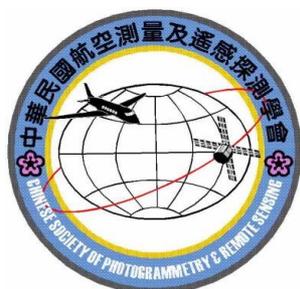


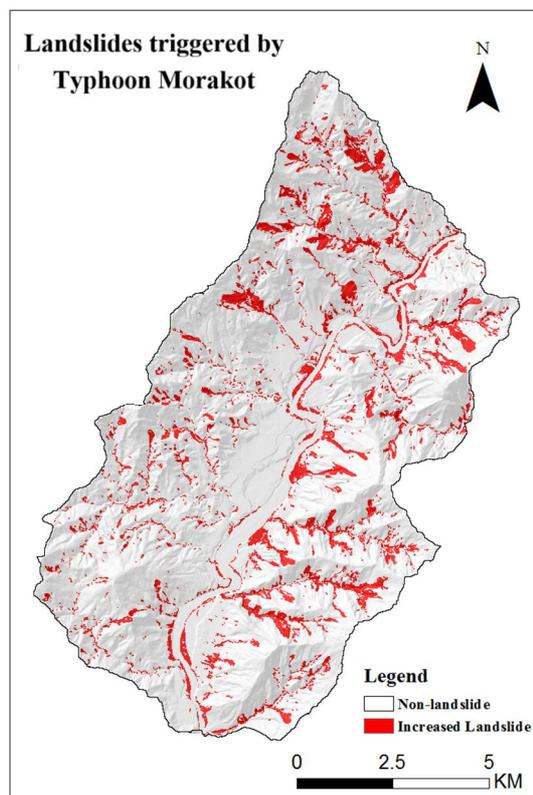
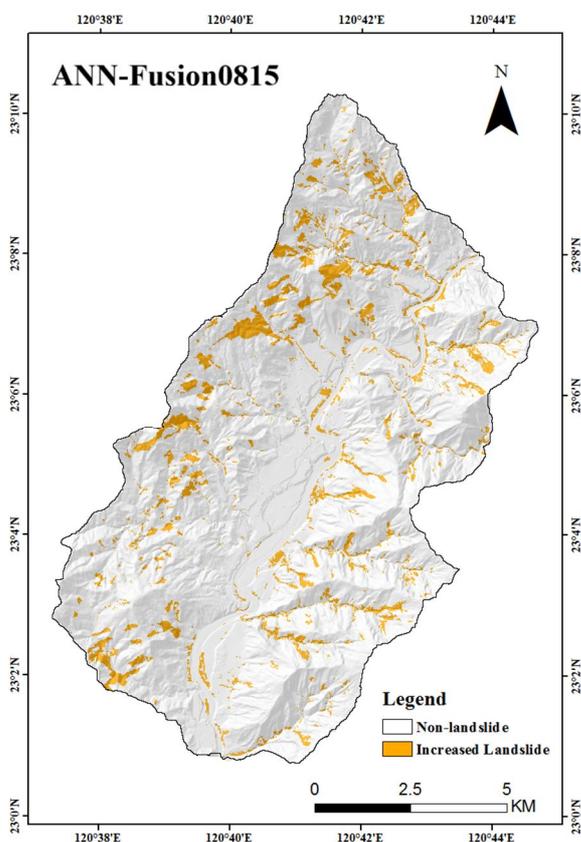
第三十一卷 第一期  
中華民國一一五年三月

ISSN 1021-8661  
DOI: 10.6574/JPRS



# 航測及遙測學刊

中華民國航空測量及遙感探測學會出版



Journal of Photogrammetry and Remote Sensing

Volume 31 No.1

March

2026

Published by Chinese Society of Photogrammetry and Remote Sensing

# 航測及遙測學刊

## Journal of Photogrammetry and Remote Sensing

發行人：蔡富安

出版者：中華民國航空測量及遙感探測學會

地址：台北市文山區羅斯福路五段 113 號三樓

電子信件：csprsmail@csprs.org.tw

網址：http://www.csprs.org.tw

總編輯：楊明德

國立中興大學土木工程學系

電話：886-4-2285-2168

電子信件：jprssubmit@gmail.com

編輯委員(依中文姓氏筆劃排列)

王素芬 (國立彰化師範大學)

王聖鐸 (國立臺灣師範大學)

史天元 (國立陽明交通大學)

何宗儒 (國立臺灣海洋大學)

林玉儂 (中央研究院)

林昭宏 (國立成功大學)

林唐煌 (國立中央大學)

邱式鴻 (國立政治大學)

洪榮宏 (國立成功大學)

徐百輝 (國立臺灣大學)

張智安 (國立陽明交通大學)

莊永忠 (逢甲大學)

郭俊麟 (國立東華大學)

陳偉堯 (國立臺北科技大學)

陳國華 (國立臺北大學)

陳朝圳 (國立屏東科技大學)

曾國欣 (國立中央大學)

黃倬英 (國立臺灣大學)

蔡富安 (國立中央大學)

蔡榮得 (國立中興大學)

蘇東青 (國立金門大學)

C. Stewart (The Ohio State University) (USA)

M. Matsuoka (Tokyo Institute of Technology) (Japan)

T. Kim (Inha University) (South Korea)

PUBLISHER: Fuan Tsai

PUBLISHED BY: Chinese Society of Photogrammetry and Remote Sensing

Address: 3F, No.113, Sec.5, Roosevelt Road, Taipei, Taiwan

E-mail: csprsmail@csprs.org.tw

WebSite: http://www.csprs.org.tw

EDITOR-IN-CHIEF: Ming-Der Yang

Department of Civil Engineering, National Chung Hsing University

Tel: 886-4-2285-2168

E-Mail: jprssubmit@gmail.com

EDITORIAL BOARD

S. F. Wang (National Changhua University of Education)

S. D. Wang (National Taiwan Normal University)

P. T. Y. Shih (National Yang Ming Chiao Tung University)

C. R. Ho (National Taiwan Ocean University)

Y. N. Lin (Academia Sinica)

C. H. Lin (National Cheng Kung University)

T. H. Lin (National Central University)

S. H. Chio (National Chengchi University)

J. H. Hong (National Cheng Kung University)

P. H. Hsu (National Taiwan University)

T. A. Teo (National Yang Ming Chiao Tung University)

Y. C. Chuang (Feng Chia University)

C. L. Kuo (National Dong Hwa University)

W. Chen (National Taipei University of Technology)

K. H. Chen (National Taipei University)

C. T. Chen (National Pingtung University of Science and Technology)

K.H. Tseng (National Central University)

C.Y. Huang (National Taiwan University)

F. Tsai (National Central University)

J. D. Tsai (National Chung Hsing University)

T. C. Su (National Quemoy University)

C. Stewart (The Ohio State University) (USA)

M. Matsuoka (Tokyo Institute of Technology) (Japan)

T. Kim (Inha University) (South Korea)

### 封面照片說明 About the Cover

Spatial comparison between the reference landslide inventory (red) and the ANN-Fusion0815 classification output (yellow). The fused result delineates most major slope failures and exhibits improved spatial continuity relative to single-source interpretations. Enhanced detection is evident in cloud-affected regions, particularly in the western sector, where radar-derived information compensates for missing optical signals. While minor discrepancies remain along narrow or fragmented features, the fusion output closely approximates the overall distribution pattern of mapped landslides, demonstrating its capability to reconstruct event-induced disturbances with high spatial consistency.

# Integrated Optical–SAR Texture Analysis for Neural Network–Based Emergency Landslide Surveying and Mapping

Yi-Keng Chen<sup>1\*</sup> Shou-Hao Chiang<sup>2</sup>

## Abstract

Optical remote sensing imagery provides critical spectral information for mapping landslide inventories but is frequently hindered by cloud cover and adverse weather. In contrast, Synthetic Aperture Radar (SAR) penetrates clouds and is highly sensitive to surface backscatter changes, offering complementary insights for detecting landslide disturbances. This study presents an integrated approach combining optical and SAR data to enable rapid and reliable landslide detection under challenging conditions. The proposed framework applies object-based image analysis (OBIA) to segment terrain into meaningful units and derives NDVI<sub>diff</sub> and NDSI indices from pre- and post-event imagery, together with six GLCM-based texture features, to characterize surface disturbances. Four classification scenarios, including optical-only, SAR-only, cloud-obstructed optical, and fusion-based models, were systematically compared. Results demonstrate that the fusion approach consistently yields more spatially coherent and complete landslide inventories, while SAR-based mapping alone successfully delineates most large-scale landslides under heavy cloud cover. These findings confirm the operational effectiveness of the proposed optical–SAR framework for rapid, event-driven landslide mapping and disaster risk assessment.

**Keywords:** Landslide Detection, Synthetic Aperture Radar, Image Texture, Image Fusion, Artificial Neural Network

## 1. Introduction

Satellite-based emergency mapping (SEM) has become a cornerstone of disaster risk reduction, enabling governments and researchers to rapidly assess hazard impacts and plan relief operations. Over recent decades, SEM has evolved into an integral component of broader resilience strategies aimed at improving the timeliness and effectiveness of emergency response and minimizing human and economic losses during and after catastrophic events.

Taiwan, located in the subtropical monsoon region, is highly susceptible to typhoon-induced landslides and debris-flow disasters. From July to September, intense rainfall frequently triggers large-scale slope landslides in its mountainous areas, dramatically transforming land cover from vegetated slopes to bare soil. Optical satellite imagery has long been used to detect such changes because of its rich spectral content and the strong radiometric contrast between vegetation and landslides under clear daylight conditions. High-resolution satellite-based sensors such as LANDSAT, SPOT, and FORMOSAT, often combined with digital elevation models (DEMs), have further enabled three-dimensional visualization and both manual and automated mapping of landslides. However, persistent

cloud cover, typhoons, and thunderstorms during Taiwan's summer can delay the availability of optical data precisely when emergency landslide mapping is most needed.

Synthetic Aperture Radar (SAR), operating in the microwave domain, mitigates this limitation by penetrating cloud cover and acquiring data both day and night. SAR thus provides complementary information on surface roughness and backscatter changes that is valuable for landslide and debris-flow detection under adverse weather. Yet, SAR imagery also poses interpretive challenges due to speckle noise, geometric distortions, and sensitivity to polarization and observation geometry. In emergency landslide mapping, optical imagery is often hindered by cloud cover and adverse weather conditions. However, cloud contamination is typically spatially heterogeneous and temporally variable, and partially cloud-free or multi-temporal optical observations can still provide useful complementary information. By integrating such heterogeneous inputs with weather-independent SAR data, fusion supports rapid initial assessment and subsequent refined delineation. When optical data are severely limited, SAR-derived features dominate the fusion process, whereas under favorable conditions, optical imagery enhances spatial accuracy.

<sup>1</sup> Ph.D. Student, Department of Civil Engineering, National Taiwan University

<sup>2</sup> Associate Professor, Center for Space and Remote Sensing Research, National Central University

\* Corresponding Author, E-mail: d10521008@ntu.edu.tw

Received Date: Oct. 15, 2025

Revised Date: Jan. 07, 2026

Accepted Date: Feb. 04, 2026

These contrasting advantages and limitations of optical and SAR imagery motivate the development of fusion approaches specifically designed for time-critical, event-based landslide mapping rather than routine monitoring. Image fusion integrates multi-source remote sensing data to improve measurement accuracy, enhance information content, and increase interpretive reliability. In terms of data quality, optical imagery generally provides higher radiometric contrast and spectral sensitivity for landslide detection under clear-sky conditions, enabling accurate delineation of disturbed surfaces. However, its performance is frequently degraded by cloud cover, terrain shadow, and illumination variability during typhoon and monsoon seasons. In contrast, SAR imagery offers stable, weather-independent observations and ensures reliable data availability under adverse conditions, although it is affected by speckle noise and geometric distortions in mountainous terrain. These complementary characteristics justify the combined use of optical and SAR data for robust emergency landslide mapping. This study advances emergency landslide mapping by introducing an Artificial Neural Network (ANN)-based image fusion framework implemented within an object-based image analysis (OBIA) environment. ANN approaches are particularly suited to this task because they can model complex, non-linear relationships and spatial heterogeneities more effectively than traditional classifiers. By leveraging ANN's capacity to learn from diverse input features, the proposed framework fuses the complementary strengths of optical and SAR datasets at the decision level, thereby avoiding the co-registration difficulties common in rugged terrain.

The objectives of this study are threefold: (1) to fuse optical and SAR remote sensing information to improve the accuracy and reliability of event-based emergency landslide and debris-flow detection; (2) to assess the applicability of incorporating SAR data into classification models when optical images are limited during crises; and (3) to provide a replicable workflow for rapid, event-driven landslide mapping that can be adapted to other disaster types and geographic settings. By focusing on the operational requirements of emergency mapping, this research contributes a novel, robust, and scalable methodology to advance satellite-based disaster response.

## 2. Literature Reviews

Landslide inventory mapping is a cornerstone of hazard assessment in Taiwan, where slopes fail under heavy rainfall, earthquakes, and anthropogenic pressures such as deforestation and intensive land use (Selby, 1993; Dai & Lee, 2002). Optical satellites, often combined with DEM, have long enabled visual delineation of landslides by exploiting spectral

contrasts between vegetation and bare soil (Haeberlin *et al.*, 2004; Nichol *et al.*, 2006). The normalized difference vegetation index, NDVI and its temporal differential NDVI<sub>diff</sub> (Chen, 2019) have proven effective in detecting fresh landslides by highlighting vegetation loss (Weier & Herring, 2000). Yet, optical systems falter under cloud cover, shadows, or low illumination—conditions common during typhoon events.

SAR offers compelling advantages for disaster mapping: all-weather, day-night imaging with sensitivity to surface roughness and backscatter changes. Conventional differential interferometry synthetic aperture radar, Differential Interferometry Synthetic Aperture Radar (DInSAR) has been effective for creeping landslides but is less reliable for abrupt, collapse-type landslides induced by heavy rainfall (Colesanti & Wasowski, 2006). Backscatter-based approaches, including Polarimetric Synthetic Aperture Radar (PolSAR) and the Sigma-naught coefficient, yield richer scattering information and, when differenced as NDSI, can delineate sharp land surface changes (Furuta & Tomiyama, 2008; Guzzetti *et al.*, 2012). However, SAR data are affected by speckle noise, geometric distortions, and sensitivity to acquisition parameters. Recent studies have further investigated the applicability of multi-sensor and multi-geometry SAR observations in complex terrain, demonstrating their potential for landslide detection under diverse acquisition conditions (Wang *et al.*, 2025).

Image fusion offers a path forward by integrating complementary strengths of optical and SAR data to enhance classification robustness (Pohl & van Genderen, 1998; Amarsaikhan & Douglas, 2004). However, pixel- and feature-level fusion demand high co-registration accuracy—untenable in complex mountainous terrains. Decision-level fusion, which integrates classification results rather than raw pixels, is more tolerant of misalignment but remains underutilized for landslide mapping. Recent advances echo these challenges. Jin *et al.* (2022) fused multisource data (DEM, spectral and backscatter features) to improve landslide classification accuracy. Zhang & Wang (2024) reviewed deep learning-based landslide detection, emphasizing difficulties in generalization and domain shift between sensors. Li & Xiao (2025) surveyed fusion of optical, SAR, and LiDAR for landslide mapping, highlighting the potential of hybrid networks but noting challenges in aligning modalities. Prakash *et al.* (2020) systematically compared deep learning and traditional machine learning approaches for landslide mapping using Earth observation data, demonstrating the advantages of deep architectures in complex environments. Chen *et al.* (2023) proposed a multi-input channel U-Net framework integrating

multisource SAR data, achieving improved detection performance through feature-level fusion. Liu *et al.* (2024) developed an attention-based Swin Transformer U-Net for landslide segmentation, highlighting the effectiveness of transformer architectures in capturing multi-scale contextual information. Tuo *et al.* (2025) further explored multimodal data fusion using an improved DeepLabv3+ model, demonstrating enhanced segmentation accuracy through integrated optical and SAR features. In addition, Fu *et al.* (2024) combined Sentinel-1 SAR and Sentinel-2 optical imagery for landslide dating, illustrating the value of multi-temporal and multi-sensor integration in post-event analysis. Coluzzi *et al.* (2025) proposed a robust change detection technique based on freely available optical imagery for rapid landslide identification, while Jiang *et al.* (2025) provided a comprehensive review of deep learning-based landslide detection, summarizing recent progress, challenges, and future opportunities.

These works indicate that while multisensor fusion and deep learning methods are advancing, most existing approaches rely on complex network structures, extensive training data, and precise preprocessing, which limit their applicability in time-critical emergency scenarios. Consequently, operational, cloud-robust frameworks that combine optical and SAR for emergency landslide mapping remain scarce. Building on this body of work, the present study introduces an ANN-driven, decision-level fusion within an OBIA framework, integrating NDVI<sub>diff</sub> and NDSI with texture features to (i) bypass strict co-registration demands, (ii) exploit complementary spectral and backscatter cues under cloudy conditions, and (iii) support rapid, reliable event-based landslide detection in steep, cloud-prone landscapes.

### 3. Methodology

In response to the need for rapid, event-driven landslide assessment, this study implements an ANN-based, decision-level image fusion within an OBIA framework to support emergency landslide mapping in the Laonong River watershed, southern Taiwan, which experienced extensive slope landslides and debris flows during Typhoon Morakot in 2009. The workflow comprises the following steps: (1) image segmentation to generate object units and texture features; (2) selection of representative training samples; and (3) supervised classification using a back-propagation multilayer perceptron. This integrated procedure operationalizes the proposed framework for rapid and reliable landslide detection during emergency conditions.

#### 3.1 Research Materials

##### 3.1.1 Pixel-based Evaluation

FORMOSAT-2, Taiwan’s first Earth observation satellite, provided high-resolution (2 m PAN, 8 m MS) multispectral imagery suitable for landslide monitoring. For this study, three geometrically corrected images were acquired to represent Typhoon Morakot conditions: a pre-event scene (9 May 2009), a post-event scene with minimal cloud (12 Sept 2009), and a cloud-obstructed scene captured soon after the typhoon (15 Aug 2009) to simulate emergency mapping scenarios (Figure 1). Because the 15 Aug 2009 image exhibited extensive cloud cover, the pre-processing of the NDVI<sub>diff</sub> image modified the cloud-covered areas by masking and assigning them a value of “0” to prevent misclassification during analysis. Leveraging these datasets, temporal changes in vegetation cover were quantified using the NDVI<sub>diff</sub> index, which highlights vegetation loss and newly exposed soil attributable to the typhoon. By differencing pre- and post-event NDVI values, NDVI<sub>diff</sub> provides a robust measure of surface change despite seasonal or illumination variations, thereby supporting event-driven landslide detection. The formulas of NDVI and NDVI<sub>diff</sub> are as follows:

$$NDVI = \frac{NIR-Red}{NIR+Red} \dots\dots\dots (1)$$

$$NDVI_{diff} = NDVI_{post} - NDVI_{pre} \dots\dots\dots (2)$$

Where NIR and Red denote the spectral reflectance in the near-infrared and red bands, respectively. NDVI<sub>post</sub> and NDVI<sub>pre</sub> represent post- and pre-event NDVI values. NDVI<sub>diff</sub> (Figure 2) highlights temporal changes in vegetation cover, enabling the detection of newly exposed soil or disturbed areas such as landslides.

##### 3.1.2 SAR Image

ALOS/PALSAR L-band SAR data from 8 Jul. 2009 (pre-event) and 23 Aug. 2009 (post-event) (Figure 3) were selected because their longer wavelength penetrates vegetation more effectively than C-band sensors, providing clearer information on ground conditions in heavily forested terrain. Radar backscatter intensity reflects surface roughness, moisture, and slope orientation, producing higher returns from areas directly facing the sensor and lower returns from opposing slopes. To extract meaningful ground signals, raw SAR images underwent radiometric calibration, geometric correction with a DEM, and speckle filtering to reduce noise. The normalized difference sigma-naught index (NDSI) was then computed from pre- and post-event backscatter coefficients to quantify changes in scattering properties, as expressed by:

$$NDSI = \frac{\sigma_m^0 - \sigma_s^0}{\sigma_m^0 + \sigma_s^0} \dots\dots\dots (3)$$

where  $\sigma_m^0$  denotes the master sigma-naught (Sigma0) image representing the post-typhoon event and  $\sigma_s^0$  denotes the slave sigma-naught image representing the pre-typhoon event. Extreme positive or negative NDSI values (Figure 4) highlight abrupt surface disturbances—such as fresh landslides—whereas stable terrain remains near zero, making NDSI a robust indicator of event-driven ground change and a valuable complement for rapid landslide detection. Because the primary focus of this study is slope landslides rather than sediment deposition within river channels, all NDSI values in river areas were assigned zero.

The original ALOS/PALSAR Level 1.1 imagery used in this study was acquired in Fine mode with a spatial resolution of approximately 14 m. To ensure spatial consistency with the digital elevation model (DEM) and optical datasets, geometric correction and terrain normalization were performed using a 10 m DEM. During this process, the original SAR images were resampled from 14 m to 10 m using bilinear interpolation. Consequently, all subsequent analyses were conducted on harmonized 10 m resolution data, thereby minimizing scale discrepancies among multi-source datasets and facilitating consistent object-based segmentation and feature extraction within the proposed fusion framework.

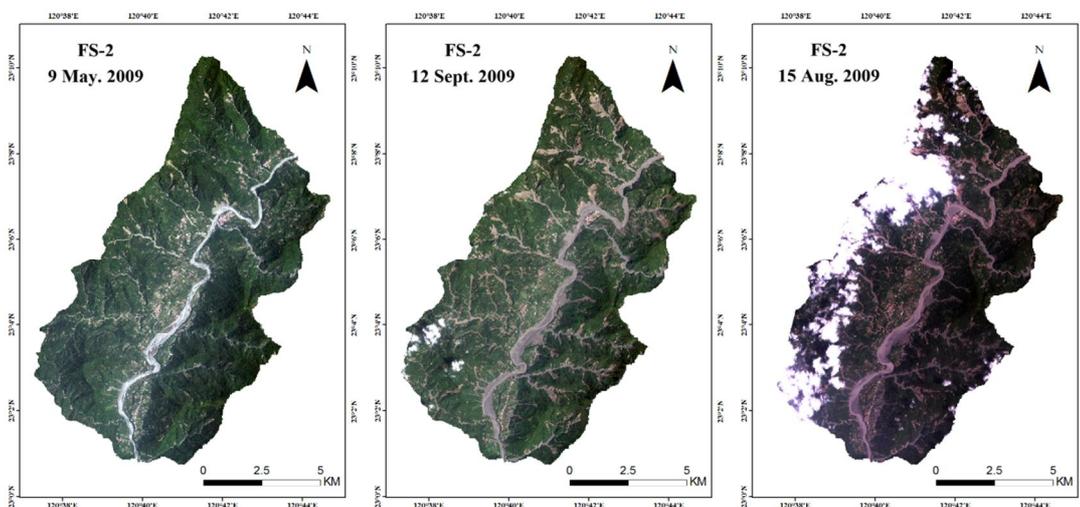


Figure 1 FORMOSAT-2 imagery. Left: 9 May 2009 (pre-Typhoon Morakot); Center: 12 Sept. 2009 (post-event); Right: 15 Aug. 2009 (post-event). The cloud cover and missing southwestern area in the right image provide conditions suitable for simulating emergency mapping

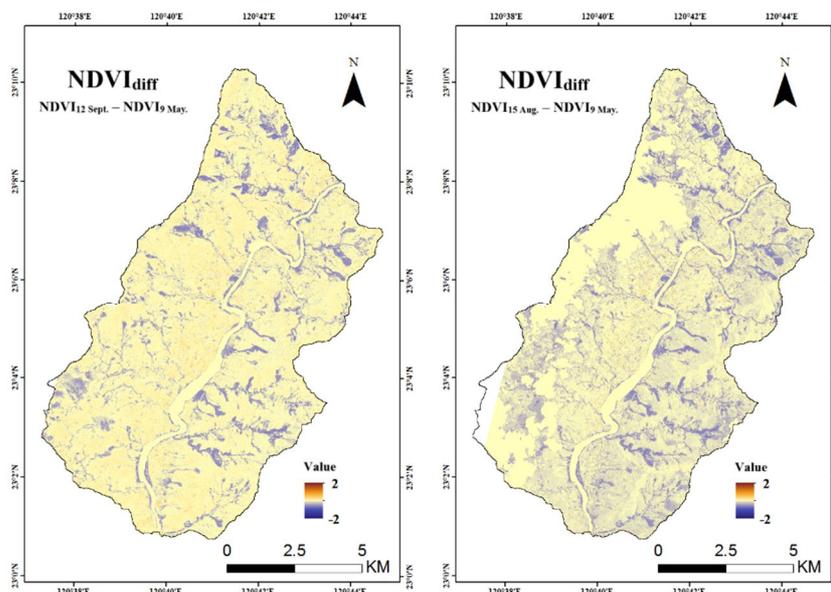


Figure 2 NDVI<sub>diff</sub> imagery. Left: NDVI<sub>diff</sub>-0912 image, generated by “NDVI<sub>12 Sept.</sub> – NDVI<sub>9 May.</sub>”; Right: NDVI<sub>diff</sub>-0815 image, generated as “NDVI<sub>15 Aug.</sub>– NDVI<sub>9 May.</sub>,” with cloud-covered areas masked and assigned a value of “0” to prevent misclassification

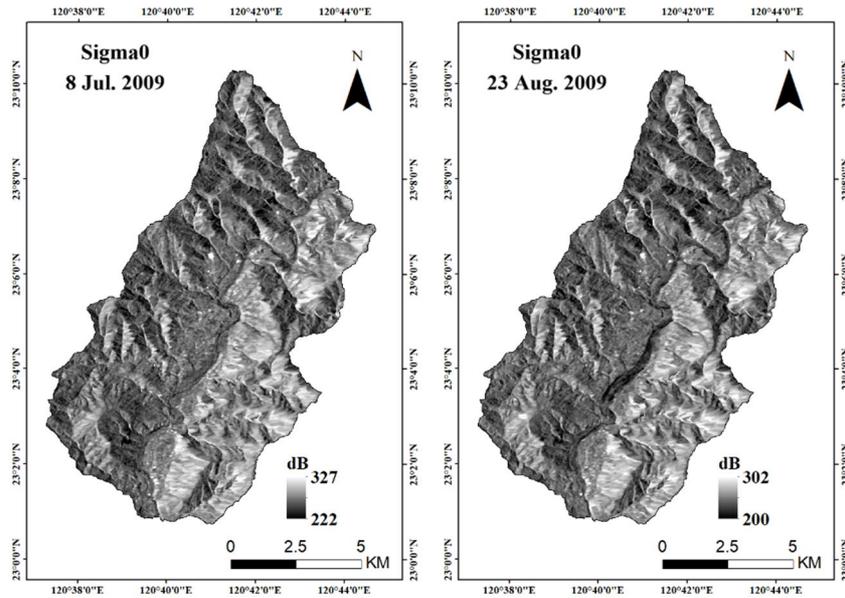


Figure 3 ALOS/PALSAR Sigma0 imagery. Left:  $\sigma_{0708}^0$  image; Right:  $\sigma_{0823}^0$  image

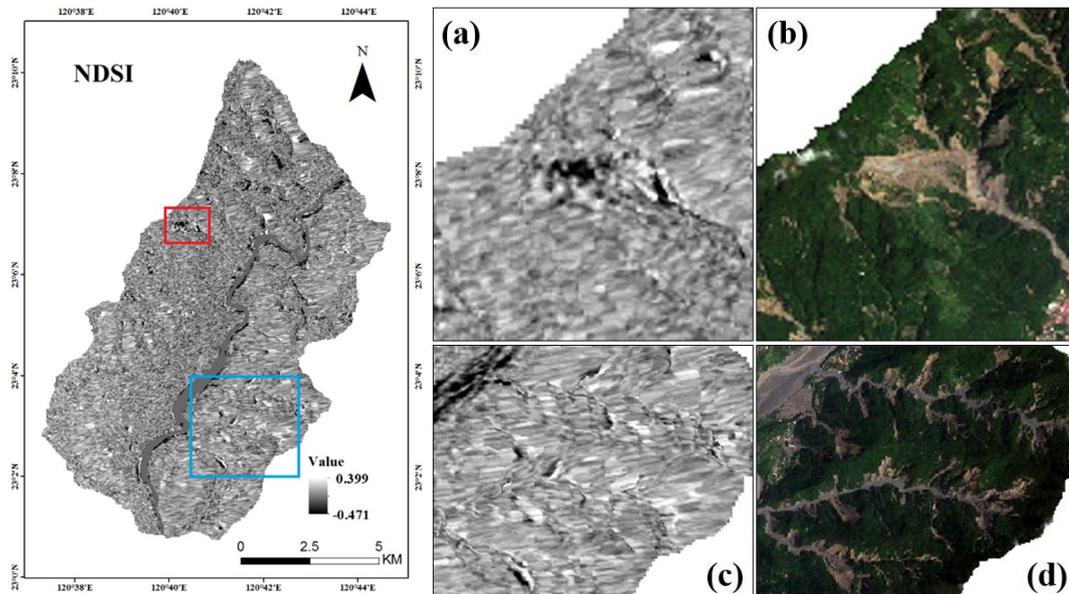


Figure 4 NDSI image. The red and blue boxes indicate areas enlarged for detailed comparison: (a) NDSI texture of landslides within the red box; (b) corresponding reference photograph of the red-box landslides; (c) NDSI texture of landslides within the blue box; and (d) corresponding reference photograph of the blue-box landslides

### 3.1.3 Landslide Inventory

The original 1:5000-scale landslide inventory map compiled from post-Morakot stereoscopic aerial photographs (SWCB, 2011) included many pre-existing landslides and thus did not accurately represent event-induced landslides. In this study, the landslide inventory (Figure 5) was revised by extracting areas of significant  $NDVI_{diff}$  decrease and cross-referencing them with the original map, FORMOSAT-2 imagery, and high-resolution Google Earth images to manually delineate and retain only landslides triggered by Typhoon Morakot.

### 3.2 Image Segmentation

Multiresolution segmentation is applied as the initial step in object-based landslide delineation. This bottom-up region-growing algorithm aggregates adjacent pixels based on scale, shape, and compactness, allowing adaptive object sizes suited to the complex geometry of mass movements (Ez-zahouani *et al.*, 2023; Tetteh *et al.*, 2023). For large scale landslides, each landslide is segmented into roughly three to five clusters, capturing internal variability without excessive fragmentation (Table 1). Object-level textures are then derived from the Gray Level Co-

occurrence Matrix (GLCM) (Haralick *et al.*, 1973), producing descriptors such as contrast, dissimilarity, entropy, and homogeneity that, combined with spectral indices, enhance representation of surface roughness and structural complexity. This integration provides robust inputs for ANN-based optical-SAR fusion and rapid emergency landslide mapping.

To enhance the discriminatory power of the segmented objects, six GLCM-derived texture features were selected to capture tonal variation, surface roughness, and internal complexity—key traits distinguishing landslide deposits from stable terrain (Table 2). These descriptors complement spectral indices by quantifying abrupt and gradual spatial changes within each segment, enhancing classification reliability. The description and formulas of texture features are as follows:

- (1) **Contrast:** highlights abrupt tonal changes, accentuating fresh landslide and rough debris surfaces.
- (2) **Dissimilarity:** linearly measures gray-level variation, capturing gradual differences within landslide bodies.
- (3) **Entropy:** reflects randomness and complexity, indicating heterogeneous disturbed slopes versus stable terrain.
- (4) **Homogeneity:** measures smoothness, distinguishing intact areas from disrupted ground.
- (5) **Mean:** represents average brightness, revealing differences between exposed soil and vegetated patches.
- (6) **Standard Deviation:** quantifies value dispersion, identifying mixed materials typical of mass-movement deposits.

