# PE&RS

March 2023 Volume 89, Number 3

*The official journal for imaging and geospatial information science and technology*

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING

# PUBLISHING OPEN-ACCESS IN *PE&RS* IS NOW EASIER!

ASPRS is changing the subscription model of our monthly journal, *PE&RS*. ASPRS is waiving open-access fees for primary authors from subscribing institutions. Additionally, primary authors who are Individual Members of ASPRS will be able to publish one open-access article per year at no cost and will receive a 50% discount on open-access fees for additional articles.

- **Open Access matters!** By providing unrestricted access to research we can advance the geospatial industry and provide research that is available to everyone.
- **Institutions and authors receive more recognition!** Giving permission to everyone to read, share, reuse the research without asking for permission, as long as the author is credited.
- **Reputation matters!** Known for its high standards, *PE&RS* is the industry leading peer-review journal. Adding open access increases authors' visibility and reputation for quality research.
- **Fostering the geospatial industry!** Open access allows for sharing without restriction. Research is freely available to everyone without an embargo period.

*Under the previous subscription model, authors and institutions paid $1500 or more in open-access fees per article. This will represent a significant cost savings. Open-access publications benefit authors through greater visibility of their work and conformance with open science mandates of funding agencies.*

**Subscriptions** asprs.org/subscribe
**Membership** asprs.org/membership

asprs THE IMAGING & GEOSPATIAL INFORMATION SOCIETY

## ANNOUNCEMENTS

**GeoCue** announces the European launch of the 2023 LiDARental Program. This new rental program provides easy access to Innovative TrueView 3D Imaging Systems and the complete LP360 Drone data processing and visualization software.

Drone lidar has solidified a position as leading technology in the geospatial industry for surveying and mapping applications. However, for some new to drone lidar, the cost of lidar hardware, software, workflow, training, and support can be intimidating. The 2023 LiDARental Program provides easy access to the drone lidar/imagery arena.

GeoCue CEO, Frank Darmayan explains, "We understand that many surveying, engineering, and construction companies are eager to implement drone lidar, but hesitant to adopt the technology due to costs, support, software, and adaptation. That's why we've made it easier than ever to try this technology at low risk and low cost through our LiDARental Program."

The program includes training, support, data processing and visualization software through affordable weekly, monthly, quarterly, and annual rental plans. GeoCue understands that some projects take more or less time than others, therefore, rental periods are structured to meet your needs; as short as 1 week, up to 12 months. Rental also includes access to LP360 Drone processing software including Strip Align and Photo options for the rental period including one month after rental end date.

Participants in the 2023 LiDARental Program can choose from a full range of professional-grade geomatics sensors to select the right drone mapping solution for their project needs. Options include the TrueView 435, the TrueView 515, and the TrueView 660. Processing is included through LP360 Drone, transforming lidar and imagery data with visualization, quality check, classification, 3D editing, and analysis capabilities.

Conveniently, should a customer decide that a traditional purchase is more practical than renting, GeoCue can easily upgrade the plan to full ownership. The program is built around the needs of the customer. According to Vincent Legrand, Vice President of Sales for GeoCue, this new rental option will help open the drone lidar market to a wider range of geospatial companies. "Our goal is to meet customers on their terms with technology adoption, budget, and resources. The LiDARental Program gives them the option to affordably complete their projects with all the benefits of our TrueView drone lidar equipment and LP360 software split into convenient monthly payments."

For more information, contact GeoCue at https://geocue.com/.

A RIEGL VZ-400i was used in a project showing the workflow and results of lidar data acquisition on urban infrastructure in Santiago, Chile. The project was led by RIEGL distribution partner GEOCOM S.A. and supported by **RIEGL** USA's Regional Manager for LATAM, Claudio Avello.

Urban development entails the coexistence between existing elements and new constructions. The electricity transmission systems, visible through high-voltage power lines and manifested in the form of high-voltage poles, characterize the area and mark the public space. This is expressed on one side of the "Parque Bicentenario" (Bicentennial Park, municipality of Vitacura ), where along Avenida Alonso de Córdoba one can see the existing relationship between the urban environment and the overhead lines and poles that continue from their source at Santiago Municipal Park along Avenida Alonso de Córdova, which in turn merges with Avenida IV Centenario between the municipalities of Vitacura, Las Condes and La Reina.

In this type of infrastructure project, where there are urban elements, electrical infrastructure, framed in a complex topographic context such as hills and parks, it is possible to capture the reality through geospatial techniques such as lidar, using a RIEGL terrestrial laser scanner for acquisition of highly precise and highly informative data in extremely productive execution times.

For more information, visit https://newsroom.riegl.international/.

Bluesky Geospatial Ltd. has announced the launch of its MetroVista 3D aerial mapping program in the United States. The service employs a hybrid imaging-lidar airborne sensor to capture highly detailed 3D data, including 360-degree views of buildings and street-level features, in urban areas for use in creating digital twins, visualizations, and simulations.

Bluesky's sister company, Bluesky International of Leicestershire, England, developed and introduced the MetroVista program in the United Kingdom in 2018 and has mapped more than 20 cities across the England and Scotland in 3D. For the U.S. program, Boston is the first metropolitan area completed by Bluesky Geospatial, which will feature in project presentations this week at Geo Week 2023 in Denver.

"For many clients, the MetroVista 3D data set serves as the foundation for highly detailed digital twins at the heart of Smart City initiatives," said Bluesky Chief Commercial Officer, Ralph Coleman. "The rich information content of the 3D data enables accurate visualization of existing features for

informed decision making and realistic simulation of future development for comparative analysis."

The MetroVista product is already being used in city management and urban planning for agencies as diverse as emergency services, transportation, environmental protection, and utilities. Bluesky has seen rapid growth in the adoption of 3D data sets and digital twins in the AEC (architecture engineering construction), real estate, and insurance sectors.

For data collection, Bluesky flies the airborne Leica CityMapper-2 hybrid sensor which captures simultaneous oblique and nadir imagery along with lidar elevation measurements. Bluesky processes the data to create georeferenced 3D mesh models with 5-cm resolution and 10-cm accuracy. The 3D models are delivered in a variety of formats for direct ingest into GIS and CAD environments.

"The MetroVista 3D product allows managers and planners to measure, model, and analyze the terrain and features in their cityscapes with confidence – all without leaving the office," said Coleman.

The MetroVista U.S. program builds on a nearly 80-year legacy of delivering accurate, reliable and quality mapping projects under the Col-East name since 1946 and as part of the Bluesky brand since 2017. Along with its full line of aerial mapping services, Bluesky Geospatial will perform MetroVista city projects for clients in the Northeastern U.S. and beyond, adding to its archive of aerial imagery, photography, and lidar products.

Visit www.bluesky-world.us for more information.

Space Flight Laboratory (SFL) announced that ground control successfully established communication with three radio frequency (RF) geolocation microsatellites developed by SFL for HawkEye 360 of Herndon, Va. Cluster 6 was launched 24 January 2023 on the inaugural flight for Rocket Lab's Electron Rocket from Wallops Island, Va. This successful mission brings the number of HawkEye 360 microsatellites built by SFL and now in orbit to 18.

The HawkEye 360 Constellation detects and geolocates RF signals for maritime situational awareness, emergency response, national security, and spectrum analysis applications. To boost revisit rates over the mid-latitude regions of the globe, Cluster 6 was launched into an inclined orbit. Upon commissioning, HawkEye 360 will be able to collect RF data as frequently as every hour anywhere in the world.

HawkEye 360 selected SFL due to the importance of formation flying by multiple satellites for successful RF geolocation. Clusters 2, 3, 4 and 5 have all been built on SFL's space-proven 30 kg DEFIANT microsatellite bus. SFL's formation flying technology enables the Virginia company to offer the most timely and actionable RF data and data analytics available on the market.

"Cluster 6 launched as planned to support expansion of the HawkEye 360 constellation as it scales to meet the growing worldwide demand for its commercial RF data and analytics services," said SFL Director Dr. Robert E. Zee. "SFL's trusted attitude control and formation-flying capabilities deliver the stability and accuracy required for precise RF geolocation."

SFL is a unique microspace provider that offers a complete suite of nano-, micro- and small satellites – including high-performance, low-cost CubeSats – that satisfy the needs of a broad range of mission types from 3 to 500 kilograms. Dating from 1998, SFL's heritage includes 64 operational successes with more than 225 cumulative years of operation in orbit and 28 currently under construction or awaiting launch.

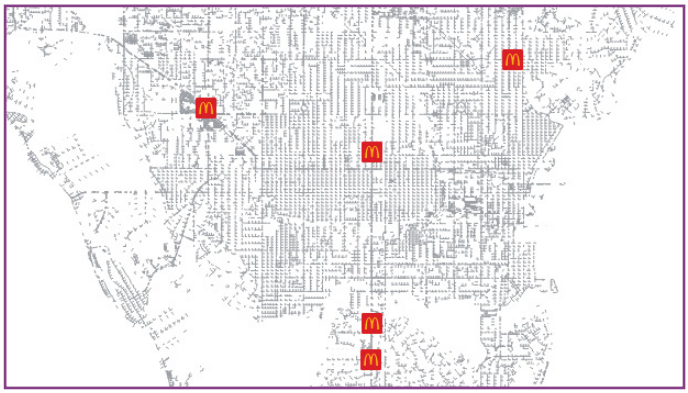
These missions relate to Earth observation, atmospheric monitoring, ship tracking, communication, radio frequency (RF) geolocation, technology demonstration, space astronomy, solar physics, space plasma, and other scientific research, utilizing core SFL-developed components such as modular (scalable) power systems, onboard radios, flight computers, and control software.

Visit www.utias-sfl.net for more information.

## CALENDAR

- 3-5 April, **RSCy2023**, Cyprus. For more information, visit www.cyprusremotesensing.com/rscy2023/ .
- 5 May, **ASPRS GeoByte — SeaSketch 2.0: A New, Free and Open Source software Service for Map-based Surveys and Collaborative Geodesign.** For more information, visit https://www.asprs.org/geobytes.html.
- 12-16 June, **ASPRS 2023 International Technical Symposium.** For more information, visit https://my.asprs.org/2023Symposium/

PE&RS
PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING
The official journal for imaging and geospatial information science and technology
March 2023 Volume 89 Number 3





## COLUMNS



## ANNOUNCEMENTS



## DEPARTMENTS





**See the Cover Description on Page 128**

# COVER DESCRIPTION

At roughly 325 square kilometers, the Ebro Delta on the northeastern coast of Spain is one of the largest wetlands along the Mediterranean Sea coast. It is an important habitat for wildlife, including flamingos and birds using the wetlands as a stopover on migratory journeys. The site in southern Catalonia has been designated a UNESCO Biosphere Reserve.

The 50-kilometer-long coastline features two sand spits: El Fangar on the north shore and La Banya on the south. These appendages are the remnants of the river's previous deltas, which were reworked when the river changed course over the past few thousand years.

The delta, which is home to 62,000 people, has also been greatly modified by human use. In the past 150 years, wetlands have been converted into fields of rice, which now cover up to 80 percent of the delta. To supply water for irrigation and to generate hydroelectric power, more than 187 dams have been built on Ebro River and its tributaries—development that trapped most of the sediment supply in Spain's largest river in reservoirs and behind dams. Erosion and land subsidence followed downstream.

The two natural-color images, acquired in 1984 (top image) and 2021 (bottom image), by Landsat 5 and Landsat 8 (respectively), show the erosion of the triangular island at the mouth of the Ebro River near Riumar has retreated by several hundred meters. Note that the differences in color between the images could be attributed to differences in the satellite sensors, changes in the landscape, and differences in the timing of tides.

Today, the shape and form of the delta is no longer controlled by the river, but by sea waves. And with sea-level rise and more frequent and intense storms, those waves are getting bigger, leading to further shoreline retreat. In January 2020, the narrow sandbar that connects the southern spit to the main delta was flooded by storm Gloria, along with 3,000 hectares of rice fields. Storms also exacerbate the shrinking and loss of dune fields on the beaches.

The Ebro Delta illustrates the hard choices to come for communities facing rising seas—try to hold back the ocean or manage the retreat.

The Spanish government recently announced a plan to buy coastal land to create a buffer zone. If the plan is adopted, the purchase would constitute the largest land buyout in Europe so far due to climate change. But it is opposed by many of the delta's inhabitants, some of whom instead favor beach nourishment, pumping, and seawalls to protect the coast. Some farmers are experimenting with strains of rice that can better withstand saltwater intrusion.

For more information, visit https://landsat.visibleearth.nasa.gov/view.php?id=149170.

NASA Earth Observatory images by Lauren Dauphin, using Landsat data from the U.S. Geological Survey. Story by Sara E. Pratt.




asprs
THE IMAGING & GEOSPATIAL INFORMATION SOCIETY

## PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING

JOURNAL STAFF

Publisher ASPRS
Editor-In-Chief Alper Yilmaz
Director of Publications Rae Kelley
Electronic Publications Manager/Graphic Artist Matthew Austin

*Photogrammetric Engineering & Remote Sensing* is the official journal of the American Society for Photogrammetry and Remote Sensing. It is devoted to the exchange of ideas and information about the applications of photogrammetry, remote sensing, and geographic information systems. The technical activities of the Society are conducted through the following Technical Divisions: Geographic Information Systems, Photogrammetric Applications, Lidar, Primary Data Acquisition, Professional Practice, Remote Sensing Applications, and Unmanned Autonomous Systems. Additional information on the functioning of the Technical Divisions and the Society can be found in the Yearbook issue of *PE&RS*.

All written correspondence should be directed to the American Society for Photogrammetry and Remote Sensing, PO Box 14713, Baton Rouge, LA 70898, including general inquiries, memberships, subscriptions, business and editorial matters, changes in address, manuscripts for publication, advertising, back issues, and publications. The telephone number of the Society Headquarters is 301-493-0290; the fax number is 225-408-4422; web address is www.asprs.org.

**PE&RS.** *PE&RS* (ISSN0099-1112) is published monthly by the American Society for Photogrammetry and Remote Sensing, 8550 United Plaza Blvd, Suite 1001, Baton Rouge, Louisiana 70809 . Periodicals postage paid at Bethesda, Maryland and at additional mailing offices.

**SUBSCRIPTION.** *PE&RS* is available as an e-Subscription (single-site and multi-site licenses) and an e-Subscription with print add-on (single-site license only). PE&RS subscriptions are on a calendar-year, beginning in January and ending in December.

The rate for a single-site e-Subscription for the USA/Non-USA is $1040 USD, for Canadian* is $1092 USD.

The rate for a multi-site e-Subscription for the USA/Non-USA is $1040 USD plus $250 USD for each additional license, for Canadian* is $1092 USD plus $263 for each additional license.

The rate for e-Subscription with print add-on for the USA is $1525 USD, for Canadian* is $1612 USD, and for Non-USA is $1565 USD.

*Note: Subscription prices for Canada includes 5% of the total amount for Canada's Goods and Services Tax (GST #135123065). **PLEASE NOTE: All Subscription Agencies receive a 20.00 USD discount.**

**POSTMASTER.** Send address changes to PE&RS, ASPRS, PO Box 14713, Baton Rouge, LA 70898. CDN CPM #(40020812).

**MEMBERSHIP.** Membership is open to any person actively engaged in the practice of photogrammetry, photointerpretation, remote sensing and geographic information systems; or who by means of education or profession is interested in the application or development of these arts and sciences. Membership is for one year, with renewal based on the anniversary date of the month joined. Membership Dues include a 12-month electronic subscription to *PE&RS*. Annual Individual Membership dues are $150.00 USD and Student Membership dues are $50.00 USD. A tax of 5% for Canada's Goods and Service Tax (GST #135123065) is applied to all members residing in Canada.

# TECHNOLOGY CHANGES DURING MY 60-YEAR MAPPING CAREER

BY DR. DAVID F. MAUNE, COLONEL USA (RET)

I am retiring at the end of 2022, after two wonderful 30-year careers in the mapping sciences – 30 years as a topographic engineer officer in the U.S. Army Corps of Engineers (USACE) (Figure 1) and 30 years as a geospatial senior project manager with Dewberry (Figure 2). Both careers included photogrammetric and topographic mapping and production of Digital Elevation Models (DEMs) – my specialty, for which I served as editor/co-author of three editions of the DEM Users Manual (Figure 3).

Figure 1. Colonel in the U.S. Army Corps of Engineers, 1961-1991. Source: Personal image.

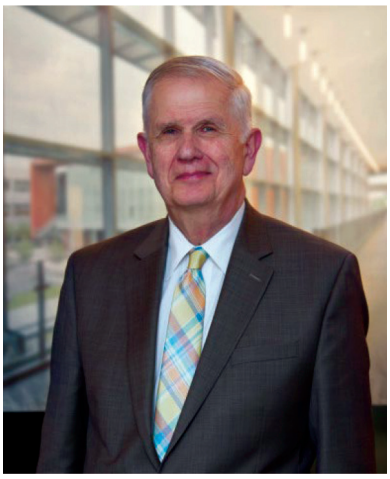


Figure 2. Senior Project Manager at Dewberry, 1992-2022. Source: Personal image.

In addition to Army schools, assignments to topographic engineering battalions in Germany and Hawaii, the Army sent me to The Ohio State University (OSU) to get M.Sc. and PhD degrees in geodetic science and photogrammetry. Prior to the advent of GPS, we were taught how to use T-3 theodolites, star catalogs and precise celestial navigation techniques to determine latitude and longitude; it took me weeks to complete all calculations and adjustments for a single position; I knew there had to be a better way. Similarly, I studied analog and analytical photogrammetry, again knowing there had to be a better way. My PhD dissertation on photogrammetric self-calibration won ASPRS' Talbert Abrams Grand Award in 1976 for what was then considered to be pioneering research in digital photogrammetry. Self-calibration has evolved into today's Structure from Motion (SfM) photogrammetry.

As an Army topographic engineer, we helped to map allied countries that asked for America's help in mapping for nation building, and we mapped and performed terrain analyses of countries where we might potentially go to war. I retired from the Army in 1991 as the Commander and Director, U.S. Army Topographic Engineering Center (TEC) – now the U.S. Army Geospatial Center (AGC) – where we developed many mapping technologies in common use today.

While at Dewberry from 1992 to the present, I'm best known as Project Manager for our major geospatial contracts with the U.S. Geological Survey (USGS) and the National Oceanic and Atmospheric Administration (NOAA) Office for Coastal Management (OCM) and National Geodetic Survey (NGS). I also supported the Federal Emergency Management Agency (FEMA) for the National Flood Insurance Program (NFIP) and emergency response contracts. All these contracts gave me experience with lidar and other elevation technologies and applications which led me to serve as editor and principal author of three editions of *Digital Elevation Model*

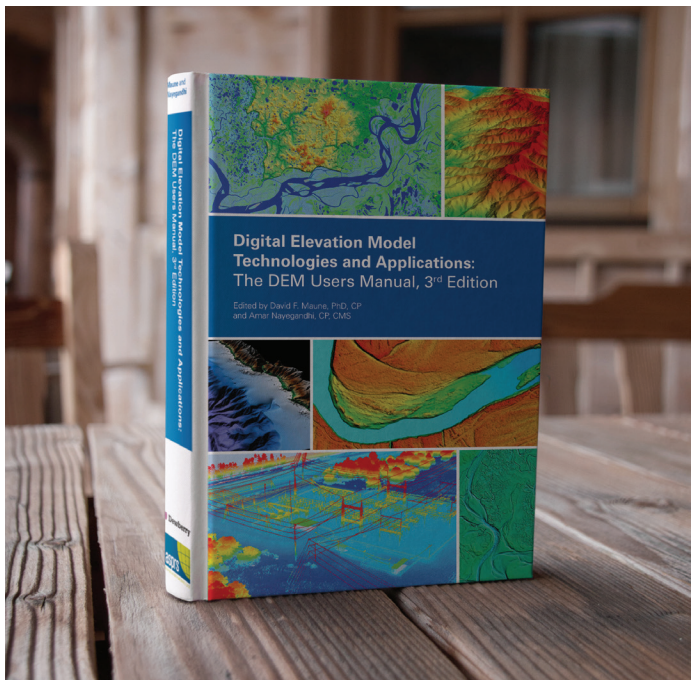



Figure 3. Editor of three editions of the *DEM Users Manual*, 2001, 2007, 2018. Source, Dewberry.

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 3, March 2023, pp. 129–142.
0099-1112/22/129–142

doi: 10.14358/PERS.89.3.129

*Technologies and Applications: The DEM Users Manual* (Figure 3), published by ASPRS.

I have been intrigued by the changes in surveying and mapping technologies during my 60-year career, summarized in the ten technology sections below. I do not look back to the "good old days" but cherish the progress made during my 60 years as a mapper.

# CHANGES IN CARTOGRAPHY

In his testimony to Congress on December 5, 1884, John Wesley Powell, 2nd Director of the USGS, stated: "A Government cannot do any scientific work of more value to the people at large than by causing the construction of proper topographic maps of the country." Until the 21st century, USGS' 7.5-minute topographic quad maps (Figure 4) were America's standard mapping product. U.S. military organizations produced similar topographic maps of foreign countries at different scales. Cartographic features are either mapped as points, lines, or polygons. "Feature separates" are merged into "color separates" for each color used on the printed map, normally black, red, blue, green, and brown, and sometimes magenta to show map updates.

For vector "feature separates", scribing was used to produce lines for cartographic map compilations before the use of computer-based geographic information systems (GIS). Lines produced by manual scribing are sharp, clear and even. Using large walk-in cameras (Figure 5), pencil manuscripts were photographed onto scribe sheet material. Using a light table (Figure 6), lines on the scribe sheet were traced with a metal scribe tool to remove thin lines of translucent coating to produce a negative image (compared with drafting that produces a positive image). Scribing produced a result superior to drafting, but it is more time-consuming. A separate stylus was required for each thickness of line required, and some were used for parallel double-line road casings. Scribing was so time-consuming that a cartographer could spend a year or more scribing linear features for a single map sheet.

Text was set by using a Leroy lettering set, a popular mechanical lettering template used by cartographers and draftsmen; laminated plastic templates had characters engraved on the front, in different sizes, with the lines serving as guide grooves for the cartographer or draftsman to ink letters, numbers, and characters consistently.

Using "peel coats" to open windows for polygons, area features were screened with patterns to depict swamps, forests, lakes, built-up areas, etc., or to lighten colors by dotted screens with differing amounts of white between the colored dots. For example, a river centerline might be printed with dark blue


Figure 4. Topographic quadrangle maps produced by USGS until about 2002. Source: USGS.

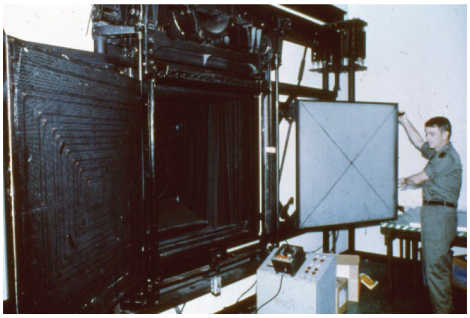


Figure 5. Walk-in mapping camera used by my topographic engineer battalion in Hawaii to produce the L653 series of topographic maps for U.S. Forces, Korea. Personal image.

Figure 6. Although these scribed lines are thick for visualization purposes, most scribed lines are a small fraction of a mm thick and require a steady hand scribing beneath a microscope. Personal image.

ink, but a lake or double-line stream would be screened so that the same dark blue ink appears to be light blue on the map.

Army topographic engineer units did much more than produce topographic maps. My topographic engineer battalion in Germany had three terrain intelligence detachments whose mission was to perform terrain analyses of countries where NATO units may go to war. Soldiers would need to determine the bridge-bearing capacity of bridges in potential enemy territory, the soil-bearing capacity for armored units travelling off-road, and obstacles in moving through cities, towns, and villages. Soldiers would need to know potential parachute landing zones, landing beaches, obstacles to cross-country movement, areas for cover and concealment, etc. My battalion produced such map-based terrain analysis studies of eastern European countries, and most of those studies were either Confidential or Secret. Thus, security of printed terrain analysis studies in our map warehouses was always a challenge.

GIS is like cartography but with key differences. Both include base maps to which additional data are added, but a GIS has no limit on the amount of supporting data that can be added, including massive geodatabases for example, and a GIS can automate analyses of connectivity, adjacency, and/or proximity. National mapping agencies, including the Defense Mapping Agency (DMA) and USGS, started adopting GIS technology and best practices in the 1980s, but they also continued using traditional cartography until the turn of this century.

Soldiers now use computers in the field with GIS software for visualizing the terrain and performing geospatial intelligence tasks in combat brigades and divisions. Most data come from the National Geospatial-Intelligence Agency (NGA), but Army geospatial intelligence analysts perform on-demand terrain analysis tasks in combat zones, tailored to satisfy tactical requirements.

For the National Map (TNM) in the U.S., US Topo is USGS' current topographic map series, modeled on the legacy 7.5-minute topographic quad maps but mass-produced quickly from GIS databases and published as digital documents. US Topo has various digital layers that overlay and can be turned on and off, including elevation, imagery, hydrography, transportation, structures, land cover, boundaries, and geographic names.

The *PE&RS GIS Tips and Tricks* column by Dewberry Senior GIS Professional Al Karlin routinely stimulates thinking on the evolving use of GIS technology.

# CHANGES IN MAP PRODUCTION AND DISTRIBUTION

Five-color topographic maps were prepared from feature separates, then merged into five color separates. For example, the black color separate might consist of (1) road vectors scribed with single or double-lines for different types of roads, (2) scribed buildings to be mapped in black, (3) black text throughout the map containing street names, highway numbers, city/town names, etc. (4) open-window "peelcoats" for screening with dots to create different shades of gray on the map; (5) map graticule, and (6) map marginalia including scale bar; north arrow/magnetic declination diagram; legend explaining the various colors and symbols used, accuracy information, and information about the map that today is called metadata. All six of these black feature separates would be photo-copied onto a black color separate, a film negative, that would subsequently be burned onto a press plate for that color. Similar color separates would be prepared for all features to be mapped with red, blue, green, or brown ink. Color proofs would then be made for quality control to see if any color incorrectly overlays another color. When the color-proofing process indicated everything was within specifications, only then could the press plates be sent for printing on single- or multi-color presses. With single-color presses, the entire system would be inked separately for the five different press plates – the paper running through the press five different times with different colors.

Map paper cutting and trimming was also a science. Blank paper map sheets need to be trimmed prior to printing to ensure that edges are perfectly square and straight, to avoid press jams and other mechanical press problems. In binding and finishing operations, cutting and trimming are performed to reduce large-size press sheets to the desired trim size and to remove extraneous edges containing registration marks. Most cutting and trimming are performed on a powerful guillotine cutter, a large device consisting of a flat bed on which the paper is stacked, and a wide, sharp steel or steel-carbide knife, which is lowered through the paper mechanically. Side and back gauges on the cutter bed also helped position the paper accurately and squarely, allowing for the ability to trim to a very accurate size. A cutter clamp holds the paper securely beneath the knife and expels air from the stack of sheets, eliminating distortion of sheets which can result in improper cutting. Cutting and trimming were also performed using cropmarks, lines in the trim area of the sheets which indicate the proper size of the finished stock. When stacks of maps are properly trimmed, automated counting machines could be used for accurate inventory audits.

