg \min & -\frac{1}{n} \sum_{i=1}^{n_l} \sum_{k=1}^K 1(y_i = k) \\ & \cdot \ln \left(\frac{\exp \left(w_k^{(2)T} h_k^{(1)} \right)}{\sum_{l=1}^K \exp \left(w_l^{(2)T} h_l^{(1)} \right)} \right) + \frac{\lambda_1}{2} \\ & \cdot \text{Trace} \left(\left(\mathbf{W}^{(1)} \mathbf{X}^{(u)} \right) \mathcal{L} \left(\mathbf{W}^{(1)} \mathbf{X}^{(u)} \right)^T \right). \end{aligned} \quad (10)$$

2.3.2. Spatial Regularization of the Output Layer. Another approach to exploit the graph Laplacian is an output regularizer based on the random Walker (RW) algorithm. This technique is similar to a Markov random field (MRF) approach; however, unlike the MRF, it yields a closed-form solution as opposed to an iterative one [32]. This regularization step needs a spatial graph Laplacian and posterior probability estimates as initial inputs. Then the RW algorithm can improve the classification result by minimizing an energy function of the graph Laplacian, defined as follows:

$$\min_f \sum_{e_{ij} \in E} \omega_{ij} \|\mathbf{f}_i - \mathbf{f}_j\|^2 + \lambda_2 \sum_{\mathbf{v}_i \in \mathbf{V}} \mu_i \|\mathbf{f}_i - \mathbf{f}_i^*\|^2, \quad (11)$$

where \mathbf{f}_i and \mathbf{f}_j are now probability vectors (of dimension C) associated with vertices (tiles) \mathbf{v}_i and \mathbf{v}_j , respectively. The first term is related to image smoothness since it computes the norm between pixels. The second term is related to data fidelity since it computes a norm between the initial estimation \mathbf{f}_i^* (provided by the fully connected NN) and the new label estimates \mathbf{f}_i . Finally, μ_i and λ_2 are local and global weights, respectively, enforcing that fidelity. Moreover, the local weights μ_i could be set equal to d_i , but for simplicity, we suppose $\mu_i = 1$ for all tiles in the image.

The set of vertices of the graph can be written as $\mathbf{V} = \mathbf{V}_T \cup \mathbf{V}_U$, where \mathbf{V}_T and \mathbf{V}_U are the set of labelled and unlabelled vertices, respectively. In this case, the Laplacian matrix \mathbf{L} can be written as $\mathbf{L} = \begin{bmatrix} \mathbf{L}_T & \mathbf{B} \\ \mathbf{B} & \mathbf{L}_U \end{bmatrix}$. After differentiating, setting the result to 0 and solving the matrix