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}} \dots\dots\dots (4)$$

$$\text{Contrast} = \sum_{i,j=0}^{N-1} P_{i,j}(i - j)^2 \dots\dots\dots (5)$$

$$\text{Dissimilarity} = \sum_{i,j=0}^{N-1} P_{i,j}|i - j| \dots\dots\dots (6)$$

$$\text{Entropy} = \sum_{i,j=0}^{N-1} P_{i,j}(-\ln P_{i,j}) \dots\dots\dots (7)$$

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \dots\dots\dots (8)$$

$$\text{Mean } \mu = \frac{1}{N} \sum_{i,j=1}^N V_{i,j} \dots\dots\dots (9)$$

$$\text{Standard Deviation} = \sqrt{\frac{1}{N} \sum_{i,j=1}^N (V_{i,j} - \mu)^2} \dots\dots\dots (10)$$

Where, **P** denotes the gray level co-occurrence matrix; *i* and *j* represent the row and column indices, respectively; *V<sub>ij</sub>* is the pixel value at cell *i, j* within the image segment; *P<sub>ij</sub>* is the normalized value at cell *i, j*; and *N* is the total number of rows or columns.

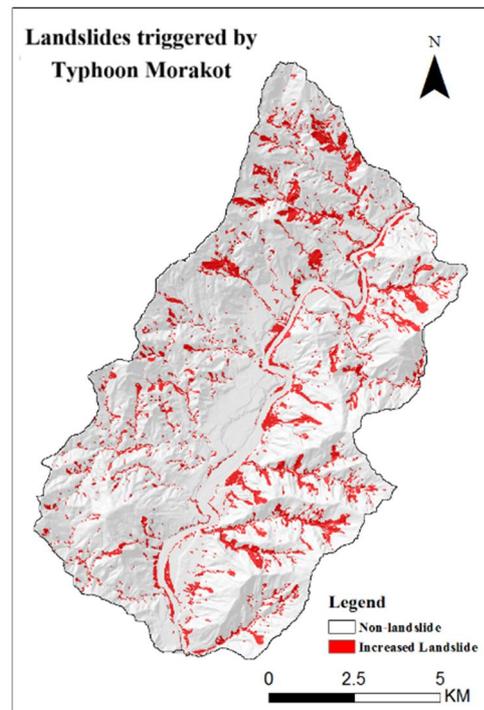


Figure 5 Landslide inventory map

Table 1 OBIA Segmentation Parameters

Parameter	Setting	Description
Segmentation Method	Multiresolution	Bottom-up region growing
Scale Parameter	30 (object size control)	Controls merging threshold
Shape	0.1	Controls spectral vs. shape weighting
Compactness	0.5	Controls object compactness
Typical Object Size	3–5 clusters per landslide	Captures internal variability

Table 2 GLCM Texture Extraction Parameters

Parameter	Setting	Description
Quantization Level	5 bits (32 levels)	Original images compressed to reduce sparse matrices
Window Size	3 × 3 pixels	Preserves local texture characteristics
Pixel Distance	1 pixel	Based on 3 × 3 window
Directions	0°, 45°, 90°, 135°	All four orientations averaged

### 3.3 Training Sample Selection

Sampling was designed to ensure that training data capture the characteristic signatures of event-induced landslides and stable terrain, with the Normalized Difference Sigma-naught Index (NDSI) serving as the primary criterion. NDSI quantifies the relative difference in backscatter between pre- and post-event SAR images and is particularly sensitive to abrupt changes in surface scattering associated with slope landslides. Landslide samples were therefore selected from areas exhibiting pronounced positive or negative NDSI values, indicating strong deviations in radar return consistent with ground disturbance. Conversely, non-landslide samples were extracted from zones where NDSI values remained close to zero, signifying stable surfaces with minimal backscatter change (Table 3).

NDVI<sub>diff</sub> was used as a secondary criterion to corroborate vegetation loss and surface exposure. Negative NDVI<sub>diff</sub> values, reflecting a transition from vegetated cover to bare soil, were required to coincide with the NDSI anomalies to classify landslide samples, while non-landslide samples were confirmed in areas showing negligible NDVI<sub>diff</sub> variation. This dual-index approach reduced the influence of noise, atmospheric effects, and geometric distortions inherent to SAR data, thereby improving the robustness of sample selection.

Using NDSI as the primary discriminant and NDVI<sub>diff</sub> as a complementary measure produces a balanced set of landslide and non-landslide samples that reflect both radar backscatter and spectral vegetation change. This sampling strategy ensures that the training data emphasize the most diagnostic characteristics of event-driven slope landslides and provide a strong foundation for subsequent classification and image-fusion analyses aimed at rapid emergency mapping.

### 3.4 Supervised Classification

An artificial neural network (ANN) framework (Figure 6) was employed to classify landslide and non-landslide objects derived from the OBIA and decision-level image fusion stages. Image segmentation combined with GLCM-derived textures generated object-level attributes from both optical (NDVI<sub>diff</sub>) and SAR (NDSI) datasets, thereby capturing spectral changes, backscatter variation, and textural complexity within each landslide cluster. These heterogeneous features exhibit non-linear interactions that cannot be effectively modeled using traditional parametric classifiers, necessitating a more flexible approach.

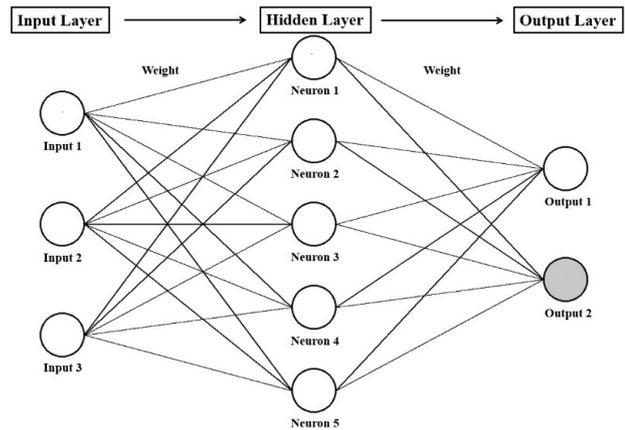


Figure 6 Schematic diagram of the Artificial Neural Network architecture

A back-propagation neural network (BPNN) implemented as a multilayer perceptron was selected due to its capacity to learn complex, non-linear decision boundaries from limited training samples (Table 4). During the forward pass, OBIA-derived features were propagated through the input and hidden layers to produce classification outputs. The backward pass computed the gradient of the loss function with respect to network weights, which were iteratively updated to minimize classification error between predicted and reference labels. This supervised learning process continued until convergence criteria were met, allowing the network to establish optimal internal representations for distinguishing landslide from non-landslide objects (Figure 7). By applying the BPNN to fused NDVI<sub>diff</sub> and NDSI features, the complementary strengths of optical and SAR data were integrated into a single probabilistic classification. This approach provided a robust and scalable means of producing rapid, event-driven landslide maps under conditions where optical imagery alone is insufficient due to cloud cover or adverse weather.

In the proposed framework, OBIA serves as an intermediate feature extraction and spatial aggregation stage that bridges raw image measurements and ANN-based classification. Specifically, multiresolution segmentation first partitions the fused NDVI<sub>diff</sub> and NDSI imagery into homogeneous objects representing candidate landslide units. For each object, statistical, spectral, and GLCM-based texture features are computed and summarized to form an object-level feature vector. These feature vectors, rather than pixel-level values, are subsequently used as input to the BPNN classifier. In this manner, OBIA transforms spatially correlated pixel information into meaningful analysis units, while the ANN learns nonlinear decision boundaries from the resulting object attributes. This object-to-network linkage enables efficient integration of spatial context, textural information, and multisource indices within a unified classification framework.

Table 3 Signatures of increased landslide and non-landslide areas in NDVI<sub>diff</sub> and NDSI images

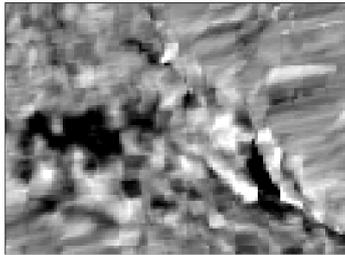
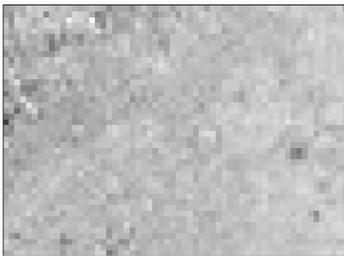
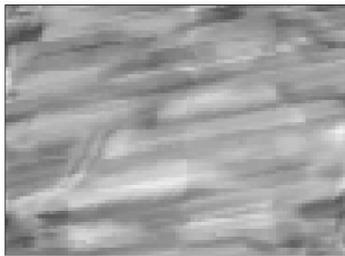
Categories	NDVI <sub>diff</sub>	NDSI
<b>Increased Landslides</b>		
	Increased landslides are characterized by dark zones (negative values) in the NDVI <sub>diff</sub> image and pronounced signal variations (extreme values) in the NDSI image.	
<b>Non-Landslide</b>		
	In contrast to increased landslides, non-landslide areas exhibit minimal index variation, with values remaining close to zero in both NDVI <sub>diff</sub> and NDSI images.	

Table 4 ANN (BPNN) Classification Parameters

Parameter	Setting	Description
Network Type	Multilayer Perceptron	Supervised back-propagation network
Training Algorithm	Levenberg–Marquardt (LM)	Fast convergence for nonlinear optimization
Transfer Function	Sigmoid	Used in hidden and output layers
Hidden Layers	1	Single hidden layer
Hidden Neurons	6 / 12	Depends on input variables
Training Epochs	1000 / 2000	Based on model complexity
Output	Probability (0–1)	Object-wise landslide likelihood
Thresholding	Jenks Natural Breaks	Data-driven classification threshold

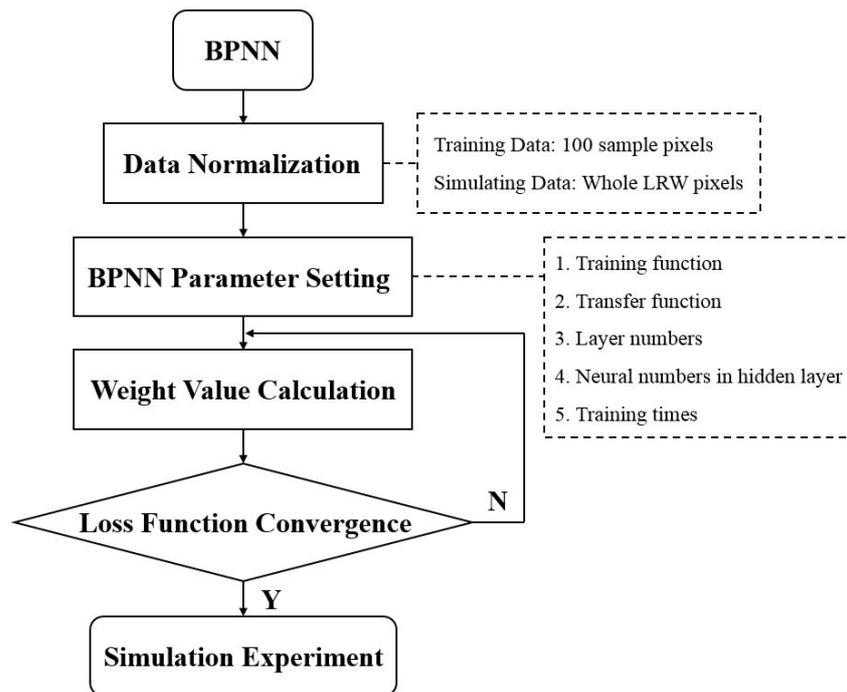


Figure 7 BPNN flow chart

The ANN produces object-wise probabilities ranging from 0 to 1, which require thresholding to classify landslide and non-landslide areas. To emulate emergency conditions without a reference inventory, the Jenks Natural Breaks (JNB) method (Jenks, 1967) was applied, as it iteratively minimizes within-class variance and maximizes between-class variance, yielding data-driven thresholds for robust landslide delineation.

## 4. Results and Discussion

Building on the object-based image analysis (OBIA) framework and the decision-level fusion strategy outlined above, four artificial neural network (ANN) classification models were developed to evaluate and simulate event-driven landslide mapping. (1) ANN-NDVI<sub>diff</sub> assesses the performance of cloud-free optical data for landslide detection, trained exclusively on six NDVI<sub>diff</sub>-0912 texture images; (2) ANN-Cloud simulates emergency conditions where cloud-free optical imagery is unavailable, using six NDVI<sub>diff</sub>-0815 texture images to test classification under adverse weather scenarios; (3) ANN-NDSI evaluates the capability of SAR data alone, trained on six NDSI texture images to capture backscatter-based changes; and (4) ANN-Fusion integrates six NDSI texture images with six NDVI<sub>diff</sub>-0815 texture images to test decision-level fusion performance under conditions mimicking delayed acquisition of clear optical data.

Although stand-alone SAR imagery can effectively delineate large-scale landslides, the integration of optical and SAR observations was conducted to improve the robustness of emergency landslide mapping. Optical vegetation indices, such as NDVI<sub>diff</sub>, are sensitive to vegetation loss and soil exposure, whereas SAR backscattering variations reflect changes in surface roughness and moisture conditions. These complementary characteristics enable more comprehensive characterization of landslide-induced surface disturbances. Moreover, SAR detection in mountainous terrain may be affected by geometric distortions, while optical imagery is often degraded by cloud cover. By integrating NDVI<sub>diff</sub> and NDSI within a decision-level fusion framework, the proposed approach enhances the spatial continuity and operational reliability of post-disaster landslide inventories.

### 4.1 Study Area and Typhoon Event

The Laonung River Watershed in Kaohsiung (Figure 8), southern Taiwan, covers about 117 km<sup>2</sup> and originates from the southwestern foothills of the Central Mountain Range. Steep, narrow valleys and highly fractured bedrock formed by folds, faults, and cleavage systems dominate the area (Mondini & Chang, 2014). Elevations range from 257 to 1673 m (mean 716 m), with 82% below 1000 m. Slopes vary from 0° to 71.2° (mean 26.9°, SD 12.6°), indicating rugged terrain. The Laonung River Watershed has a tropical monsoon climate with a mean annual temperature of 23 °C and ~3400 mm of highly seasonal rainfall (Chiang, 2017). Typhoon Morakot triggered extensive landslides and debris flows in this watershed.

Typhoon Morakot (2009) was the most destructive typhoon to hit southern Taiwan in recent decades. From 6–10 August it crossed the island from Hualien to Taoyuan, delivering a record 2749 mm of rainfall at Alishan—far exceeding the 1749 mm set by Typhoon Herb in 1996 (Lin & Jeng, 2000). The extreme precipitation caused widespread flooding, massive landslides, and debris flows. According to Tsai & Liu, 2015, Morakot isolated villages, left 693 dead and 97 missing, and inflicted about NT\$16.5 billion in agricultural and infrastructure losses, highlighting its unprecedented destructive power.

### 4.2 ANN-NDVI<sub>diff</sub>

The ANN-NDVI<sub>diff</sub> classification result (Figure 9) demonstrates the strong potential of optical imagery for mapping landslides triggered by Typhoon Morakot. Differences in spectral characteristics between disturbed and undisturbed terrain, combined with the temporal variation of NDVI, provide a powerful basis for detecting event-induced landslides. By calculating the NDVI difference between pre- (9 May 2009) and post-event (12 Sept. 2009) FORMOSAT-2 images, areas of vegetation loss are clearly distinguished from stable surfaces, enabling effective delineation of newly exposed soil and small or fragmented landslides, particularly in the southwestern area of the study area.

Despite these strengths, classification accuracy is constrained by atmospheric and illumination effects inherent to optical data. Cloud cover and terrain shadows in the 12 Sept. 2009 image produced NDVI values similar to bare soil, leading to false positives (Figure 10). For example, cloud-obscured areas C1 and C2 contained roughly 25 ha and 10 ha of redundant classified landslides (Figure 11–12), while shadowed forested slopes S1 and S2 were misclassified because topographic shading lowered NDVI values (Figure 13–14). These errors illustrate how weather conditions, solar angle, and terrain morphology can mask vegetation reflectance and confound spectral indices.

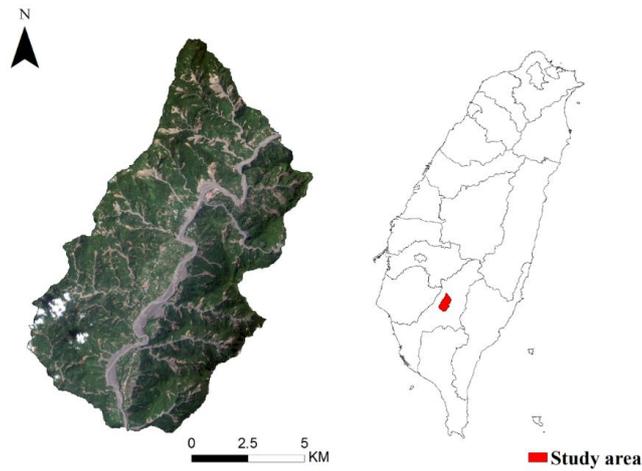


Figure 8 Study area: Laonung River Watershed overview (FORMOSAT-2, 12 Sept. 2009)

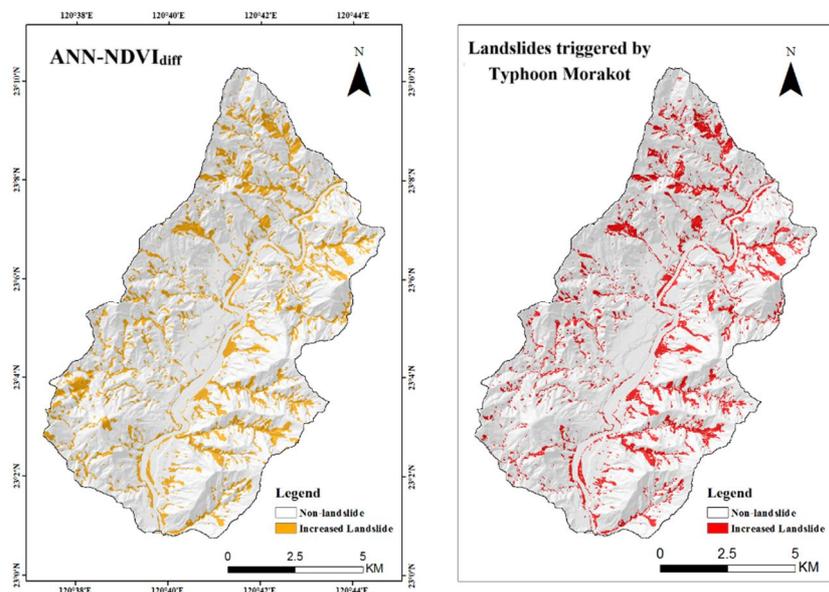


Figure 9 Comparison of the ANN-NDVI<sub>diff</sub> classification result with the ground truth

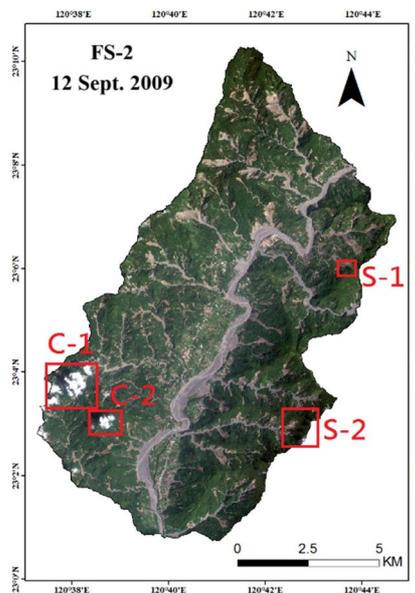


Figure 10 Locations of misclassification in the ANN-NDVI<sub>diff</sub> results: C-1 and C-2 correspond to areas affected by cloud cover, while S-1 and S-2 correspond to areas affected by terrain shadow

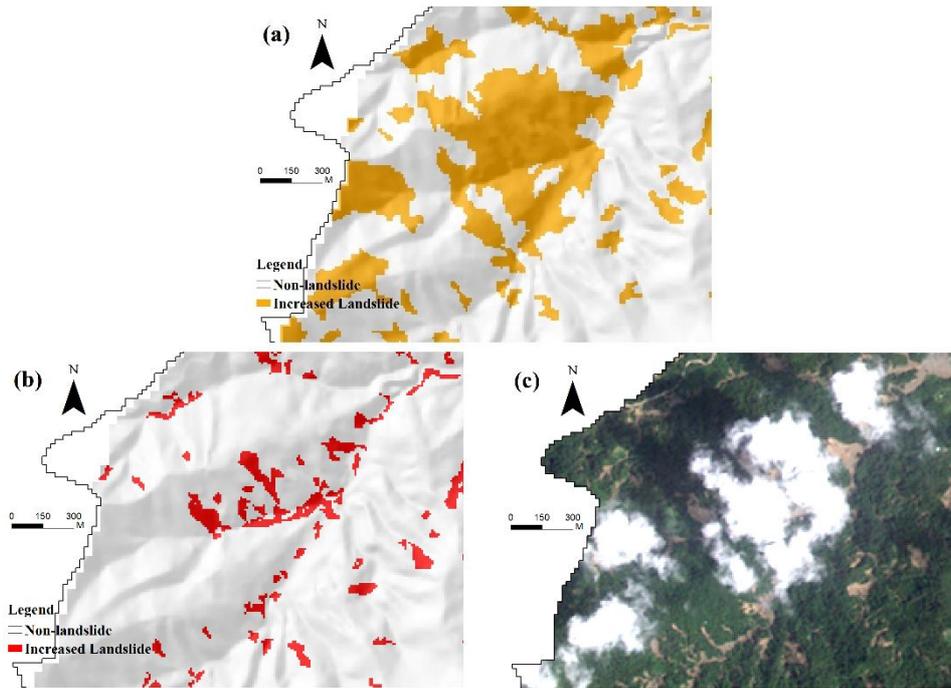


Figure 11 C-1 area. (a) ANN-NDVI<sub>diff</sub> classification result, where clouds are misclassified as landslides; (b) corresponding landslide inventory map; (c) FORMOSAT-2 image acquired on 12 Sept. 2009

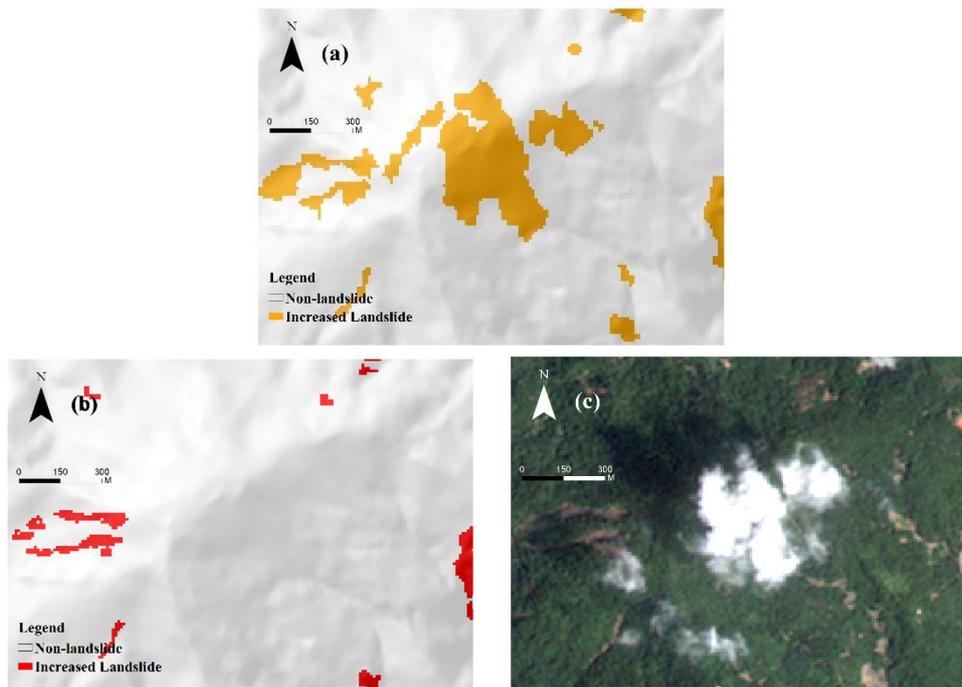


Figure 12 C-2 area. (a) ANN-NDVI<sub>diff</sub> classification result, where clouds and shadowed areas are misclassified as landslides; (b) corresponding landslide inventory map; (c) FORMOSAT-2 image acquired on 12 Sept. 2009

Overall, the object-based ANN-NDVI<sub>diff</sub> approach performs well in identifying landslides from high-resolution optical data, successfully capturing both large and small failures through combined spectral and temporal analysis of NDVI. However, its performance depends on timely cloud-free acquisitions. Because most Typhoon Morakot landslides occurred on 8 Aug.

2009 but the post-event image was not acquired until 12 Sept. 2009 due to persistent cloud cover, some temporal mismatches may have affected classification. While ANN-NDVI<sub>diff</sub> can produce accurate landslide maps under favorable conditions, these limitations underscore the reduced suitability of optical-only methods for rapid emergency mapping when immediate post-disaster imagery is unavailable.

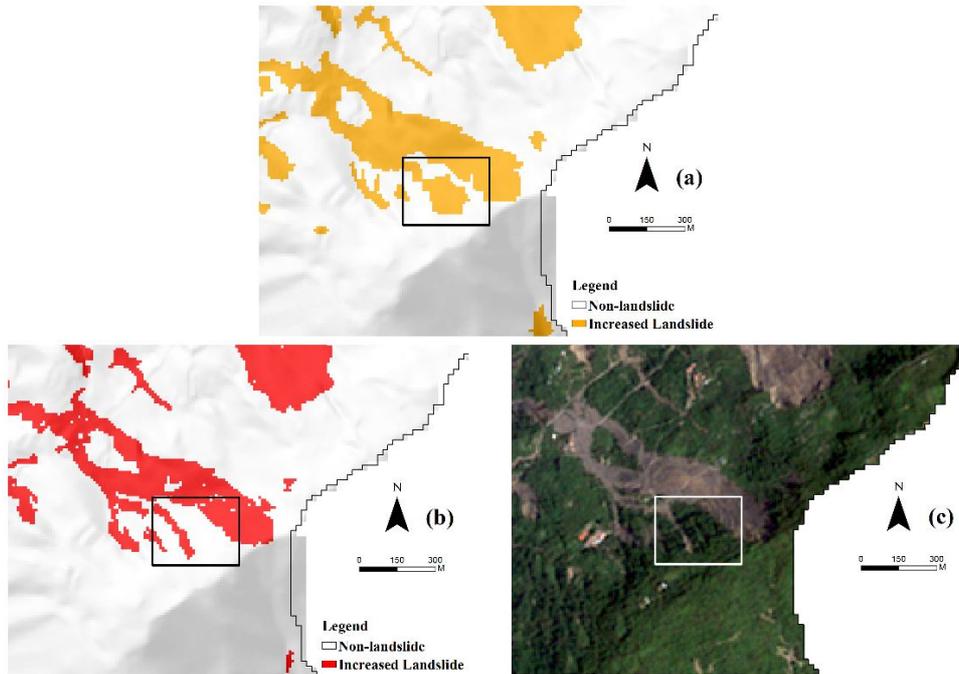


Figure 13 S-1 area. (a) ANN-NDVI<sub>diff</sub> classification result, where shadowed areas are misclassified as landslides; (b) corresponding landslide inventory map; (c) FORMOSAT-2 image acquired on 12 Sept. 2009

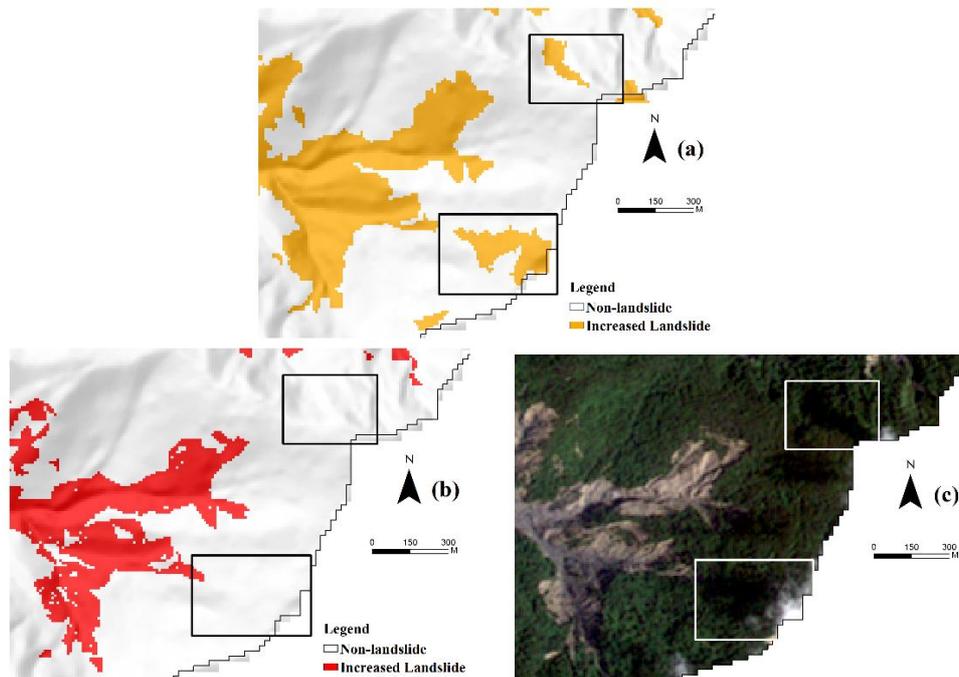


Figure 14 S-2 area. (a) ANN-NDVI<sub>diff</sub> classification result, where shadowed areas and clouds are misclassified as landslides; (b) corresponding landslide inventory map; (c) FORMOSAT-2 image acquired on 12 Sept. 2009

### 4.3 ANN-Cloud

The ANN-Cloud model was designed to replicate emergency mapping scenarios in which post-event optical images are severely obstructed by cloud cover. NDVI<sub>diff</sub> derived from the 15 Aug. 2009 scene was used, with cloud-covered areas masked and assigned a value of “0” to emulate the limited information available

during adverse weather. As shown in Figure 15, this constraint markedly reduced classification performance relative to the inventory of Typhoon Morakot-triggered landslides. While the model successfully identified broad non-landslide areas, it substantially underestimated landslide extent, classifying most of the study area as non-landslide and omitting numerous true landslides. This outcome

reflects the heavy dependence of classification on  $NDVI_{diff-0815}$  variables—particularly the mean value—under masked conditions. When large portions of an image are neutralized to zero values, the ANN is forced to rely on insufficient spectral cues, diminishing its ability to discriminate disturbed from undisturbed terrain. Although the resulting accuracy is weak, it underscores a critical limitation of optical imagery for emergency mapping, where the unavailability of clear post-event images inevitably compromises rapid analysis.

#### 4.4 ANN-NDSI

The ANN-NDSI results (Figure 16) show that SAR-derived NDSI can serve as a stand-alone tool for rapid landslide mapping under emergency conditions. Exploiting L-band backscatter sensitivity to surface roughness, moisture, and slope orientation, NDSI effectively distinguishes disturbed from stable

ground—even when clouds or low illumination prevent optical observation—allowing the model to delineate the extent and location of large landslides without post-event optical input.

Nevertheless, the results also highlight radar’s inherent constraints. Foreshortening, layover, and shadowing in steep terrain distort or displace signals from small or fragmented landslides, which are often merged with adjacent stable areas during segmentation. These effects are strongest in the south-western area (Figure 17), where slopes face away from the ascending radar beam and failures are small and discontinuous, leading to underestimation. In contrast, the model performed better in the north-west (Figure 18), north-east (Figure 19), and south-east (Figure 20) areas, where larger, more coherent landslides—such as the ~50 ha event in the north-west—produced strong, contiguous backscatter signatures that the ANN classified and outlined more accurately.