These printing and finishing processes continued, worldwide, until the early 2000s when maps became digital.

During most of my 30-year Army career, topographic engineer battalions had Areas of Responsibility for which they maintained map reproducibles for all maps of countries for which they might need to provide maps to combat units; and they also maintained map warehouses for storing and shipping maps to users. Some maps were prepositioned in caves or bunkers in potential combat zones. The Army Map Service (AMS), and subsequently the Defense Mapping Agency (DMA), provided these services worldwide; and USGS did so nationwide with their topographic quadrangles. These

larger agencies all maintained their libraries of map reproducibles and map warehouses. To the best of my knowledge, such brick and mortar facilities are no longer needed with the advent of digital mapping products.

Figures 7 and 8 show the map reproducible library and one of our printing presses in my topographic engineer battalion in Hawaii. Figure 9 shows a row of map warehouses in my topographic engineer battalion in Germany where trucks were coming and going daily, moving freshly printed maps in from the printing presses and moving new maps out to the combat units throughout Germany. Of course, as truckloads of new maps were delivered, truckloads of old maps were continuously being returned and normally destroyed. In later years, we reused old map stock by printing lines on the reverse side so recycled maps could be guillotined into 8.5"x11" writing pads.

Today, the Army no longer has topographic engineer battalions supporting Army potential theaters of operation; instead, many forms of imagery and digital data are analyzed by geospatial-intelligence specialists at Army division and brigade level, using computers with specialized GIS software and various forms of digital datasets to perform terrain analyses in support of the Army doctrine known as Intelligence Preparation of the Battlefield (IPB).

Figure 7. Map feature separates and color separates maintained by my topographic engineer battalion in Hawaii. Personal image.

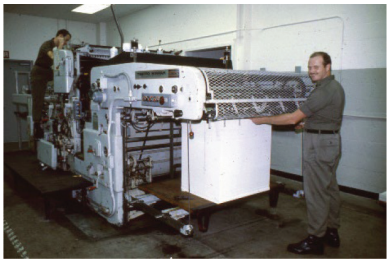


Figure 8. Printing press at the base plant operated by my topographic engineer battalion in Hawaii. Personal image.

Figure 9. Map production & warehouse buildings (four large buildings in the rear) in my topographic engineer battalion in Germany. Personal image.

# CHANGES IN MAP ACCURACY STANDARDS

The N*ational Map Accuracy Standards* (NMAS) of 1947 pertained to graphic contour maps with a published scale and contour interval, defining horizontal and vertical accuracy as follows:

- Circular Map Accuracy Standard (CMAS): "For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale; for maps on publication scales of 1:20,000 or smaller, 1/50 inch. These limits of accuracy shall apply in all cases to positions of well-defined points only."
- Vertical Map Accuracy Standard (VMAS): "Vertical accuracy, as applied to contour maps on all publication scales, shall be such that not more than 10 percent of the elevations tested shall be in error more than one-half the contour interval. In checking elevations taken from the map, the apparent vertical error may be decreased by assuming a horizontal displacement within the permissible horizontal error for a map of that scale."

The NMAS had no limits on the magnitude of errors for the 10 percent outliers.

In 1990, ASPRS published its Accuracy Standards for Large-Scale Maps, again focused on printed maps:

- "Horizontal map accuracy is defined as the root-mean-square (rms) error in terms of the project's planimetric survey coordinates (X, Y) for checked points as determined at full (ground) scale of the map. The rms error is the cumulative result of all errors including those introduced by the processes of ground control surveys, map compilation and final extraction of ground dimensions from the map. The limiting rms errors are the maximum permissible rms errors established by this standard." The limiting rms errors for Class 1 maps were tabulated in tables, along with typical map scales associated with the limiting errors.
- "Vertical map accuracy is defined as the rms error in elevation in terms of the project's elevation datum for well-defined points only. For Class 1 maps the limiting rms error in elevation is set by the standard at one-third the indicated contour interval for well-defined points only. Spot heights shall be shown on the map within a limiting rms error of one-sixth of the contour interval."
- Class 2 and Class 3 maps could have errors 2 or 3 times larger, respectively, than Class 1 maps.

In 1998, the Federal Geographic Data Committee (FGDC), assuming all mapping errors follow a normal error distribution, published the National Standard for Spatial Data Accuracy (NSSDA), specifying that horizontal and vertical errors should be reported at the 95% confidence level, based on RMSEx and RMSEy, translated into radial RMSEr for horizontal accuracy, and RMSEz for vertical accuracy.

- Horizonal accuracy at the 95% confidence level (ACCURACYr) = 1.7308 × RMSEr
- Vertical accuracy at the 95% confidence level (ACCURACYz) = 1.9600 × RMSEz

I subsequently performed extensive research into lidar errors and determined that: (1) lidar bare-earth DTM errors in vegetated terrain do not follow a normal error distribution; (2) the use of RMSEz in vegetated terrain significantly overstates the vertical errors; and (c) vertical errors in vegetated terrain should be defined in terms of the 95th percentile, rather than RMSEz. This resulted in publication of the National Digital Elevation Program (NDEP) *Guidelines for Digital Elevation Data*, Version 1.0, as well as the *ASPRS Guidelines, Vertical Accuracy Reporting for Lidar Data*, Version 1.0, both published in 2004.

In 2014, the *ASPRS Positional Accuracy Standards for Digital Elevation Data* were published. I chaired the committee that also included Dr. Qassim Abdullah, Karl Heidemann, and Doug Smith. These new standards replaced the existing *ASPRS Accuracy Standards for Large-Scale Maps* (1990) and the *ASPRS Guidelines, Vertical Accuracy Reporting for Lidar Data* (2004) to better address current digital mapping technologies. Map accuracy classes 1, 2, and 3 no longer exist.

- Recognizing that many applications of horizontal accuracy cannot be tied directly to compilation scale, resolution of digital source imagery, or final pixel resolution, and that geospatial data does not suddenly get more accurate just because an analyst on a computer can display digital data at higher resolution, horizontal accuracy is defined in terms of horizontal accuracy classes based on RMSEx and RMSEy, from which RMSEr and horizontal accuracy at the 95% confidence level can be computed. Tables included horizontal accuracy classes between 0.63 cm and 10 meters. The new standard also specifies the allowable size of orthoimagery mosaic seamline mismatches, in terms of horizontal accuracy classes for RMSEx and RMSEy.
- Vertical accuracy is computed using RMSEz statistics in non-vegetated terrain and 95th percentile statistics in vegetated terrain. Tables included vertical accuracy classes between 1 cm and 3.33 meters and included standards for Non-Vegetated Vertical Accuracy (NVA) at the 95% confidence level and Vegetated Vertical Accuracy (VVA) at the 95$^{th}$ percentile.

The 2014 standards provided additional guidance on checkpoint density and distribution, accuracy reporting, designation of low confidence areas, and other factors.

# CHANGES IN SURVEYING AND GEODESY

At Ethiopia's request, my topographic engineer battalion in Germany was tasked to establish geodetic control monuments in Ethiopia in the 1960s as the foundation for topographic mapping. Like national mapping agencies worldwide, Army surveyors used T-3 theodolites to measure vertical and horizontal angles for triangulation from Bilby towers, also used by the U.S. Coast and Geodetic Survey (USC&GS) through the 1980s (Figures 10 and 11). We used surveyor tapes to measure distances prior to the introduction of electronic distance measuring equipment (EDME). We used Bilby towers to see above obstacles and establish line-of-sight to survey targets at long distances; on a clear day, T-3 theodolites could accurately measure horizontal and vertical angles to survey targets over 50 miles away.

Bilby towers had two unconnected parts – an internal tower for mounting surveying instruments and an external tower for surveyors. This separation allowed for isolating the instruments from vibrations caused by people, increasing the precision of measurements. These survey techniques formed the backbone of America's spatial reference framework. Military and civilian surveyors also use(d) differential leveling (Figure 12) to measure vertical offsets between two points to transfer an elevation from a benchmark (BM, with known elevation) to another point (unknown elevation) by a series of

Figure 10. Bilby tower with internal tower (for instrument) isolated from external tower (for surveyor). NOAA Photo Library.

Figure 11. T-3 triangulation measurement from a USC&GS Bilby tower. NOAA Photo Library.

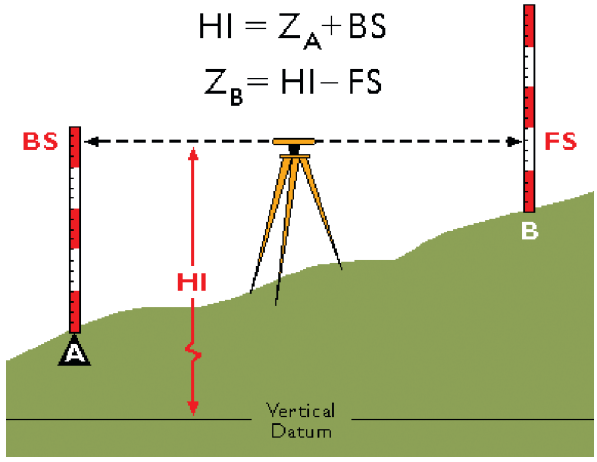



Figure 12. Differential leveling used levels that took foresight (FS) and backsight (BS) elevation differences measured on marked survey leveling rods. NOAA Photo Library.

foresights (FS) and backsights (BS). It took considerable time and expense to establish horizontal monuments and vertical benchmarks nationwide using such techniques.

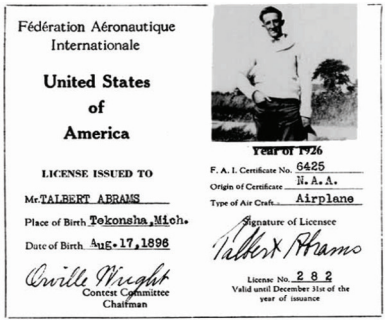
The Department of Defense's NAVSTAR, now the Global Positioning System (GPS), became fully operational in 1995, though originally intended for military purposes. In 1998, under contract with NGS, I authored the *National Height Modernization Study: Report to Congress* and documented the costs and benefits of modernizing the national height system in the U.S. based on differential GPS measurements relative to Continuously Operating Reference Stations (CORS). Initially seen as a quick way to determine horizontal positions accurately, NGS proved that accurate elevations could also be obtained from high quality GPS receivers and rigorous procedures. Today, following procedures in NOAA Technical Manual NOS NGS-58, *Guidelines for Establishing GPS-Derived Ellipsoid Heights (Standards: 2cm and 5cm)*, GPS is routinely used to transfer elevations from the nearest CORS to local survey points to 2 or 5 cm at the 95% confidence level, negating the need for benchmarks which are subject to subsidence and may otherwise be unstable. GPS revolutionized the surveying and geodesy professions. Airborne GPS is also vital for all types of aerial surveys.

Readers are invited to read "The Evolution of GPS" by Adam Goetsch, at https://illumin.usc.edu/the-evolution-of-gps/

# CHANGES IN AERIAL AND SATELLITE IMAGERY


Figure 13. Ted Abrams founded Abrams Aerial Surveys and designed his plane so photos would not be fogged by engine smoke. Source: The Abrams Foundation.

Figure 14. In 1977, Dr. Abrams presented me the Talbert Abrams Grand Award for my research in photogrammetric self-calibration. Source: Personal image.


Fédération Aéronautique Internationale

United States of America

LICENSE ISSUED TO

Mr. TALBERT ABRAMS

Place of Birth Tokonsha, Mich.

Date of Birth Aug. 17, 1898

Orville Wright
Contest Committee Chairman

Year of 1926

F. A. I. Certificate No. 6425

Origin of Certificate N. A. A.

Type of Air Craft Airplane

Signature of Licensee

Talbert Abrams

License No. 282

Valid until December 31st of the year of issuance

Figure 15. Dr. Abrams bought me a beer and showed me his pilot's license signed by Orville Wright. Source: The Abrams Foundation.

Dr. Talbert "Ted" Abrams was named the *Father of Aerial Photography* for his innovations in aerial photography. As a Marine Corps aerial photographer in Germany during WWI, he took pictures over the side of an open airplane and knew there had to be a better way. He founded Abrams Aerial Surveys and designed an airplane (Figure 13) so that the engine smoke did not obscure aerial images taken with a mounted camera. Film cameras, normally with 6" focal length, acquired 9" × 9" aerial film negatives for decades until large format metric digital cameras were introduced in the early 2000s. He also founded ASPRS' Talbert Abrams Award. When I won his top award in 1977 for my research in photogrammetric self-calibration (Figure 14), he showed me his pilot's license signed by Orville Wright (Figure 15).

I was introduced to satellite imagery in the early 1970s when the US military used Hexagon KH-9 reconnaissance satellites (Figure 16) to map countries with or without their knowledge or consent. I was the Officer-in-Charge of Production for a NATO unit responsible for mapping the Soviet Union, and I was the "sanitation board" authority for determining the security classification of NATO products


Figure 16. Now on display at the National Museum of the United States Air Force near Dayton, OH, the KH-9 Hexagon reconnaissance satellite was declassified on 9/17/2011. At 60-foot length and weighing 15 tons at launch, 19 KH-9 satellites were launched between 1971 and 1986. Panoramic images were film, not digital. Personal image.

produced from Top Secret imagery. The KH-9 used film, not digital imagery. Operators ejected the undeveloped film towards earth in reentry vehicles deployed by parachutes. When the object entered the upper atmosphere, the parachute would open and was then "snatched" mid-air by an

airplane, with a hooking apparatus beneath the plane, sent to the parachute's expected point of entry. This seemingly difficult feat was remarkably successful, with only one reentry vehicle lost in 15 years of operation. For photogrammetric compilation, the KH-9 panoramic images required rectification to remove distortions caused by tilt. During the Cold War, 19 Hexagon missions imaged 877 million square miles of the earth's surface between 1971 and 1986.

NASA's Earth Resources Technology Satellite (ERTS) was launched July 23, 1972, collecting digital multispectral imagery. Later renamed, Landsat 1 became the first earth-observing satellite explicitly designed to study planet earth. In 1993, the U.S. Department of Commerce granted DigitalGlobe the first license for private enterprise to build and operate a satellite system to gather high-resolution digital imagery of earth for commercial sale. Today, there are dozens of options for collecting aerial or satellite panchromatic, natural-color, multispectral, hyperspectral, or radar digital imagery optimized for a large variety of user applications.

In 2001, Leica introduced the first large format, calibrated digital mapping camera with its ADS40 pushbroom camera; Dewberry was the first to produce digital orthophotos for USGS using a digital mapping camera. Other calibrated metric digital mapping cameras soon followed. All digital mapping cameras have continued to improve to this day, widely used for federal, state, and local mapping projects. With the recent popularity of Structure from Motion (SfM) photogrammetry, small-format consumer-grade non-calibrated cameras are now used for image acquisition for small mapping projects where redundant observations by multiple look angles allow for camera self-calibration.

The *PE&RS Mapping Matters* monthly column by Dr. Qassim Abdullah, has chronicled the recent advances in aerial imaging and photogrammetry.

# CHANGES IN PHOTOGRAMMETRY

The May 2021 issue of *PE&RS* included my Tips & Tricks article on aerial triangulation over the years with what I call four generations of photogrammetry: analog, analytical, digital, and SfM. I first learned to be a photogrammetrist using first-generation analog stereo plotters, which attempted to physically replicate the geometry that existed when aerial film photos were taken. Stereoplotters used glass stereo diapositives to compile topographic maps.

The Army trained me to use the Multiplex (Figure 17), which had a series of projectors with reduced-scale 2" × 2" diapositives in individual projectors for each photo in a flight line. The long projector bar could be lengthened with more projectors added for longer flight lines. Based primarily on optics, Multiplexes were still used by the U.S. Army in the 1960s, unchanged from what was used during WW II. Subsequently, Kelsh Plotters were widely used with full-size 9"x9" diapositives. When I went to OSU, I learned analog photogrammetry on the Wild A-7 (Figure 18), an optical-mechanical stereo plotter with hand cranks to move in *x* and *y* directions and a foot petal to change elevations for contouring; by keeping a "floating dot" on the ground for a set elevation, the photogrammetrist would trace a contour line of equal elevation. We also learned 2nd-generation analytical photogrammetry which mathematically replicated the physical geometry when stereo hardcopy photos were taken by metric film cameras. My PhD dissertation was on photogrammetric self-calibration, which became relevant with the latest 4$^{th}$-generation SfM photogrammetry that does not require calibrated metric cameras. Dewberry has used SfM on several Alaska airfield mapping projects.

As fate would have it, I became heavily involved with the development of third-generation digital photogrammetry. As Commander and Director of the U.S. Army Engineer Topographic Laboratories (ETL) and Topographic Engineering Center (TEC) between 1988 and 1991, my organization developed the first Digital Stereo Photogrammetric Workstation (DSPW), now marketed as SocetSet, as well as the first high-resolution scanner to convert film images into

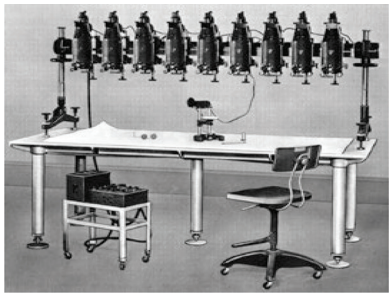


Figure 17. Multiplex projectors where I first learned to perform stereo photogrammetric map compilation in the Army. This was the easiest to visualize how aerial geometry was replicated at reduced scale to perform relative and absolute orientation. Source: US Army.

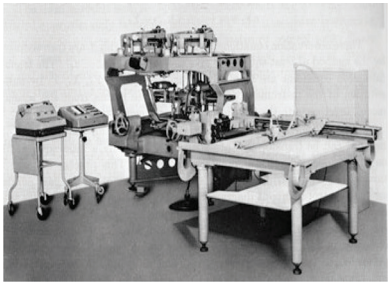


Figure 18. Wild A-7 stereo plotter with drafting table that I used to study photogrammetry at OSU. Here, only one stereo pair at a time underwent relative orientation. Although this had advantages, absolute orientation was harder to perform after bridging between multiple stereo pairs, one at a time. Source: Personal image.

digital images, subsequently marketed as PhotoScan-1. These two systems were part of the Army's Terrain Information Extraction System (TIES) introduced in 1991. Dewberry now uses SocetSet as our prime digital photogrammetric software. Thus, in one way or another, I have worked with all four generations of photogrammetry during my 60-year career.

Founded in 1934 as the American Society of Photogrammetry (ASP), and renamed the American Society for Photogrammetry and Remote Sensing (ASPRS) in 1985, the society has now published six editions of the *Manual of Photogrammetry* as well as many other publications dealing with various forms of remote sensing, imaging, and geospatial information, as well as the *ASPRS Positional Accuracy Standards for Digital Geospatial Data*.

# CHANGES IN RADAR

During my Army career, we used many forms of aerial and satellite radar systems for mapping and intelligence purposes. Several Synthetic Aperture Radar (SAR) exploitation systems were developed by ETL/TEC. During the first Gulf War in 1991, General Normal Schwarzkopf named several of our Army SAR systems as being instrumental in his ability to see the total battlefield in Iraq better than Saddam Hussein could see in his own back yard, or words to that effect. Of course, the main advantage of radar is that it enabled allied forces to perform our mapping and surveillance operations in all weather conditions, as radar maps through clouds, fog, and haze.

During my Dewberry career, one of my major achievements was in mapping all of Alaska with aerial Interferometric Synthetic Aperture Radar (IfSAR). I chose this technology because it mapped through clouds, a persistent problem in Alaska that, until 2008, had prevented Alaska from being mapped to established mapping standards at any scale. Figure 19 shows the advantage of aerial IfSAR in Alaska that also had the advantage of showing hydrographic features loud and clear. These IfSAR datasets are now used by Dewberry and others for Elevation Derived Hydrography (EDH) of Alaska.

Today, SAR satellites are extremely common for continental-scale mapping and change detection, and Differential Interferometric Synthetic Aperture Radar (DInSAR) is ideal for mapping the annual rates of land subsidence that compounds the effects of sea level rise worldwide.

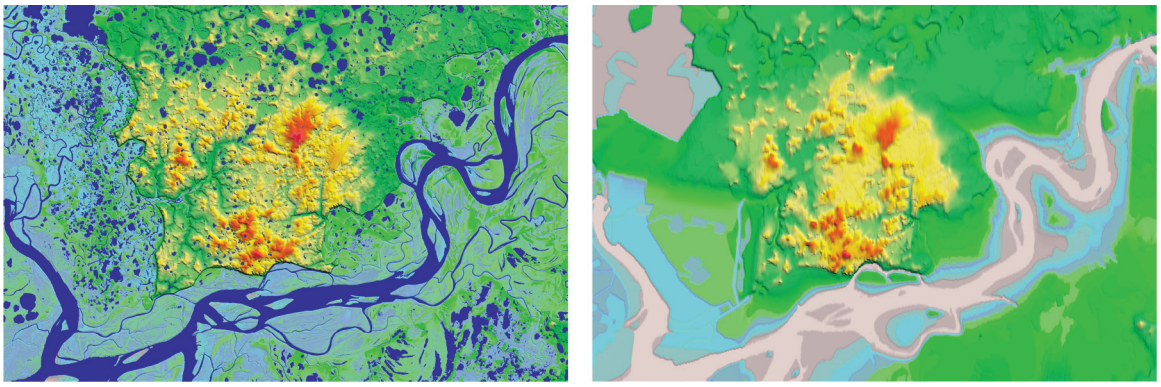


Figure 19. The aerial IfSAR Digital Terrain Model (DTM)(left) was vastly superior to the prior photogrammetric DTM (right) produced from satellite imagery for the National Elevation Dataset (NED). Source: Intermap

# CHANGES IN LIDAR

In 1997, for the Federal Emergency Management Agency (FEMA), I evaluated the use of lidar and IfSAR for floodplain mapping and modeling for the National Flood Insurance Program (NFIP), and I wrote all of FEMA's lidar guidelines and specifications between 1997 and 2010 when USGS published its first Lidar Base Specifications. In 1998, for the National Geodetic Survey (NGS), I authored the *National Height Modernization Study: Report to Congress* on how to modernize the National Height System in the U.S. based on GPS surveys relative to CORS (in lieu of differential leveling), and nationwide elevation mapping with lidar and IfSAR (in lieu of photogrammetric mapping of DEMs).

In 2000, ASPRS asked me to write a book on lidar and IfSAR, and I agreed to edit and co-author the first edition of *Digital Elevation Model Technologies and Applications: the DEM Users Manual* (published in 2001) with chapters on

photogrammetry, IfSAR, topographic lidar, bathymetric lidar, and sonar, as I've always visualized a worldwide DEM from the tops of the mountains to the depths of the seas, including inland bathymetry.

Whenever I autograph a copy of the DEM Users Manual, I write "May all your DEMs come true!" When the second edition was published in 2007, I had three basic dreams: (1) Development of high-accuracy, affordable elevation technologies for betterment of society; (2) Development of DEM technology standards, guidelines and specifications, and (3) Implementation of a nationwide program to produce and maintain standardized high-quality DEMs used by all. These three dreams have largely been realized and documented in the third edition published in 2018. My lidar dreams for the future are documented in the third edition.

Although I had almost nothing to do with the development of lidar or IfSAR (my dream #1), I was the major champion of lidar for 20 years and ended up being named the Father of Lidar by the International Lidar Mapping Forum (ILMF) and *LiDAR Magazine* in 2018 for my roles in dreams #2 and #3, having authored the major lidar standards, guidelines and specifications, as well as the National Enhanced Elevation Assessment (NEEA) that led directly to today's 3D Elevation Program (3DEP).

The 3DEP is widely heralded as a major success, having acquired Quality Level 2 (QL2) or better topographic lidar for most of the country, and QL5 IfSAR of Alaska. The map at Figure 20 shows the geographic extent of completion of the first-ever national baseline of consistent high-resolution elevation data – both bare earth and lidar point clouds – as of 2022.

Recognizing the importance of dual-frequency (red/green) topographic-bathymetric (topobathy) lidar, my boss, Amar Nayegandhi, has been an industry leader and authored the topographic lidar and bathymetric lidar chapters in the third edition of the *DEM Users Manual*, as well as in USACE EM 1110-1-1000, *Photogrammetric and Lidar Mapping*, for the U.S. Army Corps of Engineers. Dewberry now owns and operates two state-of-the-art topobathy lidar sensors, as well as a topographic lidar sensor.

When water clarity allows, topobathy lidar does an outstanding job of mapping both the topographic and bathymetric surfaces, as shown in Figure 21 which revealed the previously unknown bathymetric surface beneath waters in the Potomac River. When water clarity is poor, because of water turbidity and/or presence of significant aquatic vegetation, as shown in Figure 22, topobathy lidar data voids will occur that require sonar to fill in the gaps.

In September of 2022, the 3D Nation Elevation Requirements and Benefits Study (Figure 23) was completed by Dewberry for NOAA/NGS and USGS, primarily authored by Sue Hoegberg of Dewberry. I was pleased to see that the "3D Nation" vision included inland topography and inland bathymetry (for USGS), and nearshore and offshore bathymetry (for NOAA) – each with their own technologies and user applications. The USGS link to the study is at https://

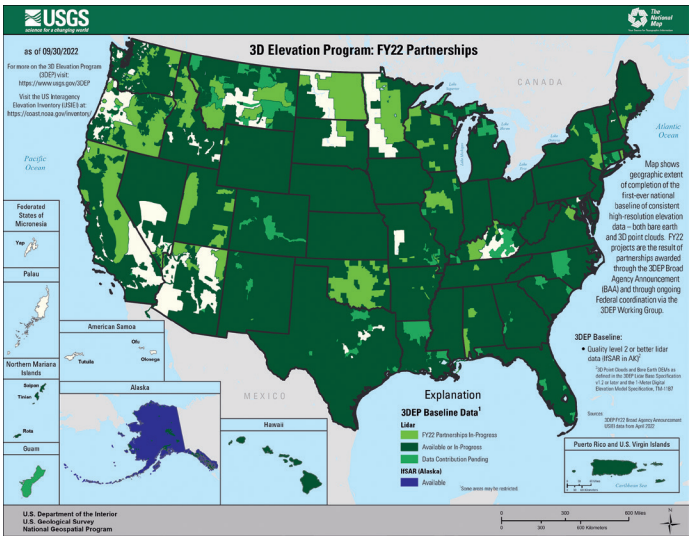



Figure 20. Geographic extents of QL2 or better lidar for 49 states and US territories and QL5 IfSAR of Alaska. Source: USGS.