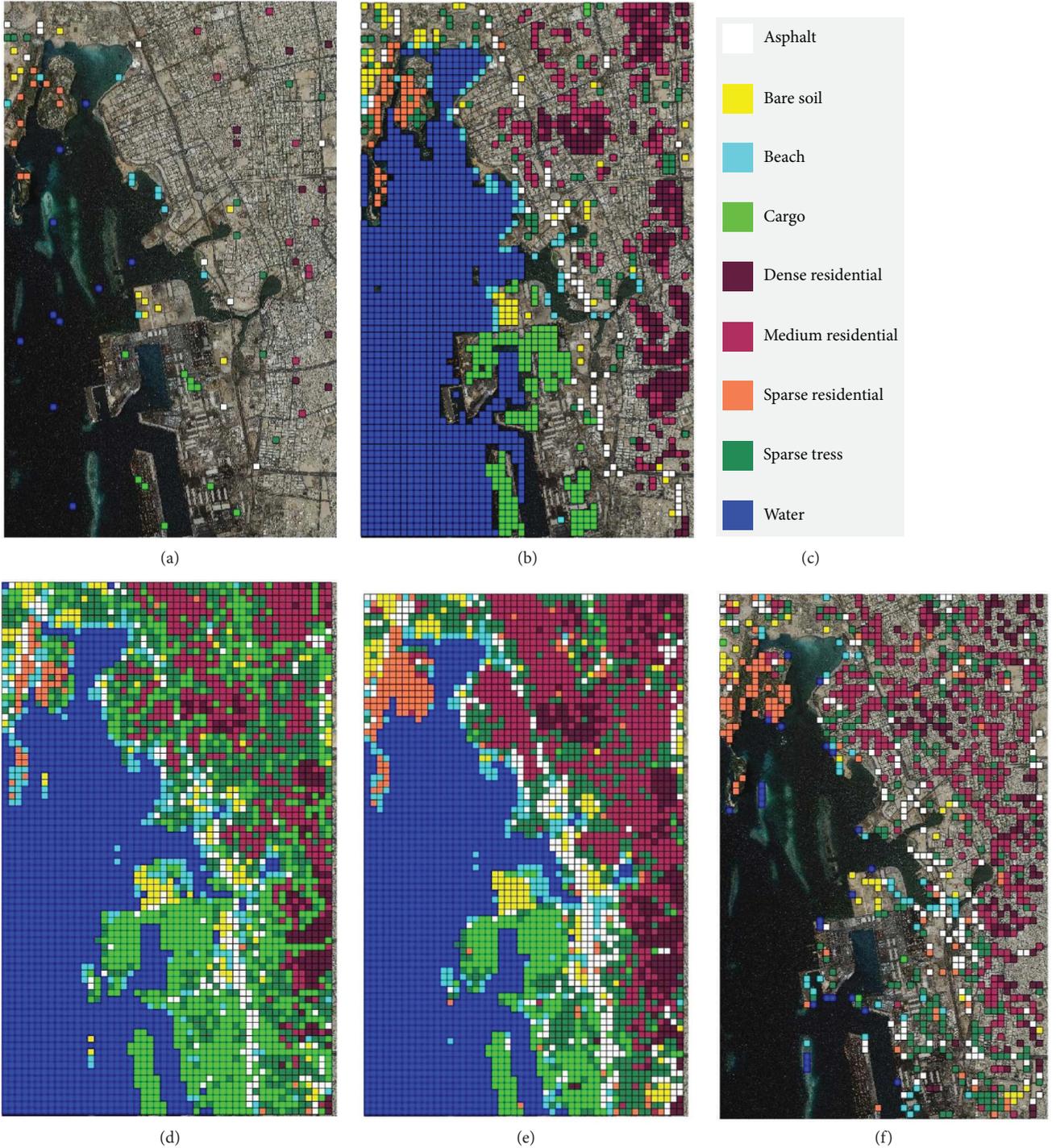


FIGURE 3: Large VHR dataset used and its qualitative results. (a) Original image of size 6500×10400 with example training tiles (10 per class), (b) testing map (ground truth), (c) legend, (d) NN classification result (OA = 89%), (e) final map (OA = 92.2%), and (f) tiles corrected by semisupervised learning.

equation, we found that the probabilities associated with the unlabelled tiles are obtained by the following closed-form solution:

$$\mathbf{F}_U = (\mathbf{L}_U + \lambda_2 \mathbf{I})^{-1} (-\mathbf{B}^T \mathbf{F}_T + \lambda \mathbf{F}_U^*), \quad (12)$$

where

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_T \\ \mathbf{F}_U \end{bmatrix} \quad (13)$$

and \mathbf{F}_U^* are the initial estimated probabilities (posterior probabilities of the fully connected NN classifier) associated with the unlabelled tiles. Note that $\mathbf{L}_U + \lambda_2 \mathbf{I}$ is a positive semidefinite matrix, and \mathbf{L} is a sparse (Laplacian) matrix.

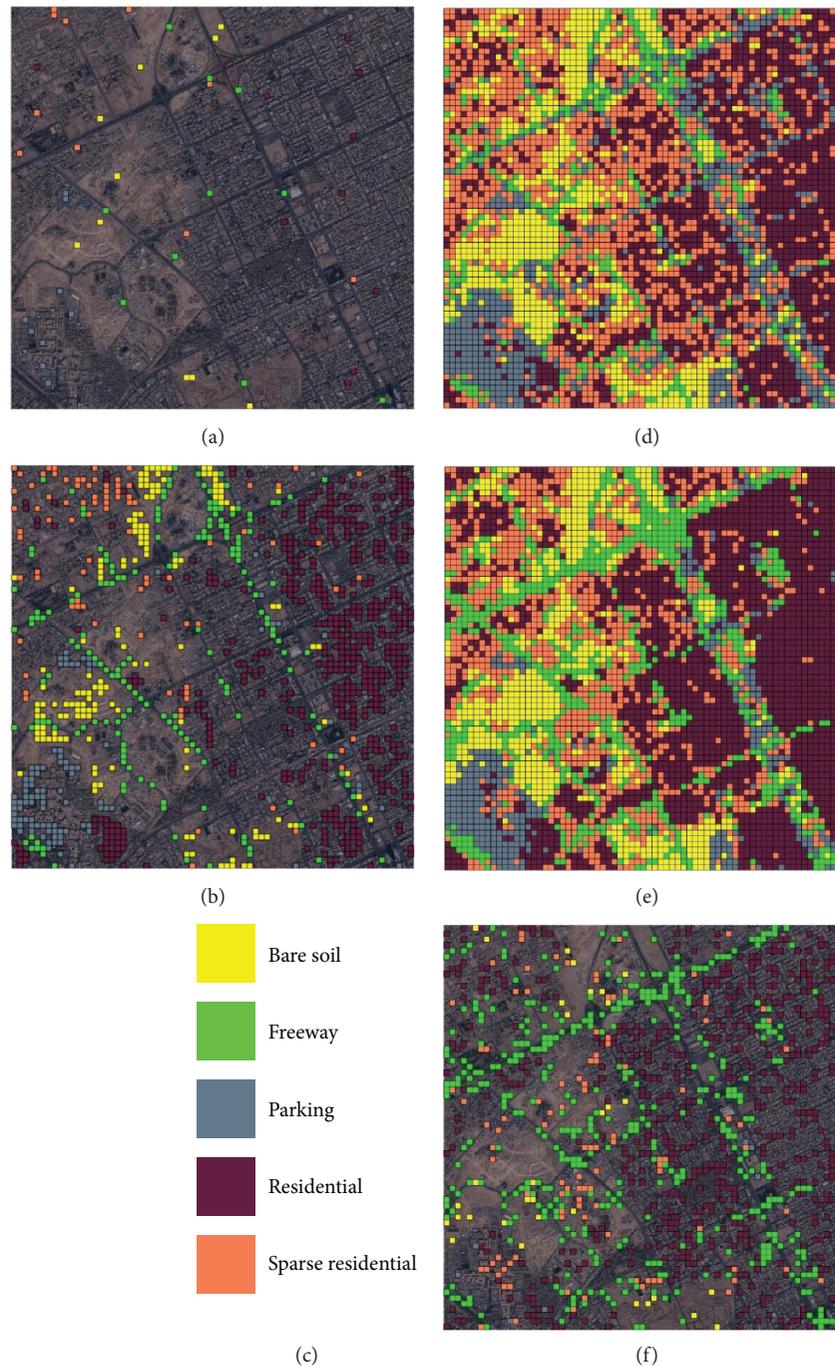


FIGURE 4: Large VHR dataset used and its qualitative results. (a) Original image of size 13440×13440 with example training tiles (10 per class), (b) testing map (ground truth), (c) legend, (d) NN classification result (OA = 91%), (e) final map (OA = 96.5%), and (f) tiles corrected by semisupervised learning.