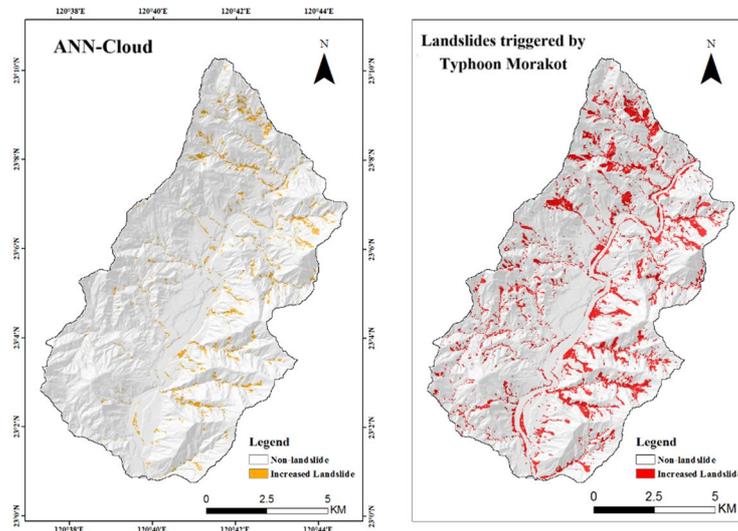


Figure 15 Comparison of the ANN-Cloud classification result with the ground truth

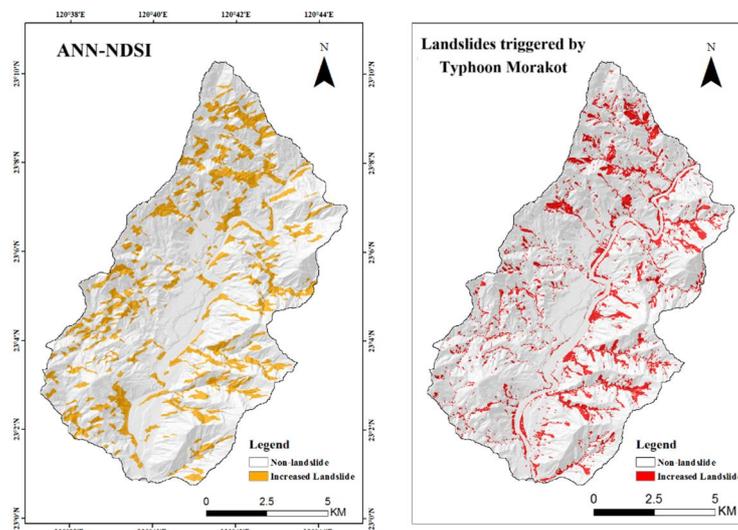


Figure 16 Comparison of the ANN-NDSI classification result with the ground truth

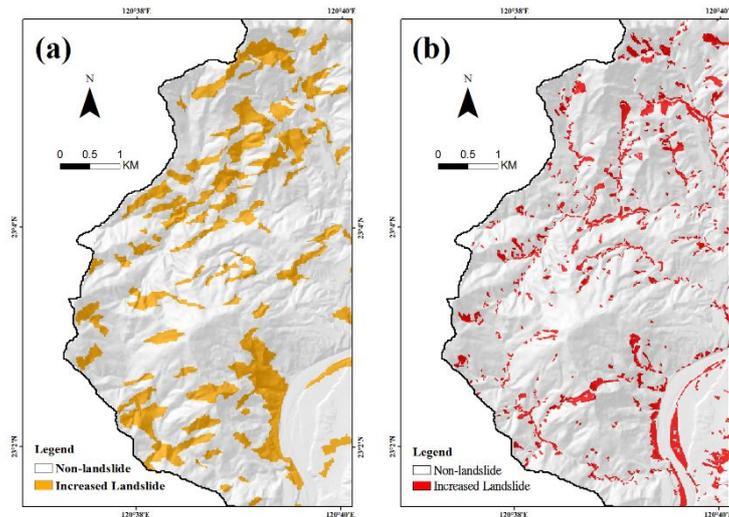


Figure 17 Southwestern area of the Laonong River Watershed: (a) ANN-NDSI classification result; (b) corresponding landslide inventory map

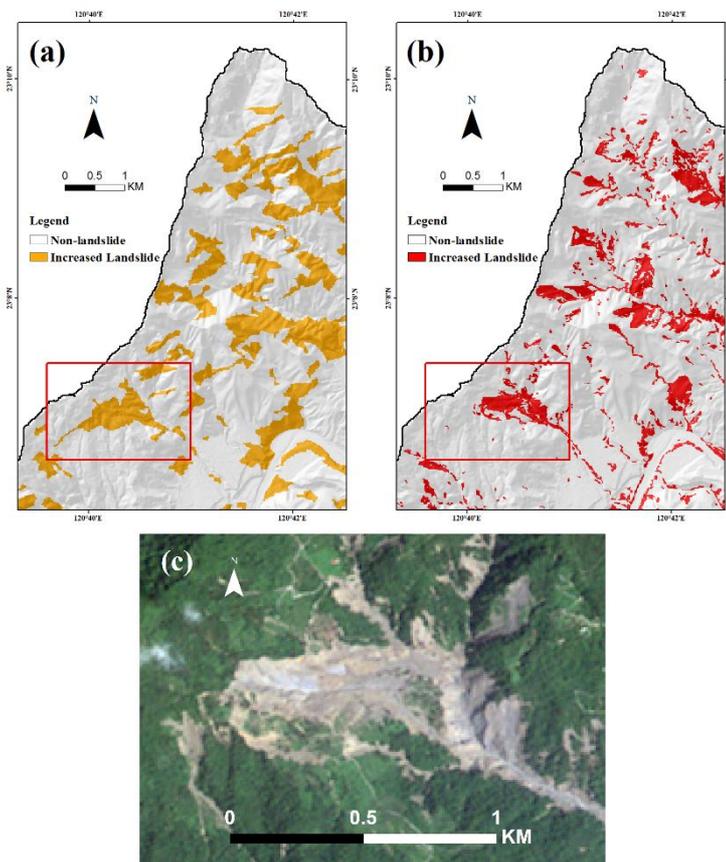


Figure 18 Northwestern area of the Laonong River Watershed: (a) ANN-NDSI classification result; (b) corresponding landslide inventory map; (c) FORMOSAT-2 image acquired on 12 Sept. 2009 showing the example landslide highlighted in the red box

Although the resulting footprints cannot perfectly coincide with ground-truth boundaries because of the inherent physical constraints of radar imaging, the overall performance confirms that NDSI provides a viable substitute for optical indices in emergency contexts. Under such conditions, SAR imagery enables rapid delineation of the approximate location and

general extent of landslides owing to its all-weather and day–night acquisition capability. Nevertheless, its side-looking imaging geometry inevitably introduces distortions, including foreshortening, layover, and shadowing in mountainous terrain, which hinder precise spatial correspondence with reference inventories and constrain the applicability of

conventional statistical evaluation methods. As a result, rigorous quantitative accuracy assessment based on traditional metrics remains challenging for SAR-dominated results in complex topographic environments. Consequently, this study did not rely solely on conventional statistical indicators to validate fusion performance. Instead, the effectiveness of the proposed framework was primarily examined through comparative analysis of landslide location, areal extent, and overall spatial distribution patterns across optical-only, SAR-only, and integrated results. These analyses demonstrate that data fusion enhances mapping reliability by combining the rapid, weather-independent detection capability of SAR imagery with the high spatial fidelity and boundary definition provided by optical observations under favorable conditions.

### 4.5 ANN-Fusion

The ANN-Fusion model integrates  $NDVI_{diff-0815}$  and NDSI texture features to simulate rapid post-event landslide mapping under conditions where cloud-free optical data are unavailable. As shown in the comparison between the fused classification output and the ground-truth inventory (Figure 21), the combined approach markedly improves emergency mapping performance relative to single-sensor models. By blending the spectral sensitivity of  $NDVI_{diff}$  with the backscatter sensitivity of NDSI, the model captures both vegetation loss and radar-derived surface changes, allowing more complete delineation of disturbed slopes immediately after the typhoon. Although the fusion is dominated by the  $NDVI_{diff}$  component—reflecting the stronger spectral contrast in optical data—the NDSI contribution becomes especially evident in the western

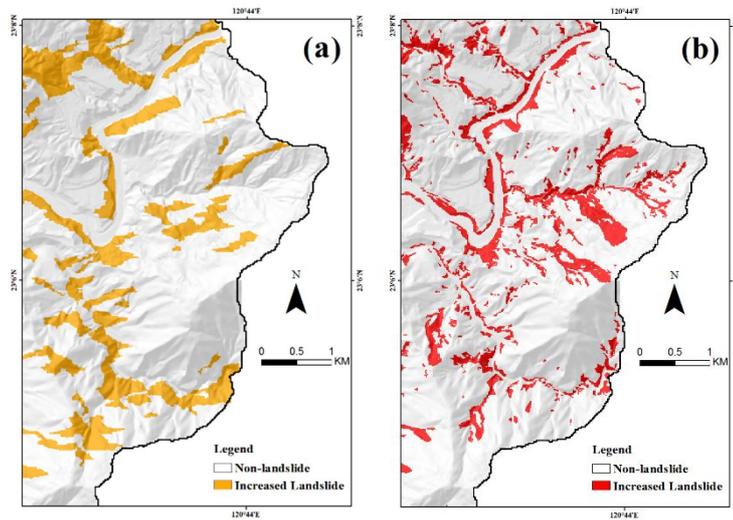


Figure 19 Southeastern area of the Laonong River Watershed: (a) ANN-NDSI classification result; (b) corresponding landslide inventory map

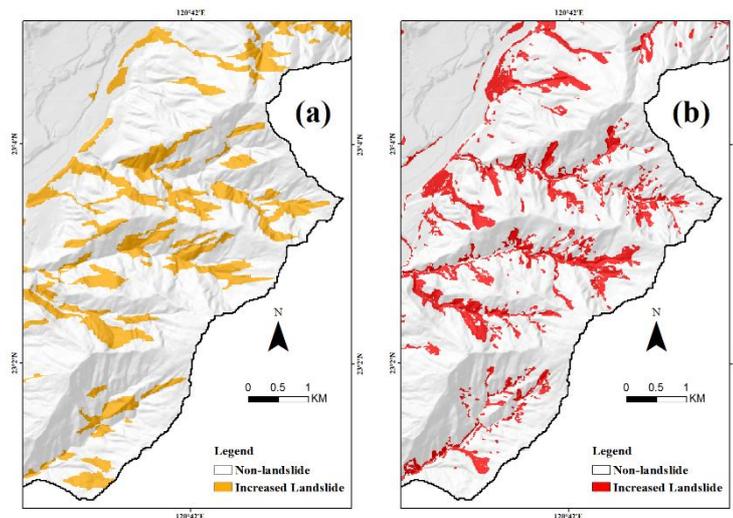


Figure 20 Southeastern area of the Laonong River Watershed: (a) ANN-NDSI classification result; (b) corresponding landslide inventory map

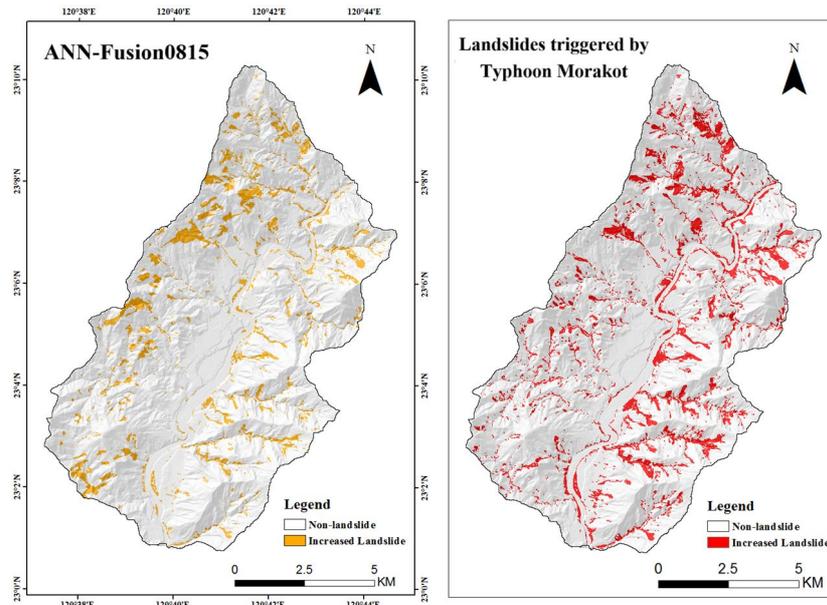


Figure 21 Comparison of the ANN-Fusion0815 classification result with the ground truth

portion of the study area where cloud cover masks large areas of the optical image. In these areas, radar backscatter differences supplement the missing optical signal and help recover landslide outlines that would otherwise be obscured. This complementary effect between optical and radar inputs enhances the robustness of the classification, yielding a more spatially coherent and realistic representation of landslide extent than either index alone. Consequently, ANN-Fusion0815 demonstrates the potential of decision-level image fusion to balance the strengths of optical and SAR data, providing a practical framework for rapid and reliable landslide inventory generation in emergency scenarios when clear optical imagery cannot be obtained.

## 5. Conclusions and Recommendation

This study presents an ANN-based decision-level image fusion framework within an OBIA environment for rapid landslide mapping after Typhoon Morakot. By combining  $NDVI_{diff}$  from optical imagery with  $NDSI$  from L-band SAR backscatter, the approach exploits complementary spectral and radar information while avoiding typical co-registration challenges in rugged terrain. OBIA segmentation and texture extraction generate object-level features that improve classification robustness and reduce noise sensitivity.

Results show that  $NDVI_{diff}$  performs well under clear conditions,  $NDSI$  provides a viable substitute when optical data are cloud-obstructed, and their fusion yields the most complete and coherent delineation of landslides under emergency conditions. Although SAR

distortions and temporal gaps in optical imagery limit perfect agreement with ground truth, the fused ANN model effectively captures the extent and location of major landslides even under heavy cloud cover. This framework delivers timely, high-resolution landslide inventories to support disaster response, post-event assessment, and risk-reduction planning in other mountainous, cloud-prone regions.

Future research should refine the ANN-OBIA fusion framework to further enhance its operational value, integrate near-real-time SAR and optical data from multiple satellite constellations to reduce latency, develop algorithms to mitigate SAR distortions and improve small-landslide delineation, and embed the workflow into official disaster-response systems for broader application to other rapid-onset hazards.

## References

- Amarsaikhan, D., and Douglas, T., 2004. Data fusion and multisource image classification, *International Journal of Remote Sensing*, 25(17): 3529-3539, DOI: 10.1080/0143116031000115111.
- Chen, H., He, Y., Zhang, L., Yang, W., Liu, Y., Gao, B., Zhang, Q., and Lu, J., 2023. A multi-input channel U-Net landslide detection method fusing SAR multisource remote sensing data, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 1215-1232, DOI: 10.1109/JSTARS.2023.3339294.
- Chen, Y.K., 2019. An artificial neural network approach for event landslide detection based on the integration of optical and SAR textures, Master Thesis, National Central University, Taiwan, ROC.

- Chiang, S.H., 2017. The application of Sentinel-1 Synthetic Aperture Radar (SAR) imagery to the assessment of landslide reactivation, Soil and Water Conservation Bureau, Taiwan, ROC.
- Colesanti, C., and Wasowski, J., 2006. Investigating landslides with space-borne Synthetic Aperture Radar (SAR) interferometry, *Engineering Geology*, 88(3-4): 173-199, DOI: 10.1016/j.enggeo.2006.09.013.
- Coluzzi, R., Perrone, A., Samela, C., Imbrenda, V., Manfreda, S., Pace, L., and Lanfredi, M., 2025. Rapid landslide detection from free optical satellite imagery using a robust change detection technique, *Scientific Reports*, 15(1): 4697, DOI: 10.1038/s41598-025-89542-8.
- Dai, F.C., and Lee, C.F., 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong, *Geomorphology*, 42(3-4): 213-228, DOI: 10.1016/S0169-555X(01)00087-3.
- Ez-zahouani, B., Teodoro, A., Kharki, O.E., Jianhua, L., Kotaridis, I., Yuan, X., and Ma, L., 2023. Remote sensing imagery segmentation in object-based analysis: A review of methods, optimization, and quality evaluation over the past 20 years, *Remote Sensing Applications: Society and Environment*, 32: 101031, DOI: 10.1016/j.rsase.2023.101031.
- Fu, S., de Jong, S.M., Hou, X., de Vries, J., Deijns, A., and de Haas, T., 2024. A landslide dating framework using a combination of Sentinel-1 SAR and-2 optical imagery, *Engineering Geology*, 329: 107388, DOI: 10.1016/j.enggeo.2023.107388.
- Furuta, R., and Tomiyama, N., 2008. A study of detection of Landslide disasters due to the Pakistan earthquake using ALOS data, in *Proceedings of the 34th International Symposium on Remote Sensing of Environment*, Sydney, Australia.
- Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., and Chang, K.T. 2012. Landslide inventory maps: New tools for an old problem, *Earth-Science Reviews*, 112(1-2): 42-66, DOI: 10.1016/j.earscirev.2012.02.001.
- Haeberlin, Y., Turberg, P., Retière, A., Senegas, O., and Parriaux, A., 2004. Validation of Spot-5 satellite imagery for geological hazard identification and risk assessment for landslides, mud and debris flows in Matagalpa, Nicaragua, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXV-B3:273-278.
- Haralick, R.M., Shanmugam, K., and Dinstein, I., 1973. Textural features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC3(6): 610-621, DOI: 10.1109/TSMC.1973.4309314.
- Jenks, G.F., 1967. The data model concept in statistical mapping, *International Yearbook of Cartography*, 7: 186-190.
- Jiang, P., Ma, Z., and Mei, G., 2025. Deep learning for potential landslide identification: Data, models, applications, challenges, and opportunities, *EGU sphere*, DOI: 10.5194/egusphere-2025-2158.
- Jin, Y., Li, X., Zhu, S., Tong, B., Chen, F., Cui, R., and Huang, J., 2022. Accurate landslide identification by multisource data fusion analysis with improved feature extraction backbone network, *Geomatics, Natural Hazards and Risk*, 13(1): 2313-2332, DOI: 10.1080/19475705.2022.2116357.
- Li, Y., and Xiao, X., 2025. Deep learning-based fusion of optical, radar, and LiDAR data for advancing land monitoring, *Sensors*, 25(16): 4991, DOI: 10.3390/s25164991.
- Lin, M.L., and Jeng, F.S., 2000. Characteristics of hazards induced by extremely heavy rainfall in Central Taiwan—Typhoon Herb, *Engineering Geology*, 58(2): 191-207, DOI: 10.1016/S0013-7952(00)00058-2.
- Liu, B., Wang, W., Wu, Y., and Gao, X., 2024. Attention swin transformer UNet for landslide segmentation in remotely sensed images, *Remote Sensing*, 16(23): 4464, DOI: 10.3390/rs16234464.
- Mondini, A.C., and Chang, K.T., 2014. Combining spectral and geoenvironmental information for probabilistic event landslide mapping, *Geomorphology*, 213: 183-189, DOI: 10.1016/j.geomorph.2014.01.007.
- Nichol, J.E., Shaker, A., and Wong, M.S., 2006. Application of high-resolution stereo satellite images to detailed landslide hazard assessment, *Geomorphology*, 76(1-2): 68-75, DOI: 10.1016/j.geomorph.2005.10.001.
- Pohl, C., and Van Genderen, J.L., 1998. Review article multisensor image fusion in remote sensing: Concepts, methods and applications, *International Journal of Remote Sensing*, 19(5): 823-854, DOI: 10.1080/014311698215748.
- Prakash, N., Manconi, A., and Loew, S., 2020. Mapping landslides on EO data: Performance of deep learning models vs. traditional machine learning models, *Remote Sensing*, 12(3): 346, DOI: 10.3390/rs12030346.
- Selby, M.J., 1993. *Hillslope Materials and Processes*, Oxford: Oxford university press, p. 451.
- Tetteh, G.O., Schwieder, M., Erasmi, S., Conrad, C., and Gocht, A., 2023. Comparison of an optimised

- multiresolution segmentation approach with deep neural networks for delineating agricultural fields from sentinel-2 images, *ISPRS International Journal of Geo-Information*, 14(12): 3123-3142, DOI: 10.1007/s41064-023-00247-x.
- Tsai, P.H., and Liu, C.W., 2015. Local economy growing at a steady pace: The case of the peasant farming project by the Taiwan Rural Front 1, In: *Social Economy in China and the World*, Pun, N., Ku, B.H., Yan, H., and Koo, A. (eds.), Routledge, London, pp. 248-259, DOI: 10.4324/9781315718286.
- Tuo, W., Zeng, J., Wu, F., Wu, W., Wang, X., and Zhao, X., 2025. Landslide detection using multimodal data fusion and an improved Deeplabv3+ model, *Scientific Reports*, 16(1): 1383, DOI: 10.1038/s41598-025-31208-6.
- Wang, H., Liu, X., Feng, G., Liu, P., Li, W., Liu, S., and Liao, W., 2025. Applicability of multi-sensor and multi-geometry SAR data for landslide detection in southwestern China: A case study of Qijiang, Chongqing, *Sensors*, 25(14): 4324, DOI: 10.3390/s25144324.
- Weier, J., and Herring, D., 2000. *Measuring Vegetation (NDVI & EVI)*, NASA Earth Observatory, Washington, DC, USA.
- Zhang, Q., and Wang, T., 2024. Deep learning for exploring landslides with remote sensing and geo-environmental data: Frameworks, progress, challenges, and opportunities, *Remote Sensing*, 16(8): 1344, DOI: 10.3390/rs16081344.

# 應用類神經網路整合光學-SAR 紋理分析於緊急崩塌測繪

陳以耕<sup>1\*</sup> 姜壽浩<sup>2</sup>

## 摘要

光學遙測影像可提供崩塌目錄製作所需之關鍵光譜資訊，但易受雲霧與天候影響而限制其應用。相較之下，合成孔徑雷達具全天候觀測能力且對地表後向散射變化敏感，可作為崩塌擾動之重要補充。本研究整合光學與雷達資料，發展快速且穩定之崩塌地辨識方法。方法採物件導向影像分析進行地物分割，並由多期影像推導標準化植被指數差值與標準化後向散射指數，計算六項灰階共生矩陣紋理特徵以表徵地物變化。本研究比較光學、雷達、受雲影響光學及光學-雷達融合四種情境。結果顯示，融合模式可提升崩塌地判釋之空間一致性與完整性；於高雲覆條件下，雷達亦能有效辨識大型崩塌地。顯示所提方法具良好作業效能，可應用於事件導向之快速製圖與災害評估。

**關鍵詞：**崩塌測繪、合成孔徑雷達、影像紋理、影像融合、人工類神經網路

---

<sup>1</sup> 國立臺灣大學土木工程學系 博士生

<sup>2</sup> 國立中央大學太空及遙測研究中心 副教授

\* 通訊作者, E-mail: d10521008@ntu.edu.tw

收到日期：民國 114 年 10 月 15 日

修改日期：民國 115 年 01 月 07 日

接受日期：民國 115 年 02 月 04 日

# 陸海交界處大地起伏建模精度探討：以臺灣東部為例

楊致逸<sup>1</sup> 蕭宇伸<sup>2\*</sup>

## 摘要

本研究以臺灣東部花蓮清水斷崖沿岸為研究區，探討數值地形模型(DEM)、數值海底地形模型(DBM)空間解析度與海岸線定位精度對陸海交界處重力法大地起伏建模之影響。研究採用 GSHHG 與人工數化海岸線，結合不同解析度 DEM-DBM，於去除-回復架構下，以剩餘地形模型及最小二乘配置法建立大地起伏模型，並以 GNSS/水準觀測資料進行精度檢核。結果顯示，GSHHG 海岸線於斷崖海岸與實際海岸線存在明顯差異，易造成海岸附近地形遮罩與改正誤差；人工數化海岸線可提升陸海邊界一致性與模型穩定性。解析度比較顯示，3"×3" DEM-DBM 之整體精度優於 9"×9"與 1"×1"。研究結果指出，精確海岸線與適當地形解析度為提升陸海交界區大地起伏建模品質之關鍵。

**關鍵詞：**數值高程模型、數值海底地形模型、大地起伏建模、陸海交界區

## 1. 前言

在地球科學與空間資訊領域中，大地水準面為關鍵參考面，廣泛應用在高程系統中橢球高轉成正高的高程轉換依據。尤其不進行傳統水準測量的條件下，準確大地起伏模型更成為高程基準現代化不可或缺的資料來源，應用涵蓋海平面變遷監測、工程測量與地殼變形等多個領域。然而在地形變化劇烈的區域，例如臺灣東部的花蓮沿海，地形於短距離內自高山急遽過渡至深海，組成複雜的陸海交界帶，使得大地起伏建模面臨極大挑戰。在此類區域中，除了重力資料外，地形的準確性對大地起伏建模成果有顯著影響，而數值高程模型(Digital Elevation Model, DEM)與數值海底地形模型(Digital Bathymetric Model, DBM)於陸海銜接區域的準確性，應是關鍵。

大地起伏模型的計算方法大致可分為三類：重力法、幾何法與天文大地法(Heiskanen & Moritz, 1967)。重力法依據大地起伏與重力異常之物理關係推算地形改正與起伏量，為臺灣最主要之大地起伏建模方法，近年來臺灣大地起伏模型計算均採用之，

例如許宏銳(2014)、內政部(2023、2025)等；在幾何法方面，是透過 GNSS 橢球高與水準正高之差值得到觀測大地起伏值，其精度頗高，然而此方法難以完整建構大地起伏面，故一般較常應用於檢核重力法大地起伏模型之成果；在天文大地法方面，則依靠垂線偏差與天文測量資料獲得大地起伏資訊，目前在國際上已經非常少用此法。大地起伏無論採用上述何種方法進行建模，在地形變化劇烈的地區，都需要精確的 DEM 進行地形相關改正。臺灣重力法大地起伏建模方面，崔國強 (1995) 早期採用 Stokes 積分法建立臺灣地區重力法大地起伏模型，Hwang (1997) 進一步導入最小二乘配置法(Least-Squares Collocation, LSC)獲得臺灣重力法大地起伏模型；Hwang & Hsiao (2003)亦使用 LSC 法，結合剩餘地形模型(Residual Terrain Model, RTM)理論，計算臺灣重力法大地起伏模型，RTM 原理為考慮研究區位細微地形對大地起伏模型的影響，故必須採用研究區位的 DEM 進行計算；內政部於 2005 年推動之空載重力測量作業(內政部, 2005)亦採用 RTM + LSC 結構納入建模流程，驗證該大地起伏模型在平地可達公分級精度；目前內政部公告的臺灣大地

<sup>1</sup> 國立中興大學水土保持學系 碩士

<sup>2</sup> 國立中興大學水土保持學系 副教授

\* 通訊作者, E-mail: yhsiao@nchu.edu.tw

收到日期：民國 114 年 07 月 30 日

修改日期：民國 114 年 09 月 03 日

接受日期：民國 115 年 02 月 16 日

起伏模型 TWHYGEO2014，採用 Stokes 積分法與最小二乘配置 LSC 法計算，而在 RTM 方面，採用內政部 3 秒×3 秒以及 9 秒×9 秒網格之 DEM (<https://www.nlsc.gov.tw/cp.aspx?n=1551>)，亦達平地公分級精度。近年來國際上採用 RTM + LSC 方法計算大地起伏的研究包括 Yildiz (2012)、Willberg *et al.* (2019)、Bucha *et al.* (2019)、Vu *et al.* (2019) 和 Jiang *et al.* (2020)等，也均獲得良好的成果。

無論採用 Stokes 或 LSC 計算方法，地形的準確度與解析度均非常重要，Hwang *et al.* (2006) 探討 RTM 技術(如稜柱積分與 Gaussian 積分法)對模型準確度之影響，指出地形遮罩與積分區域設定對山區與海岸區模型精度有關鍵性貢獻。地形遮罩即海陸地形融合前，依據海岸線位置所製作的 0/1 檔地形，積分區域設定即積分時，計算半徑之選擇。在大地起伏建模中，陸海交界地區長期被視為誤差高度集中的敏感區域。該區域常受限於 DEM 與 DBM 融合精度不一致、海岸線定義模糊等問題，進而影響地形遮罩生成與地形改正的準確性 (Hwang *et al.*, 2020)。隨著 DEM 空間解析度的提高，模型對邊界幾何與積分設計的敏感度亦隨之增加，使得邊界積分不足與遮罩不連續問題更為明顯 (Kuhn *et al.*, 2009)。Forsberg (1984) 在建立 RTM 理論時即指出，陸海交界不連續的地形將對 Stokes 積分產生邊界震盪效應。後續研究如 Omang & Forsberg (2000) 比較 Helmert 壓縮法與 RTM 方法在海岸區的表現，指出若高頻地形起伏未妥善處理，模型將產生局部偏移現象。由上述文獻可得知，在高解析度大地水準面建模中，海岸線定義與地形遮罩的準確性，已成為影響陸海交界區建模穩定性與誤差分布的關鍵因素。因此，針對海岸線進行人工數化修正與精細遮罩製作，是提升地形改正精度及重力場模型可靠度的重要研究方向。然而，針對 DEM - DBM 在陸海交界區之精度評估與海岸線誤差對地形改正與大地起伏模型精度的影響，國內外相關研究仍相對有限。本文以臺灣花蓮東部變化最為劇烈的清水斷崖、海岸山脈為研究區域，首先蒐集內政部 20 m DEM 與 50 m DBM，再分別搭配

Hierarchical High-resolution Geography, GSHHG)海岸線及人工數化海岸線，產製不同海岸線、不同空間解析度的 DEM-DBM 連續模型，接著以 LSC+RTM 方法計算各測試例的大地起伏模型，最後利用 GNSS/水準測量檢核點分析各測試例的準確性，具體步驟如圖 1 所示。

## 2. 研究方法

### 2.1 海岸線

海岸線的取得是本研究中建立高精度陸海邊界模型的關鍵步驟，其目的在於提升 DEM 與 DBM 融合時之邊界一致性。本文海岸線分為 GSHHG 與人工數化，GSHHG 取自通用製圖工具 (Generic Mapping Tools, GMT) 6.5.0 軟體 (Wessel *et al.*, 2019)，而人工數化海岸線是以 google earth 衛星影像作為底圖，於岸線明顯之海陸交界處，以人工點選方式逐點標示實際海岸線位置，點距約 10 m 以確保曲線連續。完成點選後，再利用 GMT 6.5.0 將點位連線形成向量化岸線。

### 2.2 大地起伏模型計算與精度評估

#### 2.2.1 重力法大地起伏模型計算原理

本研究重力法大地起伏計算步驟採用去除-回復法 (Remove-Compute-Restore, RCR) (Forsberg, 1984)，RCR 是由長波長(低頻)、中波長(中頻)和短波長(高頻)所組成。長波長使用全球大地位模式計算出長波長重力異常與大地起伏，短波長則採用 RTM 計算出短波長重力異常與大地起伏，將重力異常扣除長波長和短波長重力效應後，以 LSC 計算中波長大地起伏，再將長波長和短波長大地起伏加回，得到似大地起伏，最後進行似大地起伏改正，得到最終重力法大地起伏。詳細計算過程如下內政部 (2005)：

$$\Delta N_{quasi} = \Delta N_{long} + \Delta N_{res} + \Delta N_{short} \dots \dots \dots (1)$$

$$\Delta g = \Delta g_{long} + \Delta g_{res} + \Delta g_{short} \dots \dots \dots (2)$$