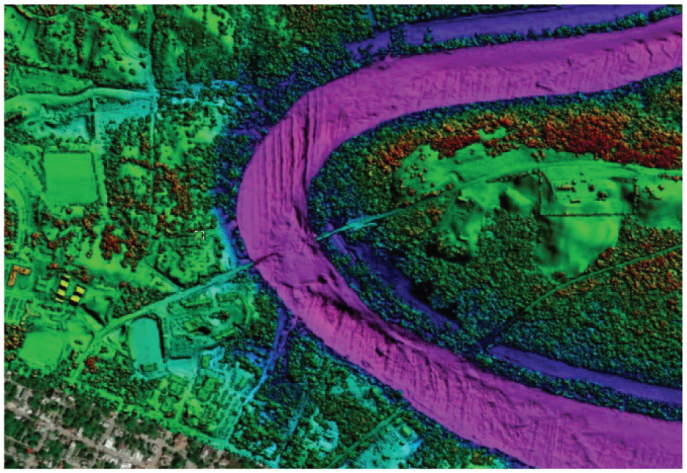


Figure 21. Topobathy lidar produced by Dewberry for USGS, showing the topo-bathy surface along the Potomac River near Shepherdstown, WV. Source: Dewberry.

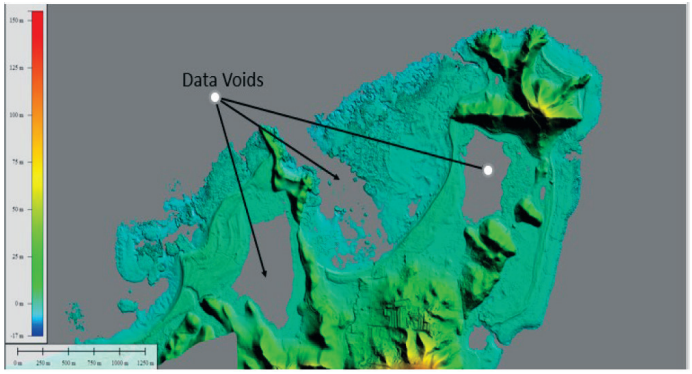



Figure 22. Example of a seamless topobathy lidar data surface in Puerto Rico, including data voids where aquatic vegetation, bioluminescence or sediments in the water prevented penetration by the green laser. The data voids in outer areas occurred where the laser extinction depth was exceeded. Source: Dewberry

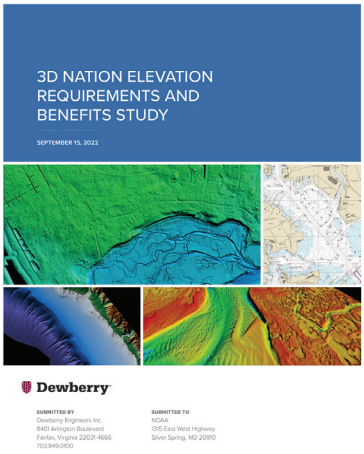



Figure 23. Front cover of the 3D Nation Elevation Requirements and Benefits Study. Source: Dewberry

www.usgs.gov/3d-elevation-program/3d-nation-elevation-requirements-and-benefits-study, which points to https://www.dewberry.com/services/geospatial-mapping-and-survey/3d-nation-elevation-requirements-and-benefits-study.

In October 2022, USGS posted a technical announcement for the 3D Nation Study, which can be found here: https://www.usgs.gov/news/technical-announcement/results-are-3d-nation-study-report-now-available.

Kevin Gallagher, USGS Associate Director for Core Science Systems said: "This study is foundational to our future direction of the 3D National Topography Model that integrates elevation and hydrography in 3D. The 3D National Topography Model will provide the terrestrial component of the 3D Nation vision we share with NOAA to build a continuous elevation and hydrography surface from the peaks of our mountains to the depths of our waters." Similarly, NOAA developed a blog post about the study.

For inland bathymetry, Figure 24 shows where topographic lidar was merged with topobathy lidar of the Lower Withlacoochee River (Florida) for a 22.5 $mi^2$ area. The CZMIL topobathy lidar coverage area was 12.5 $mi^2$. Topobathy lidar was unable to get bottom returns in the deeper parts of the river channel due to multiple bathymetric factors: depth, tannic water, and mucky bottom substrate. The areas outside the channel were shallow enough to overcome the bottom and water turbidity issues. Figure 25 shows where Multibeam Echo Sounder (MBES) sonar was collected for the deeper parts of the channel (0.3 $mi^2$ or 14 linear river miles); and Single Beam Echo Sounder (SBES) sonar with a HyDrone in the two dam spillway areas that were too shallow for MBES and too turbid for lidar. Figure 26 shows the successful merger of the topographic lidar, topobathy lidar, and sonar data to map the entire topographic-bathymetric surface. This is representative of what needs to be done for rivers and lakes nationwide in order to fully satisfy objectives of the 3D Nation initiative.

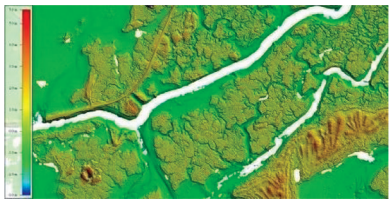


Figure 24. Topobathy lidar mapped portions of the river, but not the deeper tannic waters with mucky bottom. Image source: Dewberry.

Figure 25. Multibeam sonar mapped the deeper parts of the river not mapped with topobathy lidar. Image source: Dewberry.

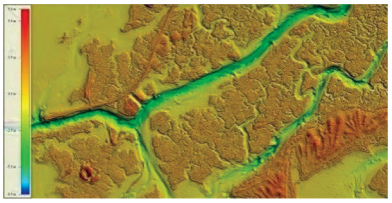


Figure 26. By merging the two datasets, the entire topographic and bathymetric surface was mapped seamlessly. Image source: Dewberry.

# CHANGES IN SONAR

During my first 30-year career, in the U.S. Army Corps of Engineers, I never managed hydrographic surveys. However, between 1988 and 1991, while serving as Commander and Director, U.S. Army Engineer Topographic Laboratories (ETL) and Topographic Engineering Center (TEC), my organization authored what was then the latest version of USACE Engineering Manual EM 1110-2-1003, *Hydrographic Surveying*, which explained technologies and best practices for SBES and MBES surveys.

During my second 30-year career with Dewberry, I specialized in DEMs from all technologies, and all three editions of my *DEM Users Manual* included chapters on photogrammetry, IfSAR, topographic and bathymetric lidar, and sonar. The sonar chapters were all authored by Captain Guy Noll (NOAA Corps, retired) and different co-authors he selected. I always considered the underwater bathymetric surface to be a continuation of the above-water topographic surface.

Between 2015 and 2022, I served as a member of NOAA's Hydrographic Services Review Panel (HSRP), which authored

numerous issue papers relevant to sonar and hydrographic surveys. In 2019, the Presidential Memorandum on *Ocean Mapping of the United States Exclusive Economic Zone (EEZ)* and *Shoreline and Nearshore of Alaska* was issued. Section 2 called for a National Ocean Mapping, Exploration, and Characterization (NOMEC) strategy. Section 3 of that memorandum directed the NOAA Administrator, in coordination with the state of Alaska and the Alaska Mapping Executive Committee (AMEC) – co-chaired by NOAA and the USGS - to develop a proposed strategy to map the shoreline and nearshore of Alaska and inform actions of the Ocean Policy Committee and relevant agencies. NOAA subsequently developed two strategies – one for the NOMEC and another for the ACMS. I was the primary author of the HSRP whitepaper with HSRP recommendations to NOAA on the implementation plan for the ACMS, including the use of Uncrewed Surface Vessels (USVs) and Autonomous Surface Vessels (ASVs).

By then, the 3D Nation Elevation Requirements and Benefits Study had already been in progress for several years, documenting how nearshore bathymetry had not been collected for most of coastal Alaska, and offshore bathymetry had not been collected for major portions of the U.S. Exclusive Economic Zone (EEZ). As stated above, the final 3D Nation Study report was released in September of 2022, including evaluations of relevant technologies that could most cost-effectively address the major unmet needs for MBES surveys. See: https://www.dewberry.com/services/geospatial-mapping-and-survey/3d-nation-elevation-requirements-and-benefits-study, which includes an analysis of technology trends and risk considerations.

Whereas there are many excellent commercial multibeam sensors available, it is the platforms for those sensors that will have the greatest impact on future benefits vs. costs. For years, crewed systems for MBES have been the norm; but uncrewed systems are now making inroads as force-multipliers, either with a crewed mothership (Figures 27 and 28) or as stand-alone collection platforms. But crewed motherships are very expensive to operate, e.g., $40,000 or more per day for the mothership shown in Figure 27.

To reduce costs dramatically for MBES surveys of large areas, several new ASV platforms have emerged that do not require crewed motherships. They are generally powered by wind-energy, solar-energy, batteries and/or micro diesel engines. They operate 24/7, often monitored and controlled from mission control thousands of miles away. They often have complex situational awareness sensors and communications for data transfer and to enable remote operators to monitor progress and take corrective actions if needed. Depending on the platform used, they can autonomously collect MBES data for hours, days, weeks, and months between services.

I see these ASVs as the future for ocean mapping, and I believe they will revolutionize the way that multibeam sonar is collected, cost-effectively, for the world's large unmapped oceans and coastlines.

Figure 27. Larger vessel (mothership) used for traditional MBES surveys in deeper water and CW5 USV (yellow) used for shallow-water surveys. Source: TerraSond.

Figure 28. With the mothership in the background, such USVs are ideal for shallow-water surveys using a variety of SBES or MBES sensors. Source: TerraSond.

# CHANGES IN GEOPHYSICAL MAPPING

Geophysical mapping involves the non-invasive investigation of subsurface conditions in the earth through measuring, analyzing, and interpreting physical properties at or close to the surface.

When geodesists like myself study geophysics, it normally boils down to measurements of gravity, needed to develop or improve geoid models to convert ellipsoid heights (from GPS observations) into orthometric heights (elevations). For years, Dewberry has been collecting airborne gravity data for NGS' Gravity for the Redefinition of the American Vertical Datum (GRAV-D). Dewberry has also used Ground Penetrating Radar (GPR) to detect underground utilities; and in 2016, we sponsored a GPR survey to measure the depth of the ice and snow on the peak of Denali, America's tallest mountain; we had previously determined the elevation at the top of the ice and snow using redundant GPS/GNSS receivers. When I started mapping, I barely understood magnetic declinations and the earth's changing magnetic field.

In 2019, USGS announced the Earth Mapping Resources Initiative (Earth MRI), apparently choosing this acronym for its similarity to a human MRI which maps inside the human brain. The Earth MRI is a geophysics initiative to map geologic features beneath the surface of the earth.

USGS' Earth MRI home page (https://www.usgs.gov/special-topics/earth-mri) states that the goal of Earth MRI is to improve our knowledge of the geologic framework in the U.S.

and to identify areas that may have the potential to contain undiscovered critical mineral resources. Enhancement of our domestic mineral supply will decrease the Nation's reliance on foreign sources of minerals that are fundamental to our security and economy. The home page provides greater details on why Earth MRI is needed; how the Earth MRI is being implemented; and Earth MRI acquisitions completed or planned.

As part of our Geospatial Products and Services Contracts (GPSC3 and GPSC4) with USGS, the Dewberry team has been awarded Earth MRI task orders for over a million line-kilometers of magnetic and radiometric (MAG-RAD) data acquired at low-altitude, typically 80 to 100 meters above the terrain, where safety is a major factor. Data are acquired at higher altitudes over cities and other sensitive locations. Two dimensional drape surfaces are established to provide vertical guidance over the terrain to an aircraft flying narrowly spaced traverse lines normally flown perpendicular to the dominant local geological trend, spaced a few hundred meters apart, and widely spaced tie or control lines flown perpendicular to the traverse lines, spaced from 5 to 10 times greater than the traverse line spacing. At the crossing points between traverse lines and tie-lines, flying heights must agree within small tolerances, requiring special aircraft guidance systems and expert pilots. When the terrain is relatively flat, specially equipped fixed wing aircraft are used, as shown at Figure 29, with the magnetic sensor housed in a "stinger" at the rear. In rugged terrain, specially equipped helicopters are required, shown equipped with a stinger in Figure 30, or alternatively with a "bird" towed beneath the helicopter as shown in Figure 31.

The airborne equipment includes cesium vapor magnetometers (used to acquire aeromagnetic data), gamma-ray spectrometers, downward facing and upward facing Nal crystals (used to acquire gamma-ray data), navigation and data acquisition systems, GNSS receivers, digital radar altimeters, laser altimeters, digital barometric altimeters, and outside air temperature sensors. Base station equipment includes cesium vapor magnetometers, GNSS receivers, and data acquisition computers.

Wikipedia explains why aeromagnetic surveys are acquired using aircraft with magnetometers housed in "stingers" or towed birds in order to separate the metallic aircraft and its magnetically noisy engine and electronics from the sensitive magnetic sensors. As the aircraft flies, the magnetometer measures and records the total intensity of the magnetic field at the sensor, which is a combination of the desired magnetic field generated in the earth as well as variations mostly due to the temporal effects of the constantly varying solar wind and the magnetic field of the survey aircraft. By subtracting the solar, regional, and aircraft effects, the resulting aeromagnetic map reflects the spatial distribution and relative abundance of magnetic minerals (most commonly the iron oxide mineral magnetite) in the upper levels of the earth's crust. Because different rock types differ in their content of magnetic minerals, the magnetic map allows a visualization of the geological structure in the subsurface, particularly the spatial geometry of bodies of rock, intrusions of volcanic material, and the presence of faults and folds. This is particularly useful where bedrock is obscured by surface sand, soil or water. Aeromagnetic data was once presented as contour plots, but now is more commonly expressed as thematic (colored) and shaded computer generated pseudo-topography images. The apparent hills, ridges, and valleys are referred to as aeromagnetic anomalies. A geophysicist can use mathematical modeling to infer the shape, depth, and properties of the rock bodies responsible for the anomalies.

In "Mineral Exploration: Principles and Applications," the author, Swapan Haldar, explains that radiometric surveys detect and map natural radioactive emanations ($\gamma$ ray) from rocks and soils. The gamma radiation occurs principally from the natural decay of isotopes of the elements U, Th, and K.

The radiometric method is capable of detecting these elements at altitudes up to 300m above the surface, or greater depending on the strength of the radiation. Some common radioactive minerals that can be detected are uraninite ($^{238}$U), monazite, thorianite ($^{232}$Th), feldspar ($^{40}$K), muscovite, and sylvite in acid igneous rocks. Exploration for these minerals by radiometric survey became important because of the demand for nuclear fuels and also for detection of associated nonradioactive deposits such as titanium and zircon. Isotopes are elements whose atomic nuclei contain the same number of protons but different number of neutrons. Certain isotopes are unstable. They disintegrate spontaneously to generate other elements. Radioactivity means disintegration of atomic nuclei by emission of energy and particles of mass. The by-products of radioactive disintegrations are in various

Figure 29. Fixed-wing geophysics aircraft with stinger in the rear. Image source: Sander Geophysics Ltd.

Figure 30. Rotary-wing geophysics aircraft with stinger in the front. Image source: Sander Geophysics Ltd.

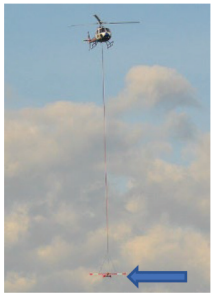


Figure 31. Helicopter with towed "bird" with MAG RAD sensors. Source: Sander Geophysics Ltd.

combinations of alpha ($\alpha$) particles of helium nuclei, beta particles ($\beta$) of electrons emitted by splitting of neutrons and gamma ($\gamma$) ray of pure EM radiation. Only the gamma ($\gamma$) ray radiation is detectable at any appreciable distance above the ground and emanates from only the top 10cm or less of the earth's surface. Even so, maps of gamma-ray radiation and their relative abundance often reflect underlying geological formations and alteration of lithology, while keeping in mind that in areas of significant overburden the character of the surface gamma-ray may reflect transported material.

When asked how USGS uses this MAG-RAD data for the Earth MRI, Anjana Shah, a Geophysics Chief Scientist at USGS, pointed me to a recently-published article on one of the first Earth MRI projects: https://www.geosociety.org/GSA/Publications/GSA_Today/GSA/GSAToday/science/G512A/article.aspx. She also pointed to upcoming articles in *Economic Geology* explaining how these MAG RAD surveys have identified previously unknown mineral deposits.

# SUMMARY – THE BIG PICTURE

I am amazed with the vast improvements in mapping technologies over my 60-year career.

Cartography: In my Army topographic engineer battalion in Germany in the 1960s, we had dozens of the finest German civilian cartographers and dozens of U.S. Army cartographers who spent two years compiling maps and preparing text, scribecoats, and peelcoats for the mapping of one city—Munich, Germany. Those maps were never completed while I was there. After two years, we were told by headquarters, US Army Europe, that new aerial photography was being flown, and we would need to update many of the features we had compiled the past two years.

Recently at Dewberry, we had a FEMA emergency response contract where we proved that, for cities or disaster areas up to 100 square miles, aerial imagery could be acquired, and high-resolution natural color digital orthophotos could be produced and delivered to FEMA within 48 hours of notice to proceed. These 1-meter orthophotos were overlaid with street and administrative boundary vectors, with names for streets and communities. Furthermore, the orthophotos were linked to a database that included location of critical infrastructure, individual building footprints, street addresses, assessed value, names of owners for potential insurance claims, and other geospatial information.

With today's GIS software, cartographic features are accompanied by geodatabases with vast amounts of supporting information. Specialized maps are produced in hours that previously would have taken years to produce.

Map Production and Distribution: With paper maps, large brick and mortar facilities were required for map production, storage, and distribution. Large quantities of new paper maps were produced daily that would replace large quantities of old paper maps that would need to be disposed of daily. Special facilities were required for paper maps and terrain analysis studies that were secret or confidential, with guards and security procedures to prevent such maps from getting into the wrong hands.

Today, with few exceptions, brick and mortar printing plants and map warehouses are obsolete. Maps and terrain analyses are prepared by geospatial specialists working at their desks with GIS or photogrammetric workstations. Digital data are stored in large databases and often served to the public via the cloud. Administrative procedures are in place to protect classified geospatial information.

Map Accuracy Standards: The National Map Accuracy Standards of 1947 defined horizontal and vertical map accuracy standards in terms of the map scale and contour interval of printed topographic maps, assuming all errors had a normal error distribution. However, with digital imagery and lidar, map scales and contour interval could be changed at the push of a button, but the data does not suddenly become more accurate just because it can be displayed at a larger scale or higher resolution on our computers. Furthermore, my personal research demonstrated that elevation errors in lidar bare-earth DEMs in vegetated terrain do not follow a normal error distribution and the 95$^{th}$ percentile should be used in lieu of RMSEz to define the vertical accuracy of lidar bare-earth DTMs in vegetated terrain. I was proud to be a co-author of the ASPRS Positional Accuracy Standards for Digital Geospatial Data, 2014, that established positional accuracy standards for digital orthoimagery, digital planimetric data, and digital elevation data from lidar, and other elevation technologies.

Surveying and Geodesy: The 1$^{st}$ edition of the *DEM Users Manual* has figures that show the large number of miles of First Order differential leveling used for the NGVD 1929 vertical adjustment, as well as the much larger vertical control network required in the NAVD 1988 adjustment. These differential leveling survey lines were extremely expensive and time-consuming. Surveyors needing to establish local vertical control would first need to identify the best and closest benchmarks in the National Spatial Reference System (NSRS); and it could then take days or weeks of differential leveling to determine an acceptable elevation for a single local benchmark or FEMA Elevation Certificate.

Today, in one hour, a surveyor equipped with a geodetic grade GPS receiver can establish local vertical control accurate to 2-cm or 5-cm at the 95% confidence level, relative to Continuously Operating Reference Stations (CORS) located nationwide. Thankfully, Bilby towers are long gone.

Aerial and Satellite Imaging: For my first decade in the mapping profession, film imagery was the only option.

Navigating largely by visual means, looking out the airplane windows and through view-finders, pilots and aerial photographers had to be extremely skilled to acquire aerial photos of the desired area with the correct position for photo centers, forward overlap, sidelap, and exposure controls; and experienced photolab personnel developed the film to produce acceptable film negatives and diapositives. Then think of the complexity in retrieving satellite film imagery from space, having to return miles of film from space to the earth for developing as we did with the KH-9 Hexagon satellites.

Today, with modern digital mapping cameras, automated exposure controls, inertial measurement units, airborne GPS, and modern flight management systems, personnel can be quickly trained to acquire digital imagery with the correct photo centers, forward overlap and sidelap, and optimal exposure controls. Of course, satellite images are now all digital and are easily transmitted back to earth for mapping and analyses.

Photogrammetry: When I first learned photogrammetry on the Multiplex, it would take up to eight hours to bring a single stereo pair of images into relative orientation. Then additional images would be added, one at a time, to bring subsequent stereo pairs into relative orientation. Then at the end of a flight line, absolute orientation would need to be performed on the entire strip in order to correctly scale the model and fit ground control on both ends of the strip. We didn't have a good way to perform block triangulation with dozens or hundreds of overlapping flight lines.

Today, with automated image correlation as well as position and orientation recorded for each image, Dewberry reviewed a large block triangulation with over 20,000 digital images acquired by 11 different cameras for a 4,000+ $mi^2$ area, and the software told us there was a problem with only one of those images. A review identified that the wrong calibration parameters had been entered for that one problem image; after correction, the block triangulation was successfully completed, in a few hours, for those 20,000+ images that each formed stereo pairs with about eight surrounding and overlapping images.

Radar: In the past, I primarily looked to radar as a low-resolution reconnaissance tool, mostly valuable to the military for broad area surveillance.

Today, I know that aerial IfSAR was the perfect tool for mapping through clouds in Alaska or anywhere worldwide where clouds are an issue. I've also learned the value of using DInSAR, with current and archived satellite SAR images, to map annual rates of subsidence at the cm- or even mm-level per year.

Lidar: For the first half of my 60-year career, neither photogrammetry nor radar could accurately map the elevations of bare earth terrain in forested areas. With aerial imagery, the ground beneath the trees could not be mapped in stereo because there were trees in the way; and radar generally mapped the top reflective surface, i.e., treetops.

Topographic lidar, with red lasers, came along in the 1990s, and we suddenly had the tool that map makers have needed for centuries to map the bare earth terrain everywhere; the 3DEP has been a huge success because of advances in topographic lidar. Bathymetric lidar, with green lasers, and topobathy lidar with both red and green lasers, came along to map subsurface bathymetry when waters are reasonably clear. The 3D Nation initiative will depend on topobathy lidar for mapping the Nation's inland and nearshore bathymetry. I will closely monitor its progress after I retire.

Sonar: The world has long needed multibeam sonar to map the bottoms of rivers and oceans; but crewed vessels with multibeam sonar are very expensive to operate, leaving vast areas of our rivers, coastlines and oceans unmapped because of affordability issues.

I see USVs and ASVs as the solution to this dilemma, and I will also closely monitor the utilization of USVs and ASVs after I am retired.

Geophysics: My prior knowledge of geophysics was minimal. My only college course in geophysics concentrated on gravity. I learned that geodesists could use satellites to gather data on gravitational changes as they pass over points on the earth's surface; on land, gravimeters measure the earth's gravitational pull on a suspended mass. Dewberry has been flying gravimeters for years for NGS' GRAV-D program needed to update our official vertical datum from NAVD88 to the upcoming North American-Pacific Geopotential Datum of 2022 (NAPGD2022); then we can more-accurately convert ellipsoid heights (from GPS or airborne GPS measurements) into orthometric heights, commonly known as elevations. I studied very little about the earth's magnetic field and knew just enough to be dangerous. I knew that the earth's magnetic field was continuously changing, and NOAA had a calculator for computing magnetic declinations shown on maps and charts. But when my boss, Sid Dewberry, asked me what Dewberry should do about the earth's changing magnetic field, I could not think of an answer.

Today, I am amazed to experience the use of magnetic and radiometric surveys for the earth MRI initiative. Being able to map the probable location of underground critical minerals and rare earth elements "blows me away." In retirement, I also plan to follow the progress of this exciting initiative so vital for our future economy and security.

During my 60-year career, we've come a long way with all these technologies. I am proud to have been actively involved in the development and/or maturation of many of these technologies, and I am so grateful to see the major improvements in our profession.

# GIS Tips & Tricks

Al Karlin, Ph.D. CMS-L, GISP

## Simple Customizations can have a Large Impact

Here is another Tips & Tricks column that comes by way of my GIS/Map Making Class at the University of Tampa. We were doing an exercise in Geocoding/Address Matching and I asked the class what local businesses they would like to geocode. Of course, fast-food restaurants were the most popular choices and we decided on McDonalds Fast Food Restaurants. So, we prepared a street locator using local street map data, opened a web-browser, navigated to the McDonald's website and found addresses for a few of the local outlets. Of course, we could have just used the Mc Donald's locator map, but that was not the point of the exercise. We copy/pasted those addresses into a spreadsheet and used ArcGIS Pro 3.0 along with our locator to geocode the restaurants.

The locator worked very well, and the resulting feature class displayed on the street map as default (4 point, green) point markers. Everyone was impressed that the technology works, but unimpressed with the display (Figure 1). This reminds me of one of my mantras… Never accept the defaults! Yes, they all knew that they could alter the display symbology, but there are just a bunch of other, generic point markers in the Esri palette. So… what to do?

### Tip — In ArcGIS Pro: Add a custom Picture Marker

ArcGIS Pro and ArcGIS Desktop can use several image formats including, .JPG, .TIF, PNG and/or .BMP as "Picture Markers". For a McDonald's picture marker, we GOOGLED "McDonald's logo", and selected a free-to-use transparent logo and saved it to a writable directory. The picture downloaded in .PNG format.

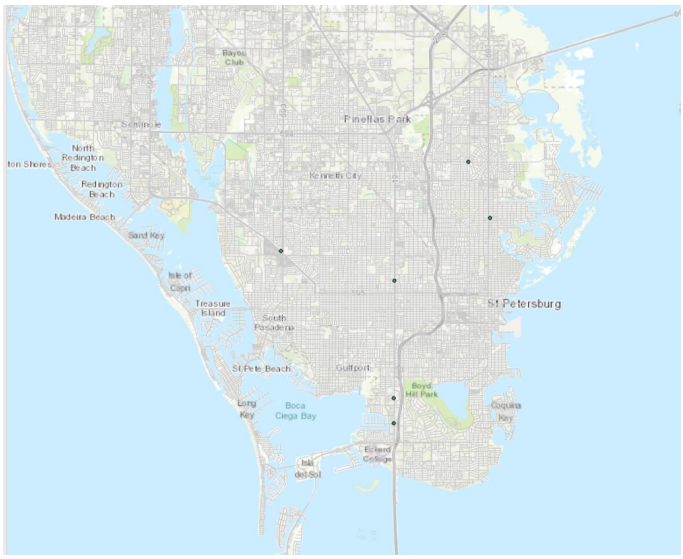



Figure 1. Default display after geocoding McDonalds restaurants in the St. Petersburg, Florida Area.

Selecting the McDonald's geocoded layer in the Contents pane, either right-click and open the Symbology Pane (or double-click on the point marker to open the Symbology Pane), and double-click on the point marker in the Symbology Pane to get to the Format Point Symbol menu, then open the Properties dialog (Figure 2).