Thus, the linear system of equations in (12) can be solved in linear time.

3. Results

3.1. Dataset Description. We tested the proposed solution on two large-scale VHR RS images. The first image is of size 6500×10400 taken over the city of Jeddah in Saudi Arabia by the IKONOS-2 sensor in July 2004 (see Figure 3(a)). The

image has three spectral bands with a spatial resolution of 1 m. It has been divided into 3977 tiles of size 128×128 , of which 1949 tiles are unlabelled, while 2028 are labelled into nine land-cover types including asphalt (73 tiles), bare soil (61 tiles), beach (60 tiles), cargo (174 tiles), dense residential (203 tiles), medium residential (193 tiles), sparse residential (47 tiles), trees (74 tiles), and water (1216 tiles).

The second image is of size 14000×14000 and is taken from the city of Riyadh by GeoEye sensor in June 2010

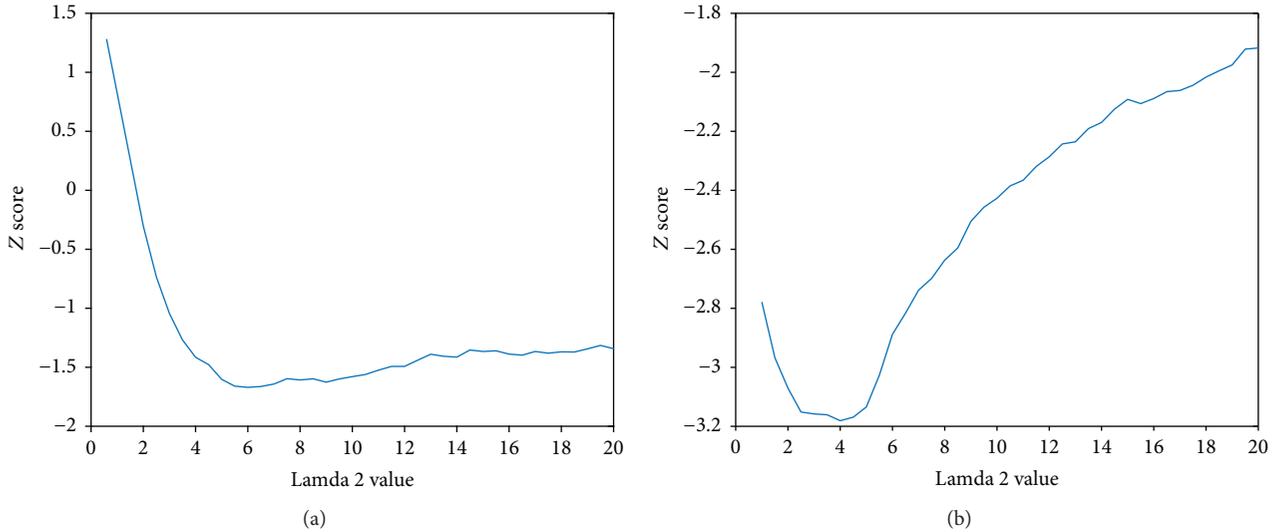


FIGURE 5: Results of output regularization via spatial graph Laplacian with sensitivity analysis of the parameter λ_2 in (11), (a) for Riyadh dataset and (b) for the Jeddah dataset.

(see Figure 4(a)). The image has three spectral bands with a spatial resolution of 0.5 m. It has been divided into 4900 tiles of which 3860 are unlabelled, while 1040 are labelled into five land-cover types including bare soil (163 tiles), freeway (185 tiles), parking (109 tiles), residential (486 tiles), and sparse residential (97 tiles).

3.2. Experimental Setup. Both large images have been subdivided into a grid of equal size tiles. The Jeddah image is divided in tiles of size 128×128 pixels, while the Riyadh image tiles are of size 196×196 pixels. We start the division from the top left corner, and any leftover pixels (in case the size of the large image is not a multiple of 128) appear at the right and bottom sides of the image. The ground truth classification maps are created manually by inspecting a random set of tiles (the test set) one by one and assigning one of the 9 labels to all tiles in the test set. Testing maps for both the Jeddah and Riyadh images are shown in Figures 5(b) and 6(b), respectively. For both images, we extract feature representations from all tiles including HOG, LBP, BOW, and several pretrained CNN architectures. The set of pretrained CNNs used includes the imagenet-vgg-very-deep-16 and imagenet-vgg-m models by the VGG group at the University of Oxford [33, 34], the GoogleNet model developed by a team from Google Inc. [35], and Microsoft's ResNet-152 model [36]. These CNNs are selected because they have won the first and second places in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC-) 2014 and 2015 challenges. Table 1 shows the sizes of each feature representation.

As for the assessment method, we use the Overall Accuracy (OA) and Average Accuracy (AA) measures. To assess the statistical significance of the improvements achieved by applying the semisupervised learning, we

computed McNemar's test, which is based on the standardized normal test statistic:

$$Z_{ij} = \frac{f_{ij} - f_{ji}}{\sqrt{f_{ij} + f_{ji}}}, \quad (14)$$

where Z_{ij} measures the pairwise statistical significance of the difference between the accuracies of the i th and j th classifiers. The variable f_{ij} stands for the number of samples classified correctly and wrongly by the i th and j th classifier, respectively. Thus, f_{ij} and f_{ji} are the counts of the classified samples on which the two classifiers disagree. At the commonly used 5% level of significance, the improvement of classifier j over i is said to be significant if $Z_{ij} < -1.96$. Finally, all experiments are repeated 10 times with different random training tiles and then the average accuracy values are presented.