其中  $\Delta N_{quasi}$  為似大地起伏； $\Delta g$  為重力異常； $\Delta N_{long}$  為長波長的大地起伏； $\Delta g_{long}$  為長波長的重力異常； $\Delta N_{short}$  為短波長的大地起伏； $\Delta g_{short}$  為短波長的重力異常； $\Delta N_{res}$  為中波長的大地起伏； $\Delta g_{res}$  為中波長的重力異常。 $\Delta g_{long}$  與  $\Delta N_{long}$  是採用全球大地位模式搭配球譜係數展開式進行計算。全球大地位模式為一個高解析度地球重力位模型，是透過多種重力觀測資料的整合並以球譜展開方法計算，而球譜係數展開式公式如下(Heiskanen & Moritz, 1967)：

$$\Delta g_{long} = \frac{GM}{R^2} \sum_{n=2}^N (n-1) \sum_{m=0}^n (\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda) \bar{P}_{nm}(\sin \varphi) \dots\dots\dots (3)$$

$$\Delta N_{long} = R \sum_{n=2}^N \sum_{m=0}^n (\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda) \bar{P}_{nm}(\sin \varphi) \dots\dots\dots (4)$$

其中  $\varphi$  和  $\lambda$  分別為緯度和經度； $G$  為萬有引力常數； $R$  為地球平均半徑； $M$  為地球質量； $\bar{C}_{nm}$  及  $\bar{S}_{nm}$  為完全正規化之地球引力位球譜級數； $\bar{P}_{nm}$  為完全正規化 Legendre 函數； $n$  和  $m$  則分別代表階數(degree)和次數(order)。

在短波長計算方面，RTM 需要一個平滑的平均

高程面作為基準(如圖 2 所示)，以刻劃地形中短波長成分的變化。RTM 的核心概念在於分離出地形高頻變化所對應的重力擾動，進而補償傳統建模方法在處理高頻重力效應上的不足。特別是在高山地區或地形劇烈起伏的區域，若僅依賴全球重力位模型與區域重力觀測資料，往往難以充分捕捉地形快速變化所引起的短波長重力訊號。這些高頻成分主要源自地表形狀的微細變化，亦即所謂的地形效應，若未妥善補償，將對大地水準面模型產生系統性偏差。透過 RTM 方法結合高解析度 DEM，可有效強化地形改正在短波長頻段的表現，進一步提升複雜地形區域大地水準面模型的精度與穩定性。

RTM 方法計算的短波長大地起伏  $N_{short}$  所採用之公式如下 (內政部，2005)：

$$N_{short}(x_p, y_p) = \frac{G\rho}{\gamma} \int_x \int_y \int_{h_{ref}}^h \frac{dxdydz}{\sqrt{(x-x_p)^2+(y-y_p)^2+(z-h_p)^2}} \dots\dots\dots (5)$$

其中  $G$  為萬有引力常數、 $\gamma$  為正常重力、 $\rho$  為地質密度、 $h$  代表真實高程表面高程、 $h_{ref}$  代表參考高程表面高程。RTM 方法計算的短波長重力異常  $\Delta g_{short}$  公式如下(內政部，2005)：

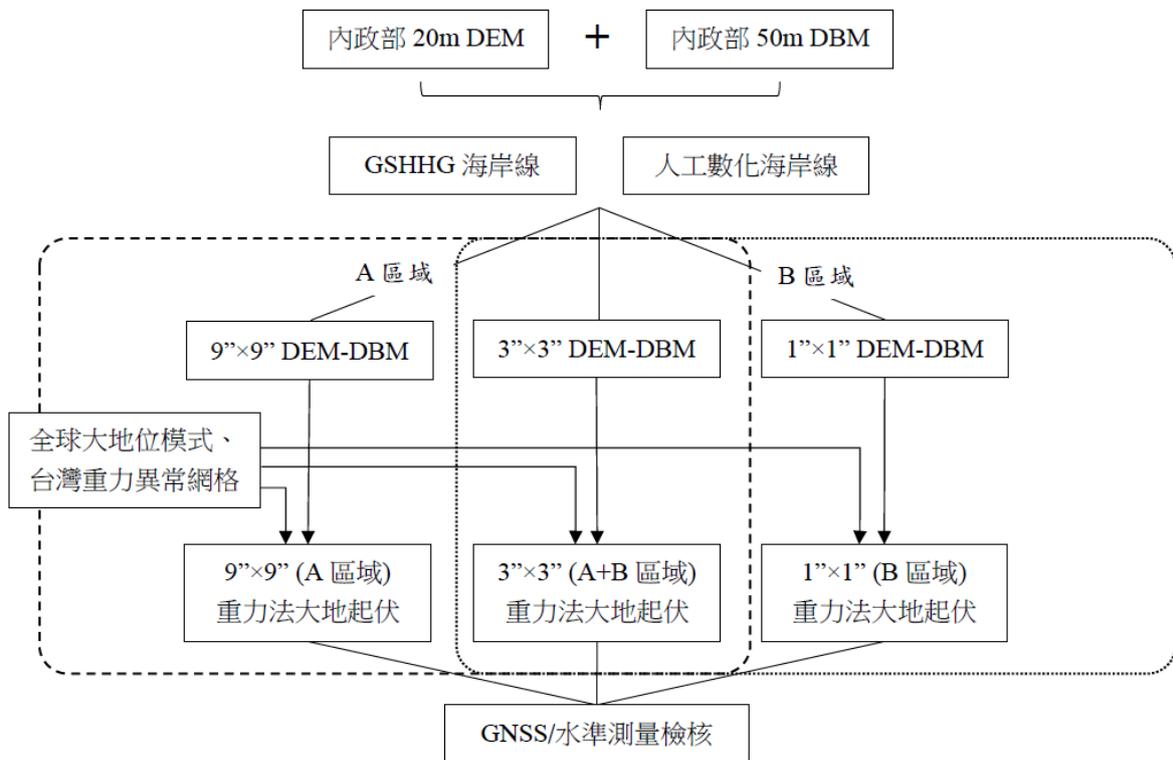


圖 1 研究流程圖

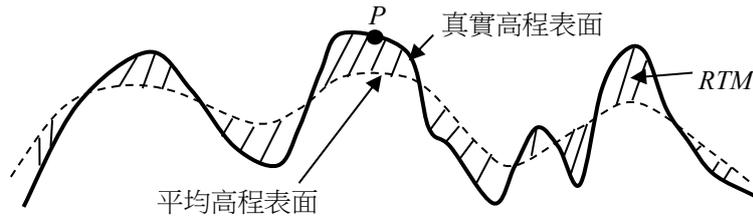


圖 2 RTM(斜線區域)示意圖，表示真實高程表面和平均高程表面之間的差異

$$\Delta g_{short}(x_p, y_p) = 2\pi G\rho(h - h_{ref}) - c(x_p, y_p) \dots (6)$$

其中  $c(x_p, y_p)$  為地形改正量。

在中波長計算方面，是大地起伏模型計算的核心。LSC 是一種基於統計推估原理的重力場建模方法，該方法結合隨機場理論與最小二乘法，具有處理觀測資料來源多樣且精度不一致以及空間分布不均的能力。此計算公式的核心在於透過協方差函數建立觀測資料與待估場變數之間的統計相關性 (Tscherning & Rapp, 1974)。透過此協方差模型，能有效選取觀測數據，並在考慮誤差與資料相關性的情況下推估未知重力場量。LSC 在中波長大地起伏  $N_{res}$  計算如下(內政部，2005)：

$$N_{res} = (C_{N\Delta g})(C_{\Delta g} + D_{\Delta g})^{-1}(\Delta g_{res}) \dots \dots \dots (7)$$

其中  $C_{N\Delta g}$  為大地起伏-重力異常間的協方差矩陣； $C_{\Delta g}$  為重力異常間的協方差矩陣； $D_{\Delta g}$  為重力異常誤差變異數的對角矩陣； $\Delta g_{res}$  為中波長的重力異常。最後將似大地起伏改正量  $\frac{2\pi G\rho}{\gamma} H^2$  加上，最後得到重力法大地起伏模型，公式如下：

$$N_{grav} = N_{quasi} - \frac{2\pi G\rho}{\gamma} H^2 \dots \dots \dots (8)$$

其中  $N_{quasi}$  為似大地起伏模型。 $N_{grav}$  為最終之重力法大地起伏模型。

### 2.2.2 精度評估方法

本文大地起伏模型精度的評估方法是利用研究區域內 GNSS/水準測量點位，內插到重力法大地起伏模型中，並比較 GNSS/水準測量觀測大地起伏值與重力法大地起伏值的差值。在 2023 年內政部之垂直基準面時變精度檢核評估工作案中(內政部，2023)，有使用臺灣本島的 14 條特選路線作為檢核大地起伏模型精度的依據。這些路線的點位均為一

等水準點，且 GNSS 觀測時間介於 12 小時到 24 小時之間，故具有高精度的觀測大地起伏值，且這些點位已經考量坐標框架變動與地殼形變速度場問題，已經修正 GNSS 和水準測量不同時期量測結果的一致性與準確性(內政部，2025)。本文所使用的 GNSS/水準測量點位，均為該計畫使用之點位。

## 3. 實驗區域與實驗資料

### 3.1 實驗區域

本研究針對臺灣東部清水斷崖、海岸山脈等地進行大地起伏模型精度影響之探討，該區域為陸海交界處中地形最為劇烈變化之區域。清水斷崖擁有山區海拔逾 3,000 m 的高山，於短距離內陡降至鄰近海域深度超過 1,000 m 之海底，具顯著的高程落差與複雜的地貌形態。此一劇烈起伏的地形條件提供了理想的環境，以分析地形效應對大地起伏與重力異常模型的影響，特別適合探討在陸海交界帶進行地形改正與模型評估之相關議題。研究區域如圖 3 所示，分為 A 區域和 B 區域。A 區域涵蓋中央山脈臨太平洋、以及大部分海岸山脈地區，B 區域僅聚焦在清水斷崖地區。

### 3.2 實驗資料

數值地形模型資料陸地部分使用內政部 2024 年全台 20 m DEM，如圖 4(a)所示。數值海洋模型則是使用內政部所提供的 50 m DBM，如圖 4(b)所示，此模型由內政部水深資料庫(CARIS Bathymetry DataBASE, BDB)、國科會航跡資料及海洋學門資料庫之 200 m 網格資料組合而成。本研究結合陸地 20 m DEM 及海洋 50 m DBM，根據不同海岸線選擇、不同空間解析度選擇產製數個 DEM-DBM 地形模型加以分析。

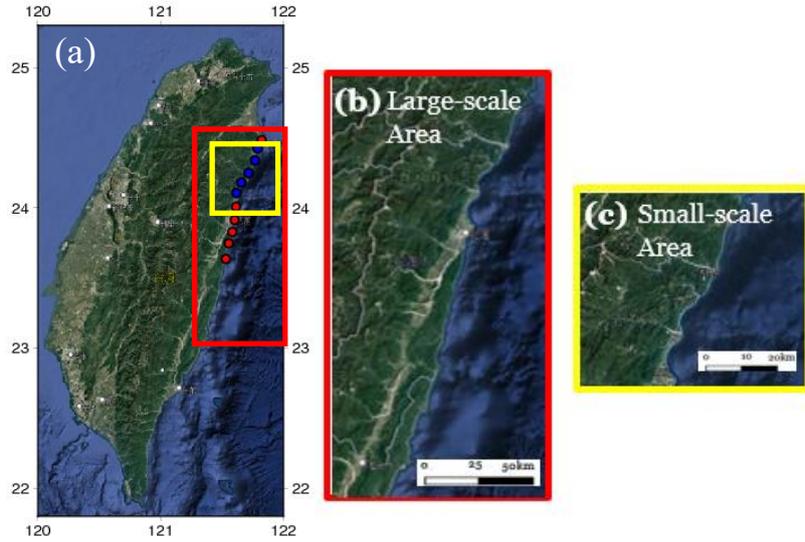


圖 3 (a) 全台衛星影像圖，紅色與藍色圓點為 GNSS/水準測量檢核點，其中紅色框為 A 區域，黃色框為 B 區域；(b) A 區域衛星影像圖；(c) B 區域衛星影像圖

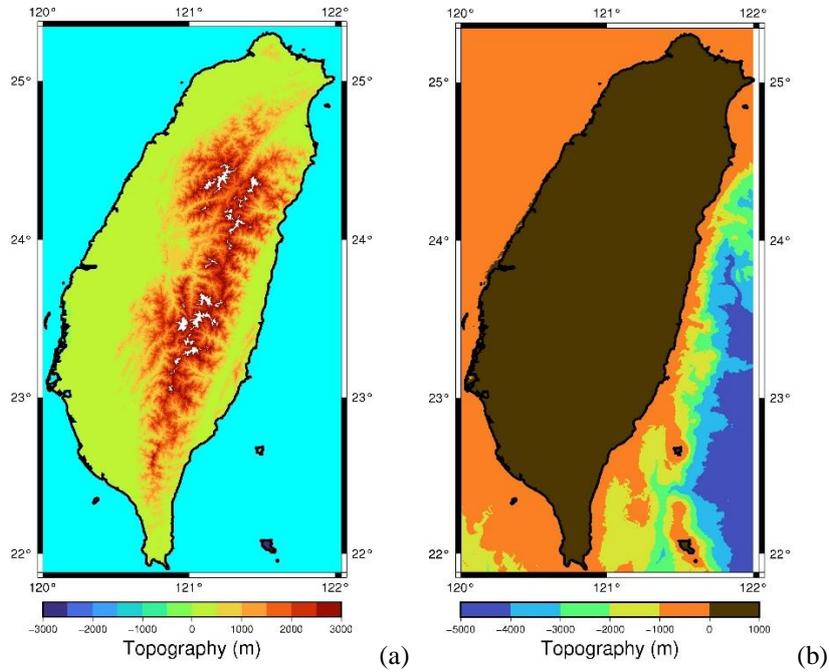


圖 4 (a) 內政部 2024 年全台 20 m DEM；(b) 內政部 50 m DBM

在長波長大地起伏計算方面，本研究使用應用最廣泛之 EGM2008 模型(Pavlis *et al.*, 2012)作為參考依據。EGM2008 球諧展開最高可達 2190 階，本研究使用 360 階來計算長波長重力異常  $\Delta g_{long}$  與長波長大地起伏  $\Delta N_{long}$ ，作為 RCR 模型中「去除」階段之基準，有助於提升後續區域模型的精度與穩定性，僅採用 360 階主要考量其對應之空間尺度約為 50 - 60 km，符合長波長重力場的典型定義，亦為區域重力與大地水準面研究中常用之分界尺度(內政部，2025)。本研究重力異常網格如圖 5 所

示，包含既有的臺灣陸測重力、船測重力及衛星測高重力組合而成，包括臺灣本島重力異常網格(內政部，2025)、最新全球海洋重力異常模型(Sandwell *et al.*, 2021)、2006 年至 2013 年的臺灣近岸船測重力資料(Hwang *et al.*, 2014)等。由圖 5 可以發現臺灣與周邊海域重力異常介於 -300 mgal 至 +300 mgal 以上。

在用於檢核重力法大地起伏模型成果的 GNSS/水準測量檢核點位置如圖 3(a)所示。A 區域和 B 區域的 GNSS/水準測量檢核點分別為 11 點(圖 3(a)紅點加藍點)和 5 點(圖 3(a)藍點)，這些位置均

具有高精度的觀測大地起伏值。觀測大地起伏值指某一地點大地水準面相對於參考橢球面之垂直距離，為建立幾何法大地起伏模型及檢核模型精度的重要依據。本研究所採用 GNSS/水準資料，乃一等一級水準點上靜態觀測 12~24 小時 GNSS。故大地起伏觀測值精度非常高，且已經考量坐標框架變動與地殼形變速度場之影響(內政部，2025)。

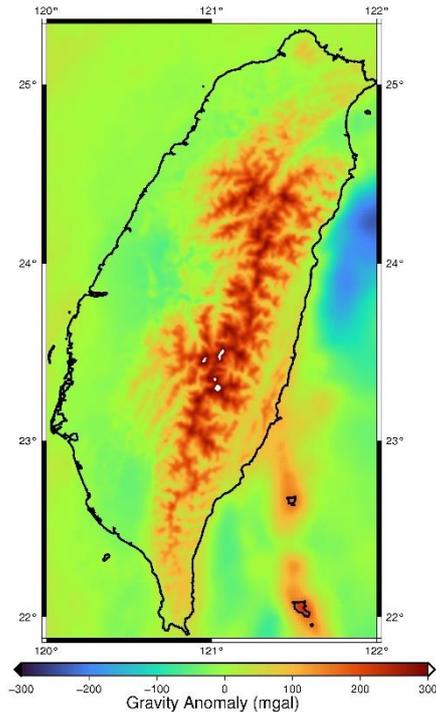
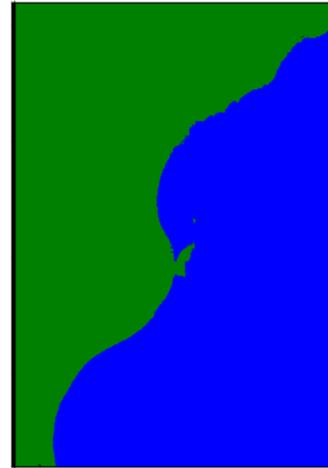


圖 5 臺灣重力異常模型

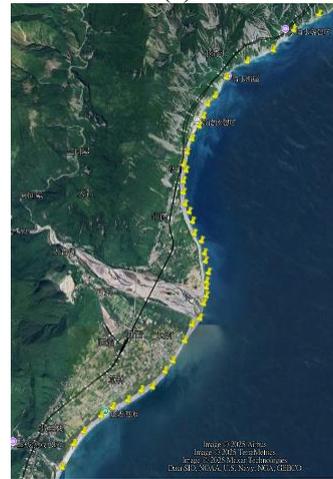
## 4. 實驗流程

### 4.1 海岸線邊界範圍

本研究大地起伏建模須整合 DEM 與 DBM，以進行後續大地起伏建模計算。為建立連續的陸海統一地形模型，必須先確認海岸線位置，本文採用的海岸線分成 GSHHG 海岸線和人工數化海岸線。GSHHG 海岸線方面，GMT 6.5 軟體具有 GSHHG 資料庫，故本研究是從 GMT 6.5 取得，圖 6(a)顯示 B 區域 GSHHG 海岸線位置；人工數化海岸線方面，是以人工方式在 google earth 衛星影像中點選實際海岸線位置，再進行數化處理，圖 6(b)顯示 B 區域中人工點選的位置。



(a)



(b)

圖 6 (a) B 區域中 GSHHG 海岸線，其中綠色為陸地、藍色為海洋，綠藍色交界處為 GSHHG 海岸線；(b) B 區域中 google earth 衛星影像中點選實際海岸線的範圍。黃色點為手動點選真實海岸線的位置

### 4.2 數值地形模型

本研究蒐集內政部陸地 20 m DEM 與海洋 50 m DBM，接著使用 GMT6.5.0 中的程式 `grdlandmask` 製作出陸地和海洋的 0/1 檔地形，圖 7 顯示分別採用 GSHHG 海岸線，製作陸地數值為 1 海洋為 0(圖 7(a))、和陸地數值為 0 海洋為 1(圖 7 (b))的全台地形。接著分別將兩個 0/1 檔地形與陸地 DEM 及海洋 DBM 相乘，結果如圖 7(c)、(d)所示。圖 7(c)為陸地地形為 DEM 值，海洋地形為 0，而圖 7(d)海洋地形為 DBM 值，陸地地形則為 0。接著將新計算之陸地模型和海洋模型合併融合，其結果如圖 7(e)所示。圖 7(e)代表以 GSHHG 海岸線融合陸地 20 m DEM 與海洋 50 m DBM 的全台地形成果。本研究

依據使用之海岸線資料類型(GSHHG 或數化海岸線)、融合後地形的空間解析度、不同研究區域等，分為 8 個大地起伏模型計算的 Case，如表 1 所示。Case 1 到 4 屬於 A 區域，DEM-DBM 解析度有 9"x

9"與 3"x3"兩種；Case 5 到 8 屬於 B 區域，DEM-DBM 解析度有 3"x3"與 1"x1"兩種。9"x9"、3"x3"和 1"x1"對應到的空間解析度分別為 270 m、90 m 與 30 m。

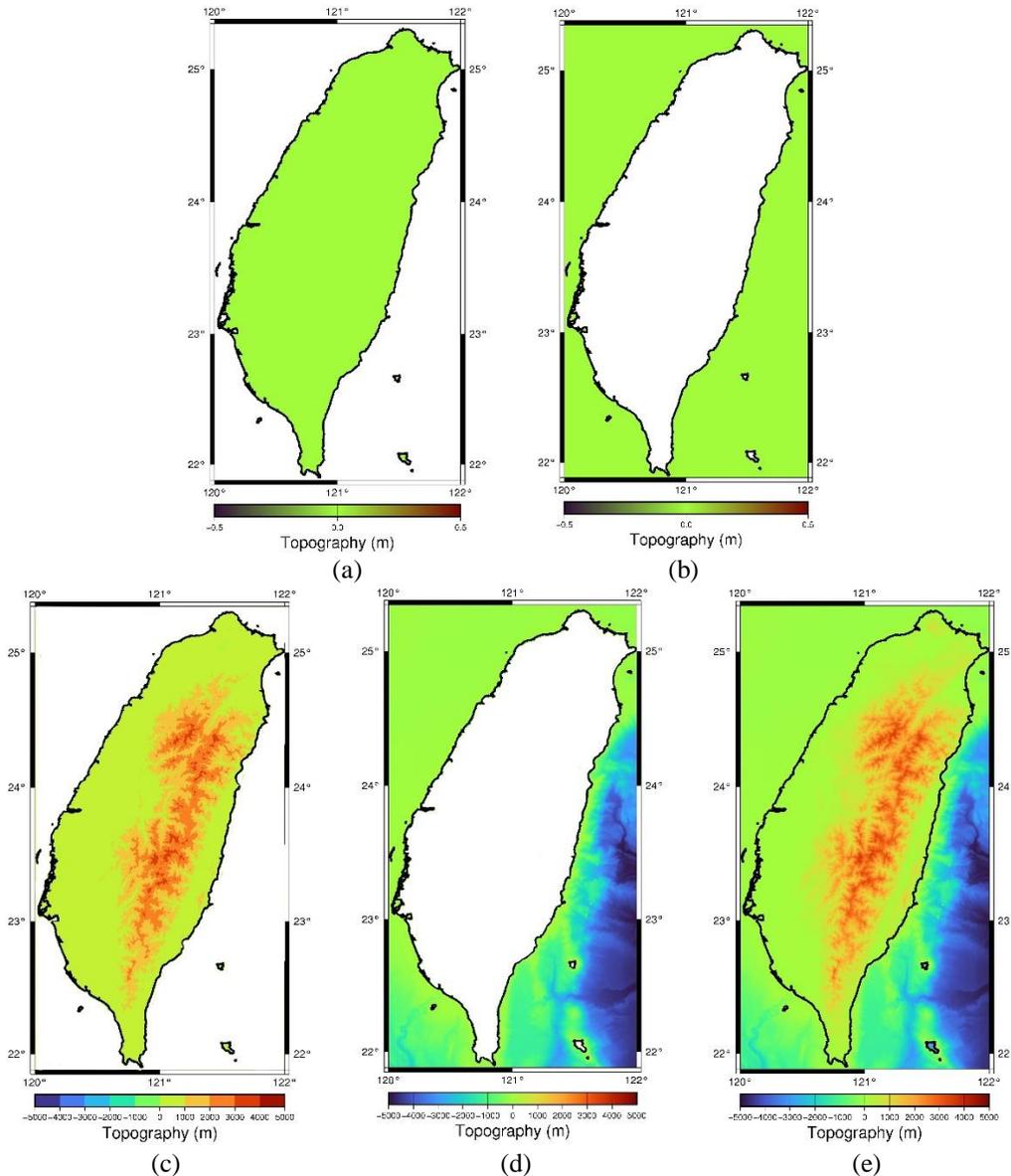


圖 7 (a)陸地數值為 1 海洋為 0 之地形；(b)陸地數值為 0 海洋為 1 之地形；(c)陸地為 20m DEM，海洋地形為 0 之地形；(d)海洋為 50m DBM，陸地地形為 0 之地形；(e) 陸地和海洋地形融合後的地形

表 1 研究測試例(Case)分類

Case	研究區域	DEM-DBM 空間解析度	海岸線
1	A 區域	270 m (9"x9")DEM-DBM	GSHHG
2	A 區域	270 m (9"x9")DEM-DBM	數化海岸線
3	A 區域	90 m (3"x3")DEM-DBM	GSHHG
4	A 區域	90 m (3"x3")DEM-DBM	數化海岸線
5	B 區域	90 m (3"x3")DEM-50 m DBM	GSHHG
6	B 區域	90 m (3"x3")DEM-50 m DBM	數化海岸線
7	B 區域	30 m (1"x1")DEM-DBM	GSHHG
8	B 區域	30 m (1"x1")DEM-DBM	數化海岸線

### 4.3 大地起伏模型

本研究採用 RCR 架構處理長波長、中波長與短波長大地起伏計算成果，並將三者合併建立完整的大地起伏模型。首先以全球重力場模型 EGM2008 進行球譜展開，計算出研究區域之長波長重力異常  $\Delta G_{long}$  與長波長大地起伏  $\Delta N_{long}$ 。接著使用 RTM 方法搭配不同解析度之 DEM(9"x9"、3"x3"、1"x1")計算短波長重力異常  $\Delta G_{short}$  與短波長大地起伏貢獻  $\Delta N_{short}$ ，有效補償地形變化劇烈區之高頻訊號。中波長部分則利用以 LSC 進行計算，將已去除長波與短波效應後之重力異常資料，建立協方差矩陣，推估中波長大地起伏  $\Delta N_{res}$ 。三組波長最終進行網格加總，建立完整的似大地起伏模型。最後藉由似大地起伏改正得到最終大地起伏模型。圖 8 為利用此方法計算得到的全臺灣與周邊海域的重力法大地起伏模型(空間解析度 30"x30")。本文的 8 個 Case 的重力法大地起伏計算，皆採用此步驟計算完成。

## 5. 結果與討論

### 5.1 海陸交接處精度分析

圖 9 顯示 B 區域 GSHHG 海岸線與人工數化海岸線(由 Google Earth 高解析影像數化而得)之差異。比較結果顯示，GSHHG 海岸線在臺灣地區，尤其於東部花蓮清水斷崖等地形起伏劇烈區域，與實際海岸線位置存在顯著偏差。與人工數化海岸線相比，其最大平面誤差可達約 700 m。此一偏差主要源自 GSHHG 海岸線資料之空間解析度限制，使其難以精確描述高起伏、近乎垂直之斷崖型海岸地形；此外，在地形變化快速的區域，海陸交界線本身即具有不連續與不穩定特性，進一步降低海岸線定位的準確性。以 A 區域 9"x9" DEM-DBM 為例，圖 10 呈現分別採用 GSHHG 海岸線與人工數化海岸線進行 DEM-DBM 融合後之差值分布情形。結果顯示，海岸線鄰近區域的 DEM 高程誤差最為明顯，最大差值約達  $\pm 100$  m，且誤差主要集中於陡峭斷崖與海陸轉換帶，顯示海岸線定位精度對 DEM-

DBM 融合成果具有關鍵影響。

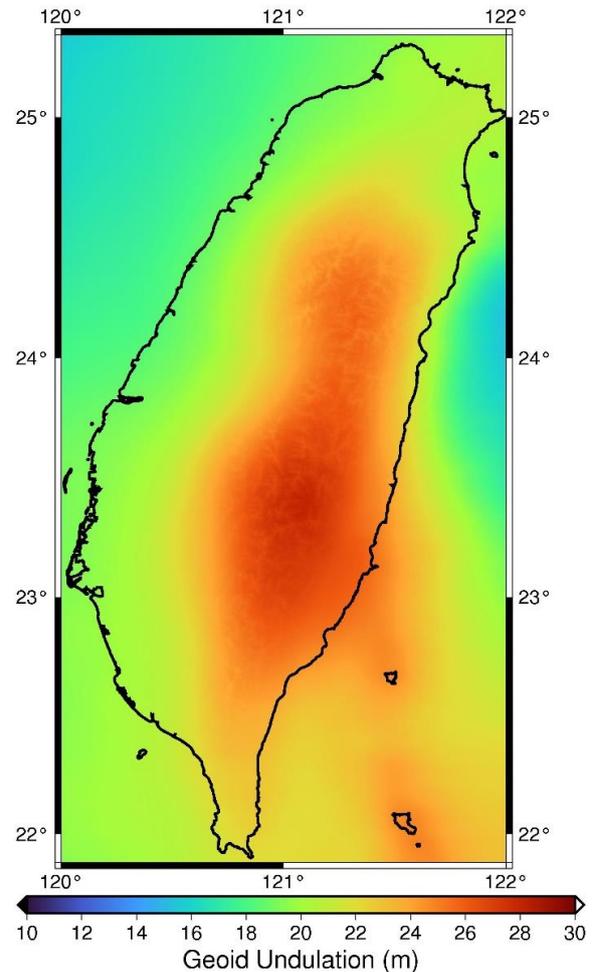


圖 8 重力法大地起伏模型

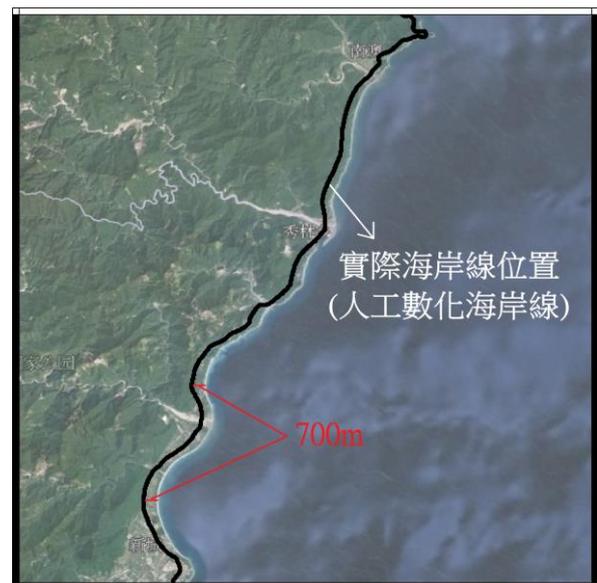


圖 9 B 區域 Google Earth 衛星影像圖。黑色線為 GSHHG 海岸線，海陸交界處即為實際海岸線(人工數化海岸線)位置

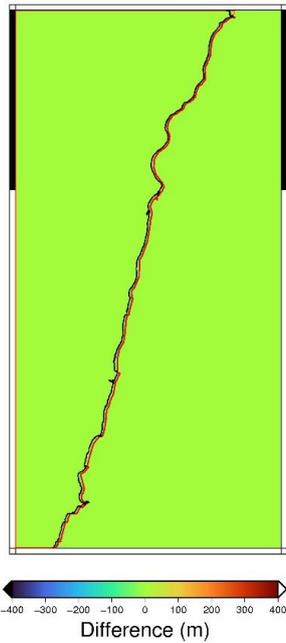


圖 10 GSHHG 海岸線與人工數化海岸線進行 DEM-DBM 融合後之差值

### 5.2 A 區位(Case 1~4)

圖 11 呈現 Case 1 與 Case 2 之相關成果。由圖 11(a)可觀察到，採用 GSHHG 與人工數化海岸線所產製之 DEM-DBM，在大地起伏模型計算結果上的差異主要集中於海岸線鄰近區域，顯示海陸邊界定位精度對 DEM-DBM 融合成果具有顯著影響，其中在地形起伏劇烈之清水斷崖一帶，兩者差異可超過 5 cm。此現象反映出在陡峭斷崖與快速地形轉換區域，海岸線微小的平面位移即可能導致 DEM 與 DBM 高程配置明顯不同，進而影響後續重力與大地起伏計算結果。圖 11(b)與圖 11(c)分別為 Case 1 與 Case 2 之大地起伏模型與 GNSS/水準測量成果之檢核結果，兩圖皆以重力法大地起伏模型作為底圖，並以柱狀長條圖呈現各檢核點重力法大地起伏與觀測大地起伏值(GNSS/水準測量)之差值絕對值。由 11 個檢核點之統計結果可知，兩案例在標準偏差表現上相當接近，顯示整體模型穩定性相似，惟 Case 2 之標準偏差略小於 Case 1，顯示在相同空間解析度條件下，採用人工數化海岸線仍可帶來輕微但一致的精度提升。在 Case 3 與 Case 4