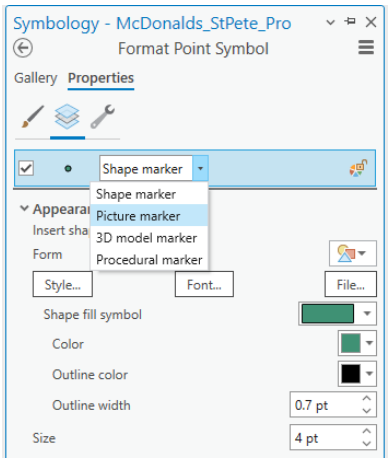



Figure 2. The Properties dialog box in the Symbology Pane in ArcGIS Pro.

Note that the default is "Shape marker", use the arrow to drop down the choices and select "Picture marker" (see Figure 2). This will open a new dialog box (Figure 3). Choose "File…" and navigate to your picture marker. Once selected, you can change the size rotation, etc. on this menu; we made the picture marker 15 points. Pressing "Apply" at the bottom of this dialog box will apply your new Picture marker, to make a more customized map as in Figure 4.

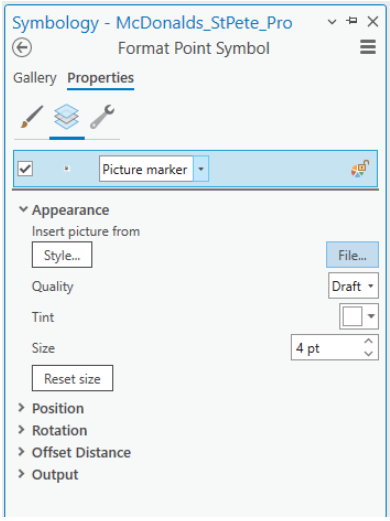



Figure 3. The Picture marker dialog box allows you to select a picture marker and manipulate the properties in ArcGIS Pro.

### Tip — In QGIS: Add a custom Picture Marker

Adding a custom picture marker in a QGIS project uses a very similar workflow to that described above. QGIS, also recognizes the common graphic formats, .JPG, .TIF, .PNG and/or .BMP.

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 3, March 2023, pp. 143-144.
0099-1112/22/143-144

doi: 10.14358/PERS.89.3.143

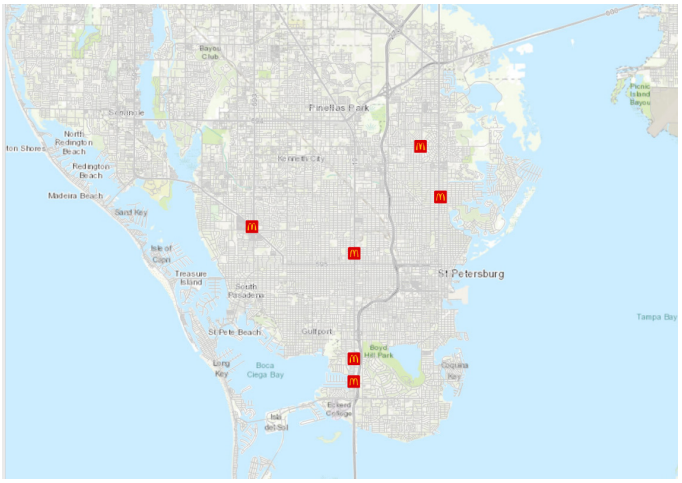



Figure 4. The same map as in Figure 1, but with customized McDonald's logo Picture markers.

To add the custom picture marker, double-click on the symbol in the Layers panel to open the QGIS Symbology panel. Here, notice that the default symbol marker is "Simple Marker", click on the "Simple Marker" to open the Symbol Layer type menu (Figure 5), use the dropdown to select Raster Image Marker (Figure 6).

Use the ellipses (…) or the dropdown arrow to navigate to the picture file to use as a marker (Figure 7) and adjust the size, and other properties as needed, then click "Apply" to change the marker and then "OK" to dismiss the Symbology Marker panel. The resulting map is identical to the ArcGIS Pro Map as in Figure 8.

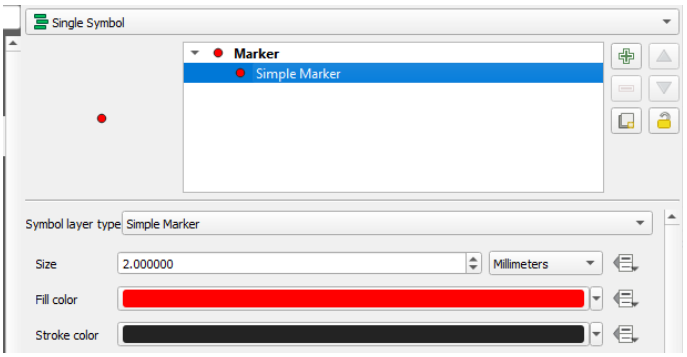



Figure 5. The Symbol Layer dialog in the Symbology panel in QGIS.

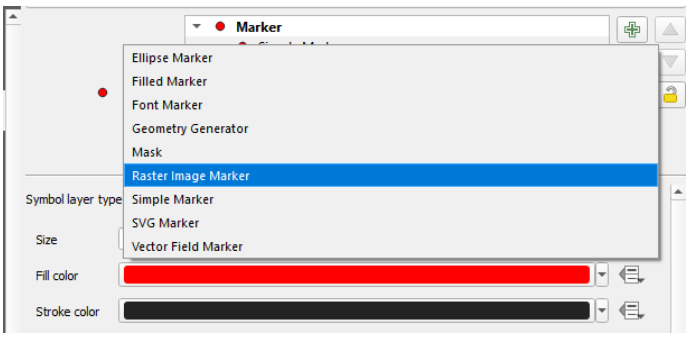



Figure 6. Selecting the Raster Image Marker from the Symbol Layer dialog in QGIS.

We had so much fun with this little trick, students went on to construct maps with gopher tortoise, starfish, sharks, and octopi, not to mention Burger Kings, Chipotle's, Starbucks, and other favorite places with little picture markers.

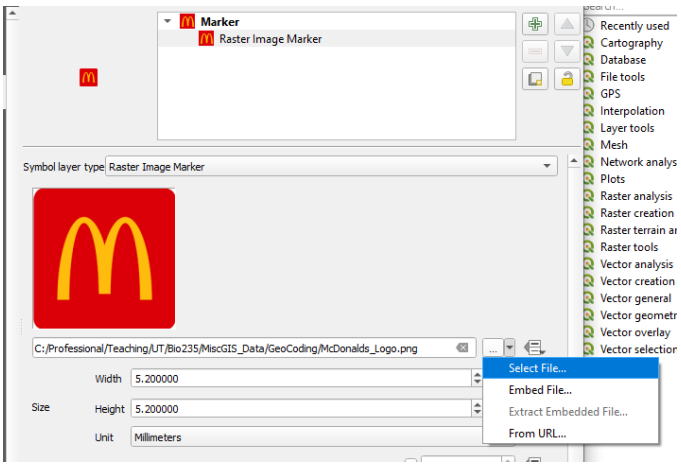



Figure 7. Navigating, selecting and applying a Raster Image Marker file in QGIS.

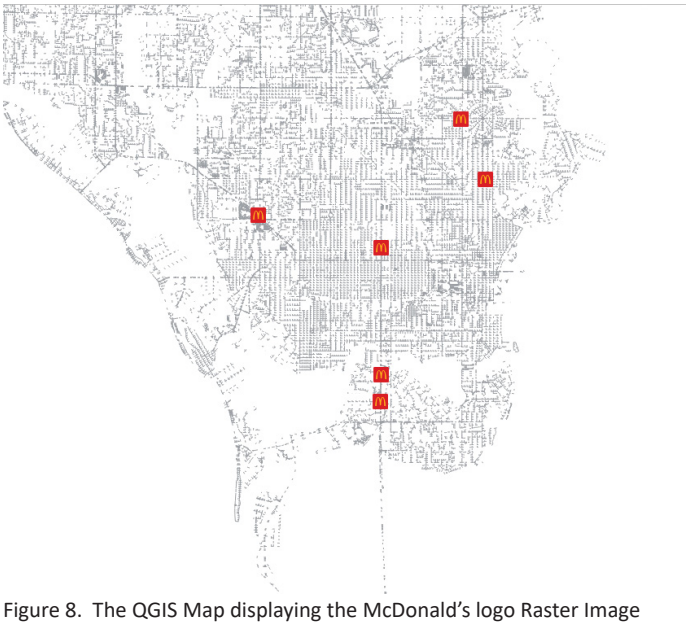
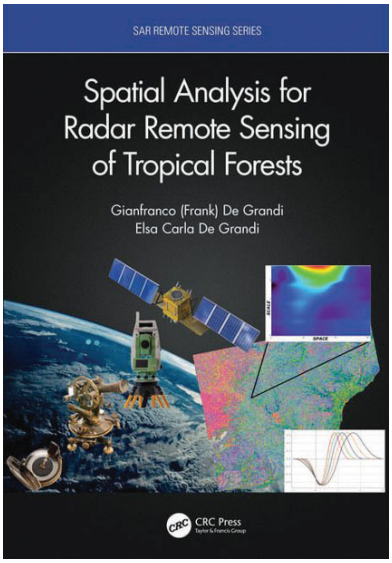


Figure 8. The QGIS Map displaying the McDonald's logo Raster Image Marker file.

CAUTION: Many images found on the internet are copywritten by their owners. It is illegal to use some images without the direct, written consent of that owner, so be aware.

Send your questions, comments, and tips to GISTT@ASPRS.org.

*Al Karlin, Ph.D., CMS-L, GISP is with Dewberry's Geospatial and Technology Services group in Tampa, FL. As a senior geospatial scientist, Al works with all aspects of Lidar, remote sensing, photogrammetry, and GIS-related projects. He also teaches beginning map making at the University of Tampa.*




## Spatial Analysis for Radar Remote Sensing of Tropical Forests

Gianfranco D. Grandi and Elsa Carla De Grandi.
Available from VitalSource Bookshelf, Taylor & Francis, 2021.

**Reviewed by** Konrad Kern, Mostly retired, image scientist who develops national technical means for the government of the US.

The title might lead one to expect a description and assessment of results found in their study of Tropical Forests. However, this is more of a "how-to" book than a "what-we-learned" book. As a result, the book has infinitely more value to an aspiring remote sensing technologist. The book addressed a synoptic study of the tropical African forest. The magnitude of the problem becomes evident when the size of the study area is considered. The greatest width of the African continent of interest is about 2800 nautical miles with an area of nearly 2 million square nautical miles. The book presents a detailed approach to how to select, prepare, and use your remote sensing tool suite for a monstrous problem.

Dr's De Grandi wrote an extremely lucid book on how to conduct a remote sensing campaign. They describe their professional journey through multiple generations (decades) of Synthetic Aperture Radar (SAR) development. Their purpose was a synoptic study of the African Tropical Jungle with a specific focus on anthropological impacts. Their approach was a monster-mosaic map on the scale described and supervised, well vetted computer analysis and classification. This is clearly unique, offering many benefits and challenges.

The first four chapters describe the tools and products to achieve these ends. Hearing Dr's De Gandi elaborate on the benefits and limitations of the various SAR satellites employed is a great lesson in knowing your tools. In image chain analysis, this understanding is referred to as radiation / material interaction. The earlier SAR's were medium to long wavelength. These long wavelength radars had real canopy penetration and showed geological patterns otherwise obscured by canopy cover. The later SAR's were phased array radars with polarization selectability. The use of polarization information in the scene demonstrates feature discrimination not otherwise available. Finally, interferometric SAR is described with its unique signatures and capabilities for coherence matching.

Continuing the theme of know your tools, the authors describe their need for scene revisit, management of noise in a SAR system and correction of the 1/R4 energy loss in an active imaging system.

- The SAR satellites they could access all had a polar orbit. While this orbit is great for optical systems, the revisit rate for a SAR sensor to mid-latitude areas of interest is horrible; the author cited 30 days. Temporal changes on this scale had to be ignored.
- Regarding noise, SAR systems have two noise contributors: Additive and multiplicative noise. The authors refer to additive noise as thermal noise and deal with characterizing and managing additive noise very well. Multiplicative noise is generally a force majeure of SAR processing and addressed with various apodization techniques. Multiplicative noise is the result of transform functions on the phase history. A radar return results in a main lobe and sidelobes. The sidelobes can be large and distracting from the image interpretation. Apodization is the process of squashing these sidelobes. As this study took place over a couple decades, apodization technology underwent huge changes. Based upon the study scale, apodization probably did not impact their analysis methods. This study involved very large impulse responses (IPR) the need for apodization could have been ignored. IPR is radar speak for smallest resolvable size,
- The radar equation gives received power as proportional to the inverse of the range from the transmitter to scene to the fourth power. Occasionally, this can be reduced to 1/R^3 from phase history compression. This difference of

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 3, March 2023, pp. 145-146.
0099-1112/22/145-146

doi: 10.14358/PERS.89.3.145

the energy loss across the huge swaths generated by these radar collections required a radiometric equalization. Without correction, the scene brightness differences across a swath would be a detriment to interpretation. Their radiometric compensation was well described and effective.

The remaining six chapters are often analysis doublets. The first chapter in the pair describes the theoretical basis for an analytic technique. The second chapter in the pair shows the validation of the analytic technique using known ground truth and field application. Finally, the authors provide an appendix discussing wavelet theory and analysis. Read this first before enjoying the final six chapters. It is very good and will help you to understand the analysis chapters.

This was a fun book to read. For the Dr Gianfranco in the writer combination, this appears to be his lifework sayonara piece. As a result, his narrative is pleasantly light; but the math and analysis is rigorous. He offers a number of concessions to the those who only understand optical systems; Resolution vs. IPR, Ground Plane presentation vs. Slant Plane collection, etc. I would have enjoyed more discussion of the magic in the middle of SAR; image formation processing from the phase history. Showing the integration of their analytical methods with image formation processing would have been interesting, but vastly increased the size of the book. Perspective is useful. Compared to the overall value of the book, my critical observations are inconsequential. If you are serious about practicing Remote Sensing, this is a must read.

---

The contents of this Book Review reflect the views of the author. The contents do not necessarily reflect the official views or policies of the American Society for Photogrammetry and Remote Sensing

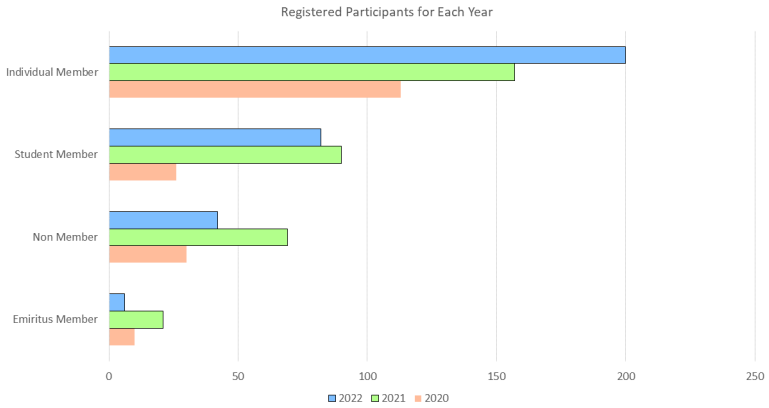
asprs THE IMAGING & GEOSPATIAL INFORMATION SOCIETY
STUDENT ADVISORY COUNCIL

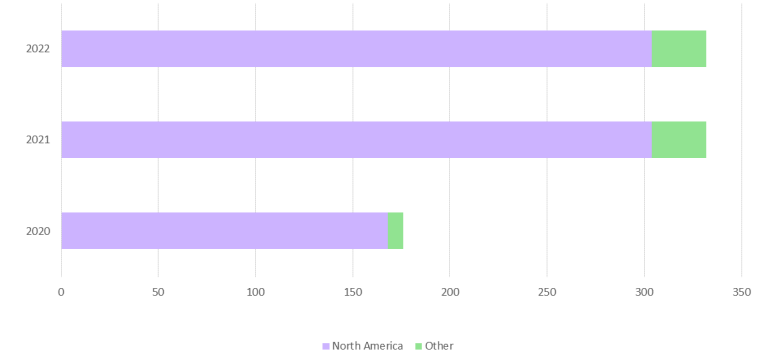
Signatures

# FROM LOCAL TO GLOBAL — THE INCREASED INVOLVEMENT IN THE ASPRS VIRTUAL CONFERENCE

**In the wake of the COVID-19 pandemic, the shift towards virtual events drastically changed the way conferences are conducted. With physical gatherings no longer being a viable option, virtual events became the primary means for professional and academic exchange, leading to an unprecedented surge in registrations. According to the registrations for the 2021 and 2022 ASPRS Virtual Conferences, the number of registrations in the student and individual membership levels has increased. These virtual conferences have been particularly beneficial for students, especially those from developing countries who previously faced barriers to attending international conferences due to travel and financial constraints. With virtual events, international students are now able to participate and network with professionals from around the world without incurring the costs of travel.**

While virtual events offer many benefits, in order to ensure a seamless and engaging experience for participants, there are technical challenges that must be addressed. However, with careful planning and preparation, these challenges can be effectively overcome. Conference organizers have successfully navigated these obstacles by implementing strong technical infrastructure, providing clear guidelines and technical support for participants, and adopting innovative solutions to improve the virtual event experience. These efforts have resulted in successful ASPRS virtual conferences with high levels of participant engagement and satisfaction.




The results of these conferences have been overwhelmingly positive, with total registrations more than doubling compared to the previous year. This is a clear indication of the growing popularity and success of virtual events, particularly given that the 2020 conference was initially planned as an in-person event. One of the most notable outcomes of the virtual format is the increased participation from participants from all over the world. It was thrilling to see such a diverse group of attendees from countries such as Brazil, the UK, Norway,

ASPRS STUDENT ADVISORY COUNCIL

OSCAR DURAN
COUNCIL CHAIR

ALI ALRUZUQ
DEPUTY COUNCIL CHAIR

RABIA MUNSAF KHAN
COMMUNICATIONS COUNCIL CHAIR

LAUREN MCKINNEY-WISE
PAST CHAIR

Spain, France, Germany, Nigeria, Egypt, Australia, New Zealand, the UAE, Iran, India, China, Pakistan, Indonesia, the Philippines, Korea, and Taiwan. This broad representation highlights the global reach of virtual events and the opportunities they provide for cross-cultural exchange and collaboration.

Another trend observed in the increased conference registration numbers is the dominance of participants from the United States, despite an increase in participants from other countries. Although the virtual format has made it easier for individuals from diverse locations to join in, the majority of participants still come from North America. This highlights the importance of continuing to expand our outreach and encourage participation from a wider range of countries. To encourage participation, we are delighted to announce that the call for abstracts for the 2023 ASPRS International Technical Symposium is open for submissions. More information can be found on the ASPRS website, www.asprs.org or at https://my.asprs.org/2023Symposium/2023-Symposium/Call-for-Abstracts.aspx.

It is also important to mention that in 2021 and 2022, ASPRS gave out $5000 worth of student presentation grants that included the conference registrations and if a registrant became a member, it also covered the cost of one-year membership to ASPRS. More information about the student grant for 2023 can be found at https://my.asprs.org/2023Symposium/2023-Symposium/Student-Grants.aspx.

In conclusion, the shift towards virtual events due to the COVID-19 pandemic has led to a significant increase in virtual participation, particularly among students. This shift has opened up new opportunities for students and professionals alike, breaking down barriers to participation and allowing individuals to connect and engage with their professional communities from anywhere in the world. While physical events will undoubtedly make a comeback, the virtual event trend is here to stay and is likely to continue to play an important role in the world of conferences and professional development.

## NEW ASPRS MEMBERS

ASPRS would like to welcome the following new members!

Andrew Aceves
Gazali Oluwasegun Agboola
Itiya P. Aneece
Lorilei Angle
Mostafa Arastounia, Ph.D., PE
Edwin Oluoch Awino
Amir Ayyash
Ray E. Barnes, III
Megan Blaskovich
Daira Brayley
Marc A. Canas
Jeremy Dancer
Jacob Davenport
Brett Edwards
Tayler Engel
Taylor Frye
Anthony Gaskill
Pedro Grijalba
Peaceibisia Jack
Justinn J. Jones
Jimmie Lee Kenyon, Jr.
Adrienne E. Kicker
Julie Lazor
Gary McKissick
William Nadal
Neil Pinto
Stephen Rector
Danny Roark
Michelle Roberts
Andrew Schwartz
Miguel Antonio Silva
Faezeh Soleimani Vostikolaei
Victoria Gyameraah Tanoh
George Varghese
Andrew Verville
Eden Tsehaye Wasehun
Jun Yu

FOR MORE INFORMATION ON ASPRS MEMBERSHIP, VISIT
**HTTP://WWW.ASPRS.ORG/JOIN-NOW**

## *PE&RS* Correction

The title of the January 2023 Feature Article contained an error.

### A Different Point-of-View
### Using Aerial Imagery to Build Stronger Cities

By Shelly Carroll, Vice President and General Manager, Public Sector, Nearmap

When it comes to aerial imagery, city governments often make do with sporadically updated captures that fail to provide the most up-to-date information about the changing landscape of their urban areas. While they understand and appreciate the value of collecting this data, the update cadence is often too infrequent to reveal vital changes. As a result, information that is essential to successful planning and ongoing operations can go unrecorded and unanticipated.

This era of cities having to "make do" with self-collected aerial data may be coming to an end, thanks to new and enhanced photogrammetric products and possibilities. Increasingly, municipalities are realizing that frequently refreshed and consistent aerial imagery is a requirement if they are to keep pace with fast-growing and quickly evolving communities.

To see the full article, visit https://doi.org/10.14358/PERS.89.1.5.

# 2023 ASPRS INTERNATIONAL TECHNICAL SYMPOSIUM

**ASPRS is happy to announce the dates of its virtual conference. The 2023 ASPRS International Technical Symposium will take place June 12-16, 2023.**

### The symposium will consist of:

- 15-minute oral presentations
- 5-minute Ignite-style presentations
- Poster Gallery
- Sustaining Member Vendor Spotlights
- ASPRS Society Highlights

Sessions will run each day from 10:00 AM to 6:00 PM Eastern Daylight Time (UTC - 4). All sessions will be recorded and made available on-demand to conference registrants. Presenters are eligible to submit full manuscripts for publication in the ISPRS Archives.

> ***"We are happy to offer this educational opportunity to the geospatial community. Virtual events are an excellent way to exchanammunity without the cost and time constraints of travel," said Karen Schuckman, ASPRS Executive Director***

**Interested in Presenting?** For more information or to submit an abstract visit https://my.asprs.org/2023Symposium/2023-Symposium/Call-for-Abstracts.aspx

- Submission deadline is May 1, 2023
- Presenters will be notified of acceptance by May 8, 2023
- Presenters must be registered for the conference by May 22, 2023 to be included in the conference program

### Registration Fees

| | |
|---|---|
| • ASPRS Member | $150 USD |
| • ASPRS Student Member | $ 50 USD |
| • ASPRS Emeritus Member | $ 25 USD |
| • Non Member | $250 USD |

### Sponsorship Opportunities

- Vendor Spotlight/Product Demo
- Day Sponsor
- Session Sponsor
- Workshop Sponsor

# 2022 *PE&RS* Reviewers

**Thank you for making *PE&RS* possible.**

Karthikeyan A.
Sundar Raj A.
Ghasem Abdi
Fanar M. Abed
Muge Agca
Salman Ahmadi
Devrim Akca
Burak Akpinar
Perpetual Akwensi
Hessah Albanwan
Ugur Alganci
Shahzad Ali
Muhammed Alshehri
Antony Dasel Amaraselvam
Itiya Aneece
Kutubuddin Ansari
Ozan Arslan
Leo John Baptist
Agnes Begue
Timea Bezdan
Dimitri Bulatov
Gurcan Buyuksalih
Michael J. Campbell
Mariana Campos
Levent Candan
Jorge A.S. Centeno
Min Chen
Wei Chen
Yanming Chen
Russell G. Congalton
Michael Cramer
Balasubramaniyan D.
Justin Xavier D.
Mahesh Kumar D.
Nagesh Kumar D.
Ashraf Dewan
Tejas Dhote
Kaichang Di
Mehmet Dikmen
Qing Ding
Mohamed A S Doheir
Jinwei Dong
Srinath Doss
Abdelrafe E.
Ahmed Elaraby
Mostafa Elhashash
Esra Erten
Vagner Goncalves Ferreira
Sagi Filin
Suresh G.
Suseendran G.
Mahendra Babu G.R
Wenxia Gan
Xun Geng
Mohsen Ghanea
Jianyu Gu
Murali Krishna Gumma
Bo Guo
Jianhua Guo
Barry N Haack
Ayman F. Habib
Andrew Hansen
Saeid Homayouni
Gang Hong
Weizhen Hou
Bin Hu
Bin Hu
Han Hu
Qiong Hu
Md Enamul Huq
Md. Enamul Huq
Akhtar Jamil
Muath Jarrah
San Jiang
Awais Khan Jumani
Subashini K.
Sridhar K.P
Michael Koelle
Divya Vishambhar Kumbhakar
Junhak Lee
Huapeng Li
Jiayi Li
Jun Li
Peng Li
Yuan Li
Zhaojin Li
Bingqing Liang
Xiao Ling
Man Liu
Wan Luoma
Iyapparaja M.
Fadam M Abdoon
Burhanuddin Ma
Hussein Mahdi
Adhiyaman Manickam
Adhiyaman Manickam
Izidor Mlakar
Noorayisahbe Binti Mohd Yaacob
Yousif Mousa
Mustafa MR
Riadh Munjy
Fatih Nar
Adam Ali Shah Noor Mohamed
Ilyas Nurmemet
Gabriel A. Ogunmola
Tomasz Olszak
Coskun Ozkan
Mohamed P. S.
Sidike Paheding
Atul Kant Piyoosh
Jean-Jacques Ponciano
Yubao Qiu
Arun Kumar R.
Dr. Veerandra Kumar R.
Kaniezhil R.
Mahendra R.
Sathyabama R.
Selvaraj Rajalakshmi
Sowmipriya Rajendiran
Mary Jansi Rani
Abel Andres Reyes-Angulo
Aiswarya RS
Ramesh Kumar S.
Selvakumar S.
Ali A.abdul Saeed
Rinesh Sahadevan
Ashraf Saleem
Dinesh Jackson Samuel
Fusun Balik Sanli
Ivan Sanz Prieto
Dursun Zafer Seker
Asma Kashif Shahzad
Qian Shi
Mohammad Faridul Haque Siddiqui
Chunqiao Song
Angelin Sophy
Shengjun Tang
Lai Tao
Pardhasaradhi Teluguntla
Xiaohua Tong
Jason A. Tullis
Indraraj Upadhyaya
Murat Uysal
Sarala V.
Sathishkumar V E
Thanjai Vadivel
Bruno Vallet
Chengyi Wang
Mi Wang
Peijuan Wang
Puzuo Wang
Yongbo Wang
Timothy A Warner
Lianhuan Wei
Martin Weinmann
Wenfu Wu
Yuxing Xie
Yanyan Xu
Yusheng Xu
Jie Yang
Zhehui Yang
H. Yilmaz
Alper Yilmaz
Ibrahim Yilmaz
Mustafa Zeybek
Guo Zhang
Hongping Zhang
Qi Zhang
Su Zhang
Tianqi Zhang
Yunsheng Zhang
Weixun Zhou
Ningning Zhu
Jiaqi Zou

# Robust Guardrail Instantiation and Trajectory Optimization of Complex Highways Based on Mobile Laser Scanning Point Clouds

Xin Jia, Qing Zhu, Xuming Ge, Ruifeng Ma, Daiwei Zhang, and Tao Liu

## Abstract

*As a basic asset of highways, guardrails are essential objects in the digital modeling of highways. Therefore, generating the vectorial 3D trajectory of a guardrail from mobile laser scanning (MLS) point clouds is required for real digital modeling. However, most methods limit straight-line guardrails without considering the continuity and accuracy of the guardrails in turnoff and bend areas; thus, a completed 3D trajectory of a guardrail is not available. We use RANDLA-Net for extracting guardrails as preprocessing of MLS point clouds. We perform a region growth strategy based on linear constraints to obtain correct instantiations and a forward direction. The improved Douglas–Puke algorithm is used to simplify the center points of guardrail, and the 3D trajectory of every guardrail can be vectorized using cubic spline curve fitting. The proposed approach is validated on two 3-km case data sets that can completely instantiate MLS point clouds with remarkable effects. Quantitative evaluations demonstrate that the proposed guardrail instantiation algorithm achieves an overall precision and recall of 98.80% and 97.5%, respectively. The generated 3D trajectory can provide a high-precision design standard for the 3D modeling of the guardrail and has been applied to a long highway scene.*

## Introduction

Accurate 3D highway information is important in the fields of transportation, road network construction, 3D digital modeling, and intelligent vehicles (Darnel 2012; Gomes 2013; Xiong *et al.* 2016). Guardrails are a fundamental asset of highways that provide information about road boundaries to improve traffic order and reduce traffic accidents (Loprencipe *et al.* 2018). Currently, design modeling does not accurately describe the asset elements of highways, making them difficult to directly apply to the operation, maintenance, and asset management of highways. Mobile laser scanning (MLS), as a rapid development of high-tech surveying and mapping technology, can quickly describe the high-precision geometric information and texture information of the target object, providing more reliable and valuable measurements to extract highway information (Gargoum *et al.* 2017; Zeybek 2021). Lidar point clouds can provide the spatial constraint (Jokkola *et al.* 2008) of 3D trajectory information (Graham 2010; Zhou *et al.* 2021) for digital highway modeling. The digital model constructed by vector trajectory will markedly improve the authenticity and reliability of the model. How to obtain the real 3D trajectory of the guardrail from point clouds has been an urgent demand for digital modeling of highways. Therefore, we directly extract the vector trajectory of guardrails from MLS data to effectively improve the quality of digital modeling rather than design modeling.