3.3. Comparison of Various Pretrained CNN and Finding the Best Hidden Layer Size. In this first experiment, we compare the performance of feature representations obtained by the pretrained CNNs and the ones obtained by BOW, LBP, and HOG. We also investigate in this experiment the sensitivity of the results with respect to the hidden layer size for all the types of feature descriptors considered in this study. We train a fully connected NN with the following hidden layer sizes 32, 64, 128, 256, 512, and 1024. We perform the experiment 10 times; each time, we select a different set of training sample tiles per class randomly and average the accuracy results.

Figure 7 shows the overall accuracies (OA) achieved for each type of feature descriptor and with different hidden layer sizes. It is clear from this experiment that CNN features are quite superior to all other types of handcrafted features regardless of the hidden layer size. This confirms other

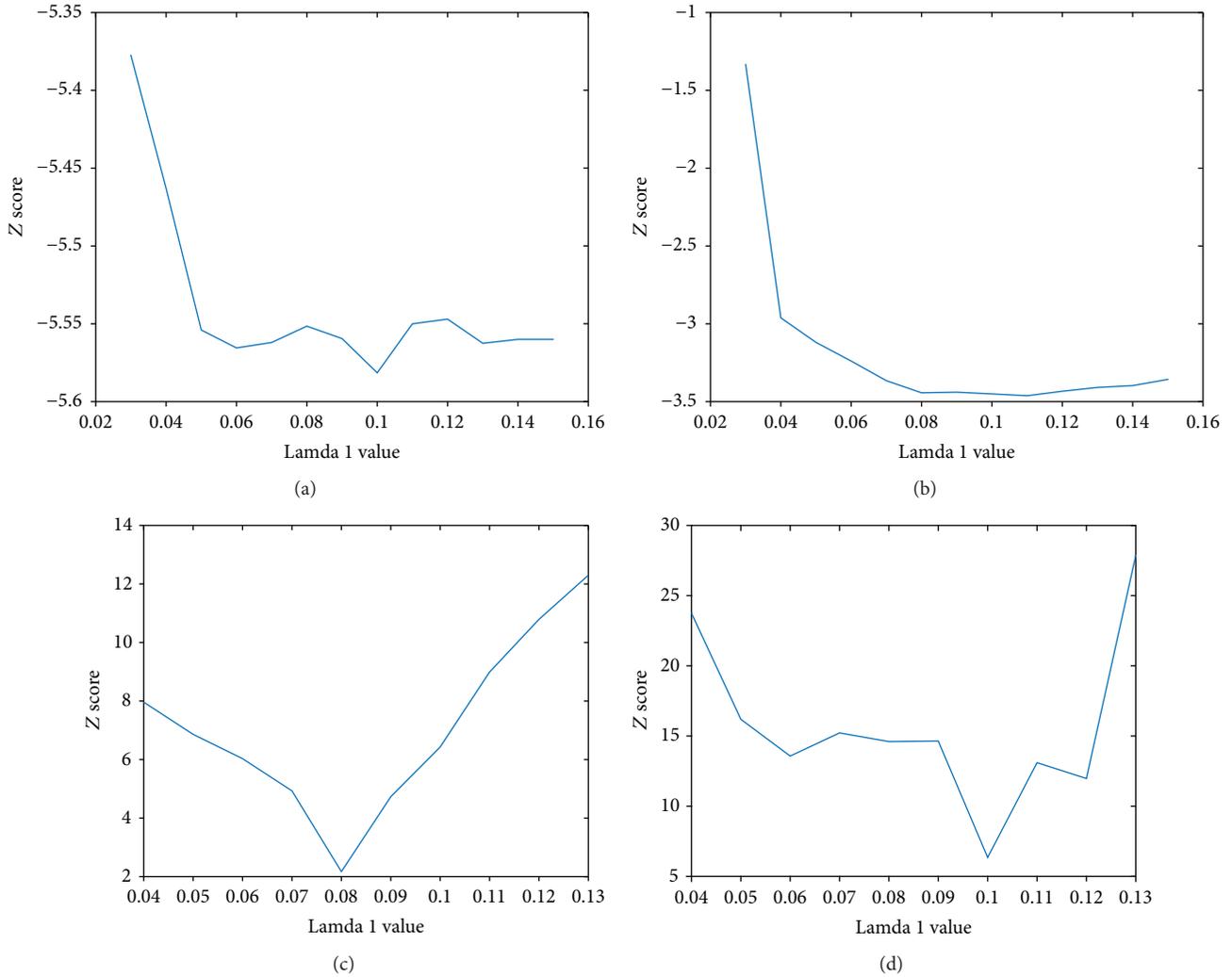


FIGURE 6: Results of embedding a graph Laplacian in the hidden layer with sensitivity analysis of the parameter λ_1 in (10). (a) Riyadh dataset with spectral graph. (b) Jeddah dataset with spectral graph. (c) Riyadh dataset with spatial graph. (d) Jeddah dataset with spatial graph.

TABLE 1: Parameters used by various feature extraction techniques.

Feature type	Feature vector length	Note
HOG	4365	
LBP	256	8 neighbourhood, uniform pattern
BOW	4000	Vocabulary size = 4000
CNN—vgg-very-deep-16	4096	
CNN—GoogLeNet	1024	
CNN—vgg-m	1024	
CNN—ResNet-152	2048	

results about CNN that is reported in the object recognition literature.

Furthermore, if we consider CNN type features only, VGG very-deep-16 pretrained CNN, in particular, provided the best OA. We can also conclude that in general the range [64, 256] constitutes the best values for the

hidden layer size. For the remainder of this paper, the hidden layer size of 256 and the VGG very-deep-16 pretrained CNN will be used in all experiments.

3.4. Semisupervised Learning Results. In this set of experiments, we present the results of applying semisupervised learning via graph Laplacian. We first implement embedding a spatial graph Laplacian as an output regularizer based on the RW algorithm without embedding any graph in the hidden layer. This step has a regularization parameter λ_2 as shown in (11). In Figure 5, we show the Z score achieved for different values of the regularization parameter in the range of [1 ... 20]. From these results, we learn that output regularization does not provide significant improvement for the Riyadh dataset as the Z score is greater than -1.96 .