之比較結果(圖 12)中，可發現大地起伏模型之最大差距同樣集中於海陸交界且地形變化劇烈之區域，最大差異亦接近 5 cm，顯示海岸線定位誤差在不同解析度條件下皆會對局部大地起伏成果造成影響。由圖 12(b)與圖 12(c)之 GNSS/水準測量檢核結果可知，Case 3 與 Case 4 之標準偏差表現亦相近，但 Case 4 整體仍略優於 Case 3，呈現與 Case 1、Case 2 相同的趨勢。所有 Case 1 至 Case 4 之檢核成果彙整於表 2。

整體而言，無論於 9"x9" 或 3"x3" 空間解析度條件下，使用人工數化海岸線所建立之 DEM-DBM，其大地起伏計算成果均略優於使用 GSHHG 海岸線者，整體改善幅度約為 2 - 3 mm。進一步比較 Case 1 - 2 與 Case 3 - 4 可發現，採用空間解析度 3"x3" 之 Case 3 - 4，其成果明顯優於採用 9"x9" 解析度之 Case 1 - 2，標準偏差由約 0.32 m 降至約 0.19 m。由於 9"x9" 格網單元約對應 270 m 的空間解析度，於海陸邊界及地形變化複雜區域時，地形與海岸線細部特徵易在重取樣與平均化過程中被過度平滑，進而削弱 DEM-DBM 對真實地形變化的描述能力，導致大地起伏計算成果精度降低。

另外值得注意的是，人工數化海岸線於大地起伏模型中的改善效果，主要應集中於海陸交界且地形變化劇烈之區域；然而，本研究所採用之 GNSS/水準測量檢核點位置，並未剛好落在上述大地起伏差值較大的區域，因此即使人工數化海岸線能有效改善海岸線鄰近區域之 DEM-DBM 表現，其優勢在整體統計指標(如標準偏差)上仍不易被充分反映，導致 Case 2 與 Case 4 相較於 Case 1 與 Case 3 的精度提升幅度僅呈現為數毫米(mm)等級。

表 2 GNSS/水準測量點位檢核成果(Case 1~4)

統計值	最大值	最小值	平均值	標準偏差
Case	(m)	(m)	(m)	(m)
1	0.921	-0.031	0.329	0.317
2	0.908	-0.033	0.328	0.315
3	0.424	-0.304	0.190	0.164
4	0.425	-0.033	0.191	0.161

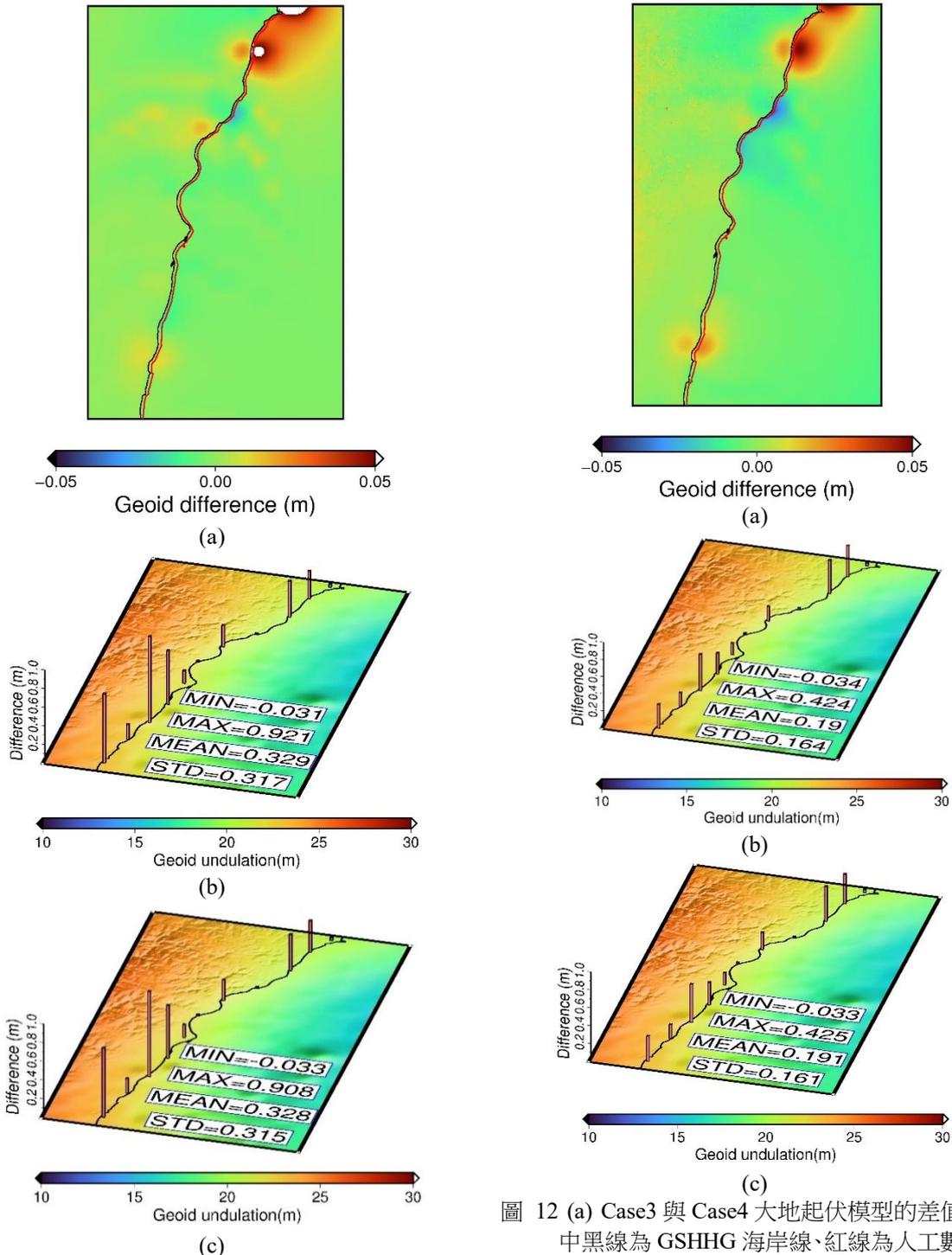


圖 11 (a) Case1 與 Case2 大地起伏模型的差值，其中黑線為 GSHHG 海岸線、紅線為人工數化海岸線；(b) Case 1 大地起伏模型與 GNSS/水準測量點位檢核成果；(c) Case 2 大地起伏模型與 GNSS/水準測量點位檢核成果。圖(b)與(c)柱狀長條圖顯示差值絕對值

圖 12 (a) Case3 與 Case4 大地起伏模型的差值，其中黑線為 GSHHG 海岸線、紅線為人工數化海岸線；(b) Case3 大地起伏模型與 GNSS/水準測量點位檢核成果；(c) Case 4 大地起伏模型與 GNSS/水準測量點位檢核成果。圖(b)與(c)柱狀長條圖顯示差值絕對值

### 5.3 B 區位(Case 5~8)

為進一步探討地形的空間解析度與海岸線精度對大地起伏模型在海陸交界處的影響，本研究於面積較小的 B 區域進行高解析度建模分析。該區域

聚焦於清水斷崖地形劇烈變化帶，分別測試空間解析度 3"x3"與 1"x1"解析度之 DEM-DBM 組合，搭配 GSHHG 與人工數化海岸線進行比較。圖 13 為 Case 5 與 Case 6 的相關成果，由圖 13(a)可發現 GSHHG 和人工數化海岸線所產製之 DEM-DBM，在計算大地起伏模型上，差異主要集中於海岸線周圍，其中清水斷崖區差異達將近 5 cm。圖 13(b)和 13(c)為 Case 5 和 Case 6 的大地起伏模型和 GNSS/水準測量檢核成果，底圖均為大地起伏模型計算成果，圖中柱狀長條圖顯示重力法大地起伏模型和觀測大地起伏值(GNSS/水準測量)之間的差值絕對值。由圖 13(b)和 13(c)可發現，若由 5 個檢核點的統計標準偏差來看，成果相近，但 Case 6 略優於 Case 5。

在 Case 7 和 Case 8 的比較方面(圖 14)，可發現大地起伏模型最大差距依然在海陸交界地形變化劇烈處，差距最大亦將近 5 cm。若由圖 14(b)和 14(c) GNSS/水準測量檢核成果標準偏差來看，Case 7 和 Case 8 成果也相近，但 Case 8 還是略優於 Case 7。Case 5 到 8 的檢核成果整理於表 3。整體而言，使用人工數化海岸線 DEM-DBM 的大地起伏計算成果，還是略優於使用 GSHHG 海岸線 DEM-DBM 的大地起伏計算成果，大約改善 2~3 mm。另外若比較 Case 5~6 和 Case 7~8，可看出採用空間解析度 3"x3"的 Case 5~6，成果明顯優於採用空間解析度 1"x1"的 Case 7~8，標準偏差從約 0.28 m 降到約 0.15 m。儘管高解析度 DEM 能更精細描述地形細部，但其所包含的大量高頻地形訊號及量測誤差，並不完全對應於真實的重力場變化。當使用 1"x1" DEM-DBM 進行大地起伏計算時 (Case 7~8)，高頻地形雜訊在 RTM 積分過程中被放大，導致計算結果不穩定且精度降低。相較之下，3"x3" DEM 透過自然的空間平均效果，有效抑制高頻雜訊，使得大地起伏計算成果顯著改善。

如同 A 區域，B 區域顯示在 Case 6 與 Case 8 中，人工數化海岸線對大地起伏成果所帶來的整體精度提升幅度並不顯著，其原因並非海岸線改進本身效果有限，而是受限於 GNSS/水準測量檢核資料的空間配置。B 區域內所使用的 5 個檢核點位置，與大地起伏差異最為集中的清水斷崖沿岸區域並

未完全重疊，導致模型於海陸交界處所產生的局部改善，無法在統計評估指標中充分展現。因而在標準偏差比較結果上，Case 6 與 Case 8 相較於 Case 5 與 Case 7 僅呈現有限的數值差異。

表 3 GNSS/水準測量點位檢核成果(Case 5~8)

統計值	最大值	最小值	平均值	標準偏差
Case	(m)	(m)	(m)	(m)
5	0.390	0.016	0.240	0.159
6	0.389	0.018	0.243	0.155
7	1.730	1.075	1.425	0.284
8	1.729	1.073	1.430	0.275

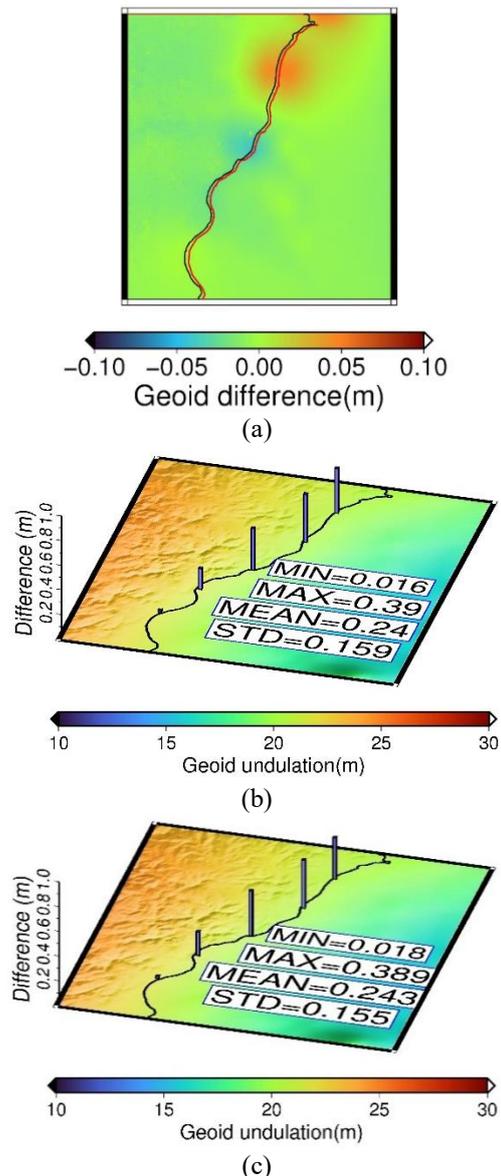


圖 13 (a) Case 5 與 Case 6 大地起伏模型的差值，其中黑線為 GSHHG 海岸線、紅線為人工數化海岸線；(b) Case 5 大地起伏模型與 GNSS/水準測量點位檢核成果；(c) Case 6 大地起伏模型與 GNSS/水準測量點位檢核成果。圖(b)與(c)柱狀長條圖顯示差值絕對值

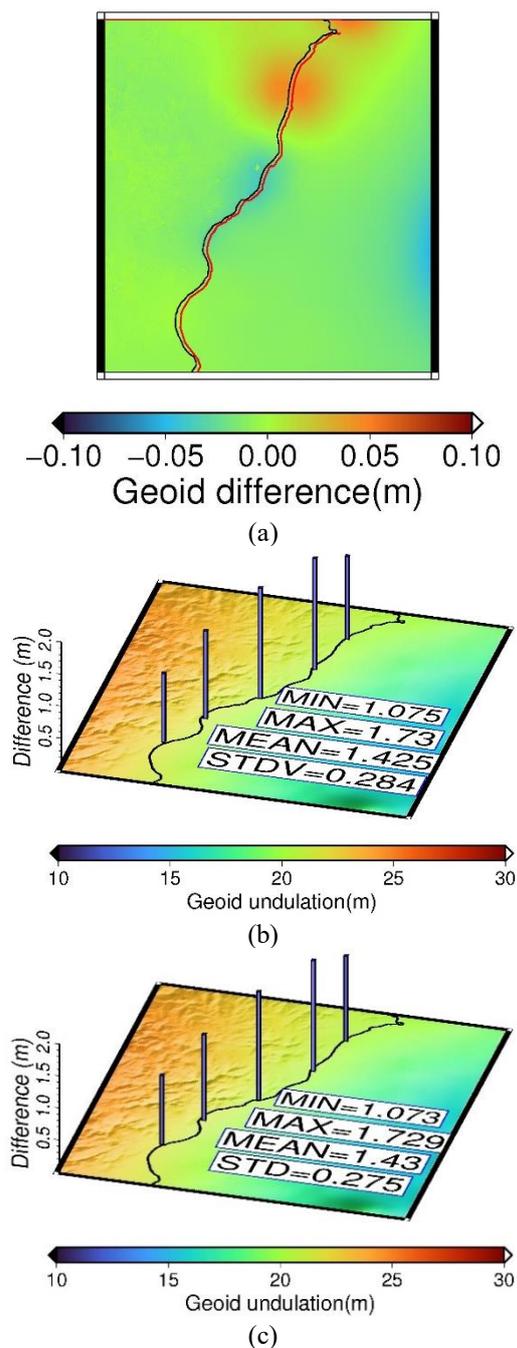


圖 14 (a) Case 7 與 Case 8 大地起伏模型的差值，其中黑線為 GSHHG 海岸線、紅線為人工數化海岸線；(b) Case 7 大地起伏模型與 GNSS/水準測量點位檢核成果；(c) Case 8 大地起伏模型與 GNSS/水準測量點位檢核成果。圖(b)與(c)柱狀長條圖顯示差值絕對值

## 6. 結論與建議

本研究以臺灣東部花蓮沿岸為研究對象，聚焦於地形起伏極為劇烈的清水斷崖一帶，探討不同 DEM、DBM 空間解析度以及海岸線定位精度，對

陸海交界處大地起伏建模成果所造成的影響。研究過程中結合 RCR 架構下之 RTM 與 LSC 方法進行大地起伏建模，並以 GNSS/水準測量資料作為外部檢核依據，使得不同資料組合對模型精度與穩定性的影響得以具體量化與比較。分析結果顯示，海岸線定位誤差在陸海交界區域中扮演相當關鍵的角色，對大地起伏模型的局部表現具有直接影響。以全球通用的 GSHHG 海岸線資料而言，其在地形相對平緩的區域仍能維持良好的一致性；然而在清水斷崖等屬於斷崖型海岸的地區，與實際海岸線之間仍存在數百公尺等級的平面差異。此類偏差會在 DEM 與 DBM 融合階段反映為邊界位置的不確定性，進而影響地形遮罩的建立，使地形改正誤差集中於海岸線附近，並在大地起伏模型中形成明顯的局部差異。相較之下，透過高解析度衛星影像進行人工數化的海岸線，能較貼近實際海陸邊界位置，使 DEM-DBM 融合後的地形連續性較為合理，並在各測試案例中呈現出穩定且一致的精度改善趨勢。在地形資料的空間解析度方面，研究結果顯示，提高 DEM-DBM 的解析度確實有助於改善地形起伏劇烈區域的大地起伏建模成果，但其改善效果並非隨解析度提升而持續增加。由 A 區域與 B 區域的比較可知，3"×3"解析度在模型穩定性與 GNSS/水準檢核結果上，整體表現明顯優於 9"×9"解析度，顯示解析度過低時，海陸交界處的地形細節容易在平滑化過程中被削弱。然而，當解析度進一步提高至 1"×1"時，模型精度反而出現下降，其主要原因在於高解析度 DEM 所包含的大量高頻地形訊號與量測雜訊，未必能對應實際的重力場變化，並在 RTM 積分過程中被放大，導致大地起伏計算結果變得不穩定。此一現象顯示，在地形劇烈變化的陸海交界區域中，解析度的選擇需兼顧訊號表現與雜訊控制，而非單純追求更細的網格式度。此外，本研究亦觀察到，人工數化海岸線所帶來的改善效果，主要集中於海陸交界及斷崖等地形快速轉換的局部區域；然而，由於 GNSS/水準測量檢核點多分布於內陸或地形相對平緩的區域，並未完全涵蓋大地起伏差異最為顯著的沿岸帶，使得岸線精化在整體統計指標(如標準偏差)上的反映僅呈現毫米等級差

異。這也說明，檢核點的空間配置對模型評估結果具有實質影響，若僅依賴整體統計量進行判斷，可能低估模型在關鍵地形區域中的實際改善幅度。

綜合上述成果，本研究建議在未來進行陸海交界大地起伏建模時，應優先確保海岸線定位的幾何正確性，並於地形劇烈變化區域適度採用人工或半自動化方式精化海岸線資料。在地形資料解析度選擇上，建議採用約 3"x3"等級之 DEM-DBM，以兼顧地形細節表現與模型穩定性，避免因過高解析度導致高頻雜訊主導計算結果。整體而言，本研究成果可作為臺灣及其他山坡地形起伏海岸區域進行精密大地水準面建模之重要參考。

## 參考文獻

- 內政部，2023。垂直基準面時變精度檢核評估工作案。[Ministry of the Interior, 2023. Project for Accuracy Assessment and Verification of Time-Variant Vertical Datum. (in Chinese)]
- 內政部，2025。大地起伏模型精進維護工作案。[Ministry of the Interior, 2025. Project for Geoid Model Improvement and Maintenance. (in Chinese)]
- 崔國強，1995。臺灣地區數值大地水準面模型計算之研究，國立中央大學地球物理學系博士論文。[Tsuei, G.C., 1995. The investigation of digital geoid model over Taiwan, PhD Dissertation, National Central University, Taiwan, ROC. (in Chinese)]
- 許宏銳，2014。最佳化重力資料處理及整合以建構公分級臺灣大地起伏模式，國立交通大學土木工程學系博士論文。[Hsu, H.-J., 2014. Optimizing gravity data processing and combination for cm-geoid modeling in Taiwan, PhD Dissertation, National Chiao Tung University, Taiwan, ROC. (in Chinese)]
- Bucha, B., Hirt, C., Yang, M., Kuhn, M., and Rexer, M., 2019. Residual terrain modelling (RTM) in terms of the cap-modified spectral technique: RTM from a new perspective, *Journal of Geodesy*, 93: 2089-2108, DOI: 10.1007/s00190-019-01303-4.
- Forsberg, R., 1984. A study of terrain reductions, density anomalies and geophysical inversion methods in gravity field modelling, PhD Dissertation, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, DOI: 10.21236/ADA150788.
- Heiskanen, W.A., and Moritz, H., 1967. *Physical Geodesy*, W.H. Fredman and company, San Francisco, CA, USA.
- Hwang, C., 1997. Analysis of some systematic errors affecting altimeter-derived sea surface gradient with application to geoid determination over Taiwan, *Journal of Geodesy*, 71: 113-130, DOI: 10.1007/s001900050080.
- Hwang, C., and Hsiao, Y.S., 2003. Orthometric corrections from leveling, gravity, density and elevation data: a case study in Taiwan, *Journal of Geodesy*, 77: 279-291, DOI: 10.1007/s00190-003-0325-6.
- Hwang, C., Ellmann, A., and Hsiao, Y.S., 2006. Toward a one-cm geoid in Taiwan: Contribution of airborne gravity data and high-resolution DEM, in *Proceedings of the International Workshop on Height Systems, Geoid and Gravity of the Asia-Pacific*, Ulaanbaatar, Mongolia.
- Hwang, C., Hsu, H.J., Chang, E.T., Featherstone, W.E., Tenzer, R., Lien, T., Hsiao, Y.S., Shih, H.C., and Jai, P.H., 2014. New free-air and Bouguer gravity fields of Taiwan from multiple platforms and sensors, *Tectonophysics*, 611: 83-93, DOI: 10.1016/j.tecto.2013.11.027.
- Hwang, C., Hsu, H.J., Featherstone, W.E., Cheng, C.C., Yang, M., Huang, W.H., Wang, C.Y., Huang, J.F., Chen, K.H., Huang, C.H., Chen, H.C., and Su, W.Y., 2020. New gravimetric-only and hybrid geoid models of Taiwan for height modernisation, cross-island datum connection and airborne

- LiDAR mapping, *Journal of Geodesy*, 94: 83, DOI: 10.1007/s00190-020-01412-5.
- Jiang, T., Dang, Y., and Zhang, C., 2020. Gravimetric geoid modeling from the combination of satellite gravity model, terrestrial and airborne gravity data: A case study in the mountainous area, Colorado, *Earth Planets Space*, 72: 189, DOI: 10.1186/s40623-020-01287-y.
- Kuhn, M., Featherstone, W.E., and Kirby, J.F., 2009. Complete spherical Bouguer gravity anomalies over Australia, *Australian Journal of Earth Sciences*, 56(2): 213-223, DOI: 10.1080/08120090802547041.
- Omang, O.C.D., and Forsberg, R., 2000. How to handle topography in practical geoid determination: Three examples, *Journal of Geodesy*, 74: 458-466, DOI: 10.1007/s001900000107.
- Pavlis, N.K., Holmes, S.A., Kenyon, S.C., and Factor, J.K., 2012. The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), *Journal of Geophysical Research: Solid Earth*, 117(B4): B04406, DOI: 10.1029/2011JB008916.
- Sandwell, D.T., Harper, H., Tozer, B., and Smith, W.H.F., 2021. Gravity field recovery from geodetic altimeter missions, *Advances in Space Research*, 68(2): 1059-1072, DOI: 10.1016/j.asr.2019.09.011.
- Tscherning, C.C., and Rapp, R.H., 1974. Closed covariance expressions for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models, *Reports of the Department of Geodetic Science*, No.208, Ohio State University, Columbus, OH, USA.
- Vu, D.T., Bruinsma, S., and Bonvalot, S., 2019. A high-resolution gravimetric quasigeoid model for Vietnam, *Earth Planets Space*, 71: 65, DOI: 10.1186/s40623-019-1045-3.
- Wessel, P., Luis, J.F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W.H.F., and Tian, D., 2019. The generic mapping tools version 6, *Geochemistry, Geophysics, Geosystems*, 20(11): 5556-5564, DOI: 10.1029/2019GC008515.
- Willberg, M., Zingerle, P., and Pail, R., 2019. Residual least-squares collocation: Use of covariance matrices from high-resolution global geopotential models, *Journal of Geodesy*, 93: 1739-1757, DOI: 10.1007/s00190-019-01279-1.
- Yildiz, H., 2012. A study of regional gravity field recovery from GOCE vertical gravity gradient data in the Auvergne test area using collocation, *Studia Geophysica et Geodaetica*, 56: 171-184, DOI: 10.1007/s11200-011-9030-8.

# Accuracy Assessment of Geoid Modeling at the Land-Sea Interface: A Case Study of Eastern Taiwan

Zhi-Yi Yang<sup>1</sup>      Yu-Shen Hsiao<sup>2\*</sup>

## Abstract

This study investigates the effects of Digital Elevation Model (DEM) and Digital Bathymetric Model (DBM) resolutions, as well as coastline positioning accuracy, on gravimetric geoid modeling across the land–sea interface along the Qingshui Cliff coast in eastern Taiwan. Two coastline datasets, namely the Global Self-consistent, Hierarchical, High-resolution Geography (GSHHG) database and manually digitized coastlines, were combined with DEM-DBM datasets of different spatial resolutions. Under the remove–compute–restore framework, gravimetric geoid models were constructed using the Residual Terrain Model and Least-Squares Collocation methods, and their accuracy was evaluated using GNSS/leveling-derived geoid heights. The results show that the GSHHG coastline deviates noticeably from the actual shoreline in cliff-type coastal areas, leading to terrain masking and correction errors near the coast. In contrast, manually digitized coastlines improve boundary consistency and model stability. Resolution comparisons further indicate that the 3"×3" DEM-DBM provides better overall accuracy than the 9"×9" and 1"×1" models. These findings suggest that accurate coastline definition and appropriate terrain resolution are essential for improving geoid modeling in complex land–sea transition zones.

**Keywords: DEM, DBM, Geoid Modeling, Land–Sea Transition Zone**

---

<sup>1</sup> Master, Department of Soil and Water Conservation, National Chung Hsing University

<sup>2</sup> Associate Professor, Department of Soil and Water Conservation, National Chung Hsing University

\* Corresponding Author, E-mail: yshsiao@nchu.edu.tw

Received Date: Jul. 30, 2025

Revised Date: Sep. 03, 2025

Accepted Date: Feb. 16, 2026

# 使用深度學習進行快速無參考標的無人機影像 品質評估

林亞立<sup>1\*</sup> 蘇冠秦<sup>1</sup> 鄒來翰<sup>1</sup> 林昭宏<sup>2</sup> 饒見有<sup>2</sup>  
賴威伸<sup>3</sup> 胡智超<sup>3</sup>

## 摘要

隨著無人飛行載具(UAV)應用於基礎設施監測,影像品質穩定性成為影響深度學習與攝影測量精度的關鍵。然而 UAV 影像常受環境干擾影響,現行品質篩選多仰賴人工檢視,且傳統結構相似度指標(SSIM)需參考影像並受限於空間對齊,難以實務應用。為此,本研究提出一套基於 Swin-Unet 的快速無參考影像品質評估方法。首先設計改良型 CLIP-SSIM 結合 Swin-Transformer,建立高精度影像品質圖(RMSE = 0.0193),再以該品質圖作為標註資料訓練 Swin-Unet 模型,使單張影像推論時間降至 0.3 秒,並維持良好準確度(RMSE = 0.04)。結果顯示,本方法可有效取代人工檢視流程,滿足高頻 UAV 影像應用需求。

**關鍵詞：**影像品質評估、深度學習、無人機影像、影像結構相似度指標

## 1. 前言

無人機因具備高效率、低成本、高機動性與低風險等優勢,已廣泛應用於基礎建設監測、地形測繪、災後勘查與環境監控等多元任務(Rakha & Gorodetsky, 2018)。尤其在橋梁、道路、水壩與電塔等難以親近或具潛在危險性的設施上,無人機可大幅降低作業人力需求與風險負擔,提升檢測作業的時效性與空間覆蓋率。伴隨影像感測與控制技術之進步,無人機得以快速擷取高解析度圖像,為智慧基礎設施管理提供豐富的視覺資料來源。然而,無人機所拍攝之影像品質常受到外部環境因素干擾,例如強風造成飛行震動、光照變化導致曝光不均,以及拍攝角度變化引發畫面偏移等,皆可能造成影像模糊、變形與對比不足,進而影響後續結構健康監測、三維重建與劣化偵測等任務之精度與可靠性(Sieberth *et al.*, 2015)。特別是在進行攝影測量與深

度學習推論等任務時,影像品質的微小變異即可能對成果產生顯著影響。因此,建立一套客觀、快速且可自動化執行之影像品質評估(Image Quality Assessment, IQA)機制,作為影像前處理與篩選之依據,已成為實務應用中亟需解決的重要課題。

傳統影像品質評估方法如峰值訊雜比(Peak Signal-to-Noise Ratio, PSNR)(Horé & Ziou, 2010)與結構相似性指數(Structural Similarity Index Measure, SSIM)(Wang *et al.*, 2004)雖被廣泛應用於影像品質評估研究中,但相關文獻亦指出其在實務應用上仍存在明顯限制。PSNR 主要基於像素強度差異進行計算,對影像結構變化與人眼主觀感知之關聯性有限,已被證實在多種影像失真情境下與人類視覺評價之相關性不足。相較之下,SSIM 雖進一步考量亮度、對比與結構資訊,在模糊與幾何失真評估上具備較佳表現,然而該方法仍需仰賴參考影像進行比較,且對影像尺寸、裁切位置與像素對齊條件高

<sup>1</sup> 國立成功大學測量及空間資訊學系 博士生

<sup>2</sup> 國立成功大學測量及空間資訊學系 教授

<sup>3</sup> 交通部運輸研究所運輸工程組 研究員

\* 通訊作者, E-mail: alecfree2@gmail.com

收到日期: 民國 114 年 11 月 19 日

修改日期: 民國 114 年 12 月 18 日

接受日期: 民國 115 年 01 月 30 日

度敏感。此一特性使得 SSIM 難以直接應用於無人機巡檢等缺乏對應參考影像，且拍攝視角與尺度變化頻繁之實務情境，進而限制其實際可行性。為解決此問題，近年來興起的深度學習技術提供了無需參考影像的評估方案。透過卷積神經網路 (Convolutional Neural Network, CNN)、殘差網路 (Residual Network, ResNet)(Kang *et al.*, 2014、Ma *et al.*, 2017)，以及具備全局特徵建模能力之 Transformer 架構(Vaswani *et al.*, 2017)，影像品質評估得以轉化為具備語意理解能力的分類或回歸任務，具備更高的判別力與自適應能力。然而，深度學習模型之訓練效果高度仰賴大量且高品質之標註資料，而無人機影像常處於環境多變、光線與構圖條件差異大的情境下，造成標註資料取得困難，進而限制模型於實務領域的泛化能力與應用潛力。

為解決上述挑戰，本研究提出一個二階段影像品質評估架構。首先以基於 Swin-Transformer (Liu *et al.*, 2021)的機率加權模型產生高精度影像品質圖，作為訓練標註；再透過 Swin-Unet 架構(Cao *et al.*, 2021)，進行全影像推論取代逐像素預測，兼顧推論速度與準確性，使本架構能有效支援無人機快速影像品質評估任務，具有實務應用潛力。本研究聚焦於橋梁檢測任務中無人機影像品質評估的實務挑戰。橋梁檢測作業往往需於高風速、背光或遮蔽等複雜環境下執行飛行任務，致使所拍攝影像品質參差不齊。由於橋梁結構細節眾多，裂縫、剝落等劣化特徵常需仰賴高品質影像才能準確辨識，影像品質的不穩定將直接影響後續深度學習模型之偵測準確性與三維建模結果的可靠性。因此，本研究旨在發展一套具效率、無需參考影像且適用於橋梁檢測環境的自動化影像品質評估架構，期能有效支援橋梁劣化檢測與結構安全監測工作，提升無人機應用於基礎建設巡檢作業的整體效能。