## Related Work

Accurately instantiating guardrails is the basis of generating a trajectory. Many studies have focused on extracting highway guardrails from point clouds rather than instantiating them. These methods only segment guardrail objects from images or point clouds. Some scholars projected point clouds onto images for extraction using image processing techniques such as the sliding window threshold (Broggi *et al.* 2005; Matsumoto *et al.* 2019; Seibert *et al.* 2013; Scharwächter *et al.* 2014), but these methods are prone to interference from image noise, yielding low accuracy. Based on the spatial characteristics of the guardrail, segmentation methods can be divided into traditional feature cluster methods and deep learning methods. The guardrail is extracted by setting specific geometric parameters of the guardrail (Jiang *et al.* 2016; Zhu and Guo 2018; Vidal *et al.* 2020). However, those methods can be used only for a guardrail of a specific shape and must set multiple thresholds and thus are not universal. Justo *et al.* (2021) proposed a pointwise supervised machine learning model for the classification of guardrails based on the work of Gressin *et al.* (2013). That method was effective only for simple highway scenarios. With the scan line information of the point clouds, researchers extracted the guardrail from the mutation point between the guardrail point and the road points (Lehtomaki *et al.* 2010; Zhu and Guo 2018). However, these methods rely on scan line information and do not fully consider the morphological characteristics of guardrails, making them vulnerable to noise. Some scholars have proposed the binary coding voxel segmentation method (Huang *et al.* 2020). The method of cluster slicing has also been used to identify the guardrail target in the highway scene, but the guardrail can be interrupted due to vehicle occlusion, which causes one guardrail to be detected as multiple guardrails, which cannot guarantee the completeness and continuity of the guardrail (Pu *et al.* 2011). Gao *et al.* (2020) detected guardrails from point clouds using multiple filters and modified DBSCAN clustering. However, their algorithm is not suitable for curved guardrail detection due to the use of straight-line fitting. Based on the work of Gao *et al.* (2020), Yue *et al.* (2021) added the refinement based on verticality to find candidate guardrails. These methods can extract guardrails on both sides of the road but cannot distinguish between different kinds of guardrails. Recently, the deep learning method has been widely used in point cloud processing, in which point cloud semantic segmentation is an important step. Qi *et al.* (2017) proposed the novel method of PointNet, which directly used irregular primitive point clouds, which

Xin Jia is with the Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China, and the Academician Expert Workstation of Gansu Dayu Jiuzhou Space Information Technology CO., Ltd..

Qing Zhu is with the Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China, and the Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, China.

Xuming Ge is with the Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, China (xuming.ge@swjtu.edu.cn).

Ruifeng Ma, Daiwei Zhang, and Tao Liu are with the Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China..

*Contributed by Bo Wu, July 8, 2022 (sent for review August 12, 2022; reviewed by Mustafa Zeybek, Levent Candan, Yuan Li).*

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 3, March 2023, pp. 151–161.
0099-1112/22/151–161

doi: 10.14358/PERS.22-00100R2

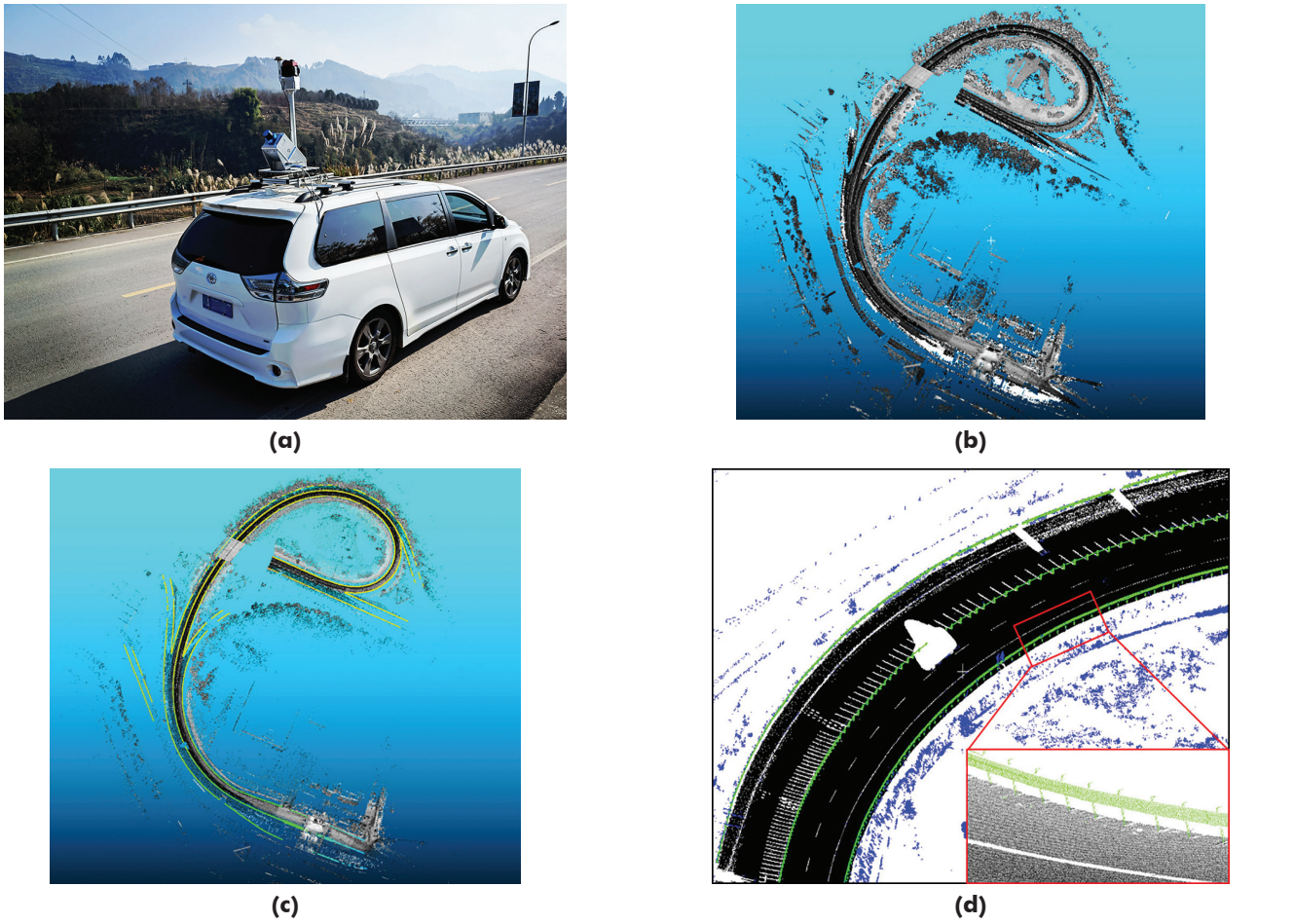



Figure 1. Point clouds of a portion of the Du Wen highway. (a) Point cloud data collection. (b) Raw point clouds of the highway. (c) Guardrail point cloud annotation. (d) Partial guardrail point cloud display.

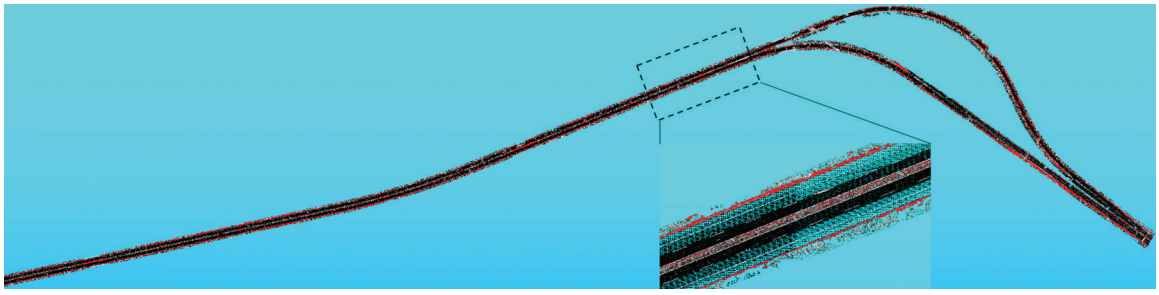


Figure 2. Point clouds of a portion of the Ya Xi highway.

improved computing efficiency, memory overhead, and precision beyond the vast majority of other methods that existed at that time. However, point features extracted by shared MLPs cannot capture the local geometry of the point clouds and the interaction between points and are not suitable for fine field spot cloud segmentation. Balado *et al.* (2019) used PointNet (Qi *et al.* 2017) to segment road elements, including pavement, scarp, guardrail, fence, road marking, and road boundary. Researchers also proposed RandLA-net (Hu *et al.* 2020), which is more suitable for semantic segmentation of point clouds in large-scale scenes. However, oversegmentation still occurs, and each guardrail cannot be identified accurately. The above method has been applied in many practical applications to segment guardrails. Most existing methods focus on straight lines without describing a guardrail's continuity and accurately instantiating in turnoff and bend areas. Thus, these results cannot meet the demand of digital modeling.

Fine vector lines and 3D models of highway guardrail are essential for various intelligent transportation applications. Similarly, there are

Table 1. Specifications of the mobile laser scanning system.

| Components | Specifications | |
|---|---|---|
| Laser scanner component | Laser scanning dot frequency: | 500 000/s |
| | Measurement range: | 300 m |
| | Scanner FOV: | 360° |
| | Scanner frequency: | 100 lines/s |
| | Ranging accuracy: | 6 mm at 40 m |
| | Laser type: | Level 1 laser |
| Position and orientation system component | Heading: | 0.012° |
| | Roll and pitch: | 0.005° |
| | Horizontal/vertical: | 0.02 m/s |
| | Vehicle speed: | 30 km/hr |

few studies on guardrail vectorization. We can learn from and refer to the relevant methods, which include vectorization of highway road marking and alignment. Atia *et al.* (2015) proposed a method about road centerlines that were extracted at first with the assistance of trajectory data. Subsequently, the parameters of horizontal features were determined according to a least-squares optimization of characteristic curves. Additionally, Gikas and Stratakos (2011) put an automated solution to estimate curvature diagrams and analyzed horizontal geometric features by using trajectory data. However, these methods depend mainly on the localization precision of GNSS signals, which can be impacted in GNSS weak/denied environments. Zhou *et al.* (2021) proposed an effective framework to extract highway alignments by minimizing an elaborate energy function and to reconstruct highway 3D models with the restrictions of alignments, such as straight lines, circular arcs, and clothoids. Meanwhile, Holgado-Barco *et al.* (2014) also did similar work, dividing different alignments according to road axis modeling and road design standards. Ma *et al.* (2019) put the development of a semiautomated driving line generation method based on MLS point clouds; estimating such road horizontal parameters based on the generated mathematical equations can remarkably improve the estimation precision and reduce the labor cost. Yue *et al.* (2021) made an inventory map, plotting guardrails on Google Earth using their coordinates. The start and end points of each guardrail were correctly verified in Google Street View, but they did not take into account the guardrails of complex highways, including the vectorization of turnoff areas. Processing large MLS point clouds in highway scenes is challenging. Additionally, during mobile measurement, a discontinuous guardrail caused by the movement and occlusions of other vehicles makes detection and instantiation challenging. Other limitations arise from the structure of the guardrails during center point extraction and curve fitting. Such challenges add to the difficulty in generating guardrail 3D trajectories using MLS point clouds.

These considerations show that accurately describing the spatial distribution of a guardrail has essential research value and importance for the precise digital application demand of highway guardrails. To address these challenges, we summarize the limitations and shortcomings of existing methods and propose an effective framework for instantiating highway guardrails and generating high-quality trajectories from segmented point clouds. The primary contributions of this study are as follows:

1. We develop an efficient method for instantiating variant guardrails in complex highways, including straights, turnoffs, and bends of highways.
2. An optimized automated algorithm to generate a 3D trajectory is proposed according to instance labels and geometric features of guardrails.

## Method Flow

### Data Set

We use the Du Wen (DW) and Ya Xi (YX) highway data sets to verify and assess the practical properties in terms of the feasibility and robustness of the proposed approach (see Figures 1 and 2). Point clouds were collected by an MLS system. More information about the MLS system used in this study is shown in Table 1.

### Robust Instantiation of the Guardrails

The strategy used in this study includes an instantiating process based on semantic segmentation and then identifying the guardrails with instance labels to fit curves. Although each point has a corresponding label after the semantic segmentation, these points of guardrails still belong to the same objects without accurate instantiation (see Figure 3a). The points of guardrails with the semantic label cannot provide data support for generating trajectory lines accurately. The four guardrails in the two-way highway have not been instantiated, and the vector lines of each guardrails cannot be generated finely (see Figure 3b). Thus, we propose a local region growth strategy based on linear constraints to assign accurate instance labels.

The primary steps of the proposed process are as follows: (1) extract the scan lines and normalize them, (2) establish a region growth strategy for instantiation, (3) eliminate noise around the guardrail (e.g., pavement, central guardrail platform, and so on) after semantic segmentation, and (4) perform guardrail 3D trajectory fitting and optimization. A flowchart of this process is shown in Figure 4.

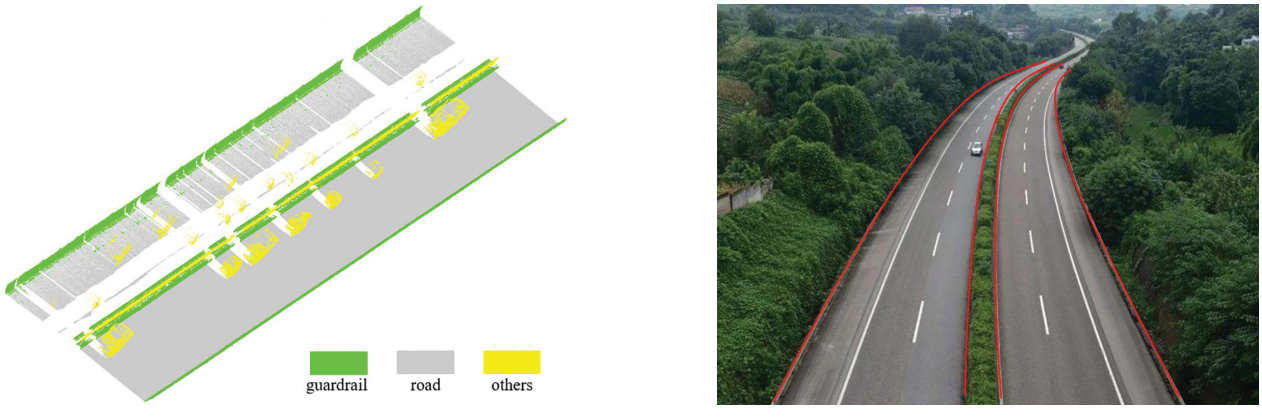



Figure 3. Sematic segment of highway point clouds. (a) Highway point cloud semantic segment. (b) Guardrail lines without instantiation.

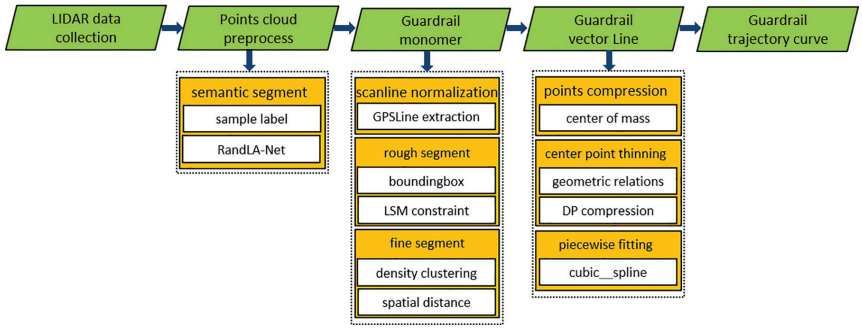



Figure 4. Flowchart of point cloud guardrail instantiation and vectorization method.

### Extracting the Scan Lines

The mobile laser scanner is a linear array and continuous scanning according to the scanning principle. Although the scanning prism rotates 360 degrees, it emits the laser signal only within the field angle range. Therefore, the same scan line has point clouds corresponding only to the field angle. The scanning angle difference of adjacent points in each scan line is fixed, related to the scanning angle resolution (Manandhar and Shibasaki 2001).

The proposed method divides MLS point clouds into a set of scanning lines, each consisting of a highway cross section (Fang and Yang 2013; Yang *et al.* 2013; Yan *et al.* 2016). The scanning lines $L_i - L_n$ are closely and continuously distributed. There is a relatively large time interval ($\Delta T$) from the end of one scanning line to the beginning of another scanning line, and different colors correspond to each scanning line. We analyze the actual collected point clouds, setting $\Delta T = 2000$. According to the $\Delta T$, we determined the two break points in each scan line so as to extract the scan line information of each point and organize the discrete scan points into a series of ordered 2D strips. The GPS time information normalizes according to the time interval (see Figure 5)

$$GPSTk_{i+1} - GPSTk_i > \Delta T \tag{1}$$

$$GPSTk_i = (GPSTk_i - GPST_{MIN})/(GPST_{MAX} - GPST_{MIN}) \tag{2}$$

where $GPSTk_i$ is the instantaneous time of point $k_i$ and $\Delta T$ is the GPS time difference between the last point of one scan line and the first point of the following scan line (as Equations 1 and 2). Each scan line of the point clouds can be extracted through experimental examples by taking four decimal places of the experimental data $L_i - L_n$.

### Region Growing Based on the Linear Constraints

Due to the spatial distribution characteristics of guardrails in different strips of road, there can be large spatial distance changes between two adjacent guardrails. The features of each scanning line are as follows: when there is a large distance between point $k_i$ and point $k_{i+1}$, point $k_i$ is the dividing point, and points $k_i$ and point $k_{i+1}$ belong to two different ground objects. Therefore, Euclidean distance segmentation is performed using the spatial distribution form of the guardrails. We set the distance threshold $d = 8$ m, which can be set according to the road width; divide each scan line into several subblocks through the threshold d; and then start from the initial subblock to establish regional growth constraints. When the conditions are not met, the growth of the guardrail ends, the guardrail that has been grown is removed, and the growth steps of the proposed process are repeated from the remaining guardrail point clusters until all point cloud subblocks are traversed. Due to the shielding of other vehicles in the process of onboard laser movement measurement, the guardrail data are missing or broken, which affects the regional growth of the guardrail. The accuracy of the regional growth of the point clouds directly affects the accuracy of the monomer of the guardrail (Jiang *et al.* 2016).

Therefore, this study develops an efficient and robust regional growth strategy. First, we traverse the points from the initial subblock and then calculate the distance m between the subblock $K_i$ of centroid oi and the subblock $K_{i+1}$ of centroid $O_{i+1}$. When $m > m_t = 1$ m, we set the threshold $m_t = 1$ m according to the geometric structure characteristics of the guardrail (Xu *et al.* 2014) (Figure 5), which indicates that subblock $K_i$ is different from subblock $K_{i+1}$. When m < mT , we cluster the subblock $K_{i+1}$ and the subblock $K_{i+1}$ into a cluster, retain $O_{i+1}$, and continue to compare with the subblock $K_{i+2}$ centroid $O_{i+2}$. We then repeat these steps. When the cluster has the geometric characteristics of the guardrail and the number of points is ≥5000, where 5000 is the initial number of clusters, the points of the guardrail still grow linearly in a small range. Therefore, the least-squares linear fitting of the cluster provides linear constraints (Atia *et al.* 2015; Biçici and Zeybek 2021) and improves the accuracy of point cloud area growth in Figure 6.

To accurately express the distribution characteristics and growth trend, we calculate the points of each scan line and the mean value. We set the radius $r$ to extract the points on each scan line whose distance from the centroid is less than $r$. Linear least-squares fitting is performed with these points, which can obtain the linear equation parameters $E$

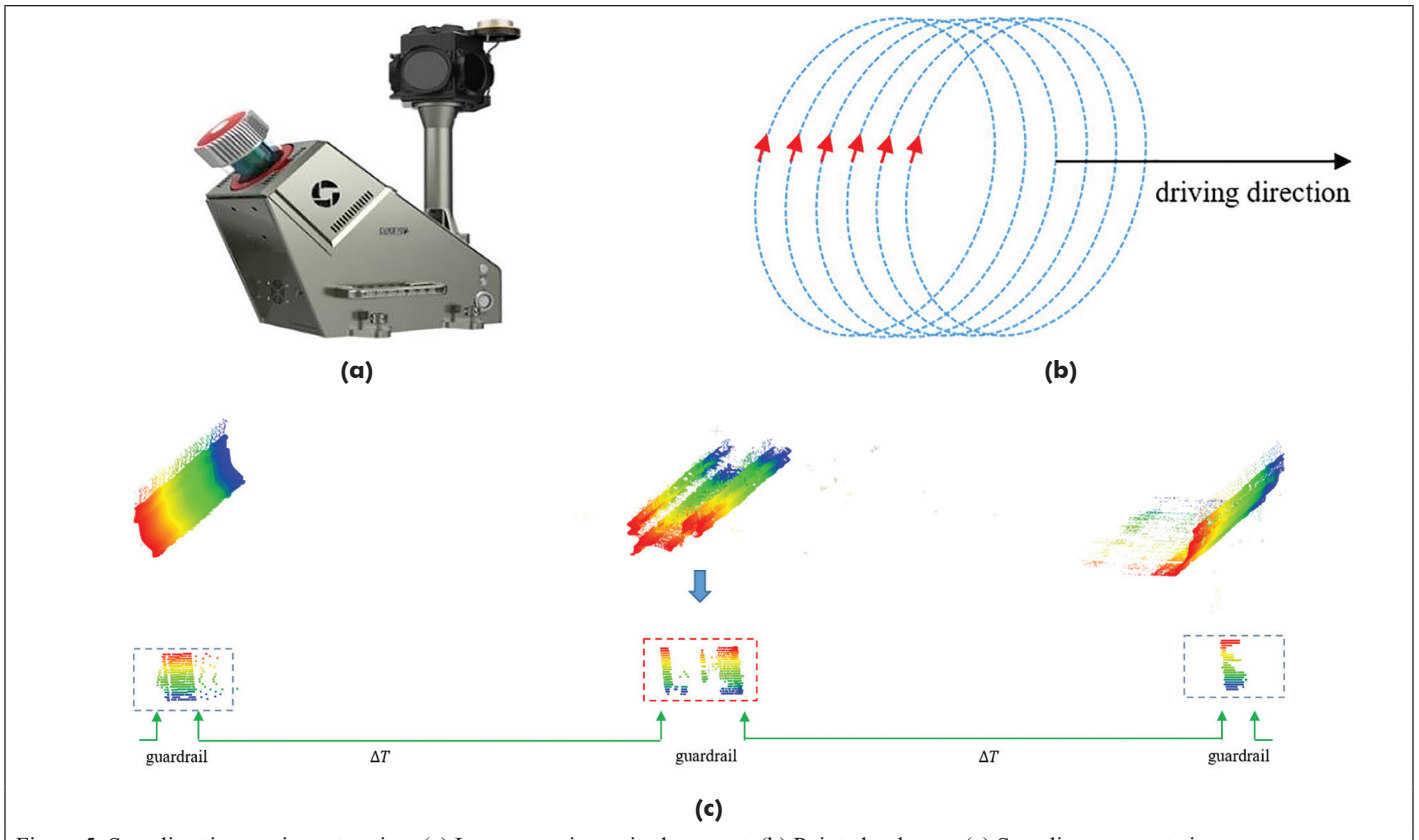



Figure 5. Scan line time-series extraction. (a) Laser scanning unit placement. (b) Point cloud scan. (c) Scan line segmentation.

and $F$. The parameter $F$ is used to build the growth boundaries (Guan *et al.* 2014; Yang *et al.* 2015). When the points of subblock $K_i$ meet the boundaries, subblock $K_i$ is classified into clusters. The points selected by the cluster are updated to obtain new parameters $E_i$ and $F_i$ and a new judgment of the next subblock. Until no new subblock within the growth limit meets the conditions, we stop the cluster growth and mark the cluster as the first split point cluster. The remaining guardrail point clouds also use the above steps to grow the guardrails. Finally, after traversing and calculating all subblocks, the process ends. A comparison of the proposed method and the DBSCAN (Ester *et al.* 1996) method shows that the process can ensure the integrity and continuity of the guardrail and provide guardrail direction (Table 2). The missing and interrupted guardrails due to the movement of other vehicles can be effectively identified (see Figure 7).