However, the improvement for the Jeddah dataset is significant (< -1.96). We also learn that the best value for the parameter λ_2 is around 4 to 6; thus, we decided to fix this parameter to 5 for optimal performance of the method.

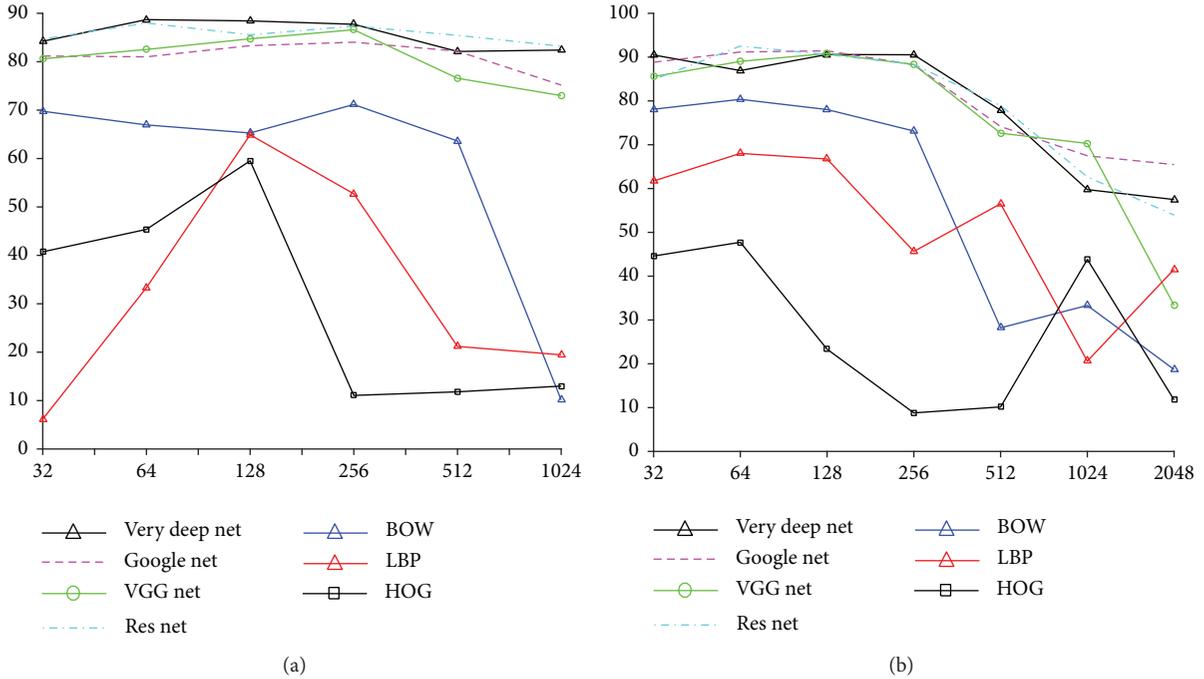


FIGURE 7: Comparison between various feature descriptors in terms of overall accuracy (OA). (a) OA for the Jeddah-IKONOS2 dataset and (b) OA for the Riyadh-GEOEYE dataset.

In the second experiment, we present the results of embedding a graph Laplacian in the hidden layer. We investigated embedding both a spectral graph Laplacian and spatial graph Laplacian. The cost function to be optimized, shown in (10), has a regularization parameter λ_1 . Figure 6 shows the Z score achieved for different values of the regularization parameter in the range of $[0.04 \dots 0.15]$. From these results, we learn that embedding a spatial graph Laplacian in the hidden layer degrades the classification accuracy significantly as the Z scores in Figures 6(c) and 6(d) are all above -1.96 . On the other hand, embedding a spectral graph Laplacian in the hidden layer does provide significant improvement for both datasets as the Z score is less than -1.96 . As for the best value for the parameter λ_1 , it is clear, from Figures 6(a) and 6(b), that it is around the values 0.08 to 0.1; thus, we decided to fix this parameter to 0.1 for optimal performance of the method.

Next, we present the results of the complete system shown in Figure 1. The results for the Jeddah and Riyadh image are shown in Tables 2 and 3, respectively. These tables show the OA, AA, and per class accuracy for 5 different scenarios: (1) using the fully connected NN only, (2) using the NN followed by spatial regularization using RW algorithm, (3) using the NN with spatial graph Laplacian in the hidden layer, (4) using the NN with spectral graph Laplacian in the hidden layer, and finally, (5) using the NN with both spectral graph Laplacian in the hidden layer followed by spatial regularization using RW algorithm, which is the complete solution proposed in this paper. These tables also show execution times in seconds, which include both training and testing times.

The results show clearly that semisupervised learning has improved the accuracy significantly for all types of CNN. However, in agreement with an earlier experiment, the best results achieved are with the *vgg-very-deep-16* pretrained CNN, with an OA of 91.23% and an AA of 83.68%. Also, the semisupervised learning has significantly improved the accuracy of the NN as the Z score reached -4.67 . Similarly for the Riyadh image, the results shown in Table 3 indicate an even better improvement induced by semisupervised learning with a Z score below -5 for all types of CNN. Again, the best results are achieved using the *vgg-very-deep-16* pretrained CNN, with an OA of 95.26%, an AA of 92.91%, and a Z score of -5.40 .

Finally, qualitative results are also shown in Figures 3 and 4 for the Jeddah and Riyadh datasets, respectively. Here, we show the classification maps that have achieved the highest OA after semisupervised learning among the 10 successive runs of the algorithm. In Figures 3(e) and 4(e), we show the tiles whose labels are corrected by semisupervised learning. Comparing Figure 3(d) to Figure 3(c), one can see the type of improvement produced by semisupervised learning, namely, the reduced noise particularly in the water and in residential classes. As for the Riyadh classification maps shown in Figures 4(c) and 4(d), the improvements are clearly visible in the freeway class, where it is more recognizable as roads.