## 2. 研究方法

人機在每次任務中往往會拍攝大量影像，尤其是在橋梁巡檢任務中，為求完整覆蓋橋體結構，通常會以多角度、多視角方式進行連續拍攝，動輒產

生數百至上千張高解析度影像。在此情境下，若仍仰賴傳統人工方式逐張進行影像品質評估與篩選，不僅耗費大量人力與時間，更無法滿足效率與高頻率的實務需求，進而可能延誤檢測進度或降低整體資料處理效率。為解決此問題，本研究建立一套自動化的影像品質評估流程，使用實地無人機拍攝的橋梁影像資料集，並訓練三種不同目的之模型以應對實務需求，如圖 1 所示。

首先，從資料集中選取 600 張高品質影像進行直方圖均化 (histogram equalization) (Pizer *et al.*, 1987)，統一亮度與對比度，並將這些高品質影像評分為 1.0。接著，以演算法對這些高品質影像製作各種程度的退化影像，用以模擬真實無人機環境中的干擾條件，如：亮度變化、對比變化、高斯模糊、水平與垂直移動模糊。接著，則透過 CSSIM 指標對退化影像進行評分，再以 CSSIM 指標預測分數作為標註，並取對應的退化影像作為基於 Swin-Transformer 架構之機率加權模型之訓練資料。該模型會針對每像素周圍 500x500 區域進行分析，逐像素偵測 CSSIM 數值，生成高解析度的影像品質圖，成功實現無參考影像品質評估。然而，因該模型採逐像素推論方式，計算時間成本過高，不利於即時應用。為提升實用性，本研究進一步以該模型所生成的影像品質圖作為標註影像，訓練一個基於 Swin-Unet 架構之模型，讓模型可以直接推論整張影像的影像品質圖，大幅提升推論速度與實務應用彈性。

### 2.1 CSSIM 影像品質指標

SSIM 為傳統常用之客觀影像品質指標，主要根據亮度、對比與結構三項元素進行評分，其中結構分數依賴區域內兩張影像的共變異數計算。因此，SSIM 需在影像對齊的前提下進行評估，對深度學習模型因影像裁切所導致的像素錯位極為敏感，進而影響評估準確性。為解決此問題，本研究提出 CSSIM 指標，將原本依賴兩影像共變異數之結構分數，替換為透過 CLIP 模型(Radford *et al.*, 2021)之圖像編碼器(image encoder)所提取之高階特徵向量間的餘弦相似度，如式(1)所示。於本研究中，CLIP

僅作為影像特徵編碼工具使用，而非重新訓練之深度學習模型，其可將輸入之影像轉換為固定維度之特徵向量表示。由於影像在輸入編碼器前會經由模型內部之標準化處理，不同尺寸之影像皆可被轉換為相同維度之特徵向量，進而進行一致之相似度計算。

兩影像間之結構差異程度。具體而言，影像首先經由 CLIP 之圖像編碼器轉換為特徵向量表示，其中高品質影像所對應之特徵向量記為  $CLIP_x$ ，退化影像所對應之特徵向量記為  $CLIP_y$ ，並以兩者之餘弦相似度作為結構相似性量測基礎。該結構相似性結果再與亮度與對比度等影像品質相關因素進行綜合計算，以得到最終之 CSSIM 指標，並作為後續影像品質圖生成模型之訓練標註。

CSSIM 指標之計算流程如圖 2 所示，其係以高品質影像(品質分數為 1.0)與對應之退化影像作為比較對象，透過影像特徵相似度量測方式，評估

$$\text{Cosine Similarity}(CLIP_x, CLIP_y) = \frac{CLIP_x \cdot CLIP_y}{\|CLIP_x\| \cdot \|CLIP_y\|} \dots\dots\dots(1)$$

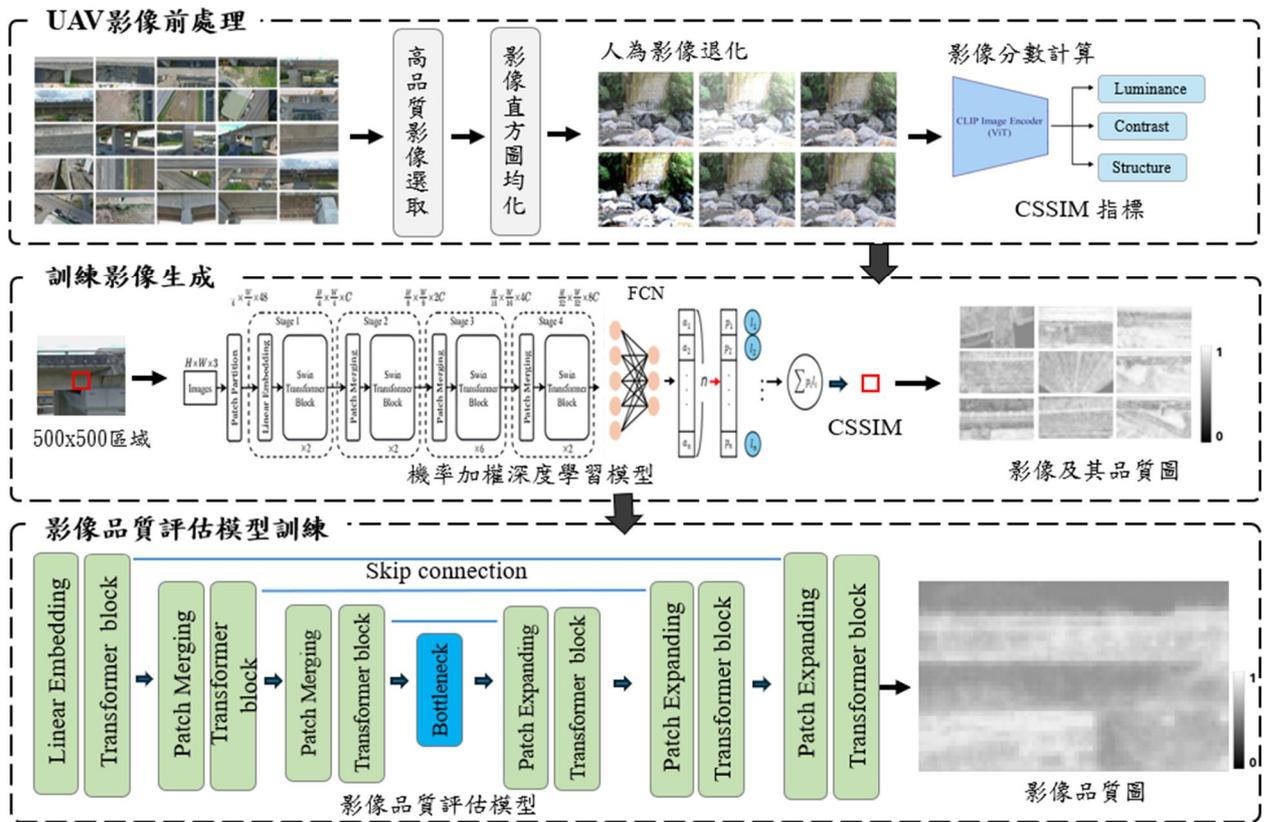


圖 1 快速無參考影像評估架構圖

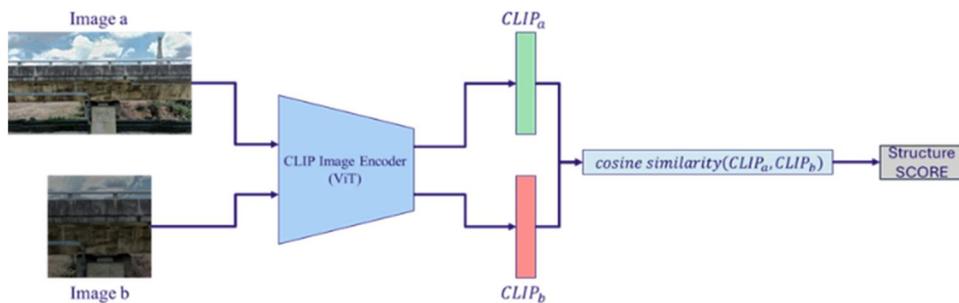


圖 2 CSSIM 結構評分計算架構圖

## 2.2 基於 Swin Transformer 影像品質圖生成模型

本研究建立一個基於 Swin-Transformer 架構的機率加權模型，以生成高品質的影像品質圖，如圖 3 所示。模型以影像中裁切之區域影像作為輸入，首先透過 Swin Transformer 骨架進行階層式特徵萃取，隨後經由全連接網路(FCN)輸出對應於不同品質區間之預測機率分佈。各品質區間代表 CSSIM 值域中之一段連續範圍，模型透過學習輸入影像特徵與品質區間之關聯性，輸出該區域落於各品質區間之機率。最終再利用各區間之代表數值與其對應機率進行加權計算，取得單一之 CSSIM 預測數值，作為該區域之影像品質評估結果。在骨架方面，Swin Transformer 模型具備自注意力機制、滑動視窗與階層式結構，可有效捕捉全局語意與多尺度特徵，提升品質預測之準確性。在解碼器設計上，本研究將影像品質預測任務由傳統回歸問題轉換為分類形式，藉由學習品質區間之機率分佈來進行數值推估。相較於直接回歸連續數值，此一機率加權策略可降低單一數值預測不穩定所造成之影響，並有助於改善回歸模型中常見之訓練不穩定與梯度消失問題。

由於 CSSIM 分數之計算需考量目標像元周圍之結構資訊，無法僅依賴單一像元之像素值進行預測，因此本研究採用區域式預測策略。模型會以目標像元為中心，擷取其周圍 500x500 影像區域作為輸入，並以該區域中心像元所對應之 CSSIM 數值作為預測目標，使模型能透過周圍結構特徵推估中心位置之影像品質。透過此設計，模型可在無需參考影像之情況下，逐一對影像中各像元進行 CSSIM 數值預測，進而生成具備空間連續性之影像品質圖。然而，此逐像元區域推論方式需對影像中大量像元重複進行預測，導致推論時間與運算資源消耗顯著增加。因此，本研究進一步利用上述模型所產出之影像品質圖作為訓練標註資料，建立一套可直接對整張影像進行品質圖生成之高效率模

型，以因應實務應用情境下對即時性與運算效率之需求。

## 2.3 即時無參考影像品質評估模型— Swin Unet 模型

為提升影像品質圖生成效率，本研究採用 Swin-Unet 作為最終影像品質評估模型架構，如圖 4 所示。模型整體採用編碼器-解碼器架構，輸入為整張無人機影像，經由編碼器逐步萃取階層式影像特徵，並透過跳接(skip connection)將編碼階段之局部結構資訊傳遞至解碼器，以於解碼階段回復空間解析度，最終輸出與輸入影像同尺度之影像品質圖。

在此架構下，模型所輸出之每一像素值係對應於該位置之影像品質評估結果，其數值範圍與 CSSIM 指標一致。此一對應關係係透過訓練階段以退化影像及其對應之影像品質圖(由前述逐像元 CSSIM 預測模型產生)作為監督訊號所建立，使模型得以學習 CSSIM 在影像空間中的分佈特性，而非重新定義評估指標本身。因此，本研究將影像品質圖生成問題視為類語意分割任務，透過像素層級的品質值回歸，直接推估整張影像中各區域之品質分佈情形，並可進一步透過空間整合方式反算整張影像之 CSSIM 指標。此策略不僅顯著降低推論時間，亦維持良好預測穩定性，有效支援無人機應用場域中對於快速影像品質評估之實際需求。

## 3. 成果與討論

本研究所使用之影像資料為橋梁檢測任務中所取得之橋梁表面影像，皆由研究團隊實地操作無人機進行拍攝。影像皆採用高解析度相機取得，原始解析度為 5472 x 3076 像素，具備足夠細節以支援後續劣化辨識與三維建模等應用。資料內容涵蓋多種橋梁構件表面材質與拍攝角度，並反映實務巡檢作業中常見之光照變化、視角偏移與局部遮蔽等情境，如圖 5 所示。

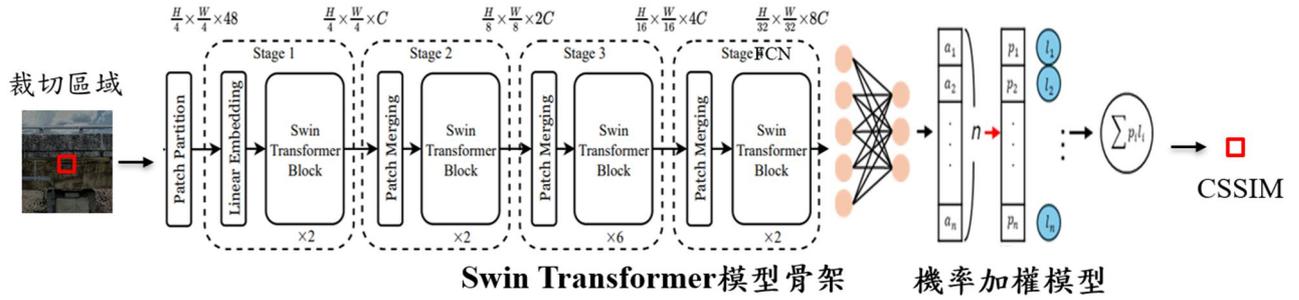


圖 3 基於 Swin Transformer 模型架構圖

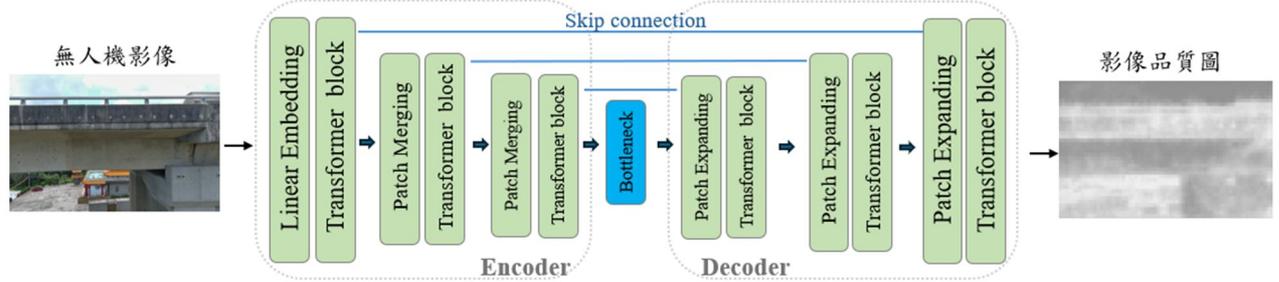


圖 4 Swin-Unet 影像品質評估模型架構圖



圖 5 實地無人機拍攝的橋梁影像資料集

在資料前處理階段，首先由原始影像中篩選出 600 張影像結構清晰、無明顯模糊或過曝現象之影像，作為高品質影像樣本。高品質影像為依據橋梁構件可清楚辨識且不影響後續缺失判釋之工程判準，而非主觀品質評分結果。為了降低初始影像亮度與對比差異對後續品質退化分析之影響，本研究對所選取之高品質影像進行直方圖均化處理，以建立一致之影像品質基準。在完成基準對齊後，透過演算法方式對上述高品質影像施加不同程度之模擬退化處理，以產生對應之退化影像，模擬無人機實際巡檢環境中可能遭遇之光照變化與影像干擾情形。透過此方式，可在受控條件下建立具參考對應關係之影像品質資料，進一步用於影像品質評估

模型之訓練與驗證。所有模型訓練與實驗皆於配備 NVIDIA GeForce RTX 4060 Ti GPU 與 64 GB 記憶體之本地工作站上進行，作業系統為 Windows 平台。

### 3.1 基於 Swin Transformer 的機率加權模型

基於 Swin Transformer 的機率加權模型之訓練參數設定如下：訓練圈數(epoch)為 50，學習率為  $10e-5$ ，優化器為 ADAM，損失函數為 RMSE。該模型由骨架與解碼器部分所構成，為評估模型效能，本研究比較多種不同的骨架，如：CNN、ResNet-18、Vision Transformer 與 Swin Transformer，並針對本

研究所設計之四種不同預測策略進行比較，包括傳統數值回歸模型，以及三種機率加權分類模型。後者係將連續之影像品質分數離散為不同解析度之品質區間，分別對應 11、21 與 41 個分類區間，用以評估不同品質分級解析度設定下之預測表現，其成果詳列於表 1。實驗結果顯示，Swin Transformer 架構在所有預測策略中皆優於其他骨架模型，且機率加權分類方法明顯優於傳統回歸模型，能更靈活捕捉品質變化細節，提升整體預測準確度。分類區間數由 11 提升至 21 時，預測效果明顯提高；惟進一步提升至 41 時，誤差僅略有降低(約 0.004)，顯示模型已擷取主要資訊，再增加分類數效益有限。因此，本研究選用使用 21 類分類區間之 Swin Transformer 模型作為產出高精度影像品質圖之最佳模型。

### 3.2 Swin Unet 影像品質評估模型

為降低逐像素推論所需之運算成本，本研究以高精度影像品質圖作為訓練資料，並採用 Swin-Unet 架構作為影像品質圖預測模型。訓練參數設定如下：訓練圈數為 150，學習率為  $10e-5$ ，優化器採用隨機梯度下降法(Stochastic Gradient Descent, SGD)，以確保在模型訓練過程中具備穩定且可控之

收斂行為，損失函數同樣採用 RMSE。該模型與基於 Swin Transformer 的機率加權模型的模型效能比較如表 2。結果顯示基於 Swin Transformer 的機率加權模型雖可達最高預測精度(RMSE = 0.019)，但平均每張影像推論時間為 532.3 秒，實務應用受限；相較之下，Swin-Unet 影像品質評估模型雖預測精度略降(RMSE = 0.044)，但推論速度大幅提升，每張影像僅需 0.3 秒，展現極佳即時性，適用於無人機巡檢等快速應用情境。

此外 Swin-Unet 影像品質評估模型的預測成果如圖 6 所示。此外，Swin-Unet 影像品質評估模型之預測成果如圖 6 所示。圖中之顏色條表示模型所預測之影像品質分數分佈，其中顏色由淡色至深色分別對應由低至高之影像品質數值，較低品質區域代表影像中可能存在亮度不足、過曝或模糊等失真情形。從品質圖中可以具體呈現畫面中品質下降的區域，並能對應原始影像中的失真類型。例如，案例 1 右下角柱子亮度過高，存在過曝現象；案例 2 在橋梁底下拍攝整體亮度較暗；案例 3 因背光造成橋側亮度過低；案例 4 與案例 3 為同一張，但加上移動模糊，導致影像的品質較低；而案例 5 與案例 6 整體無明顯問題，可排除為低品質影像。

表 1 不同模型骨架與預測策略下的影像品質圖精度比較

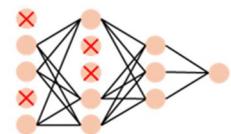
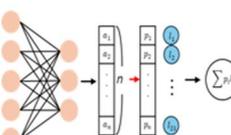
預測策略	骨架	RMSE
 回歸模型	CNN	0.1648
	ResNet-18	0.1571
	Vision Transformer	0.0751
	<b>Swin Transformer</b>	<b>0.0523</b>
 機率加權模型 (11個分類區間)	CNN	0.0855
	ResNet-18	0.0794
	Vision Transformer	0.0444
	<b>Swin Transformer</b>	<b>0.0362</b>
 機率加權模型 (21個分類區間)	CNN	0.0704
	ResNet-18	0.0578
	Vision Transformer	0.0337
	<b>Swin Transformer</b>	<b>0.0193</b>
 機率加權模型 (41個分類區間)	CNN	0.0695
	ResNet-18	0.0561
	Vision Transformer	0.0314
	<b>Swin Transformer</b>	<b>0.0189</b>

表 2 兩模型影預測精度與平均推論時間比較

模型	RMSE	平均推論時間
基於 Swin-Transformer 的機率加權模型	0.0019	532.3 秒
Swin-Unet 即時無參考影像品質評估模型	0.0044	0.3 秒

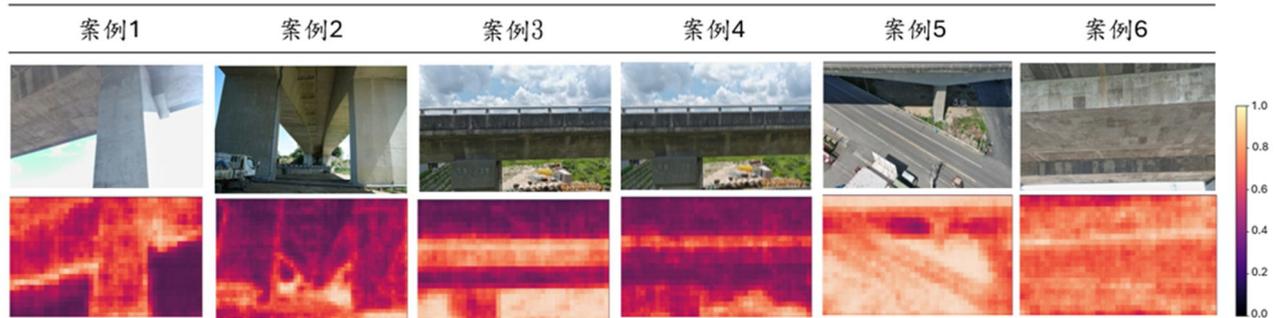


圖 6 Swin-Unet 影像品質評估模型架構圖

## 4. 結論

本研究提出一套基於 Swin-Unet 架構之即時無參考影像品質評估模型，能於 0.3 秒內完成單張影像品質圖之生成，並達到  $RMSE = 0.044$  之預測精度，展現良好之效率與準確性。為解決傳統 SSIM 指標於影像裁切後易產生像素錯位之問題，本研究進一步設計結合 CLIP 圖像編碼器與 SSIM 概念之 CSSIM 指標，並搭配基於 Swin Transformer 之機率加權模型，有效克服無參考條件下品質評估精度與標註資料不足的雙重挑戰。實驗結果顯示，Swin Transformer 機率加權模型(21 分類區間設計)可達最佳預測準確度( $RMSE = 0.0193$ )，惟其逐像元推論機制導致單張影像推論時間高達 532.3 秒，限制其於即時應用情境中之可行性。為提升實務應用性，本研究進一步將上述高精度模型所產生之影像品質圖作為標註資料，訓練一套基於 Swin-Unet 架構之整張影像品質評估模型，以在合理精度下大幅縮短推論時間，成功實現一套可應用於無人機即時影像篩選之高效率、低成本、無參考影像品質評估方法。於實務構造物巡檢流程中，該模型可作為影像品質即時判斷與篩選之決策支援工具，協助辨識因模糊、過曝或光照不足而不適合後續分析之影像，

並提供即時重拍或補充拍攝之依據，以避免低品質影像進入後續缺失判釋或建模流程。

未來研究將進一步整合影像品質評估結果與影像補救或後處理機制，例如針對低品質影像進行自動化亮度調整、去模糊或雜訊抑制處理，或作為觸發重拍與任務調度之依據，以強化系統於實際巡檢作業中之完整性與實用性。同時亦將擴充資料集所涵蓋之影像失真類型，如色偏、雜訊與壓縮失真，以提升模型在多樣現場環境下之泛化能力，並持續優化系統效率與可部署性，推動自動化 UAV 檢測任務之品質管控邁向實務應用。

## 參考文獻

- Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., and Wang, M., 2021. Swin-Unet: Unet-like pure transformer for medical image segmentation, in Proceedings of the 24<sup>th</sup> International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Strasbourg, France, DOI: 10.48550/arXiv.2105.05537.
- Horé, A., and Ziou, D., 2010. Image quality metrics: PSNR vs. SSIM, in Proceedings of the 20<sup>th</sup> International Conference on Pattern Recognition,

- Istanbul, Turkey, pp. 2366–2369, IEEE, DOI: 10.1109/ICPR.2010.579.
- Kang, L., Ye, P., Li, Y., and Doermann, D., 2014. Convolutional neural networks for no-reference image quality assessment, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, pp. 1733–1740, IEEE, DOI: 10.1109/CVPR.2014.224.
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B., 2021. Swin transformer: Hierarchical vision transformer using shifted windows, in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, pp. 10012–10022.
- Ma, K., Liu, W., Zhang, K., Duanmu, Z., Wang, Z., and Zuo, W., 2017. End-to-end blind image quality assessment using deep neural networks, IEEE Transactions on Image Processing, 27(3): 1202–1213, DOI: 10.1109/TIP.2017.2774045.
- Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., Terhaarromeny, B., Zimmerman, J.B., and Zuiderveld, K., 1987. Adaptive histogram equalization and its variations, Computer Vision, Graphics, and Image Processing, 39(3): 355–368, DOI: 10.1016/S0734-189X(87)80186-X.
- Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I., 2021. Learning transferable visual models from natural language supervision, In Proceedings of the 38<sup>th</sup> International Conference on Machine Learning (ICML), pp. 8748–8763.
- Rakha, T., and Gorodetsky, A., 2018. Review of Unmanned Aerial System applications in the built environment: Towards automated building inspection procedures using drones, Automation in Construction, 93: 252–264, DOI: 10.1016/j.autcon.2018.05.002.
- Sieberth, T., Wackrow, R., and Chandler, J. H., 2015. UAV image blur—Its influence and ways to correct it, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-1/W4: 33–39, DOI: 10.5194/isprsarchives-XL-1-W4-33-2015.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I., 2017. Attention is all you need, In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, Vol. 30, pp. 5998–6008.
- Wang, Z., Bovik, A.C., Sheikh, H.R., and Simoncelli, E.P., 2004. Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing, 13(4): 600–612, DOI: 10.1109/TIP.2003.819861.

# Real-time and Reference-free UAV Image Quality Assessment using Deep Learning

Ya-Li Lin<sup>1\*</sup>   Guan-Chin Su<sup>1</sup>   Lai-Han Zou<sup>1</sup>   Chao-Hung Lin<sup>2</sup>  
Jiann-Yeou Rau<sup>2</sup>   Wei-Shen Lai<sup>3</sup>   Chih-Chao Hu<sup>3</sup>

## Abstract

With the increasing use of unmanned aerial vehicles (UAVs) in infrastructure monitoring and environmental inspection, stable image quality has become critical for deep learning and photogrammetry applications. However, UAV images are often degraded by environmental disturbances, while existing quality filtering still relies on manual inspection, making it unsuitable for high-frequency or real-time deployment. This study proposes a real-time, reference-free image quality assessment (IQA) framework based on a Swin-Unet architecture to improve screening efficiency and ensure data quality stability, while simultaneously generating image quality maps (IQMs) for downstream applications. To overcome limitations of traditional SSIM-based methods, including the requirement for reference images and sensitivity to pixel misalignment, an improved metric, termed CLIP-SSIM (CSSIM), is introduced to construct an image scoring model. A probability-weighted Swin-Transformer is first employed to generate high-accuracy IQMs (RMSE = 0.0193); however, its pixel-wise inference is computationally expensive (532 seconds per image). Therefore, the generated IQMs are used as supervisory labels to train a Swin-Unet model, enabling real-time inference (0.3 s per image) with acceptable accuracy (RMSE = 0.04). The proposed approach provides an efficient, accurate, and scalable solution for UAV image screening, effectively replacing manual inspection in high-frequency UAV applications.

**Keywords:** Image Quality Assessment, Deep Learning, UAV Imagery, Structural Similarity

---

<sup>1</sup> Ph.D. Student, Department of Geomatics, National Cheng-Kung University

<sup>2</sup> Professor, Department of Geomatics, National Cheng-Kung University

<sup>3</sup> Researcher, Transportation Engineering Division, Institute of Transportation

\* Corresponding Author, E-mail: alecfree2@gmail.com

Received Date: Nov. 19, 2025

Revised Date: Dec. 18, 2025

Accepted Date: Jan. 30, 2026

# 應用深度資訊於混凝土橋梁結構物影像拼接之研究

林于婷<sup>1\*</sup> 高書屏<sup>2</sup> 王豐良<sup>3</sup> 林志憲<sup>4</sup>

## 摘要

影像拼接可擴展視野、消除盲區，但場景深度差異容易導致視差與重影。為此，本研究整合單影像深度估計與語義分割模型，建立橋梁立面影像拼接流程，重建完整結構外觀圖作為損壞分析和管理底圖。透過遷移學習建置橋側影像數據集，沿用預訓練參數訓練 RGB-D 語義分割模型，mIoU 達 86.44%、mAcc 91.24%、召回率 92.11%、F1-score 91.56%，展現穩定性與泛化能力，並藉其成果間接驗證深度估計模型準確性。針對影像傾斜導致的幾何錯位，利用深度圖重建點雲校正。拼接精度比較顯示，結合分割模型與校正影像之平均 SSIM 為 0.6807 高於傳統方法 0.5081，證實本研究方法在精度與視覺一致性上的優勢。

**關鍵詞：**橋梁立面檢測、影像拼接、深度估計、語義分割、影像幾何校正

## 1. 前言

臺灣地形多山谷、河川縱橫，陸路交通高度依賴橋梁，使其成為重要交通設施。混凝土橋梁長期受氣候與載重影響易老化，需定期巡檢以確保安全。目前檢測多仰賴目視與簡易工具，檢測人員需搭乘作業車近距離觀察，過程耗時且具主觀性與檢測盲區。近年無人機(UAV)影像已成為重要輔助，可多角度進入難以接近區域，具便利、高效與低成本等優勢，提升檢測全面性與精準度。然而，單張影像僅呈現局部外觀，難以全面掌握橋梁狀態，需透過影像拼接重建完整結構外觀圖。然而，當影像場景深度差異顯著時，拼接結果容易在深度方向產生幾何錯位，進而出現偽影或拉伸現象，尤其在幾何結構複雜的混凝土橋梁中更難完全消除。解決此類問題的關鍵在於提升拼接過程中的特徵點檢測與對齊能力，以克服視差造成的錯位，確保拼接成果的幾何一致性與視覺品質。

無人機是目前被廣泛應用的技術之一，在建築監測與檢查中備受矚目。然而，在災後評估與建築

管理中，準確偵測建物立面的損毀情形與類型仍是關鍵課題，對提升分類準確度與決策支援具有顯著影響(Tu *et al.*, 2017)。為提升損壞判讀的準確性與效率，目前多數研究以二維影像或三維點雲為基礎，結合深度學習方法進行建築物缺陷識別與結構損壞判斷，支援自動化的立面目視檢查與建築狀況評估。影像拼接則可透過整合多張影像所獲得的視覺資訊，產生視覺感知更完整的合成影像，提供比單一影像更豐富的場景描述。影像對齊是拼接演算法的核心，即便使用低成本的商用相機，也能產生自然連續的全景影像。(Kekec *et al.*, 2014)

基於特徵匹配的影像拼接透過兩張影像中對應的特徵點，計算用於投影變形的單應性矩陣(Homography Matrix)，以完成影像對齊與融合。其中，Lowe(2004)提出的 SIFT(尺度不變特徵轉換)為經典演算法，其描述子在現有局部描述子中表現優異(Liao *et al.*, 2013)，具備良好性能與穩健性，因而在眾多影像拼接演算法中廣受採用(Tang *et al.*, 2023)。後續亦有多項改進方法被提出，主要著重於提升計算效率(Ma *et al.*, 2016)。

<sup>1</sup> 國立中興大學土木工程學系 碩士

<sup>2</sup> 國立中興大學土木工程學系 教授

<sup>3</sup> 健行科技大學應用空間資訊系 助理教授

<sup>4</sup> 國立中興大學土木工程學系 博士

\* 通訊作者，電話：0919-907-056, E-mail: venusborn2023@gmail.com

收到日期：民國 114 年 09 月 03 日

修改日期：民國 114 年 11 月 21 日

接受日期：民國 115 年 02 月 11 日

在大視差場景下，非重疊區難以兼顧對齊精度與整體變形控制(Li *et al.*, 2021)。透過全域單應性預扭曲，再以能量函數優化對齊，視為網格扭曲問題，能更有效處理視差(Xiang *et al.*, 2018)。Zaragoza *et al.* (2013) 提出的 As-Projective-As-Possible(APAP) 演算法，通過估計投影扭曲並將圖像劃分為網格，逐一估計每個網格的變換。然而，網格變形方法雖能比傳統全局變換更有效地處理視差問題，但在低紋理影像中容易因特徵不足與辨識性低導致匹配不穩與變換誤差，產生明顯錯位(Xiang *et al.*, 2018)。此外，若影像重疊區涵蓋多個幾何平面，僅依賴全域單應性推導出單一全域相似性變換並不足夠(Lin *et al.*, 2015)。