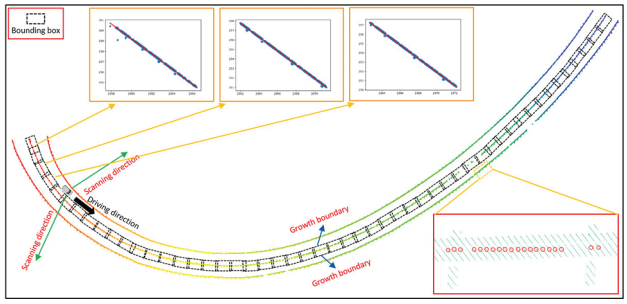


Figure 6. Example diagram of the point cloud area growth method.

Table 2. Comparing the clustering algorithm to the proposed method.

| | Clustering Number | |
|---|---|---|
| **Method** | **Section 1** | **Section 2** |
| DBSCAN method | 33 | 43 |
| Proposed method | 4 | 4 |
| Real value | 4 | 4 |

### Extra Non-Guardrail Point Filtering

Via the above steps, the point clouds after the initial instance segmentation contain some non-guardrail points, such as ground points and platform points (see Figure 8). The next process in the proposed process is to remove non-guardrail points.

There are primarily two types of noise: (1) ground points and (2) platform points (Huang *et al.* 2020). Because the guardrail has marked geometric characteristics, point clusters are arranged closely according to the scanning line. The distribution of ground points is discrete and random with no regularity between points. This article uses the statistical filtering method to eliminate marked outliers. The specific method calculates the distance distribution from the point to the adjacent point in the point clouds. We calculate the average distance from every point to all adjacent points, and then the points with an average distance outside the range can be defined as outliers and removed from the data. In this study, using the standard deviation as the judgment basis, we set the filtering parameters with experimental data. The *knn* parameter is the number of adjacent points (*knn* = 20), and the *std* parameter is the set standard deviation threshold (*std* = 1.0), which are according to the density of the point clouds. Experiments show that this method can effectively eliminate the ground points (see Figure 9a).

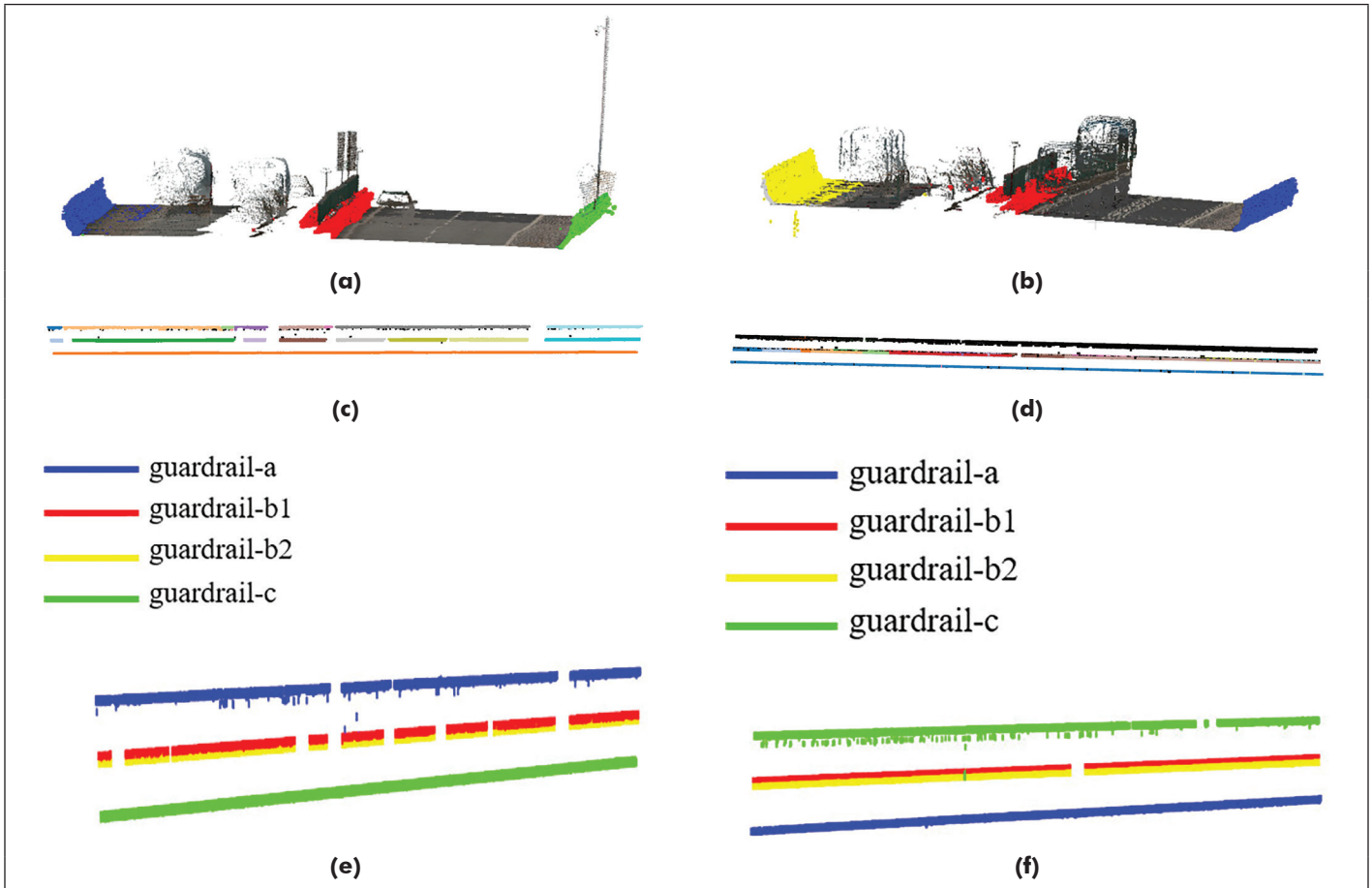
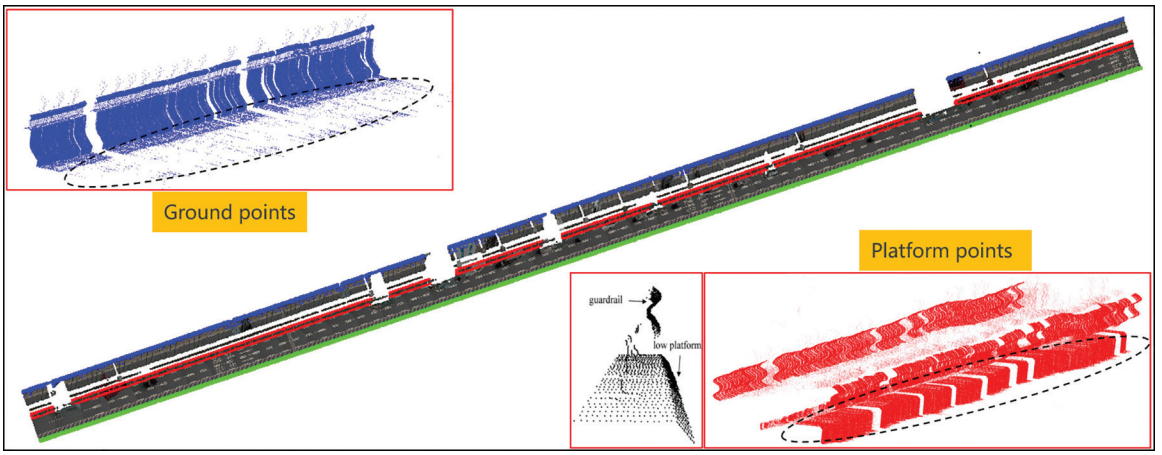
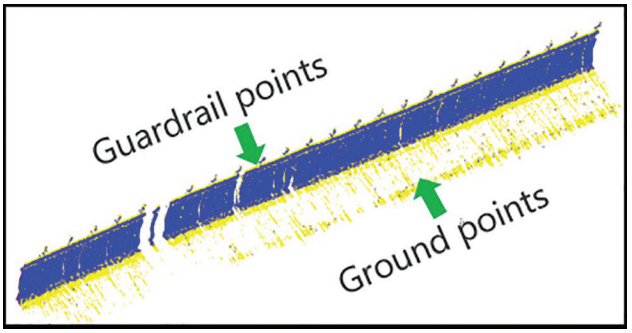


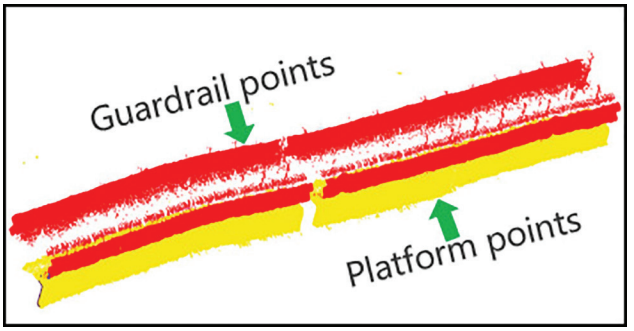
Figure 7. Segment of the guardrail in a straight line section. (a) 3D display of guardrail in Section 1. (b) 3D display of guardrail in Section 2. (c) Result 1 of the DBSCAN cluster method. (d) Result 2 of the DBSCAN cluster method. (e) Result 1 of the proposed method. (f) Result 2 of the proposed method.




(a)

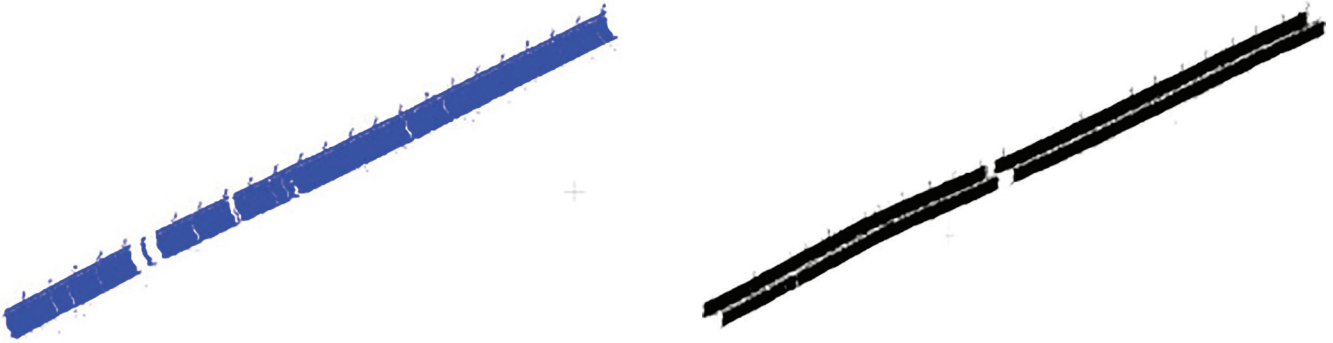



(b)



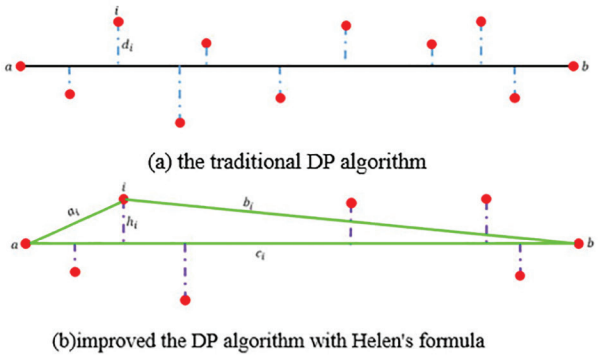

(c)

Figure 8. Two different types of noise. (a) Guardrail points contain some non-guardrail points. (b) Ground points. (c)Platform points.



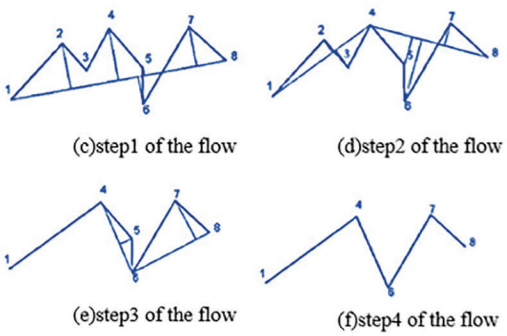
(a) (b)

Figure 9. Fine segmentation of the guardrail. (a) Ground point elimination. (b) Platform point elimination.



(a) the traditional DP algorithm

(b)improved the DP algorithm with Helen's formula



(c)step1 of the flow (d)step2 of the flow

(e)step3 of the flow (f)step4 of the flow

Figure 10. Guardrail center point simplification process based on the improved Douglas–Puke (DP) algorithm. (a) The traditional DP algorithm. (b) The improved DP algorithm with Helen's formula. (c) Step 1 of the flow. (d) Step 2 of the flow. (e) Step 3 of the flow. (f) Step 4 of the flow.

Oversegmentation of the guardrail also generates platform noise. The platform closely connects with the guardrails and is easy to cluster into guardrails. Therefore, the proposed process eliminates the platform by setting the spatial distance threshold. Because the slope of the guardrail on the same scan line is the same, we traverse the points on each scan line of the guardrail and search the boundary points by calculating the slope and slope difference of adjacent points. If the slope difference changes markedly, the point is considered a platform boundary point. The slope of adjacent points is calculated as

$$n_{(i+1)-i} = \frac{z_{i+1} - z_i}{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}} \tag{3}$$

where $n_{(i+1)-i}$ is the slope of point $k_i$ and the adjacent point $k_{i+1}$ on the scan line and $(x, y, z)$ represents the three coordinates of point $k_i$ on the scan line. The slope $k$ of adjacent points of all scan lines is traversed conversely, and the slope difference $\Delta n_i$ is calculated.

After the guardrail boundary points are determined through the above steps, the extracted lower boundary of the guardrail is determined. Therefore, we take the $z$ coordinate of the boundary point as the boundary standard and count all guardrail points with $z$ coordinates greater than the boundary point to divide the guardrail and platform (see Figure 9b).

### 3D Guardrail Trajectory Fitting and Optimization

Compared with other methods, such as that of Soilán *et al.* (2021), that project the guardrail onto the *XOY* plane to extract 2D track points, we directly extract the 3D center points (Holgado-Barco *et al.* 2014) from the guardrails. This process is time consuming using the traditional Douglas–Puke (DP) algorithm to simplify the center points. It is thus necessary to calculate the linear equation of each interval and then calculate the distance from the point to the straight line. We improve the DP algorithm with Helen's formula to improve process efficiency. We convert the distance between point $i$ and the line of two points $a$ and $b$ into the height $h_i$ of the spatial triangle $S_i$ formed by points $i$, a, and b (see Figure 10). These parameters can be defined as

$$P_i = \frac{A_i + B_i + C_i}{2} \tag{4}$$

$$h_i = \sqrt{P_i(P_i - A_i)(P_i - B_i)(P_i - C_i)} / 2A_i \tag{5}$$

In Equations 4 and 5, $A_i$, $B_i$, and $C_i$ are the lengths of the three sides of Si, and Pi is the perimeter. The specific process is as follows. The first and last points on the central line are connected with a straight line and can thus be used to calculate the distance between all points and the straight line. These calculations allow us to identify the maximum distance. Compared with Imax and tolerance $I$, if $I_{max} < I$, all the points on the central line are removed; if $I_{max} \geq I$, the point is reserved, and the central line is divided into two parts with this point as the boundary. The method is repeated for these two parts (see Figure 10). We can control the degree of simplification by controlling $I$. Due to noise, the precision of the extracted center point is poor (see Figure 11a and 11b). The precision of the extracted center points is markedly improved after removing noise points via the proposed method (see Figure 11c). Comparing the trajectory fitting through two different center points, the track of point fitting after noise removal is also smoother (see Figure 11d), while the track of point fitting without noise removal is not smooth (see Figure 11e).

We use the cubic spline function to generate the trajectory curve (Ma *et al.* 2019) to reduce data redundancy and smooth the curve (see Figure 12). Each subinterval node and the start and end break-point conditions are substituted into the matrix equation. The quadratic differential equation is obtained to solve the matrix. The coefficient matrix is decomposed into a lower triangular matrix and upper triangular matrix using the Gaussian elimination method,

$$Ax = (LU)x = (Ux) = Ly \tag{6}$$

where $A = LU$, $L$ is the lower triangular matrix, and $U$ is the upper triangular matrix.

We then calculate the coefficients of the spline curve:

$$\begin{cases} a_i = y_i \\ b_i = \dfrac{y_{i+1} - y_i}{h_i} - \dfrac{h_i}{2} m_i - \dfrac{h_i}{6}(m_{i+1} - m_i) \\ c_i = \dfrac{m_i}{2} \\ d_i = \dfrac{m_{i+1} - m_i}{6h_i} \end{cases} \tag{7}$$

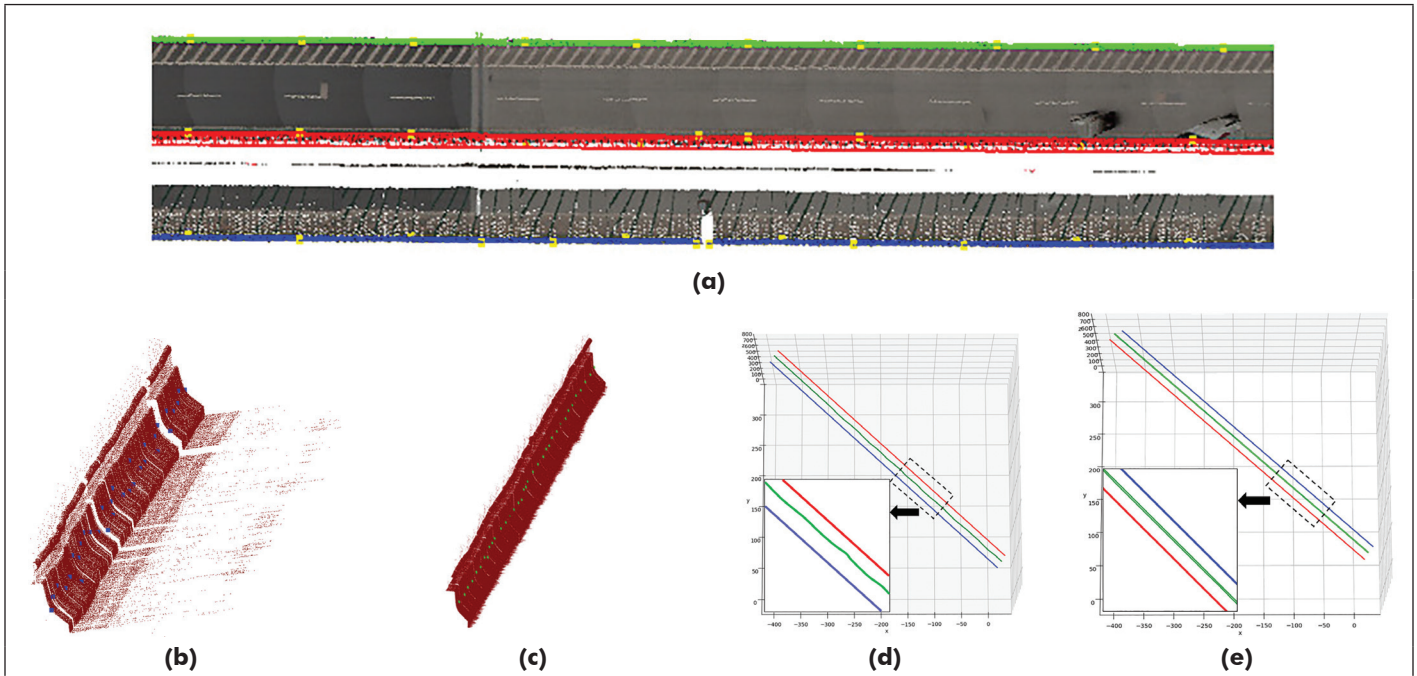



Figure 11. Guardrail center points are extracted in two different cases. (a) Extraction guardrail center points. (b) Unoptimized center points. (c) Optimized center points. (d) Distorted curve fitting. (e) Smooth curve fitting.

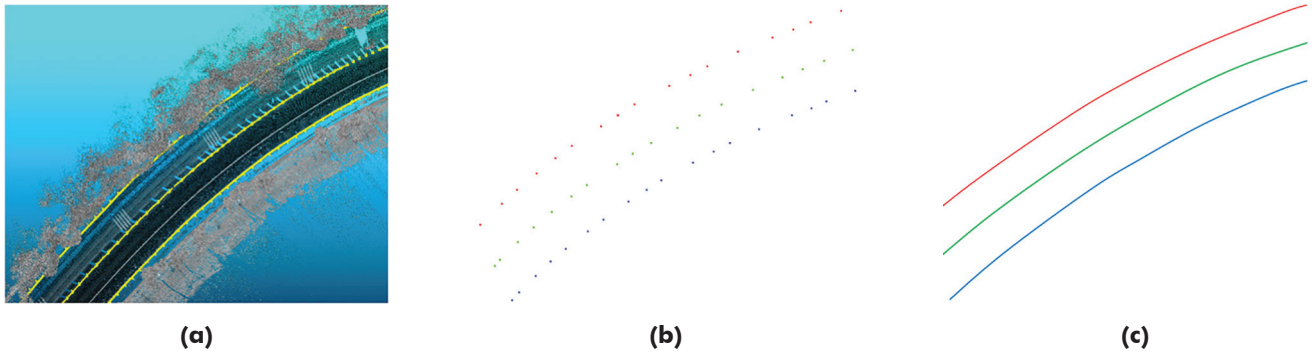



Figure 12. Curve fitting of the guardrail in the bend area of the highway. (a) Guardrail point cloud view. (b) Center point extraction. (c) Curve fitting.

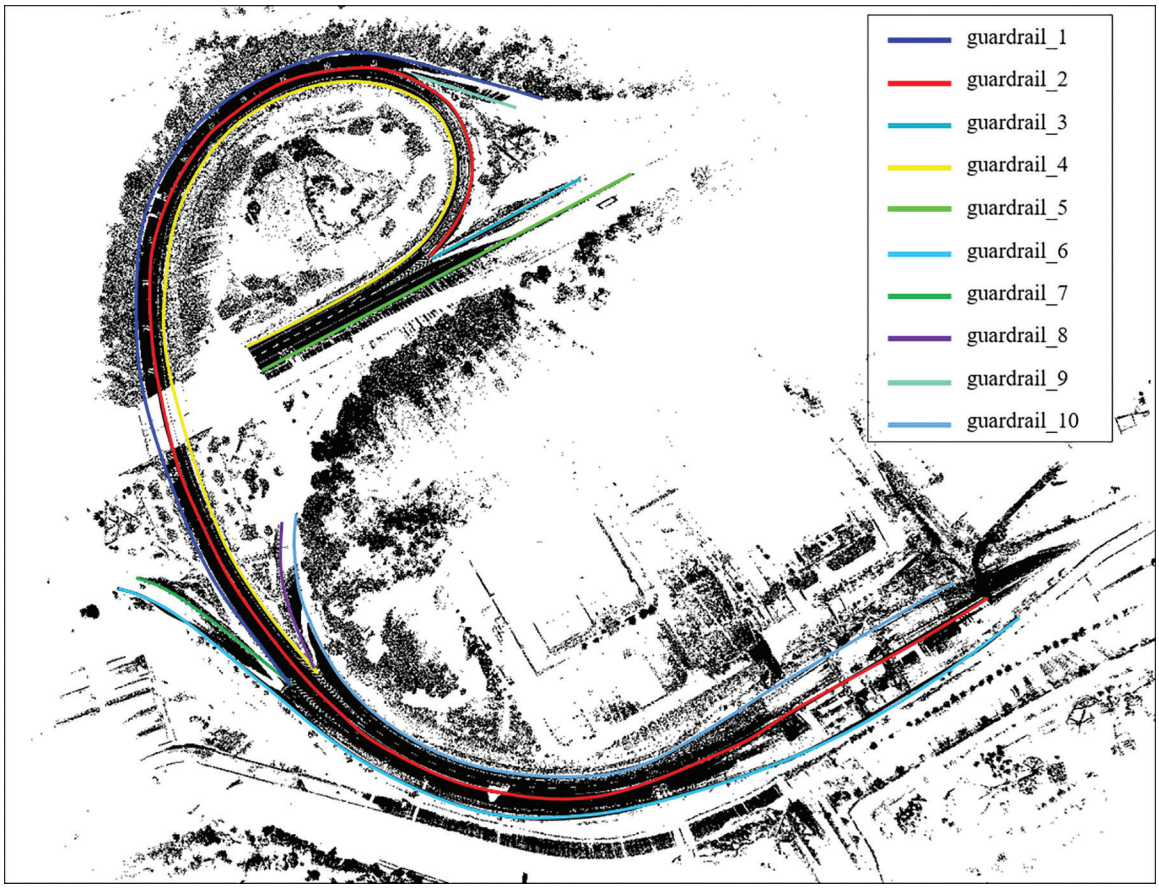



Figure 13. DUWEN highway point cloud guardrail curve fitting view.

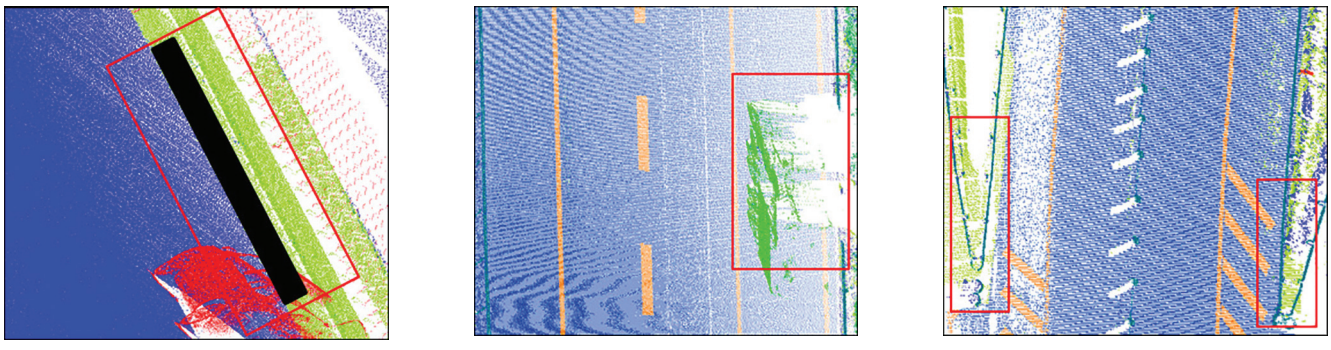



Figure 14. Three different special cases of point cloud guardrail. (a) Platform near the guardrails. (b) Gaps caused by vehicle occlusions. (c) Connection of the turnoff area.

## Experimental Results

The RANDLA-Net (Hu *et al.* 2020) used random sampling instead of farthest point sampling, greatly improving the efficiency. For the guardrails, a good segmentation effect is achieved with the local spatial coding module and the self-attention mechanism. Therefore, we use RANDLA-Net for extracting highway guardrails as preprocessing of MLS point clouds.