In this final experiment, we compare the proposed method to a state-of-the-art method in the literature called the Laplacian extreme learning machine (LapELM) method [37]. To make the comparison fairer, we have changed the features used in the original method [37] to the same CNN features used here. Thus, the two methods

TABLE 2: Classification results for the Jeddah dataset using one hidden layer of size 256.

	OA	AA	Time (s)	Z score	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
VGG-m													
NN	86.05	72.84	0.23	0.00	60.32	82.35	54.40	82.38	56.58	73.77	77.03	71.72	96.97
NN + RW	86.42	73.24	0.04	-1.23	60.32	81.76	53.60	84.45	53.68	76.56	77.57	73.91	97.28
Spatial graph in the hidden layer	58.24	36.42	19.86	20.96	17.30	61.57	18.40	8.36	21.86	52.29	36.49	34.22	77.33
Spectral graph in the hidden layer	87.75	77.41	25.29	-2.71	80.00	84.90	57.40	81.77	80.57	59.34	86.22	69.69	96.81
NN + spectral graph + RW	88.33	78.08	0.04	-3.59	80.64	84.71	56.80	83.05	81.25	61.97	85.95	71.41	97.01
GoogleNet													
NN	84.48	73.23	0.23	0.00	71.75	87.45	55.20	71.89	75.86	44.81	83.78	72.35	96.00
NN + RW	85.21	74.39	0.04	-2.44	72.86	87.45	55.20	75.80	76.94	46.12	84.59	74.38	96.14
Spatial graph in the hidden layer	77.37	58.45	20.05	6.59	47.30	87.65	11.60	78.72	70.21	35.57	58.92	45.31	90.81
Spectral graph in the hidden layer	85.97	76.82	26.38	-2.57	73.65	91.57	60.60	81.10	87.82	39.07	91.08	70.94	95.55
NN + spectral graph + RW	86.46	77.76	0.04	-3.32	74.92	91.77	60.80	84.27	89.33	37.60	91.89	73.59	95.67
ResNet													
NN	87.68	79.33	0.37	0.00	81.59	83.14	63.80	91.04	75.13	70.00	81.35	73.59	94.36
NN + RW	88.33	79.69	0.05	-2.59	81.27	84.51	62.60	91.28	75.49	70.22	82.70	73.91	95.26
Spatial graph in the hidden layer	78.82	62.59	33.94	10.27	46.03	93.34	33.80	82.38	61.40	64.86	39.73	52.97	88.78
Spectral graph in the hidden layer	88.65	82.96	38.38	-2.00	84.76	91.96	78.00	91.95	79.22	69.73	86.76	70.16	94.12
NN + spectral graph + RW	89.01	82.92	0.06	-2.63	84.29	92.16	76.60	92.44	79.64	69.67	87.30	69.53	94.70
Very-deep-16													
NN	88.41	78.52	0.62	0.00	74.76	88.63	67.60	96.34	57.67	76.45	82.16	65.63	97.45
NN + RW	89.37	78.98	0.09	-3.77	80.95	88.24	84.00	93.90	50.26	86.89	81.08	51.56	96.85
Spatial graph in the hidden layer	68.79	56.74	60.98	16.15	44.60	67.06	29.60	54.15	67.41	48.64	72.16	49.06	77.97
Spectral graph in the hidden layer	90.28	83.45	96.01	-3.00	82.86	90.59	74.80	93.96	74.77	73.28	95.67	68.28	96.86
NN + spectral graph + RW	91.23	83.68	0.09	-4.67	84.13	90.59	72.40	94.39	74.51	74.48	95.13	69.22	98.25

now use the same CNN features extracted with the help of the vgg-very-deep-16 pretrained CNN. As for the other parameters of the LapELM method, we have set the ranges for the regularization parameter C to $[0.001, 1000]$ and the RBF kernel parameter γ to $[0.01, 5]$. To estimate the best parameter values in these ranges, we use the differential evolution technique, which outperforms the grid search technique as shown in [37]. As for the proposed semisupervised method, we note that the batch size parameter is usually set to 100 in the literature. However, here, we should take care not to set the batch size to be larger than the total number of training samples (number of samples per class \times the number of classes). For this experiment, we set it equal to the number of samples per class times the number of classes.

We also study the sensitivity of the method to reduced training samples (less than 10). Figure 8 shows the OA for sample sizes from 2 to 10 for both methods being compared. Figure 8(a) shows the results for the Jeddah dataset, while Figure 8(b) shows the result for the Riyadh dataset. The first observation is that the classification accuracy for both methods degrades gracefully when the number of training samples per class decreases. However, the results confirm again the superiority of the proposed method compared to Laplacian ELM.

4. Conclusions and Future Work

In this letter, we proposed a practical and efficient solution to the semisupervised classification of large-scale VHR RS

TABLE 3: Classification results for the Riyadh dataset using one hidden layer of size 256.