不同於網格變形方法，場景分割可為每個像素賦予語義標籤，將影像劃分為具語義意義的區域。Cai *et al.* (2023) 結合語義分割辨識多幅影像中對應的平面區域，建立平面間的變換與高度關係以完成正射影像拼接；Kao *et al.* (2024) 則先辨識裂縫區域作為 ROI (Region of Interest)，僅對 ROI 偵測 SIFT 關鍵點以降低匹配數與運算時間，並以原始影像補足紋理資訊確保拼接精度。然而，在幾何結構複雜或紋理相近的場景中，僅依賴 RGB 影像進分割效果有限，故學者們提出結合深度資訊的 RGB-D 語義分割模型 (Wang *et al.*, 2021)。Yin *et al.* (2023) 基於 RGB-D 預訓練框架提出 DFormer 模型，透過 AdaBins (Bhat *et al.*, 2021) 生成大規模對應的 RGB 與深度影像對進行多模態訓練，於編碼階段強化 RGB 與深度特徵交互，解碼階段則聚合來自編碼器最後三個層級的多尺度特徵。在多組 RGB-D 分割與檢測資料集上取得領先準確率。因此，本研究導入 RGB-D 語義分割技術引導特徵對應，以提升影像拼接的匹配精度與品質。

單目深度估計僅需單台攝影機即可恢復影像深度資訊，應用更為廣泛(Khan *et al.*, 2020)。深度學習方法在單目深度估計中取得了最佳結果，其中結合語義分割更進一步提升準確性(Rajapaksha *et al.*, 2024)。Bhat *et al.* (2021) 提出的 AdaBins 採自適應深度範圍學習機制，將深度範圍劃分為可學習區間

(bins)，預測其中心值並線性組合輸出深度圖，以適應不同場景深度變化並提升準確性與穩健性。本研究採用 RGB-D 語義分割模型 DFormer(Yin *et al.*, 2023) 結合 AdaBins (Bhat *et al.*, 2021) 進行遷移學習，以增強模型在複雜場景下對幾何結構與深度變化的辨識能力。

無人機姿態誤差會影響影像幾何對齊，特別在相機位置與拍攝角度不一致時。儘管二維影像仍為最常用形式(Kluge *et al.*, 2023)，但對於幾何畸變嚴重或結構複雜的場景中，僅依賴二維影像難以修正深度方向誤差，導致傳統拼接方法失效(Liu *et al.*, 2020)。深度資訊輔助二維影像配準可有效消除錯位並保留清晰度(Wang *et al.*, 2024)。Kwon & Lee (2015) 利用深度相機資料推求主法向量並以旋轉變換修正透視變形。本研究則將此幾何校正方法應用於估計深度，以驗證其可行性。

## 2. 研究方法

### 2.1 研究流程

本研究以無人機拍攝的橋側 RGB 影像為實測對象，整合多種深度學習模型進行拼接。首先以 AdaBins 進行單張影像深度估計；再透過 DFormer 進行 RGB-D 場景分割，融合色彩與深度資訊區分不同空間區域。為提升拼接的幾何一致性與準確度，本研究設計兩項前處理：依場景分割結果分類影像內容，以及結合深度資訊進行透視投影校正，修正拍攝角度引起的三軸旋轉誤差。最後透過特徵點匹配將不同視角影像拼接，生成具完整性與幾何一致性的橋梁立面全景影像，以利後續狀況觀測與分析。整體研究流程如圖 1 所示。

### 2.2 橋側數據集建立與相機率定

本研究採用 SONY RX0 數位相機(1 吋 Exmor RS CMOS 感光元件，如圖 2)，具輕巧體積與高解析度影像擷取能力，拍攝解析度為 4800 × 3200 像素。

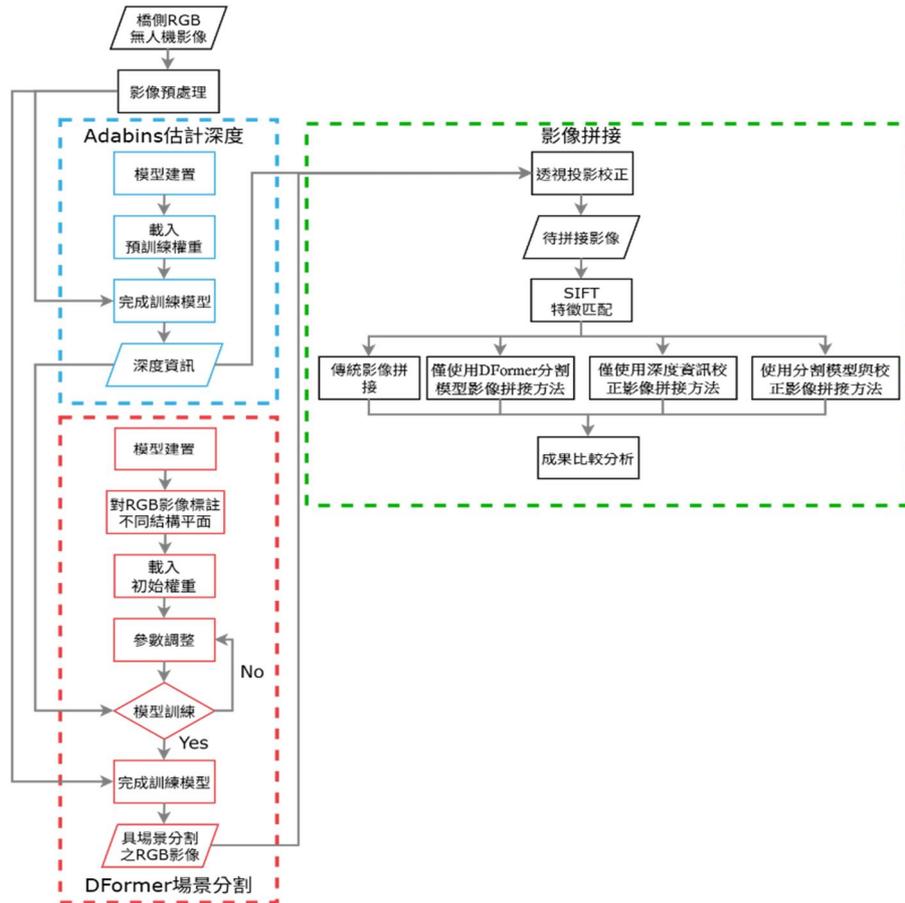


圖 1 研究流程圖



圖 2 Sony RX0 數位相機圖

實驗對象為桃園市大溪區武嶺橋，影像資料為 2021 年 3 月 31 日進行橋梁檢測作業所蒐集，涵蓋南、北兩側立面(如圖 3、4)。共取得 300 張高解析影像作為語義分割模型訓練資料，另選取 7 張側面影像作為測試資料，並應用於後續影像拼接流程研究。

相機率定旨在透過像片上像點的測量，重建拍攝瞬間光束進入鏡頭的精確幾何關係。本研究參考鄒芳諭(2010)提出的非量測型相機率定方法，採用平面棋盤格校準法，從多個視角拍攝黑白棋盤影像(圖 5)，並利用影像中棋盤格交點坐標估算相機的內部參數與畸變係數(表 1)。在取得畸變係數後，即可用來校正影像變形，使畫面更接近真實場景。



圖 3 武嶺橋南側圖



圖 4 武嶺橋北側圖



圖 5 影像率定示意圖

表 1 相機內方位參數表

Sony RX0 內方位參數	
焦距( $f_x, f_y$ )	(3325.9, 3320.81)
像主點位置( $X_0, Y_0$ )	(2344.71, 1583.13)
徑向鏡頭畸變係數( $k_1, k_2$ )	(-0.0861077, 0.328456)
偏心鏡頭畸變係數( $p_1, p_2$ )	(0.00156193, -0.00188202)

## 2.3 AdaBins 預訓練模型與數據集

Bhat *et al.* (2021) 提出 AdaBins 為單張 RGB 影像生成高品質密集深度圖的模型架構。其核心理念在於對傳統編碼器-解碼器(encoder-decoder)輸出的深度分佈進行全域統計分析，並加入學習式後處理模組 AdaBins 於最高解析度下運行，能有效優化深度估計結果。本研究在生成單影像深度資訊時，採用預訓練之 AdaBins 模型，運行環境為 Windows 10 專業版，搭載 12th Gen Intel® Core™ i9-12900F @ 2.40 GHz、NVIDIA GeForce RTX 3080Ti 與 32 GB 記憶體，並於 Python 3.8.0 下執行。AdaBins 原始模型分別於 NYU Depth v2 (Silberman *et al.*, 2012)與 KITTI (Geiger *et al.*, 2012) 資料集完成訓練，可有效由 RGB 影像估計對應深度資訊。本研究選用於 NYU Depth v2 訓練所得之預訓練權重，以生成橋側影像深度資訊。該模型訓練時之設定參數為：初始學習率  $1.4 \times 10^{-5}$ 、最大學習率  $3.5 \times 10^{-4}$ 、採用 AdamW 優化器、批次大小 16，並於 4 張 NVIDIA® V100(32GB)環境下完成訓練。

## 2.4 影像標註

RGB-D 場景分割模型的訓練資料需同時包含 RGB 影像、對應語義標註圖與深度影像，且三者必須保持像素對齊，以利模型同時學習色彩特徵、幾何結構與語義資訊，提升分割準確度與場景理解能力。

本研究針對無人機拍攝的 300 張解析度  $4800 \times 3200$  的 RGB 影像進行人工語義標註，共劃分為 16 類語義類別(表 2)。標註作業以 Photoshop 的多邊形套索工具逐一框選並標記語義區域(圖 6)，完成後將 24 位元標註影像轉換為 8 位元單通道格式，並將 16 類標註對應至像素值 0 至 15(圖 7)，以確保一致性與訓練品質。為提升辨識度，將單通道影像像素值乘以 15(圖 8)，最後再將影像縮放至  $640 \times 480$ ，以符合模型輸入規格。

表 2 各語義類別對應之 RGB 數值表

類別名稱	RGB 數值	類別名稱	RGB 數值
橋柱 2	(255,120,20)	背景車道	(20,20,20)
橋柱 3	(250,170,30)	橋面板 1	(40,40,40)
管線	(0,60,255)	橋側邊 1	(60,60,60)
背景草地	(0,255,50)	橋側邊 2	(90,90,90)
鋼板 1	(240,240,110)	橋側邊 3	(100,100,100)
鋼板 2	(255,255,0)	橋柱 1	(255,66,0)
鋼板 3	(245,245,75)	隔音板	(200,0,0)
背景天空	(0,200,255)	護欄	(140,140,140)

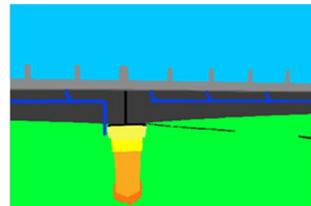
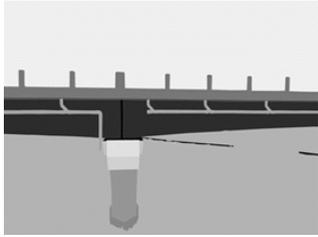


圖 6 原始標註影像



圖 7 轉換後之單通道影像

圖 8 像素值  $\times 15$  之單通道影像

## 2.5 DFormer 預訓練模型與數據集

本研究採用 DFormer 預訓練模型(Yin *et al.*, 2023)進行場景分割。透過引入 RGB-D 交互預訓練策略，模型得以克服僅以 RGB 預訓練骨幹在深度圖中易誤編碼幾何資訊的限制，並藉由遷移學習強化橋梁影像之辨識性能。初始訓練數據為 ImageNet-1K(Russakovsky *et al.*, 2015)，並結合 AdaBins 深度估計器(Bhat *et al.*, 2021)生成對應深度資訊，以增強模型的場景理解與泛化能力。

實驗過程使用 300 張橋側 RGB 影像，其中 180 張訓練影像、120 張驗證影像，進行語義標註並產製深度影像，作為遷移學習資料集(圖 9)。模型初始權重引用 DFormer 預訓練模型，並於 Ubuntu 24.04.2 LTS、Intel® Core™ i9-12900F、NVIDIA RTX 3080Ti、32 GB RAM、Python 3.11.10 環境下訓練。參數設置為初始學習率  $5 \times 10^{-6}$ 、AdamW 優化器、批次大小 8，並採隨機縮放 {0.5, 0.75, 1, 1.25, 1.5} 完成訓練。

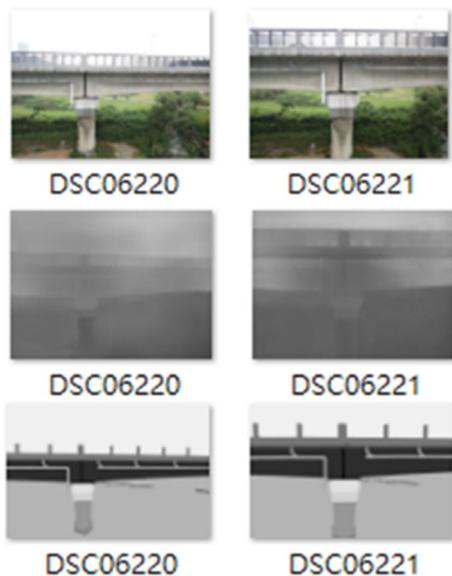


圖 9 用於遷移學習之橋側影像

## 2.6 DFormer 分割結果後處理

本研究選取 7 張待拼接橋側 RGB 影像作為測試資料，以評估所訓練 RGB-D 場景分割模型之效果。結果顯示，大部分具深度差異的區域能正確辨識，但部分仍存在誤判(圖 10)。為提升拼接中平面區分的準確性，本研究設計後處理方法，主要透過不同窗格尺寸比對測試，調整搜尋窗格與面積閾值修正誤分類，並依鄰近像素標籤重新歸類誤判區域(圖 11)。

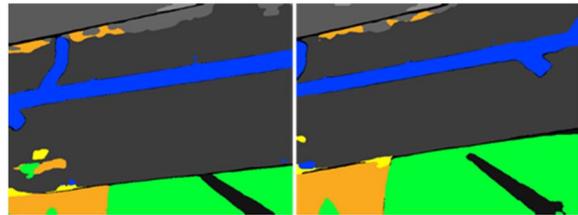


圖 10 DFormer 預測之場景分割結果圖

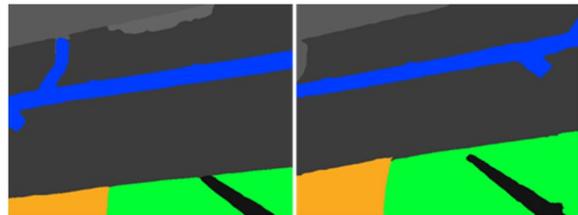


圖 11 後處理之 DFormer 場景分割結果圖

後處理參數包含兩項：

- (1) 以每像素為中心擴展為  $7 \times 7$  視窗，搜尋鄰近不同顏色像素資訊(圖 12)，並排除所有 RGB 值為零的異常區塊。

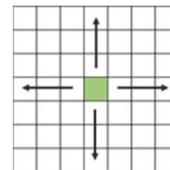


圖 12 誤判區域搜尋視窗示意圖

- (2) 測試多組面積閾值 {1000, 2000, 2500, 3000}，逐步調整，直到所有誤判之分類區塊均能正確歸類。

## 2.7 深度輔助影像校正

傾斜像片為攝影時光軸與鉛垂線形成夾角所獲得的影像。曝光瞬間，即使攝影軸並非刻意傾斜，只要偏離鉛垂線，仍會在像片上產生傾斜現象。為

修正因傾斜造成的幾何變形，需進行幾何校正，使其轉換為同一攝影站位置下的垂直像片(何維信，1995)。像片傾斜主要來自相機在曝光瞬間的姿態變化，其空間取向可由繞 X、Y、Z 軸旋轉的三個角度，即  $\omega$ (Roll)、 $\varphi$ (Pitch)、 $\kappa$ (Yaw)所組成的旋轉矩陣  $M$  表示，用以描述相機在三維空間中的姿態與方向變化，其公式如式(1)。

本研究針對待拼接之橋側 RGB 影像，結合深度資訊進行傾斜校正。由於所選橋側影像屬平面結構，無曲面或大角度起伏，因此僅需將無人機姿態恢復至正攝，使攝影光軸垂直於橋側立面。

(1) 深度影像正規化：影像採用以左上角為原點的 u-v 左手坐標系，深度則以灰階值表示相對距離。故將灰階深度圖正規化至 640x480 影像坐標(圖 13)。

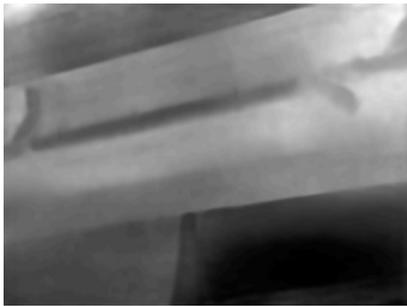


圖 13 深度影像正規化處理圖

(2) 三維點雲重建與語義分類：結合每像素坐標 (u,v) 與灰階深度值，依語義標註篩選主要結構以排除背景干擾(圖 14)。傾斜分析中以深度方向為光軸，分別針對橫軸與縱軸估算傾斜角度(圖 15)，再依據旋轉矩陣將三維點雲校正至無傾斜狀態，完成影像之三維幾何校正(圖 16)。

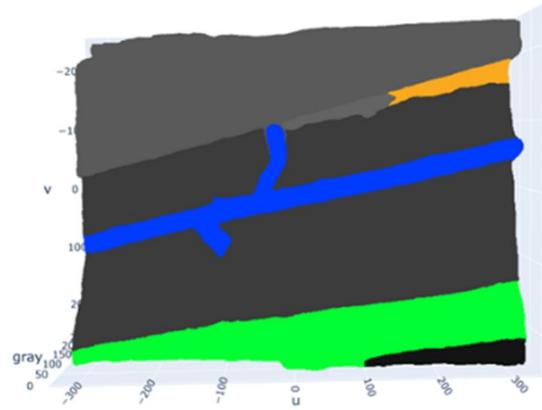


圖 14 語義分割三維點雲俯視圖

(3) 仿射變換與投影映射：利用旋轉矩陣進行反向投影，分別應用至原始 RGB 影像與語義分割影像，並以雙線性內插補齊坐標，避免重疊區域出現稀疏或留白(圖 17)。

$$M = M_{\kappa} M_{\varphi} M_{\omega} = \begin{bmatrix} \cos \kappa & \sin \kappa & 0 \\ -\sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \dots (1)$$

式中：

$$m_{11} = \cos \varphi \cos \kappa$$

$$m_{12} = \cos \omega \sin \kappa + \sin \omega \sin \varphi \cos \kappa$$

$$m_{13} = \sin \omega \sin \kappa - \cos \omega \sin \varphi \cos \kappa$$

$$m_{21} = -\cos \varphi \sin \kappa$$

$$m_{22} = \cos \omega \cos \kappa - \sin \omega \sin \varphi \sin \kappa$$

$$m_{23} = \sin \omega \cos \kappa + \cos \omega \sin \varphi \sin \kappa$$

$$m_{31} = \sin \varphi$$

$$m_{32} = -\sin \omega \cos \varphi$$

$$m_{33} = \cos \omega \cos \varphi$$

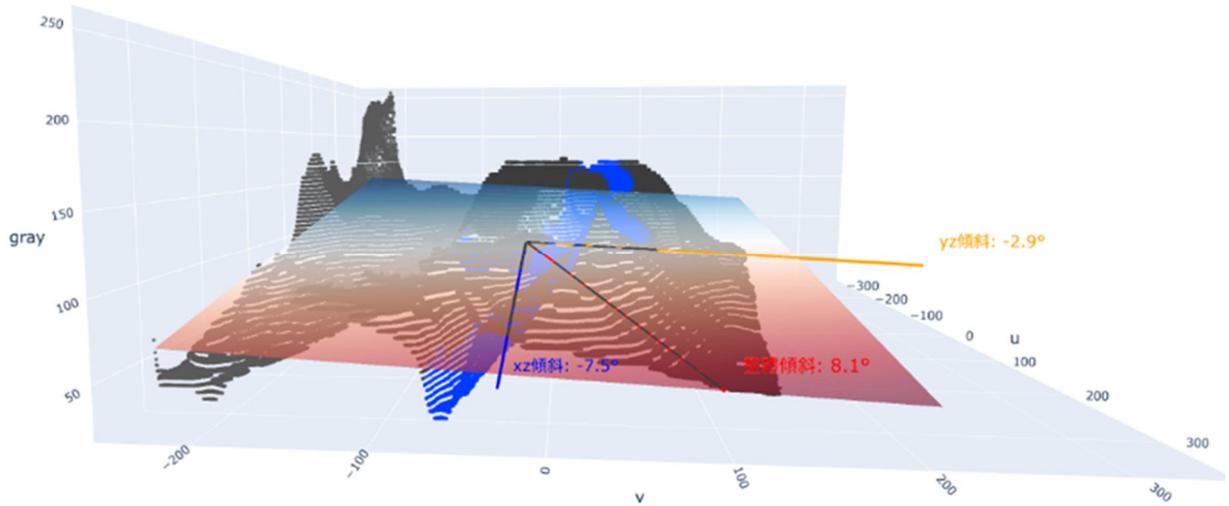


圖 15 影像整體傾斜方向與  $\omega$ (Roll)、 $\varphi$ (Pitch)圖

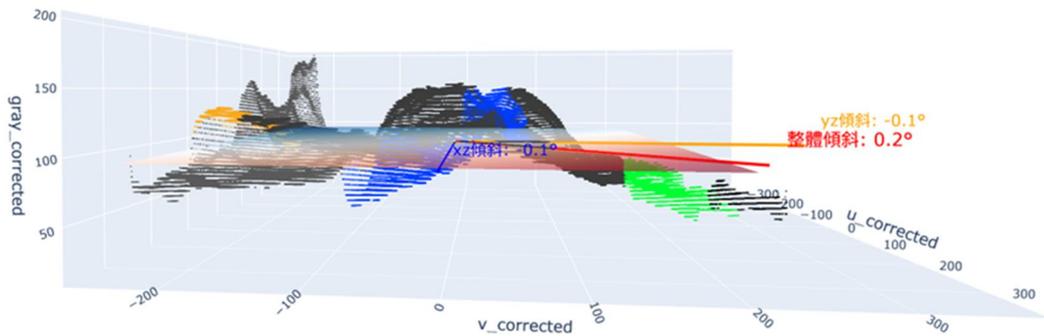
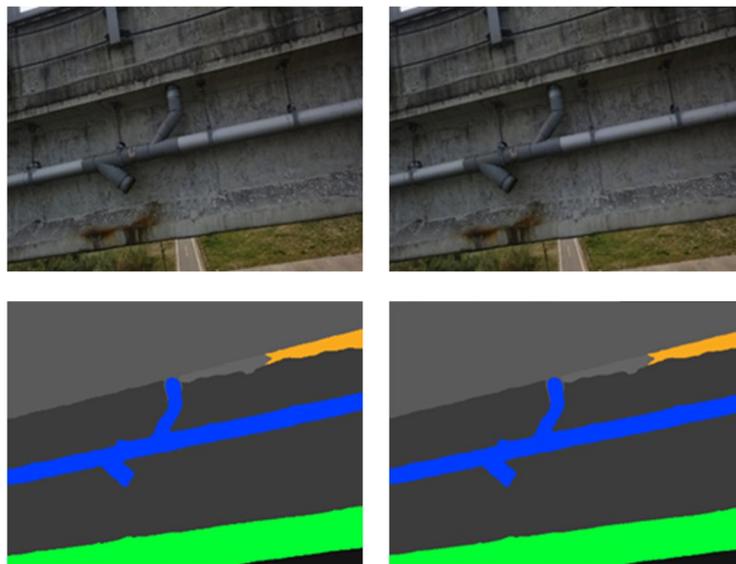


圖 16 校正後影像整體傾斜方向與  $\omega$ (Roll)、 $\varphi$ (Pitch)圖



原始影像

校正影像

圖 17 RGB 與語義分割影像傾斜校正結果圖

## 2.8 影像拼接

在多張影像的連續拼接過程中，以每組影像的左影像作為參考，固定其坐標系；尺度基準則以第

一張影像為依據。相鄰影像之間的重疊率保持在60%以上，以確保具備足夠的有效重疊區域。透過關鍵點偵測與匹配，計算單應性矩陣(Homography Matrix)，將右影像對齊至參考影像，逐步完成拼接。

### 2.8.1 投影變換

又稱為單應性變換(Homography)，指將影像由原始視角轉換至另一視角的操作，可用於產生立體感或修正因拍攝角度造成的透視畸變。在攝影變換的分解中，透視變換描述了影像於投影平面上的幾何對應關係，並可透過一組透視矩陣進行數學建模，其公式如式(2)：

$$H = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda & k0 & 0 \\ 0 & 1/\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ v1 & v2 & v \end{bmatrix} \dots\dots\dots (2)$$

也可表示為式(3)：

$$H = H_s H_A H_p = \begin{bmatrix} sR & t/v \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} k & 0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ v^T & v \end{bmatrix} = \begin{bmatrix} A & t \\ v^T & v \end{bmatrix} \dots (3)$$

### 2.8.2 SIFT 特徵點提取

尺度不變特徵轉換(Scale-Invariant Feature Transform, SIFT)由 Lowe 於 1999 年提出，並於 2004 年完整發表，用於偵測與描述影像中的局部特徵點的演算法，具備尺度與旋轉不變性。其流程包括：首先透過高斯模糊構建尺度空間，並以高斯差分(Difference of Gaussian, DoG)金字塔進行多尺度處理，在其中搜尋局部極值點作為潛在關鍵點。為提升抗噪性與穩定性，進一步以三維二次函數擬合進行位置與尺度細化。隨後根據局部特徵為每個關鍵點分配方向，確保描述子具旋轉不變性。最後，將特徵點鄰域劃分為 4x4 區塊，生成關鍵點描述子，並透過比對參考影像與觀測影像的描述子集合完成特徵匹配。

### 2.8.3 BF 特徵點匹配

完成特徵點擷取後，透過特徵匹配演算法辨識具有高相似性的點對，完成影像間的配對與拼接。本研究於特徵匹配階段中採用 OpenCV 函式庫中的 Brute-Force(BF)匹配器。

BF 匹配器，又稱暴力匹配法，是一種簡單的二維特徵匹配方式，是將一幅影像中的每個特徵點描述子，逐一與另一幅影像中所有描述子比較，並依距離度量選出匹配點。常用策略包含兩種：

- (1) 最佳匹配：選取距離最近的單一描述子作為對應點。
- (2)  $k$ -近鄰匹配：返回前  $k$  個最近的匹配點，由使用者指定  $k$  值，再透過後處理進行過濾。

暴力匹配在處理大量特徵點時效率較低，但其匹配準確度高、結果穩定，特別適合特徵點較少或對準確性要求高的應用場景。

本研究參考 Kao *et al.* (2024) 方法，在影像連續拼接流程中，先以語義分割區分不同空間平面，再利用 SIFT 偵測特徵點，於原始 RGB 影像計算描述子，並透過 BF 演算法進行匹配。針對每一張目標影像，依據與前一張參考影像的匹配點對估算投影轉換矩陣 (Homography Matrix)，完成幾何對齊與拼接(圖 18)。整體流程包括：

- (1) 影像分割：使用語義分割模型劃分不同空間平面。
- (2) 特徵點提取：在目標與參考影像的相同平面內進行 SIFT 偵測。
- (3) 描述子計算：以原始 RGB 影像生成描述子。
- (4) 特徵匹配：利用 BF 匹配器建立兩影像間的特徵點對應。
- (5) 變換參數估算：計算投影轉換矩陣。
- (6) 影像對齊與拼接：進行幾何變換，使影像逐一對齊並拼接成完整全景圖。

## 2.9 評估指標

本研究從兩個面向量化評估拼接效果：一為 RGB-D 場景分割模型預測結果與實際標註之重疊程度；另一為拼接後全局影像的一致性與對齊品質。

### 2.9.1 模型性能評估指標

針對 RGB-D 場景分割模型，透過混淆矩陣(表 3)計算精確度(Precision)、召回率(Recall)、F1-score 與交並比(IoU)，量化預測區域與標註資料的重疊一致性。

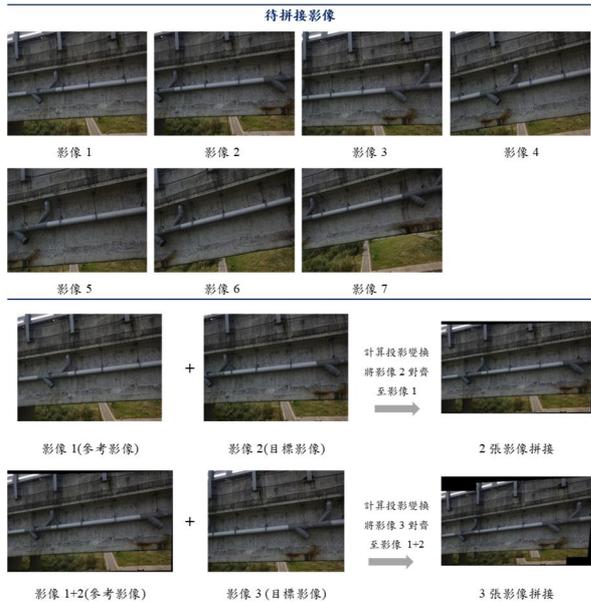


圖 18 連續影像拼接流程示意圖

表 3 分類模型之混淆矩陣評估指標表

	實際為正類	實際為負類
預測為正類	TP	FP(誤報)
預測為負類	FN(漏報)	TN

- (1) 精確度(Precision): 衡量模型在預測為正類時的正確比例, 如式(4)。

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots (4)$$

- (2) 召回率(Recall): 衡量模型對正類樣本的涵蓋能力, 如式(5)。

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots (5)$$

- (3) F1-score: 為精確率(Precision)與召回率(Recall)的調和平均數, 適用於樣本不平衡的分類任務中, 如式(6)。

$$F1 = \frac{2TP}{2TP+FP+FN} = 2 * \frac{Precision*Recall}{Precision+Recall} \dots\dots\dots (6)$$

- (4) 交並比(IoU): 衡量預測區域與真實標註區域的交集與並集比例, 介於 0 至 1, 數值越高表示相似度越高, 如圖 19 所示。

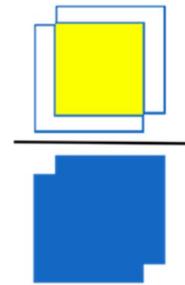


圖 19 交並比(IoU)計算示意圖

### 2.9.2 結構相似性指數

影像拼接品質評估關注錯位情形及亮度與顏色一致性(Zhou *et al.*, 2017)。本研究採用結構相似性指數(SSIM)評估全局影像的對齊與內容一致性。相較於均方誤差(MSE)與峰值訊號雜訊比(PSNR), SSIM 以人類視覺感知為基礎, 綜合考量亮度、對比與結構, 能更準確反映影像間具感知差異的誤差, 具有較高的主觀一致性與實用價值(Sara *et al.*, 2019)。SSIM 值介於 -1 至 1, 越接近 1 表示兩張影像越相似。SSIM(x,y) 為評估兩張影像相似程度的指標, 如式(7)~(10)所示。

$$SSIM(x,y) = [L(x,y)]^\alpha [C(x,y)]^\beta [S(x,y)]^\gamma \dots\dots\dots (7)$$

$$L(x,y) = \frac{2\mu_x\mu_y+c1}{\mu_x^2+\mu_y^2+c1} \dots\dots\dots (8)$$

$$C(x,y) = \frac{2\sigma_x\sigma_y+c2}{\sigma_x^2+\sigma_y^2+c2} \dots\dots\dots (9)$$

$$S(x,y) = \frac{2\mu_{xy}+c3}{\sigma_x\sigma_y+c3} \dots\dots\dots (10)$$

式中:

$L(x,y)$ 表示亮度項

$C(x,y)$ 表示對比度項

$S(x,y)$ 表示結構項

$\alpha$ 、 $\beta$ 、 $\gamma$  分別為各項的加權係數

$\mu_x$ 與 $\mu_y$ 為兩張影像的像素平均值

$\sigma_x^2$ 與 $\sigma_y^2$ 為兩張影像的方差

$\sigma_x\sigma_y$ 為兩張影像之間的協方差

$c1$ 、 $c2$ 、 $c3$  用於避免分母為零

### 3. 研究成果與分析

#### 3.1 DFormer 訓練成果與分析

本章節將分析模型訓練結果，以 300 張影像作為訓練數據，隨機劃分為訓練集 180 張與驗證集 120 張，並依檔名對應 RGB、深度 (Depth) 與語義標註 (Label) 影像，以利資料載入與訓練。模型訓練前載入預訓練權重作為初始參數，並於單張 NVIDIA RTX 3080 GPU 環境下進行。訓練影像統一調整為 640 × 480 像素，初始學習率設為 6e-5，隨訓練逐步衰減；優化器採 AdamW，動量 0.9，權重衰減 0.01。為提升泛化能力，訓練中採多尺度[0.5, 0.75, 1, 1.25, 1.5]與隨機翻轉策略，每批次含 8 張影像，共訓練 500 個 epoch。

考量 GPU 資源限制，本研究選用輕量級 DFormer-T (Tiny) 作為編碼器，以兼顧效率與可行性。圖 20 顯示訓練過程的損失函數變化。

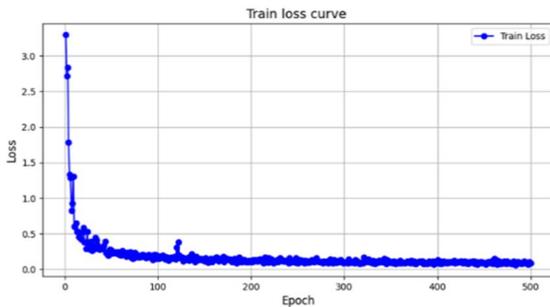


圖 20 訓練階段之損失函數曲線圖

在驗證階段，本研究以平均交並比 (mIoU) 評估模型效能。如圖 21 所示，隨著訓練過程中損失函數逐漸收斂，mIoU 指數亦呈現穩定上升趨勢，最終驗證結果達到 86.44%，顯示模型具有良好的語義分割準確性與收斂表現。

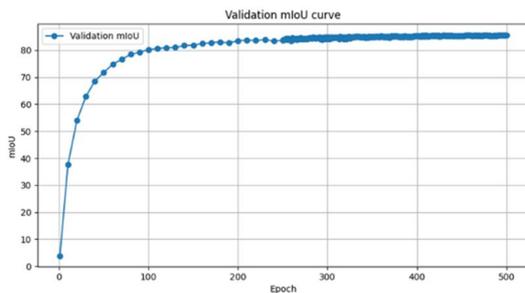


圖 21 驗證階段之 mIoU 曲線圖

並針對完成訓練之模型進行多項性能指標評估，包含各語義分割類別的交並比(IoU)(圖 22)、精確度(Precision)與召回率(Recall)(圖 23)以及 F1-score，如表 4 所示。

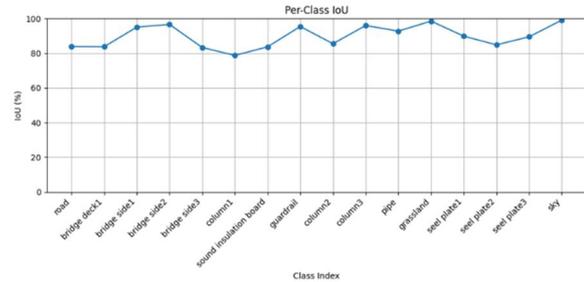


圖 22 各類別 IoU 分數曲線圖

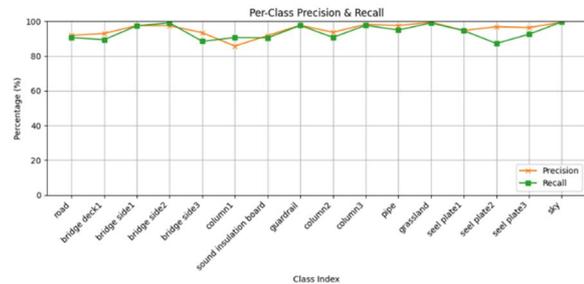


圖 23 各類別精度與召回率分數曲線圖

表 4 模型訓練各語義類別之平均精度指標

mIoU	mAcc	Mean Precision	Mean Recall	mF1
86.44%	92.1%	91.24%	92.11%	91.56%

#### 3.2 影像拼接成果比較與分析

##### 3.2.1 相鄰影像拼接成果

進行連續拼接前，先對 DFormer 分割影像與深度校正影像進行兩張影像拼接測試，並與 Zaragoza *et al.* (2013) 所提出的 APAP 演算法比較。實驗結果顯示，APAP 在低紋理區域因關鍵點不足而受限(圖 24)，使扭曲模型估算不準確，導致影像錯位(圖 25)；相較之下，結合 DFormer 分割搭配 Kao *et al.* (2024) 所提方法，利用原始 RGB 影像提取描述子，可有效提升對齊精度與拼接品質(圖 26)。進一步將語義分割資訊導入拼接流程(圖 27)，能改善大視差下的特徵對應，提升穩定性與一致性；若先進行傾斜校正再結合分割結果(圖 28)，則能進一步減少幾何變形，確保拼接的準確性與完整性。



圖 24 APAP 演算法相鄰 RGB 影像特徵點匹配展示圖



圖 25 APAP 演算法之相鄰 RGB 影像拼接展示圖

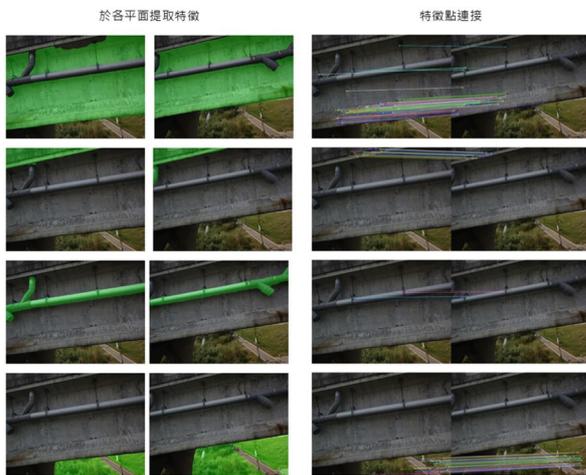


圖 26 僅使用 DFormer 分割影像特徵點匹配展示圖



圖 27 僅使用 DFormer 模型分割之相鄰 RGB 影像拼接展示圖



圖 28 使用分割模型與校正影像方法之相鄰 RGB 影像拼接展示圖

### 3.2.2 相鄰影像拼接精度評估

針對三種不同預處理方法的相鄰拼接結果進行結構相似性指數(SSIM)評估，其分析結果見表 5、圖 29。

表 5 相鄰影像拼接之評估指標比較表

影像對	傳統方法	僅使用 DFormer 分割模型方法	僅使用深度資訊校正影像方法	使用分割模型與校正影像方法
影像 1& 影像 2	0.4896	0.5719	0.5204	0.6136
影像 2& 影像 3	0.4943	0.6228	0.5305	0.6681
影像 3& 影像 4	0.4012	0.6402	0.4312	0.6850
影像 4& 影像 5	0.4768	0.6292	0.5069	0.6924
影像 5& 影像 6	0.6200	0.6664	0.6730	0.7306
影像 6& 影像 7	0.5669	0.6323	0.5977	0.6947
Average	0.5081	0.6271	0.5433	0.6807

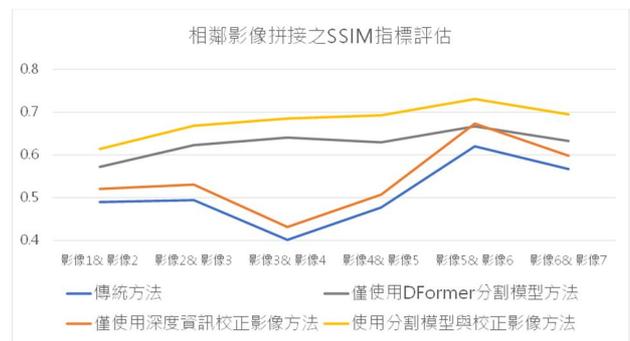


圖 29 相鄰影像拼接之評估指標比較折線圖

### 3.2.3 連續影像拼接成果

本研究以 7 張橋側影像比較三種拼接策略與傳統方法之差異，分別為：僅用 DFormer 分割模型、僅用深度校正影像，以及結合分割與校正影像，以評估橋梁影像拼接的準確性。

- (1) 傳統方法：透過 SIFT 偵測特徵點與描述子，使用 BF 匹配並結合 RANSAC 排除異常後估算單應矩陣完成拼接。但因僅依賴 RGB 資訊，當影像存在深度差異時，易產生對齊誤差，影響拼接品質(圖 30)。



圖 30 傳統方法之連續影像拼接成果圖

- (2) 僅用 DFormer 分割模型：透過語義分割辨識橋梁主體與背景，於橋梁結構上可達較高對齊精度。但因拍攝角度與深度差異，背景區域易產生視差與局部錯位。藉由分割遮罩排除不具幾何一致性的背景，可避免誤匹配並提升拼接準確度(圖 31)。



圖 31 僅使用 DFormer 分割模型方法之連續影像拼接成果圖

- (3) 僅用深度校正影像：透過深度資訊進行幾何校正，可補償無人機姿態誤差，降低拼接中的幾何變形，並提升連續拼接的對齊穩定性(圖 32)。



圖 32 僅使用深度資訊校正影像方法之連續影像拼接成果圖

- (4) 結合分割模型與校正影像：先利用深度資訊估算拍攝姿態並計算旋轉矩陣，再對 RGB 與 RGB-D 影像進行傾斜校正，修正因角度與視差造成的幾何扭曲。此方法不僅提升跨視角影像的對齊準確度，也保留分割資訊對橋梁平面的辨識能力，進而改善拼接的一致性與精度(圖 33)。



圖 33 使用分割模型與校正影像方法之連續影像拼接成果圖

### 3.2.4 連續影像拼接成果分析

本實驗採用以序列首張影像作為尺度基準，後續影像依序進行投影變換與對齊。由於未進行全局優化，誤差會隨拼接順序逐步累積，導致序列後段出現明顯的幾何變形與對位偏差，特別在長距離拼接中更為顯著，顯示缺乏尺度統一與誤差抑制機制時，拼接結果易受累積誤差影響。

## 4. 結論

### 4.1 結論

本研究提出之影像拼接方法無需繁瑣的三維建模，以橋側影像驗證其效能，有效降低建模時間

與硬體需求。本研究方法結合單影像深度估計與 RGB-D 場景分割模型，並利用深度資訊進行傾斜校正，修正拍攝姿態造成的幾何變形，提升複雜場景下的拼接品質與對齊準確性。綜合實驗結果，結論如下四點：

- (1) 透過遷移學習提升 RGB-D 場景分割模型於橋梁場景的辨識能力，訓練過程損失函數穩定收斂，驗證階段 mIoU 提升至 86.44%，最終 mAcc 達 92.1%。結合 AdaBins 生成之深度圖，有效區分不同空間深度物體並提升語義辨識準確度，另以後處理機制修正誤判區域，進一步改善分割效果。
- (2) 透過深度資訊輔助 RGB 影像進行幾何校正，修正因相機傾斜造成的影像扭曲，有效降低拼接時的幾何變形，提升對齊準確性與拼接品質。
- (3) 傳統以 SIFT 提取特徵點結合 APAP 網格扭曲方法對參數較為敏感，易導致拼接錯位；本研究透過語義分割與深度資訊輔助的影像校正策略，有效提升複雜場景下的拼接精度與穩定性。以結構相似性指數(SSIM)量化評估拼接品質，實驗結果顯示，三種預處理方法之平均 SSIM 值分別為 0.6271、0.5433、0.6807，皆顯著高於傳統方法之 0.5081，顯示本方法在視覺一致性上具明顯優勢。
- (4) 本研究產製之橋側立面全景影像可作為橋梁檢測與管理的記錄底圖，特別適用於長裂縫等無法以單張影像完整呈現的作業需求，有助提升結構診斷效率與作業系統化程度。

## 4.2 建議

- (1) 本研究針對語義分割中的誤判區域，採用逐步閾值調整結合鄰近像素分類進行修正，雖有效提升分割品質，但在大範圍或高複雜度場景中處理成本較高。建議未來導入自動化閾值選取或機器學習方法，以提升效率與泛化能力。
- (2) 本研究直接使用預訓練之 AdaBins 模型生成深度資訊，建議未來加入混凝土橋梁影像進行遷移學習，以強化其在結構場景下的深度估計準確性，進一步提升 RGB-D 語義分割精度。

- (3) 本研究所選用之待拼接影像以連續影像中的第一張作為尺度參考，惟影像間存在尺度不一致的情況，可能導致誤差累積。建議未來可透過轉換公式統一尺度，或採用等距離拍攝方式以維持影像尺度一致性，進一步提升拼接穩定性與準確性。
- (4) 本研究驗證場域為混凝土橋梁側面影像，建議未來擴展至更複雜環境與多目標語義分類任務，拓展於智慧建築監測與自動巡檢等應用場景。

## 參考文獻

- 何維信，1995。航空攝影測量學，大中國圖書公司。  
[Ho, W.H., 1995. Photogrammetry, The Great China Book Co., Ltd., Taiwan, ROC. (in Chinese)]
- 鄒芳諭，2010。以非量測性相機進行近景攝影測量，國立交通大學土木工程學系碩士論文。  
[Tsou, F.Y., 2010. Analysis of close-range photogrammetry by using non-metric camera, Master Thesis, National Chiao Tung University, Taiwan, ROC. (in Chinese)]
- Bhat, S.F., Alhashim, I., and Wonka, P., 2021. Adabins: Depth estimation using adaptive bins, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, pp. 4009-4018, DOI: 10.1109/CVPR46437.2021.00400.
- Cai, W., Du, S., and Yang, W., 2023. UAV image stitching by estimating orthograph with RGB cameras, Journal of Visual Communication and Image Representation, 94: 103835, DOI: 10.1016/j.jvcir.2023.103835.
- Geiger, A., Lenz, P., and Urtasun, R., 2012. Are we ready for autonomous driving? the KITTI vision benchmark suite, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, pp. 3354-3361, IEEE, DOI: 10.1109/CVPR.2012.6248074.

- Kao, S.P., Lin, J.S., Wang, F.L., and Hung, P.S., 2024. A large-crack image-stitching method with cracks as the regions of interest, *Infrastructures*, 9(4): 74, DOI: 10.3390/infrastructures9040074.
- Kekec, T., Yildirim, A., and Unel, M., 2014. A new approach to real-time mosaicing of aerial images, *Robotics and Autonomous Systems*, 62(12): 1755-1767, DOI: 10.1016/j.robot.2014.07.010.
- Khan, F., Salahuddin, S., and Javidnia, H., 2020. Deep learning-based monocular depth estimation methods—A state-of-the-art review, *Sensors*, 20(8): 2272, DOI: 10.3390/s20082272.
- Kluge, M., Weyrich, T., and Kolb, A., 2023. Progressive refinement imaging with depth-assisted disparity correction, *Computers & Graphics*, 115: 446-460, DOI: 10.1016/j.cag.2023.07.036.
- Kwon, S.K., and Lee, D.S., 2015. Correction of perspective distortion image using depth information, *Journal of Korea Multimedia Society*, 18(2): 106-112, DOI: 10.9717/kmms.2015.18.2.106.
- Li, J., Wu, D., Jiang, P., Li, Z., and Song, S., 2021. Locally aligned image stitching based on multi-feature and super-pixel segmentation with plane protection, *IEEE Access*, 9: 168315-168328, DOI: 10.1109/ACCESS.2021.3134887.
- Lin, C.C., Pankanti, S.U., Ramamurthy, K.N., and Aravkin, A.Y., 2015. Adaptive as-natural-as-possible image stitching, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 1155-1163, DOI: 10.1109/CVPR.2015.7298719.
- Liao, K., Liu, G., and Hui, Y., 2013. An improvement to the SIFT descriptor for image representation and matching, *Pattern Recognition Letters*, 34(11): 1211-1220, DOI: 10.1016/j.patrec.2013.03.021.
- Liu, Y.F., Nie, X., Fan, J.S., and Liu, X.G., 2020. Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction, *Computer-Aided Civil and Infrastructure Engineering*, 35(5): 511-529, DOI: 10.1111/mice.12501.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60: 91-110, DOI: 10.1023/B:VISI.0000029664.99615.94.
- Ma, W., Wen, Z., Wu, Y., Jiao, L., Gong, M., Zheng, Y., and Liu, L., 2016. Remote sensing image registration with modified SIFT and enhanced feature matching, *IEEE Geoscience and Remote Sensing Letters*, 14(1): 3-7, DOI: 10.1109/LGRS.2016.2600858.
- Rajapaksha, U., Sohel, F., Laga, H., Diepeveen, D., and Bennamoun, M., 2024. Deep learning-based depth estimation methods from monocular image and videos: A comprehensive survey, *ACM Computing Surveys*, 56(12): 315, DOI: 10.1145/3677327.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., and Fei-Fei, L., 2015. Imagenet large scale visual recognition challenge, *International Journal of Computer Vision*, 115: 211-252, DOI: 10.1007/s11263-015-0816-y.
- Sara, U., Akter, M., and Uddin, M.S., 2019. Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study, *Journal of Computer and Communications*, 7: 8-18, DOI: 10.4236/jcc.2019.73002.
- Silberman, N., Hoiem, D., Kohli, P., and Fergus, R., 2012. Indoor segmentation and support inference from rgb-d images, in *Proceedings of the Computer Vision—ECCV 2012: 12<sup>th</sup> European Conference on Computer Vision*, Florence, Italy, pp. 746-760, DOI: 10.1007/978-3-642-33715-4\_54.
- Tang, Z., Zhang, Z., Chen, W., and Yang, W., 2023. An SIFT-based fast image alignment algorithm for

- high-resolution image, *IEEE Access*, 11: 42012-42041, DOI: 10.1109/ACCESS.2023.3270911.
- Tu, J., Sui, H., Feng, W., Sun, K., Xu, C., and Han, Q., 2017. Detecting building facade damage from oblique aerial images using local symmetry feature and the Gini index, *Remote Sensing Letters*, 8(7): 676-685, DOI: 10.1080/2150704X.2017.1312027.
- Wang, C., Wang, C., Li, W., and Wang, H., 2021. A brief survey on RGB-D semantic segmentation using deep learning, *Displays*, 70: 102080, DOI: 10.1016/j.displa.2021.102080.
- Wang, T., Huang, H., Cai, Z., Song, J., and Yang, J., 2024. 360° panorama stitching method with depth information: Enhancing image quality and stitching accuracy, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48: 191-197, DOI: 10.5194/isprs-archives-XLVIII-4-W10-2024-191-2024.
- Xiang, T.Z., Xia, G.S., Bai, X., and Zhang, L., 2018. Image stitching by line-guided local warping with global similarity constraint, *Pattern Recognition*, 83: 481-497, DOI: 10.1016/j.patcog.2018.06.013.
- Yin, B., Zhang, X., Li, Z., Liu, L., Cheng, M.M., and Hou, Q., 2023. Dformer: Rethinking rgb-d representation learning for semantic segmentation, *arXiv preprint*, arXiv:2309.09668, DOI: 10.48550/arXiv.2309.09668.
- Zaragoza, J., Chin, T.J., Brown, M.S., and Suter, D., 2013. As-projective-as-possible image stitching with moving DLT, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, Portland, OR, USA, pp. 2339-2346, DOI: 10.1109/CVPR.2013.303.
- Zhou, X., Zhang, H., and Wang, Y., 2017. A multi-image stitching method and quality evaluation, in *Proceedings of the 4<sup>th</sup> International Conference on Information Science and Control Engineering (ICISCE)*, Changsha, China, pp. 46-50, DOI: 10.1109/ICISCE.2017.20.

## Research on Depth-Enhanced Image Stitching for Concrete Bridge Structures

Yu-Ting Lin <sup>1\*</sup>    Szu-Pyng Kao <sup>2</sup>    Feng-Liang Wang <sup>3</sup>    Jhih-Sian Lin <sup>4</sup>

### Abstract

Image stitching can expand the field of view and eliminate blind spots, but differences in scene depth often lead to parallax and ghosting. To address this, this study integrates single-image depth estimation and semantic segmentation models to establish a façade image stitching workflow for bridges, reconstructing a complete structural appearance map for damage analysis and management. Through transfer learning, a bridge-side image dataset was constructed, and a pretrained RGB-D semantic segmentation model was fine-tuned. The model achieved an mIoU of 86.44%, mAcc of 91.24%, recall of 92.11%, and F1-score of 91.56%, demonstrating stability and generalization capability, while indirectly validating the accuracy of the depth estimation model. To correct geometric misalignment caused by image tilt, depth maps were used to reconstruct point clouds for correction. A stitching accuracy comparison shows that the average SSIM of images corrected with the segmentation model (0.6807) was higher than that of the traditional method (0.5081), confirming the advantages of the proposed approach in terms of accuracy and visual consistency.

**Keywords:** Bridge Façade Inspection, Image Stitching, Depth Estimation, Semantic Segmentation, Image Geometric Correction

---

<sup>1</sup> Master, Department of Civil Engineering, National Chung Hsing University

<sup>2</sup> Professor, Department of Civil Engineering, National Chung Hsing University

<sup>3</sup> Assistant Professor, Department of Applied Geoinformatics, Chien Hsin University of Science and Technology

<sup>4</sup> Ph.D., Department of Civil Engineering, National Chung Hsing University

\* Corresponding Author, Tel: 886-919907056, E-mail: venusborn2023@gmail.com

Received Date: Sep. 03, 2025

Revised Date: Nov. 21, 2025

Accepted Date: Feb. 11, 2026

## 中文稿件投稿注意事項（英文稿件請參閱下頁 Author Guidelines）

本學刊定期出版攝影測量及遙感探測領域之學術論著、技術短文(按中文及英文摘要、關鍵字、前言、材料及方法、結果、討論、參考文獻等撰寫)及應用性報導。收錄之稿件 皆經本編輯委員會安排之專家學者審查程序，並由通訊作者簽署著作權讓渡書。作者可自行聲明文稿和所屬類別，本編輯委員會將依此及審查委員意見決定收錄 文稿之類別。本學刊不接受已經刊載過之文稿，於研討會發表過之論文，經適當增補其內容者為例 外。

投稿地址為 [jprsubmit@gmail.com](mailto:jprsubmit@gmail.com)，請以附件寄送稿件之 WORD 檔及 PDF 檔各一份。為便利編審作業，請於 PDF 稿件中移除作者資料(如姓名、服務機關、電話、及電子郵件信箱等)。稿件撰寫要點如下：

1. 頁面：採用 A4 規格頁面。
2. 字體：(1) 中文均以新細明體橫式打字；(2) 英文均以 Times New Roman 字體橫式打字。
3. 單位：內文、圖表及符號中之度量單位皆採用 SI 公制單位。
4. 首頁：首頁應含內容依次為「題目、作者姓名、摘要、關鍵詞、及作者服務機關及職稱（置於頁尾）」，接續文稿內文。
5. 末頁：末頁應含內容與首頁相同，但皆以英文書寫。
6. 題目：論文題目宜簡明，字體大小為 20 粗體，英文題目每字第一位字母以大寫字體打印。
7. 作者：作者姓名列於題目之下方，字體大小為 14 粗體，其服務機關及職稱以<sup>1,2,...</sup>上標編號註記於頁尾，字體大小為 9。
8. 摘要：摘要字數以 250 字為限，字體大小為 10 粗體。
9. 關鍵詞：列出三至五個關鍵詞，字體大小為 12 粗體。
10. 章節及標題：(1) 章節標題置中，編號一律為阿拉伯數字，如 1, 2, ...，標題字體大小為 18 粗體；(2) 小節標題靠左，編號一律為 1.1, 1.2, 1.1.1, 1.1.2.....，字體大小為 14 粗體。
11. 內文：內文之長度(含圖表)以不超過 22 頁為原則，各段落第一行應縮排，字體大小為 10。
12. 數學式：所有數學式均須標式號於圓括弧內，數學式內所有符號須於符號第一次出現時，應於數學式下方逐一加以定義。
13. 圖及影像：(1) 圖及影像均須數位化置於文稿中適當位置；(2) 所有圖及影像下方須附圖流水編號及說明；(3) 本刊接受彩色圖及影像，但紙版印刷一律採灰階印刷；(4) 結尾毋加「。」及「.」。
14. 表格：(1) 表格均須數位化置於文稿中適當位置；(2) 所有表格上方須附表流水編號及說明；(3) 本刊接受彩色表格，但紙版印刷一律採灰階印刷；(4) 結尾毋加「。」及「.」。
15. 參考文獻(中、日文須英譯)：(1) 參考文獻應列於文末，文稿中未直接引用者不得列入；(2) 中文文獻需加英文翻譯，於結尾加上(in Chinese)，日文書寫同中文文獻，於結尾加上(in Japan)。先排列中、日文參考文獻，再排列英文參考文獻，其他文字參考文獻列於英文參考文獻之後；(3) 中文參考文獻以第一作者 姓名筆劃順序排列，英文參考文獻以第一作者之姓依英文字母順序排列。當作者相同時，依發表年代排列，作者及發表年代完全相同時，依日期在年代後加以 a, b, c 順序。第一作者姓名相同時，先排列單一作者之參考文獻，其次為多作者文獻；(4) 參考文獻之格式如下列：
  - 期刊論文：作者姓名，年代。文章標題，期刊名稱，卷數(期數)：起始頁數 - 結束頁數。
  - 書籍：作者姓名，年代。書名，出版者，出版地。
  - 研討論文集：作者姓名，年代。篇名，論文集名稱，發表地點，卷(期)：起始頁數 - 結束頁數。
  - 網路引用：作者姓名，年代。篇名，取自網址，檢索日期。
16. 參考文獻引用：(1) 作者為一人時，以作者全名(中文文獻)或姓(英文文獻)及發表年代，如楊明德(2007)或(楊明德，2007)、Yang (2007)或(Yang, 2007)；(2) 作者為二人時，以二作者全名(中文文獻)或姓(英文文獻)及發表年代，如楊明德及蘇東青(2014) 或(楊明德及蘇東青，2014)、Yang & Su (2014)或(Yang & Su, 2014)；(3) 作者為多人時，以第一作者全名(中文文獻)或姓(英文文獻)，之後加「等」(中文文獻)或「*et al.*」(英文文獻)及發表年代，如(楊明德等，2016)或楊明德等(2016)、Yang *et al.* (2016)或(Yang *et al.*, 2016)，「*et al.*」請以斜體；(4) 作者及發表年代完全相同時，依日期在年 代後加以 a, b, c 引用，如 Yang & Su (2008a)、Yang & Su (2008b)。

# Author Guidelines

The Journal of Photogrammetry and Remote Sensing periodically publishes academic articles, technical short papers, and application reports addressing topics in photogrammetry and remote sensing. Acceptance of submitted manuscripts for publication is based on a peer-reviewing process arranged by the editorial board. A manuscript will not be formally accepted until a copyright agreement is received. Although authors can indicate their submitted manuscripts belonging to one of the categories of academic articles, technical short papers or application reports, the editorial board will make the final decision based on the suggestions of reviewers. This journal accept no paper ever been published, but a revised appropriately paper published in conference proceedings is an exception.

The manuscript submission address is [jprsubmit@gmail.com](mailto:jprsubmit@gmail.com). Please send your manuscript in both WORD and PDF formats via e-mail attachments. However, for blind reviewing, please eliminate author information in the PDF file. Instructions for authors to prepare a manuscript are as follows:

1. Typing format: Manuscripts should be typed single-column and double-space on A4 page size using font-type of 10pt **Times New Roman**. The pages should be numbered.
2. Units: Manuscripts must use SI (Metric) units in text, figures, or tables.
3. Cover page: The cover page should contain manuscript title, author names, abstract, key words, and author affiliations.
4. End page: The end page contains the same information as the cover page, but in Chinese.
5. Title: The title of the manuscripts should be concise and informative. Please use font size of 20pt, Bold and capitalize the initial letter of each word in the title.
6. Authors: Authors names (using font size of 14pt, Bold) are listed below the title, affiliations (using font size of 9pt, Bold) are listed at the end of the first page marked with <sup>1, 2, ...</sup>.
7. Abstract: The abstract should be shorter than 250 words and typed using font size of 10pt..
8. Keywords: Keywords (no more than 5 words) should be given and placed below the abstract (using font size of 12pt, Bold).
9. Sections and Sub-sections: Section heading (using font size of 18pt, Bold) should be centered and numbered in 1, 2... etc. Sub-sections (using font size of 14pt, Bold) are justified to the left in 1.1, 1.2, 1.1.1, 1.1.2...etc.
10. Paper length: Published papers are limited to 22 pages.
11. Equations: Type formulae and equations clearly in black and numbered in curved brackets. All symbols in equations should be clearly defined when they firstly appear in the manuscripts.
12. Figures and Tables: Figures and tables are numbered and described in short. When papers are accepted, original figures, tables, and photographs are expected for reproduction. Color figures and tables are acceptable in digital version, but they will be printed in gray levels in hard copies.
13. References: Only works cited in the text will be included. References should be cited in the text by giving the last name of the author(s) and the year of publication of the reference, such as Yang and Su (2014) or (Yang and Su, 2014). With three or more names use the form “Yang *et al.* (2016) or (Yang *et al.*, 2016)” in the text. When two or more references have the same text identification, please distinguish them by appending “a”, “b”, etc., to the year of publication. References should include the complete title(s), the number of the volume and the page(s) of the information cited. References should be listed in alphabetical order, according to the last name of the first author, at the end of the paper.

**Volume 31 No. 1**  
**March 2026**

# **Journal of Photogrammetry and Remote Sensing**

## **CONTENTS**

### **Regular Papers**

Integrated Optical–SAR Texture Analysis for Neural Network–Based Emergency Landslide Surveying and Mapping (Yi-Keng Chen, Shou-Hao Chiang) .....	<b>1</b>
Accuracy Assessment of Geoid Modeling at the Land-Sea Interface: A Case Study of Eastern Taiwan (Zhi-Yi Yang, Yu-Shen Hsiao).....	<b>21</b>
Real-time and Reference-free UAV Image Quality Assessment using Deep Learning (Ya-Li Lin, Guan-Chin Su, Lai-Han Zou, Chao-Hung Lin, Jiann-Yeou Rau, Wei-Shen Lai, Chih-Chao Hu) .....	<b>37</b>
Research on Depth-Enhanced Image Stitching for Concrete Bridge Structures (Yu-Ting Lin, Szu-Pyng Kao, Feng-Liang Wang, Jhih-Sian Lin).....	<b>47</b>

第 三 十 一 卷 第 一 期

中 華 民 國 一 一 五 年 三 月

# 航測及遙測學刊

## 目錄

### 學術論著

- 應用類神經網路整合光學-SAR 紋理分析於緊急崩塌測繪  
(陳以耕、姜壽浩) ..... 1
- 陸海交界處大地起伏建模精度探討：以臺灣東部為例  
(楊致逸、蕭宇伸) ..... 21
- 使用深度學習進行快速無參考標的無人機影像品質評估  
(林亞立、蘇冠秦、鄒來翰、林昭宏、饒見有、賴威伸、胡智超) ..... 37
- 應用深度資訊於混凝土橋梁結構物影像拼接之研究  
(林子婷、高書屏、王豐良、林志憲) ..... 47