The whole algorithm was implemented by the Python language and run on a PC (Intel Core CPU i7-6850K at 3.6 GHz and 64 GB ROM), and TensorFlow was used for network model training. Moreover, we used libraries such as NumPy, Open3D, SciPy, PyYAML, and pandas to construct the process framework. In data preprocessing, the voxel size was set to 0.06 m. The number of input point clouds was set to 65 536, the number of training rounds was set to 50, the initial learning rate was 0.01, each round was reduced by 5%, batch size was set to 5, and the Adam optimizer was used.

### Guardrail Instantiating and Curve Fitting

The proposed method is used to verify some sections of the DUWEN highway (see Figure 13). The point clouds of guardrails are segmented, and the problems in different areas are managed separately. The proposed method can accurately extract the trajectory of each guardrail, which can be divided into 10 types of guardrails, and achieve accurate guardrail instantiation.

The proposed method performs effectively in complex environments. For example, the platform is near the guardrails (see Figure 14a). The detected results also show that the proposed algorithm exhibits high robustness according to the traffic condition invariance. The gaps caused by vehicle occlusions (Soilán *et al.* 2020) are identified correctly by the proposed tracing algorithm (see Figure 14b). The turnoff area is well identified (see Figure 14c), and the noise from vegetation and dynamic vehicles can be filtered while generating guardrail feature points and tracing. We can quantitatively evaluate the proposed method by comparing the detection result with the manually labeled ground truth (see Table 2).

Two groups of different data verify the robustness and convergence of the proposed method: the guardrail data of lines, ramps, and curves with low or high noise. Semantic segmentation of segmented data was performed through experimental comparison, and then the guardrail was finely instantiated using the proposed method (see Figures 15 and 16).

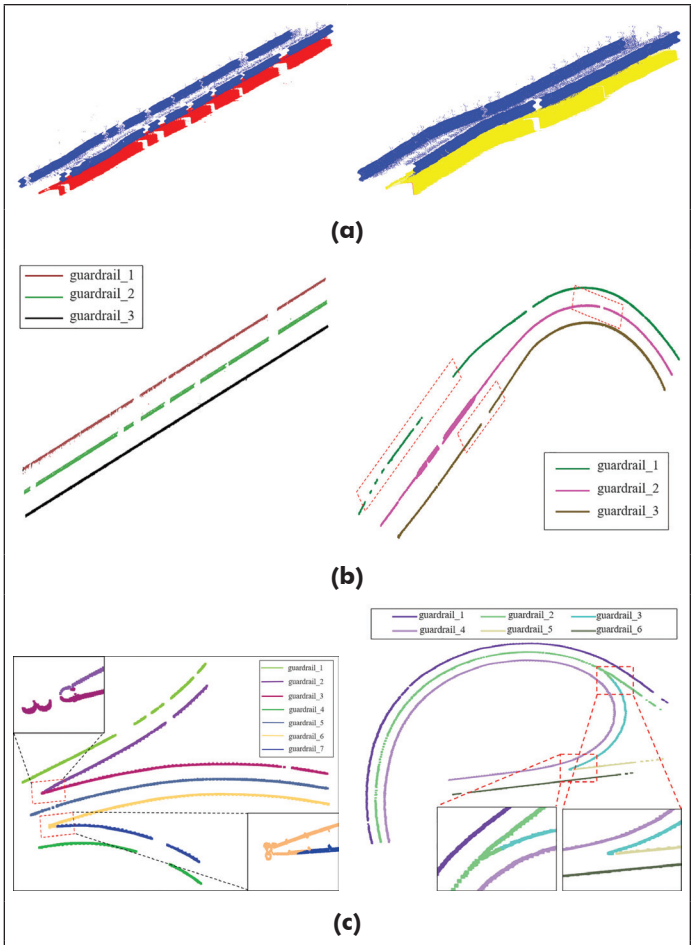



Figure 15. Fine instantiation of guardrail. (a) Platform point elimination. (b) Straight and bend instantiation of guardrail. (c) Turnoff and bend instantiation of the guardrail.

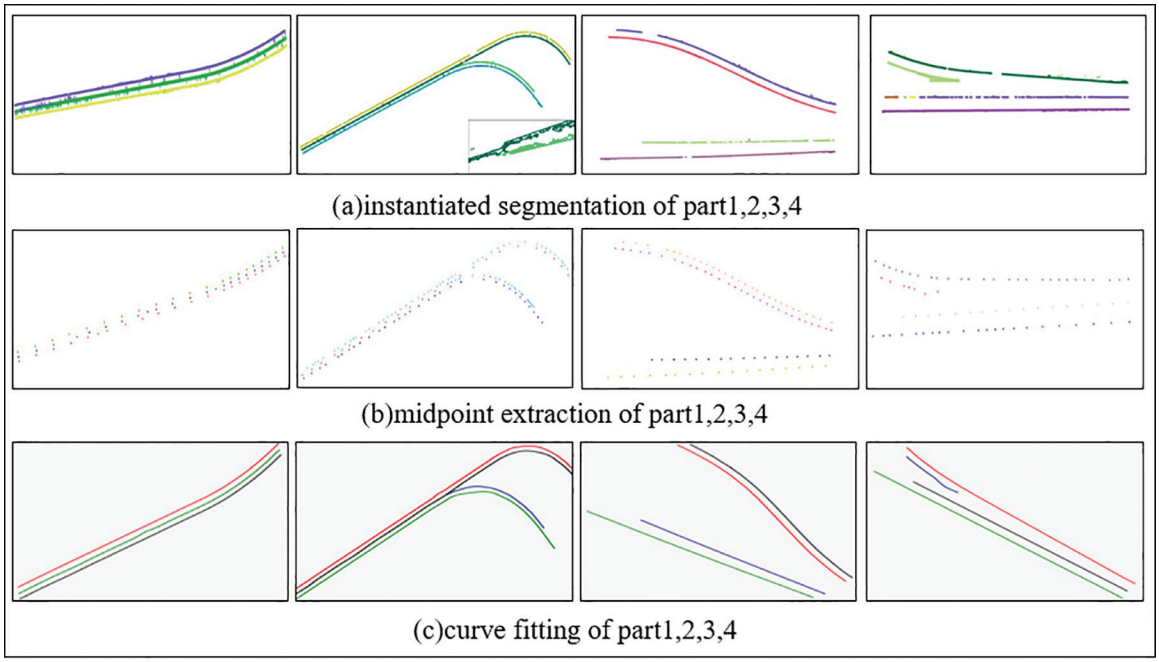



Figure 16. Guardrail with noise divided into four sections of processing steps. (a) Instantiated segmentation of parts 1–4. (b) Midpoint extraction of parts 1–4. (c) Curve fitting of parts 1–4.

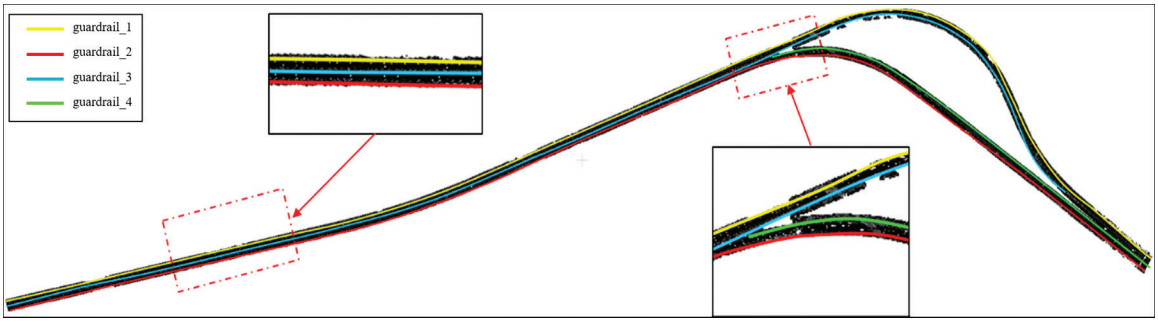


Figure 17. YAXI highway point cloud guardrail curve fitting view.

We perform a second experiment to conduct robust tests for the guardrail with noise, divide it into four sections, instantiate each section, and extract the center point by merging each guardrail to fit the 3D trajectory (see Figure 16). Results show that this method is still valid, eliminates noise accurately, and can instantiate the guardrail accurately (see Figure 17).

### Performance Analysis

We propose some performance evaluation criteria as follows:

$$\text{precision} = \frac{TP}{TP + FP} \tag{8}$$

$$\text{recall} = \frac{TP}{TP + FN} \tag{9}$$

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{10}$$

where *TP*, *FP*, *TN*, and *FN* denote true positives, false positives, true negatives, and false negatives, respectively (as Equation (8), (9), (10) Figures 1b and 2 show the original data of these two data groups. The proposed method is used to verify instance segmentation according to the technical process to evaluate the accuracy. The quantitative and qualitative results are shown in Table 3 and Figures 13 and 17, respectively.

Table 3. Quantitative accuracy evaluation of guardrail segmentation results.

| Test Data | TP+FN | TP+FP | TP | Accuracy (%) | Recall (%) |
|---|---|---|---|---|---|
| DUWEN | 1 450 123 | 1 445 203 | 1 430 952 | 99.0 | 98.6 |
| YAXI | 1 025 707 | 1 016 710 | 996 304 | 97.9 | 97.1 |

## Discussion

In the article,since our main work is for instantiating guardrail and optimizing trajectory, which does not involve and discuss the semantic segmentation of point clouds. We extract guardrail by using RANDLA-net to segment semantically MLS point clouds in highway scenes. The time complexity of the random sampling method in RANDLA-Net does not change with the increase of the number of point clouds. Using RANDLA-Net to extract guardrail points from massive point clouds can speed up time efficiency. Then the guardrail is instantiated on the basis of preprocessing process. There are some advantages to experimental verification. Despite missing point clouds caused by the movement and occlusion of other vehicles, we use a regional growth strategy with local linear constraints to improve robustness. Each guardrail has clear boundaries in straight and bent sections of highway scenes, and each guardrail can obtain specific instance information with a remarkable effect. There are also some limitations. Due to the connectivity between the guardrails, there is no clear boundary between guardrails in the turnoff areas. We can set the critical point of segmentation only according to prior knowledge, and there are some deviations in this process. Subsequent research on the instance segmentation of the guardrail at turnoffs and bends will continue. For complex highway scenes, the proposed method first divides the point cloud data into several sections, each section is instantiated, and the center point is extracted by merging each guardrail to fit the 3D trajectory. We do not use parallel processing in this process; thus, the next step is to consider parallel computing to improve efficiency.

## Conclusion

In this study, a highway guardrail instantiating and trajectory fitting method of mobile laser point clouds is designed and shown to robustly instantiate the guardrail of banded asset elements, eliminate the noise during segmentation, and effectively preserve the integrity of guardrails. In preprocessing, we use libraries such as NumPy, Open3D, SciPy, PyYAML, and pandas to construct RANDLA-Net for extracting guardrails of MLS point clouds. Then we propose a region growth strategy based on linear constraints to assign accurate instance labels. The proposed process can finely instantiate guardrails in the straight road section and has a good effect in turnoff areas, fully describing the practicability of the proposed process. Further, we improve the DP algorithm by Helen's formula to promptly extract center points of guardrails with correct instance labels. Finally, the trajectory is completely fitted according to the cubic spline function, which reduces data redundancy, makes the curve smoother, saves more storage space, and effectively improves the accuracy of data fitting. Experimental results show encouraging performance of the proposed approach.

## Acknowledgments

This study was jointly supported by the Sichuan Science and Technology Fund for Distinguished Young Scholars 22JCQN0110, National Natural Science Foundation of China under Project 42071437, and Project 62006199.

## References

Atia, M., S. Liu, H. Nematallah, T. Karamat and A. Noureldin. 2015. Integrated indoor navigation system for ground vehicles with automatic 3-D alignment and position initialization. *IEEE Transactions on Vehicular Technology* 64(4):1279–1292.

Balado, J., J. Martínez-Sánchez, P. Arias and A. Novo. 2019. Road environment semantic segmentation with deep learning from MLS point cloud data. *Sensors* 19(16):3466.

Biçici, S. and M. Zeybek. 2021. Effectiveness of training sample and features for random forest on road extraction from unmanned aerial vehicle-based point cloud. *Transportation Research Record* 2675(12):401–418.

Broggi, A., P. Cerri, F. Oleari and M. Paterlini. 2005. Guard rail detection using radar and vision data fusion for vehicle detection algorithm improvement and speed-up. *Proceedings of IEEE Intelligent Transportation Systems* 2005:552–556.

Darnel, C. 2012. Using LiDAR to solve industry challenges. *Geo: Geoconnexion International Magazine* 11(1):18–19.

Ester, M., H. Kriegel, J. Sander and X. Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Pages 226–231 in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, held in Portland, Oregon, 4-8 August. Published by The Association for the Advancement of Artificial Intelligence (AAAI) Press: Menlo Park, California,United States.

Fang, L. and B. Yang. 2013. Automated extraction structural roads from mobile laser scanning point clouds. *Acta Geodaetica et Cartographica Sinica* 42(2):260–267.

Gao, J., Y. Chen, J. Marcato, C. Wang and J. Li. 2022. Rapid extraction of urban road guardrails from mobile LiDAR point clouds. *IEEE Transactions on Intelligent Transportation Systems* 23(2):1572–1577.

Gargoum, S., K. El-Basyouny, J. Sabbagh and K. Froese. 2017. Automated highway sign extraction using lidar data. *Transportation Research Record* 2643:1–8.

Gikas, V. and J. Stratakos. 2011. A novel geodetic engineering method for accurate and automated road/railway centerline geometry extraction based on the bearing diagram and fractal behavior. *IEEE Transactions on Intelligent Transportation Systems* 13(1):115–126.

Gomes, S. 2013. The influence of the infrastructure characteristics in urban road accidents occurrence. *Accident Analysis and Prevention* 60:289–297.

Graham, L. 2010. Mobile mapping system overview. *Photogrammetric Engineering and Remote Sensing* 76(3):222–229.

Gressin, A., C. Mallet, J. Demantké and N. David. 2013. Towards 3D lidar point cloud registration improvement using optimal neighborhood knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing* 79:240–251.

Guan, H., J. Li, Y. Yu, C. Wang, M. Chapman and B. Yang. 2014. Using mobile laser scanning data for automated extraction of road markings. *ISPRS Journal of Photogrammetry and Remote Sensing* 87:93–107.

Holgado-Barco, A., D. Gonzalez-Aguilera, P. Arias-Sanchez and J. and Martinez-Sanchez. 2014. An automated approach to vertical road characterisation using mobile LiDAR systems: Longitudinal profiles and cross sections. *ISPRS Journal of Photogrammetry and Remote Sensing* 96:28–37.

Hu, Q., B. Yang, L. Xie, S. Rosa, Y. Guo, Z. Wang and N. Trigoni. 2020. RandLA-Net: Efficient semantic segmentation of large-scale point clouds. Pages 11108–11117 in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, held in Seattle,Washington,United States,13-19 June. Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway, New Jersey, United States.

Huang, M., C. Zhou, P. Huo and Z. Wu. 2020. Expressway guardrail extraction and classification based on slice method. *Sensors and Materials* 32(11):4005.

Jiang, Y., B. He, L. Liu, R. Ai and X. Lang. 2016. Effective and robust corrugated beam guardrail detection based on mobile laser scanning data. Pages 1540–1545 in *Proceedings of 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, held in Rio de Janeiro, Brazil, 1–4 November. Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway, New Jersey, United States.

Jokkola, A., J. Hyyppa and H. Hyyppa. 2008. Retrieval algorithms for road surface modeling using laser-based mobile mapping. *Sensors* 8:5238–5249.

Justo, A., M. Soilán, A. Sánchez-Rodríguez and B. Riveiro. 2021. Scan-to-BIM for the infrastructure domain: Generation of IFC-complaint models of road infrastructure assets and semantics using 3D point cloud data. *Automation in Construction* 127:103703.

Lehtomaki, M., A. Jaakkola, H. Juha, A. Kukko and H. Kaartinen. 2010. Detection of vertical pole-like objects in a road environment using vehicle-based laser scanning data. *Remote Sensing* 2(3):641–664.

Loprencipe, G., L. Moretti, G. Cantisani and P. Minati. 2018. Prioritization methodology for roadside and guardrail improvement: Quantitative calculation of safety level and optimization of resources allocation. *Journal of Traffic and Transportation Engineering* 5(5):348–360.

Ma, L., Y. Li, J. Li and Z. Zhong. 2019. Generation of horizontally curved driving lines in HD maps using mobile laser scanning point clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12(5):1572–1586.

Manandhar, D. and R. Shibasaki. 2001. Vehicle-borne laser mapping system (VLMS) for 3-D GIS. Pages 2073–2075 in *Geoscience and Remote Sensing Symposium, IGARSS '01*, held in Sydney, New South Wales, Australia,9-13 July.Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway, New Jersey, United States.

Matsumoto, H., Y. Mori and H. Masuda. 2019. Extraction and shape reconstruction of guardrails using mobile mapping data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42:1061–1068.

Pu, S., M. Rutzinger, G. Vosselman and S. Elberink. 2011. Recognizing basic structures from mobile laser scanning data for road inventory studies. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(6):S28–S39.

Qi, C., H. Su, K. Mo and L. Guibas. 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. Pages 652–660 in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, held in Honolulu, Hawaii,United States,21-26 July.Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway,New Jersey,United States.

Scharwächter, T., M. Schuler and U. Franke. 2014. Visual guard rail detection for advanced highway assistance systems. Pages 900–905 in *Proceedings of IEEE Intelligent Vehicles Symposium*, held in Dearborn, Michigan, United States, 8-11 June. Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway, New Jersey, United States.

Seibert, A., M. Hähnel, A. Tewes and R. Rojas. 2013. Camera based detection and classification of soft shoulders, curbs and guardrails. Pages 853–858 in *Proceedings of IEEE Intelligent Vehicles Symposium*, held in Gold Coast, Queensland, Australia,23-26 June. Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway, New Jersey, United States.

Soilán, M., A. Justo, A. Sánchez-Rodríguez and B. Riveiro. 2020. 3D point cloud to BIM: Semi-automated framework to define IFC alignment entities from MLS-acquired LiDAR data of highway roads. *Remote Sensing* 12(14):2301.

Soilán, M., A. Nóvoa, A. Sánchez-Rodríguez, A. Just and B. Riveiro. 2021. Fully automated methodology for the delineation of railway lanes and the generation of IFC alignment models using 3D point cloud data . *Automation in Construction* 126:103684.

Vidal, M., L. Díaz-Vilarino, P. Arias and J. Balado. 2020. Barrier and guardrail extraction and classification from point clouds. *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences* 43:157–162.

Xiong, W., B. Yang and Z. Dong. 2016. Refining and robust extraction of roads from mobile laser scanning point clouds. *Journal of Geo-Information Science* 18(3):376–385.

Xu, S., G. Vosselman and S. Elberink. 2014. Multiple-entity based classification of airborne laser scanning data in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing* 88(2):1–15.

Yan, L., H. Liu, J. Tan, Z. Li, H. Xie and C. Chen. 2016. Scan line based road marking extraction from mobile LiDAR point clouds . *Sensors (Basel)* 5(16):1–21.

Yang, B., Z. Dong, G. Zhao and W. Dai. 2015. Hierarchical extraction of urban objects from mobile laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing* 99:45–57.

Yang, B., L. Fang and J. Li. 2013. Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 79:80–93.

Yue, Y., M. Gouda and K. El-Basyouny. 2021. Automatic detection and mapping of highway guardrails from mobile LiDAR point clouds. Pages 2520–2523 in *IEEE International Geoscience and Remote Sensing Symposium IGARSS*, held in Brussels, Belgium, 1-16 July. Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway, New Jersey, United States.

Zeybek, M. 2021. Extraction of road lane markings from mobile LiDAR data. *Transportation Research Record* 2675(5):30–47.

Zhou, Y., R. Huang, T. Jiang, Z. Dong and B. Yang. 2021. Highway alignments extraction and 3D modeling from airborne laser scanning point clouds. *International Journal of Applied Earth Observation and Geoinformation* 102:102429.

Zhu, H. and B. Guo. 2018. A beam guardrail detection algorithm using lidar for intelligent vehicle. Pages 1398–1402 in *Proceedings of IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, held in Tianjin, China, 19-23 July. Published by Institute of Electrical and Electronics Engineers (IEEE): Piscataway,New Jersey,United States.

# Call for *PE&RS* Special Issue Submissions

## Innovative Methods for Geospatial Data using Remote Sensing and GIS

Internationally comparable data is a global need for managing resources, monitoring current trends and taking actions for sustainable living. Even though there has been a significant progress on geospatial data availability, extensive data gaps are still a major problem for general assessment and supervise the progress through the years. According to United Nations 2022 The Sustainable Development Goals Report, while health and energy sectors have the highest data available, limited data available for climate action.

The COVID-19 crisis has also shown that there are innovative data collection methods utilizing information and computer technologies. However, only 5% of the countries have benefit from remote sensing technologies to measure the impact of COVID-19. Additionally, novel approaches such as artificial intelligence should be used in conjunction with assessments to make sure they are put to use for critical situations.

The recent developments in remote sensing, geographic information systems and ICT have provided a wide accessibility to create geospatial data for various purposes. The proposed special issue focuses on *"Innovative Methods for Geospatial Data using Remote Sensing and GIS" for wide range of applications. This special issue aims to bring researchers to share knowledge and their expertise about innovative methods to contribute to fill data gaps around the world for a better future.*

*The proposed special issue aims to contributes ASPRS's key mission on 'Simplify and promote the use of image-based geospatial technologies for the end-user', 'Promote collaboration between end users and geospatial experts to match data and technology to applications and solutions' and 'promote the transfer of geospatial data and information technology to developing nations' by providing innovative methods to create geospatial data using remote sensing and geographic information systems utilizing state-of-the-art developments and solutions.*

**Deadline for Manuscript Submission—July 1, 2023**

**Submit your Manuscript to http://asprs-pers.edmgr.com**

### Guest Editors

**Dr. Tolga Bakirman**, bakirman@yildiz.edu.tr , *Yildiz Technical University, Department of Geomatic Engineering*, Davutpasa Campus, 34220 Esenler-Istanbul/Turkey

**Dr. George Arampatzis**, garampatzis@pem.tuc.gr, *Technical University Crete, School of Production Engineering & Management*, 73100 Chania – Crete/Greece

# Use of Artificial Intelligence Toward Climate-Neutral Cultural Heritage

Tolga Bakirman, Bahadir Kulavuz, and Bulent Bayram

## Abstract

*Cultural heritage (CH) aims to create new strategies and policies for adapting to climate change. Additionally, the goals of sustainable development aim to protect, monitor, and preserve the world's CH and to take urgent action to combat climate change and its effects. Therefore, developing efficient and accurate techniques toward making CH climate neutral and more resilient is of vital importance. This study aims to provide a holistic solution to monitor and protect CH from climate change, natural hazards, and anthropogenic effects in a sustainable way. In our study, the efficiency of deep learning using low-cost unmanned aerial vehicles and camera images for the documentation and monitoring of CH is investigated. The dense extreme inception network for edge detection and richer convolutional feature architectures have been used for the first time in the literature to extract contours and cracks from CH structures. As a result of the study, F1 scores of 61.38% and 61.50% for both architectures, respectively, were obtained. The results show that the proposed solution can aid in monitoring the protection of CH from climate change, natural disasters, and anthropogenic effects.*

## Introduction

Natural and cultural resources are fundamental heritages for human well-being. While natural resources are needed for survivorship, cultural resources provide largely nonmaterial, nonrenewable benefits to humans, such as identity and culture. Cultural heritage (CH) is an essential bridge between the past and the present in terms of the architecture, social structure, development, and lifestyle of a society (Bakirman *et al.* 2020). National or regional strategies, cohesion policies, or spatial planning consider CH as resources for development, especially in relation to tourism (World Tourism Organization 2020). Besides its historical and educational role, CH provides an immediate value creation for the cultural enrichment of the individual and local identity and provides important evidence of the effects of past climate change and how humanity has adapted to these changes (Markham *et al.* 2016). The need to develop preparedness for climate change and natural disasters, national and regional adaptation strategies, and intersectoral risk management plans has become a strategic target for climate-neutral CH.

Climate change, as revealed by gradual changes in temperature, precipitation, atmospheric moisture, water warming, and wind intensity, as well as sea-level rise, coastal erosion, and changes in the occurrence of extreme events (tsunamis, earthquakes, and so on), are affecting CH sites. Other hazards related to climate change (e.g., global warming and anthropogenic effects), such as acid rains, extended dry seasons or floods, heavy storms, and forest fires, are threatening CH around the world (Sesana *et al.* 2021). Identification of new ways of raising benefits for the public and financing climate-neutral CH research while keeping high standards of monitoring and protection is needed. In order to avoid negative natural and anthropological effects on the CH while making it accessible to the public, CH management mechanisms should be established. The protection of CH should aim to control damage from human intrusion and environmental factors. Therefore, cost-effective, reliable, and accurate monitoring of CH constitutes a necessary step toward risk reduction and the pursuit of protection and preservation strategies.

A wide range of data and instruments can be used for digital documentation and monitoring of CH, such as terrestrial laser scanners, unmanned aerial vehicles (UAVs), and cameras. Moreover, image processing techniques, such as edge detection, can be exploited to determine current conditions and characteristics in the monitored structure and to automatically detect temporal damages on the surface by extraction of contours and cracks. The edge detection problem has been investigated for many years in the field of computer vision. Many algorithms and methods have been developed for automatic edge detection in the literature (Figure 1). The Sobel operator (Sobel 1970), Laplacian of Gaussian (Marr and Hildreth 1980), zero crossing (Torre and Poggio 1986), and the Canny detector (Canny 1986) are classical edge detectors that are the pioneering methods in the field. Learning-based algorithms were proposed as technology developed. For example, the probability of boundary algorithm (Martin 2004) aims to extract edges using natural scenes (light, color, and detail) in images. The derivative of infinite symmetric exponential filter (McIlhagga 2011) and color boundary (Yang *et al.* 2013) methods were proposed over time and inspired by the boosted edge learning (Dollar *et al.* 2006), multiscale (Ren 2008), smoothing proximal gradient (Chen *et al.* 2012), and Canny methods. With the increase in the processing capacity of computers and the rise of artificial intelligence techniques, pointwise mutual information (Isola *et al.* 2014), DeepNet (Kivinen *et al.* 2014), neural network nearest neighbor fields (Ganin and Lempitsky 2014), and relaxed deep supervision (Liu and Lew 2016) algorithms using convolutional neural networks that provide deep learning (DL)–based automatic feature extraction have been developed. Many subsequent algorithms based on a particular backbone network—holistically-nested edge detection (HED) (Xie and Tu 2015), convolutional encoder-decoder network (Yang *et al.* 2016), simultaneous edge alignment and learning (Yu, Liu, *et al.* 2018), richer convolutional features (RCF) (Liu *et al.* 2017), learning to predict crisp boundaries (Deng *et al.* 2018), dense extreme inception network for edge detection (DexiNed) (Soria *et al.* 2020), bidirectional cascade network (He *et al.* 2019), and RefineContourNet (Kelm *et al.* 2019) architectures—have been developed in the literature with increasing edge detection accuracy. Architectures that do not need pretrained models—such as the fast inference network for edge detection (Wibison 2020); reflectance, illumination, normal, and depth net (Pu *et al.* 2021); pixel difference networks (Su *et al.* 2021); edge detection with transformer (Pu *et al.* 2022); diverse deep super-vision (Liu *et al.* 2022); and lightweight dense convolutional neural network (Soria *et al.* 2022) architectures—have also been designed recently (Jing *et al.* 2022).