	OA	AA	Time (s)	Z score	Class 1	Class 2	Class 3	Class 4	Class 5
VGG-m									
NN	87.96	81.98	0.22	—	99.09	71.49	86.26	96.53	56.55
NN + RW	88.07	82.21	0.05	-0.49	99.15	71.37	87.27	96.56	56.67
Spatial graph in the hidden layer	81.52	71.48	24.00	6.25	81.90	74.57	78.79	94.33	27.82
Spectral graph in the hidden layer	93.12	88.98	30.36	-5.30	99.22	85.03	91.01	98.61	71.04
NN + spectral graph + RW	93.18	89.07	0.05	-5.38	99.22	85.09	91.52	98.64	70.92
GoogleNet									
NN	89.03	88.63	0.37	—	94.51	85.89	82.22	89.48	91.03
NN + RW	89.43	89.05	0.08	-1.06	94.51	85.43	83.64	90.09	91.61
Spatial graph in the hidden layer	80.97	75.54	23.97	5.76	96.21	77.77	70.61	85.50	47.59
Spectral graph in the hidden layer	93.45	91.03	55.17	-4.80	97.91	90.34	87.98	96.26	82.64
NN + spectral graph + RW	93.84	91.40	0.09	-5.20	97.91	90.40	88.18	96.83	83.68
ResNet									
NN	89.58	86.33	0.35	—	88.30	81.09	87.78	95.40	79.08
NN + RW	89.72	86.44	0.07	-0.50	89.22	81.26	86.97	95.46	79.31
Spatial graph in the hidden layer	73.64	65.76	39.26	8.61	99.68	30.91	90.20	87.19	20.81
Spectral graph in the hidden layer	95.13	93.56	45.94	-5.86	98.76	89.54	96.36	97.50	85.63
NN + spectral graph + RW	95.29	93.80	0.07	-6.00	99.09	89.43	96.47	97.61	86.44
Very-deep-16									
NN	90.97	88.09	0.63	—	95.36	75.60	94.85	96.81	77.82
NN + RW	93.13	91.84	0.11	-1.71	97.39	82.29	95.96	96.22	87.36
Spatial graph in the hidden layer	79.73	67.40	69.42	6.74	80.65	62.29	83.64	97.19	13.22
Spectral graph in the hidden layer	94.72	92.21	91.27	-4.72	98.00	88.85	95.12	98.29	80.79
NN + spectral graph + RW	95.26	92.91	0.11	-5.40	98.50	90.06	94.45	98.55	82.99

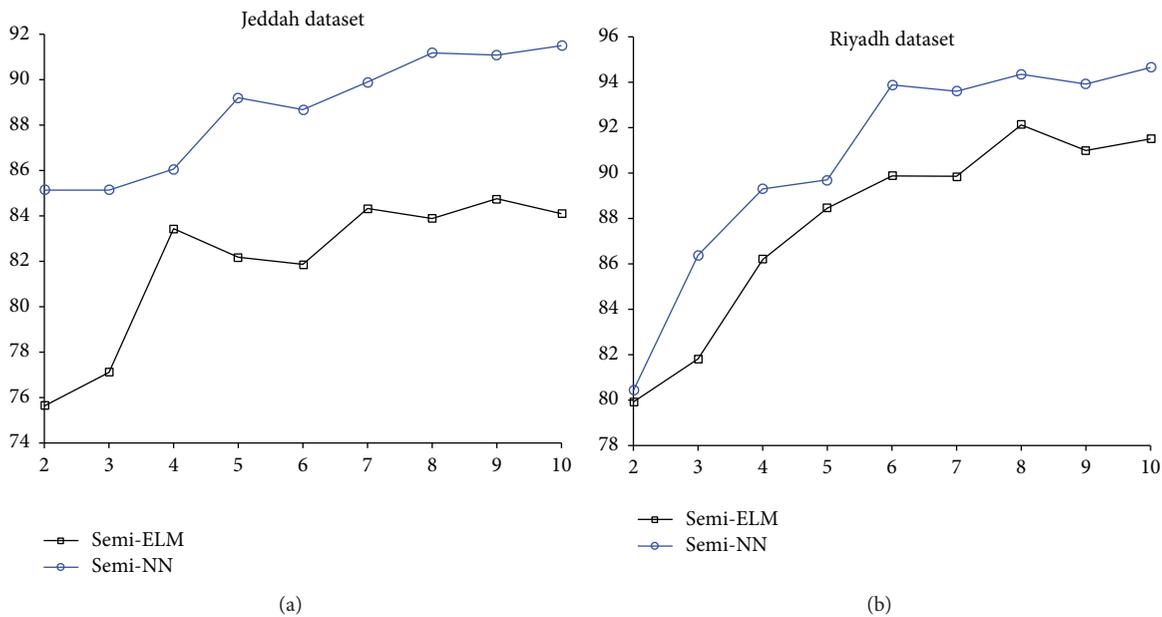


FIGURE 8: Comparison between the proposed semisupervised NN and the Laplacian ELM method. (a) Results for Jeddah dataset. (b) Results for Riyadh dataset.

images. The proposed method is based on subdividing the image into square tiles and then extracting feature descriptors on these tiles. These in turn are fed into another fully connected neural network with semisupervised learning, which exploits the structural information contained in the unlabelled tiles. Semisupervised learning is implemented in our proposed method in two ways, first by embedding a graph Laplacian in the hidden layer and second by spatial regularization of the output layer using a random walker algorithm.

Experimental results have shown that our proposed method can provide coarse classification maps for very large RS images with high accuracy. Furthermore, we have shown that our proposed semisupervised learning that combines a graph Laplacian in the hidden layer and spatial regularization of the output layer always provides significant improvement in terms of the overall accuracy, while having reasonable computational times for such large VHR images.

As a future research direction, we can employ a self-paced selection strategy of unlabelled titles similar to the techniques used in [38, 39] to further improve the performance. One can also study different fusion techniques of the feature descriptors of many pretrained CNNs. Another direction is to research ways to combine data from different datasets in order to improve the accuracy of the learning algorithm.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

The authors would like to extend their sincere appreciation to Deanship of Scientific Research at King Saud University for funding this research group (no. RG-1435-050).

References

- [1] P. Lassalle, J. Inglada, J. Michel, M. Grizonnet, and J. Malik, "A scalable tile-based framework for region-merging segmentation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 10, pp. 5473–5485, 2015.
- [2] T. Moranduzzo, F. Melgani, M. L. Mekhalfi, Y. Bazi, and N. Alajlan, "Multiclass coarse analysis for UAV imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6394–6406, 2015.
- [3] T. Moranduzzo, M. L. Mekhalfi, and F. Melgani, "LBP-based multiclass classification method for UAV imagery," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2362–2365, Milan, Italy, 2015.
- [4] X. Yao, J. Han, G. Cheng, X. Qian, and L. Guo, "Semantic annotation of high-resolution satellite images via weakly supervised learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3660–3671, 2016.
- [5] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '10*, pp. 270–279, New York, NY, USA, 2010, ACM Press.
- [6] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [7] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 886–893, San Diego, CA, USA, 2005.
- [8] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [9] K. Fukushima, "Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.
- [10] Y. LeCun, B. Boser, J. S. Denker et al., "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.
- [11] M. Hayat, M. Bennamoun, and S. An, "Learning non-linear reconstruction models for image set classification," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1915–1922, Columbus, OH, USA, 2014.
- [12] J. Bai, Y. Wu, J. Zhang, and F. Chen, "Subset based deep learning for RGB-D object recognition," *Neurocomputing*, vol. 165, pp. 280–292, 2015.
- [13] S. Gao, Y. Zhang, K. Jia, J. Lu, and Y. Zhang, "Single sample face recognition via learning deep supervised autoencoders," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2108–2118, 2015.
- [14] T. Brosch and R. Tam, "Efficient training of convolutional deep belief networks in the frequency domain for application to high-resolution 2D and 3D images," *Neural Computation*, vol. 27, no. 1, pp. 211–227, 2015.
- [15] G. Hinton, L. Deng, D. Yu et al., "Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.
- [16] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, pp. 865–873, 2015.
- [17] "Visual geometry group home page," May 2017, http://www.robots.ox.ac.uk/~vgg/research/very_deep/.
- [18] A. Anuse and V. Vyas, "A novel training algorithm for convolutional neural network," *Complex & Intelligent Systems*, vol. 2, no. 3, pp. 221–234, 2016.
- [19] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva, "Land use classification in remote sensing images by convolutional neural networks," *Computing Research Repository CoRR* 2015, <http://arxiv.org/abs/1508.00092>.
- [20] D. Marmanis, M. Datcu, T. Esch, and U. Stilla, "Deep learning earth observation classification using ImageNet pretrained networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 105–109, 2016.
- [21] X. Chen, S. Xiang, C.-L. Liu, and C.-H. Pan, "Vehicle detection in satellite images by hybrid deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 10, pp. 1797–1801, 2014.
- [22] F. Zhang, B. Du, and L. Zhang, "Scene classification via a gradient boosting random convolutional network framework,"

- IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1793–1802, 2016.
- [23] G. Cheng, Z. Li, X. Yao, L. Guo, and Z. Wei, “Remote sensing image scene classification using bag of convolutional features,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1735–1739, 2017.
- [24] G. Cheng, J. Han, and X. Lu, “Remote sensing image scene classification: benchmark and state of the art,” *Proceedings of the IEEE*, vol. 105, no. 10, pp. 1865–1883, 2017.
- [25] Y. Bazi, N. Alajlan, and F. Melgani, “Improved estimation of water chlorophyll concentration with semisupervised Gaussian process regression,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2733–2743, 2012.
- [26] X. Yin, W. Yang, G. S. Xia, and L. Dong, “Semi-supervised feature learning for remote sensing image classification,” in *2014 IEEE Geoscience and Remote Sensing Symposium*, pp. 1261–1264, Quebec City, QC, Canada, 2014.
- [27] Y. Bazi and F. Melgani, “Semisupervised PSO-SVM regression for biophysical parameter estimation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 6, pp. 1887–1895, 2007.
- [28] L. Bruzzone and C. Persello, “Recent trends in classification of remote sensing data: active and semisupervised machine learning paradigms,” in *2010 IEEE International Geoscience and Remote Sensing Symposium*, pp. 3720–3723, Honolulu, HI, USA, 2010.
- [29] Z. Wang, N. M. Nasrabadi, and T. S. Huang, “Semisupervised hyperspectral classification using task-driven dictionary learning with Laplacian regularization,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1161–1173, 2015.
- [30] O. Chapelle, B. Schölkopf, and A. Zien, *Semi-Supervised Learning*, The MIT Press, 1st edition, 2010.
- [31] J. Weston, F. Ratle, and R. Collobert, “Deep learning via semi-supervised embedding,” in *Proceedings of the 25th International Conference on Machine Learning - ICML '08*, pp. 1168–1175, New York, NY, USA, 2008, ACM.
- [32] L. Grady, “Random walks for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1768–1783, 2006.
- [33] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, “Return of the devil in the details: delving deep into convolutional nets,” Computing Research Repository CoRR 2014, <http://arxiv.org/abs/1405.3531>.
- [34] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” Computing Research Repository CoRR 2014, <http://arxiv.org/abs/1409.1556>.
- [35] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” Computing Research Repository CoRR 2014, <http://arxiv.org/abs/1409.4842>.
- [36] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Computing Research Repository CoRR 2015, <http://arxiv.org/abs/1512.03385>.
- [37] Y. Bazi, N. Alajlan, F. Melgani, H. AlHichri, S. Malek, and R. R. Yager, “Differential evolution extreme learning machine for the classification of hyperspectral images,” *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 6, pp. 1066–1070, 2014.
- [38] X. Yao, J. Han, D. Zhang, and F. Nie, “Revisiting co-saliency detection: a novel approach based on two-stage multi-view spectral rotation co-clustering,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3196–3209, 2017.
- [39] D. Zhang, D. Meng, and J. Han, “Co-saliency detection via a self-paced multiple-instance learning framework,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 5, pp. 865–878, 2017.