There are also studies focusing on DL-based edge detection applications using three-dimensional point clouds. The edge-aware point set consolidation network (Yu, Li, *et al.* 2018) is a pioneer in this

Tolga Bakirman, Bahadir Kulavuz, and Bulent Bayram are with Yildiz Technical University, Faculty of Civil Engineering, Department of Geomatics Engineering, Istanbul, Türkiye (bakirman@yildiz.edu.tr).

*Contributed by Alper Yilmaz, August 8, 2022 (sent for review October 11, 2022; reviewed by Izidor Mlakar, Dursun Zafer Seker, Akhtar Jamil).*

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 3, March 2023, pp. 163–171.
0099-1112/22/163–171

doi: 10.14358/PERS.22-00118R2

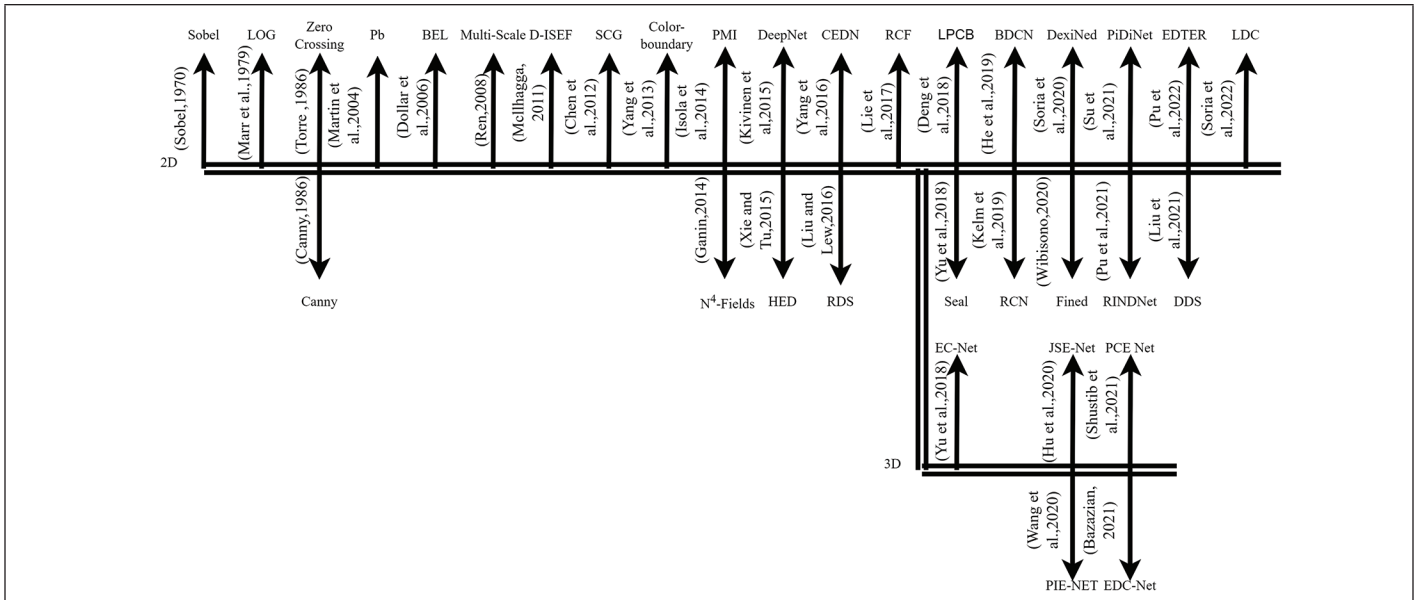


Figure 1. Chronological representation of algorithms used for edge detection.

field with its regression component and edge-sensitive loss function to strengthen point clouds. The joint semantic segmentation and edge detection network (Hu *et al.* 2020) performs the tasks of detecting semantic segmentation edge points from the point cloud. Wang *et al.* (2020) proposed a method that detects edge and corner points based on parametric curves. It works with a parameter based on a shape-dependent differential information set created at different scales around each point in the developed parametric inference of point cloud edges network. The edge detection capsule network (Bazazian and Pares 2021) network is used to structuralize the recently popular capsule networks for edge extraction. However, we do not dwell on these methods since they are out of the scope of this study.

Despite these developments in artificial intelligence and especially in the field of DL, the use of these networks on CH is quite limited in the literature (Fiorucci *et al.* 2020). Compared to the current literature, to the best of our knowledge, there are no up-to-date studies that propose DL-based edge detection to extract contours and cracks for CH monitoring. In this study, we investigate the use of state-of-the-art DL architectures, namely, DexiNed and RCF, for CH monitoring and documentation. Our main contributions to the literature can be summarized as follows:

- DL-based DexiNed and RCF architectures were exploited for the first time in the literature for the extraction of contours and cracks on CH.
- Our own labeled edge detection data set specifically for CH was generated.
- The proposed experimental setup and configuration was utilized for the first time in the literature, particularly for automatic edge detection on CH.
- In order to obtain well-detected edges, a new binarization postprocessing procedure was proposed.
- A holistic and sustainable CH monitoring solution was proposed.

This article is structured as follows. In "Materials and Methods," we explain how we created our data set and how the architectures function. In "Results and Discussion," we provide comparative accuracy metrics, postprocessing, and an independent CH example from Cappadocia. We conclude with final remarks and wider impacts on relevant directives.

## Materials and Methods

Considering the climate and anthropogenic effects, it is necessary to change priorities from focusing solely on estimating impacts and vulnerabilities to include adaptive proactive and planned mitigation approaches in the planning process, the identification of the determinants of adaptive capacity, and the barriers to adaptation. Therefore, we proposed a DL-based solution to improve the monitoring, maintenance, and preparedness for new risks on CH.

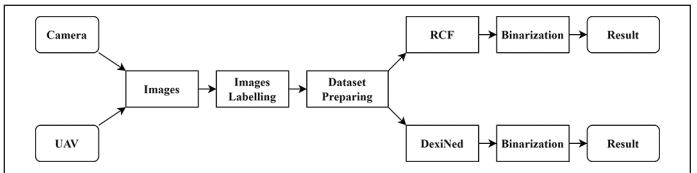


Figure 2. General workflow of the study.

The general workflow of the study is given in Figure 2. As a first step, our edge detection data set was generated specifically for CH using images from 89 historical buildings. Our data set consists of various Internet, camera, and low-cost UAV images. The images have three bands (RGB) and 8-bit radiometric resolution. Sample images from the data set can be seen in Figure 3. All images are manually digitized for contours and cracks. The images were then split into 512×512 patches since the images do not have a fixed geometric resolution. The images that have an edge detail ratio of less than 2% within the image have been eliminated from the data set, and we have obtained 5380 image patches. The data set is then split into 90% (4880 image patches) and 10% (500 image patches) for training and testing, respectively.

Most DL architectures require extensive amounts of data for training. Therefore, we have applied horizontal flip, gamma correction (0.30 and 0.60), and rotation (19°, 57°, 90°, and 180°) augmentations on the train images in order to increase data size. As a result, a total of 146 400 images (512×512) are generated for the training set.

DexiNed and RCF architectures have been utilized for the first time in the literature in terms of the concept of our study to detect CH contours and cracks on the created data set.

DexiNed architecture was created based on the Xception architecture. DexiNed consists of two basic structures: Dexi, an extremely dense initial layer, and USNet, a sampling layer (Figure 4). Dexi architecture consists of six main blocks. These blocks have different access layers. By feeding the feature maps obtained by these blocks, separate USNet architectures and six intermediate edge maps are produced. The resulting intermediate maps are combined, and a stack of learned filters is created. The features obtained at the end of the architecture are combined, and the resulting edge map is created. The sampling network consists of two blocks. Using Block-2, the feature map dimensions are increased to two times the label size. Then the new feature map obtained is transmitted to Block-1. Block-1 has a Relu activation function, and the size of the feature map at the end of the block is matched

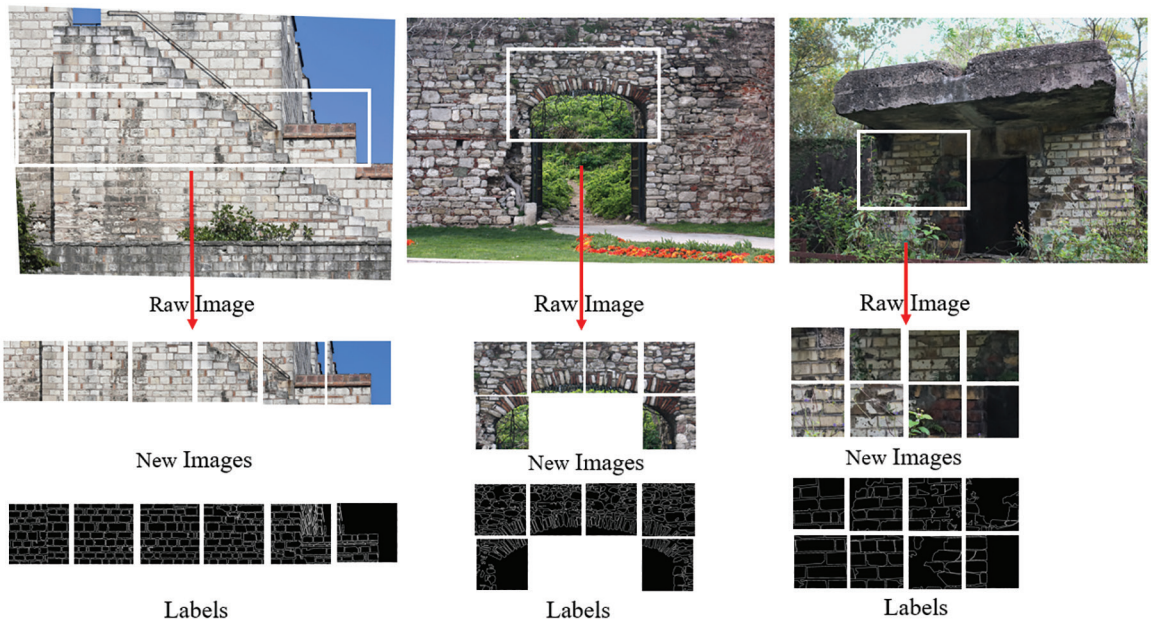



Figure 3. Editing the data set.

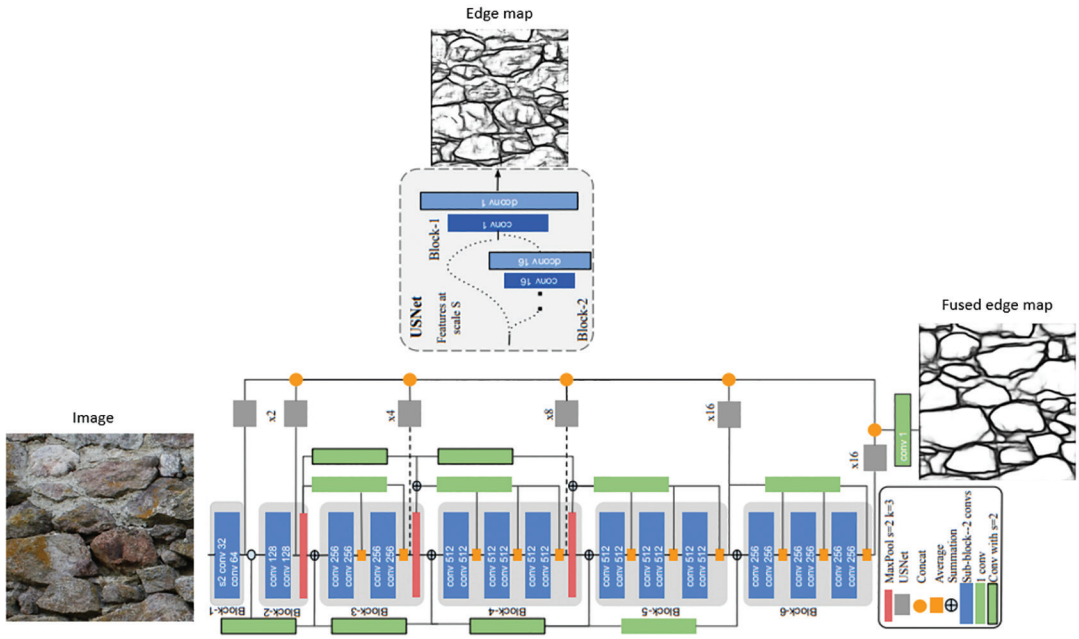



Figure 4. The general structure of the dense extreme inception network for edge detection architecture.

to the size of the label image using a 1×1 kernel. The main purpose of using sampling blocks is to obtain thin edges while visualizing edge maps (Soria *et al.* 2020).

The loss function used by DexiNed is given in Equation 1:

$$\beta = \frac{|Y^-|}{|Y^+ + Y^-|}$$
$$(1-\beta) = \frac{|Y^+|}{|Y^+ + Y^-|}$$
$$l^n(W, w^n) = -\beta \sum_{j \in Y^+} \log\sigma\left(y_j = 1 \mid X; W, w^n\right) - (1-\beta) \sum_{j \in Y^-} \log\sigma\left(y_j = 0 \mid X; W, w^n\right) \quad (1)$$
$$L(W, w) = \sum_{n=1}^{N} \delta^n \times l^n \left(W, w^n\right)$$

Here, the values $|Y^-|$ and $|Y^+|$ denote the edge and non-edge in the label, respectively, and are obtained by using the "Y" label image and the prediction image. Then the calculation is made with "X" input images. This loss function is designed based on HED architecture. In the equation, $W$ is the collection of all network parameters, $w$ is the $n$ corresponding parameter, and $\delta$ is a weight for each scale level.

RCF architecture is based on the VGG16 architecture (Figure 5). Unlike DexiNed architecture, it consists of five blocks. The feature maps obtained from each block are collected with the help of the eltwise layer. Deconv layers are used to up-sample the feature maps obtained from the layers. The up-sampling layers in the blocks are connected to the sigmoid layer. The resulting edge map is produced by combining the feature maps obtained from each interconnected block (Liu *et al.* 2017).

The loss function used by RCF is given in Equation 2:

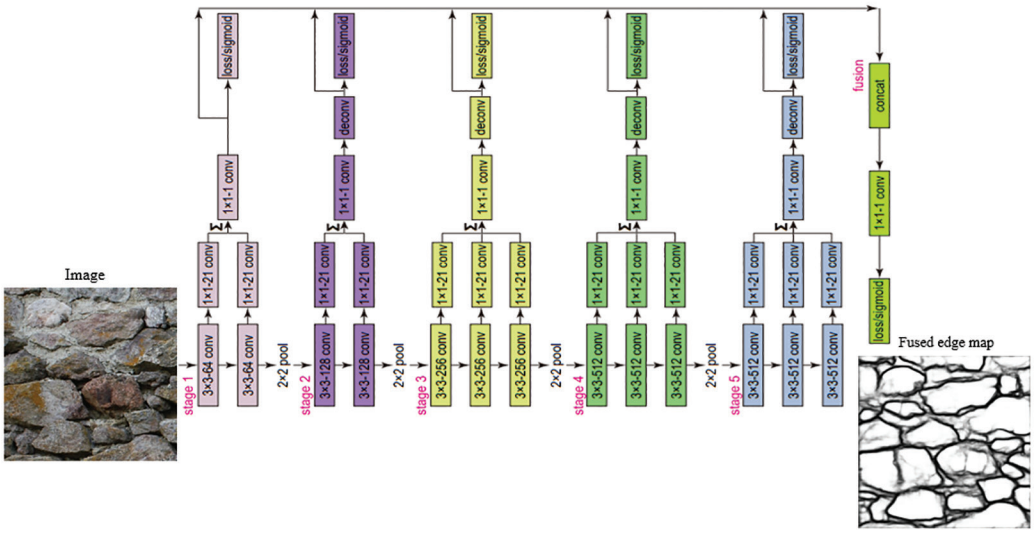



Figure 5. The general structure of the richer convolutional features architecture.

$$\alpha = \lambda \cdot \frac{|Y^+|}{|Y^+| + |Y^-|}$$

$$\beta = \frac{|Y^-|}{|Y^+| + |Y^-|}$$

$$l(Xi;W) = \begin{pmatrix} \alpha \cdot \log\left(1 - P(Xi;W)\right) & \text{if } y_i = 0 \\ 0 & if\ 0 < yi \le \cdot \\ \beta \cdot \log P(Xi;W) & \text{otherwise} \end{pmatrix} \quad (2)$$

$$L(W) = \sum_{i=1}^{|I|} \left( \sum_{k=1}^{K} l\left(X_i^{(k)};W\right) + l\left(X_i^{fuse};W\right) \right)$$

Here, The parameter $\lambda$ is a constant used to balance the number of positive and negative samples. The $\eta$ hyperparameter acts as a threshold and ignores negative samples. The values $|Y^-|$ and $|Y^+|$ denote the positive sample set and the negative sample set, respectively, and P(x) is used to represent the standard sigmoid function. The developed loss function is based on the loss function used in the HED architecture.

## Results and Discussion

In this study, a workstation with an 11th-generation Intel Core i9-11900 2.50-GHz processor and an NVIDIA Quadro RTX 5000 16-GB graphics card was used for training and testing of the DL architectures. The hyperparameters used in the training are shown in Table 1. In the training of the data set, we first investigated the effect of the used epochs on the accuracy. For this purpose, initial weights from Barcelona Images for Perceptual Edge Detection (BIPED) (BIPED 2020) and ImageNet (Deng *et al.* 2009) data sets have been used for training of DexiNed and RCF architectures on our data set for 10, 60, and 150 epochs, respectively. The results depicted that 10 epochs generated higher accuracy with a 44.12% F1 score. After the number of epochs was determined, the effect of the average RGB value that the DL architectures demand was investigated and tested. According to assessments, it was observed that the demanded average RGB value depends on the used data set. Thus, we defined unique RGB average values for our data set as 131.2563, 136.4856, and 140.4427, respectively.

The accuracy results to be obtained by using the weight files belonging to the architectures and without using the initial weights were compared in the training. First, the DexiNed architecture was trained for 10 epochs with and without initial weights. Then the same process was carried out for the RCF architecture. However, the initial weights of the RCF architecture did not match the data set, and very poor prediction results were obtained.

Table 1. Hyperparameters used in architectures.

| Hyperparameters | DexiNed | RCF |
|---|---|---|
| Epoch | 10 | 10 |
| Learning rate | 0.0001 | 1E-06 |
| Activation function | ReLu | ReLu |
| Optimizer | Adam | SGD |
| Batch size | 8 | 1 |
| Number of images | 146 400 | 146 400 |
| Image sizes | 512×512 | 512×512 |

DexiNed = dense extreme inception network for edge detection; RCF = richer convolutional features architecture.

Testing was carried out for both architectures using the weights of the 4th and 10th epochs. After the testing phase, our binarization approach was implemented on the prediction images since the obtained results are not binary (Figure 6a). First, the Non-Maximum Suppression (NMS) method was applied, and the images were converted to binary images (Figure 6b). After performing NMS, some edges that were supposed to be extracted as strong edges were extracted as thin edges. To keep the continuity of strong edges, the strong edges (with the gray value of 255) and thin edges (with the gray values less than 255) were split (Figure 6c). The noisy thin edges were eliminated by utilizing eight-neighborhood pixel analysis, and suitable edges were merged with strong edges (Figure 3d).

The accuracy values calculated with the test data set can be seen in Table 2. It should be again noted that the RCF architecture's initial weights created from ImageNet were not compatible with our data set, and transfer learning could not be applied. Table 2 shows that there is not a significant difference between the 10th and 4th epoch results. Furthermore, the 4th epoch accuracy results were slightly better in some accuracy metrics. The best F1 score values for DexiNed and RCF are 52.80% and 52.06%, respectively. It can be seen that the best accuracy values are also obtained without the initial weights of DexiNed.

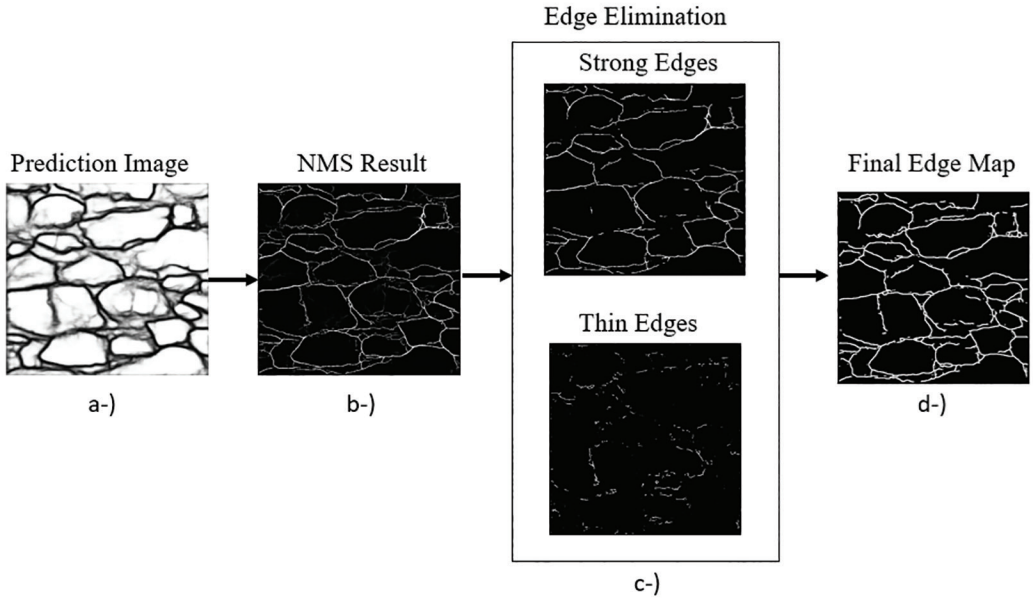



Figure 6. Binarization process. (a) The raw prediction image. (b) Non-Maximum Suppression result. (c) Splitting edges. (d) Final edge map.

Table 2. Accuracy values for both architectures with the raw test data set. The best value for each accuracy value for each architecture is indicated in bold.

| | DexiNed | | | | RCF | |
|---|---|---|---|---|---|---|
| | With Initial Weights | | Without Initial Weights | | Without Initial Weights | |
| | Epoch 4 | Epoch 10 | Epoch 4 | Epoch 10 | Epoch 4 | Epoch 10 |
| Accuracy | 0.8978 | **0.8988** | 0.8990 | **0.9001** | 0.8996 | **0.9005** |
| IoU | **0.3571** | 0.3560 | **0.3587** | 0.3580 | 0.3511 | **0.3519** |
| Precision | 0.4682 | **0.4714** | 0.4727 | **0.4769** | 0.4739 | **0.4777** |
| Recall | **0.6008** | 0.5926 | **0.5980** | 0.5894 | **0.5753** | 0.5719 |
| F1 score | **0.5263** | 0.5251 | **0.5280** | 0.5272 | 0.5197 | **0.5206** |

DexiNed = dense extreme inception network for edge detection; RCF = richer convolutional features architecture.

Prediction results that were generated with the highest F1 score for each experiment are given in Figure 7. The results show that DexiNed architecture produced more accurate results morphologically compared to the RCF architecture. On the other hand, it was observed that the edge details were more distinctive in the RCF architecture.

Image labeling was carried out by different operators with different thickness values manually. In order to investigate this effect on test accuracy, two additional test data sets were generated by balancing the edge thicknesses of the label images and their impact on the prediction images (Figure 8).

The results obtained are shown in Tables 3 and 4 for thin and thick labels, respectively. As can be seen from Table 3, while the use of thin labels marginally improved accuracy (~+0.5%) and precision (~+3.5%), it decreased IoU (~−2%), precision (~−4%) and F1 scores (~−2.5%) for both architectures. On the other hand, even though the use of thick labels marginally decreased accuracy (~−1%), it significantly improved IoU (~+9%), precision (~+16.5%), recall (~+3%), and F1 scores (~+9.5%) for both architectures. Therefore, it can be clearly stated that thick labels help produce significantly better results. Based on the accuracy values (Table 4), both architectures seem to perform with similar efficiency even though these architectures focus on different topics. Moreover, it was observed that high epoch number and training with initial weight had no significant effect on accuracy metrics.

Table 3. Accuracy values for both architectures with the thin test data set. The best value for each accuracy value for each architecture is indicated in bold.

| | DexiNed | | | | RCF | |
|---|---|---|---|---|---|---|
| | With Initial Weights | | Without Initial Weights | | Without Initial Weights | |
| | Epoch 4 | Epoch 10 | Epoch 4 | Epoch 10 | Epoch 4 | Epoch 10 |
| Accuracy | 0.9040 | **0.9052** | 0.9050 | **0.9062** | 0.9045 | **0.9053** |
| IoU | **0.3363** | 0.3358 | **0.3347** | 0.3342 | **0.3281** | 0.3278 |
| Precision | 0.5019 | **0.5085** | 0.5073 | **0.5142** | 0.5046 | **0.5091** |
| Recall | **0.5048** | 0.4970 | **0.4958** | 0.4885 | **0.4839** | 0.4794 |
| F1 score | **0.5034** | 0.5027 | **0.5015** | 0.5010 | **0.4941** | 0.4938 |

DexiNed = dense extreme inception network for edge detection; RCF = richer convolutional features architecture.

Table 4. Accuracy values for both architectures with the thick test data set. The best value for each accuracy value for each architecture is indicated in bold.

| | DexiNed | | | | RCF | |
|---|---|---|---|---|---|---|
| | With Initial Weights | | Without Initial Weights | | Without Initial Weights | |
| | Epoch 4 | Epoch 10 | Epoch 4 | Epoch 10 | Epoch 4 | Epoch 10 |
| Accuracy | 0.8851 | 0.8870 | 0.8870 | **0.8883** | 0.8873 | **0.8884** |
| IoU | 0.4409 | 0.4423 | **0.4428** | 0.4412 | **0.4441** | 0.4440 |
| Precision | 0.6159 | **0.6253** | 0.6250 | **0.6339** | 0.6262 | **0.6327** |
| Recall | **0.6082** | 0.6017 | 0.6030 | 0.5920 | **0.6042** | 0.5982 |
| F1 score | 0.6120 | **0.6133** | **0.6138** | 0.6123 | **0.6150** | **0.6150** |

DexiNed = dense extreme inception network for edge detection; RCF = richer convolutional features architecture.

We have also tested our proposed best solution with an independent low-cost UAV-derived orthoimage (Figure 9a). The orthoimage was created using 133 low-cost UAV images captured with DJI Mavic Mini 2. The UAV has a 12-MP camera with 4000×2250 pixel resolution. This

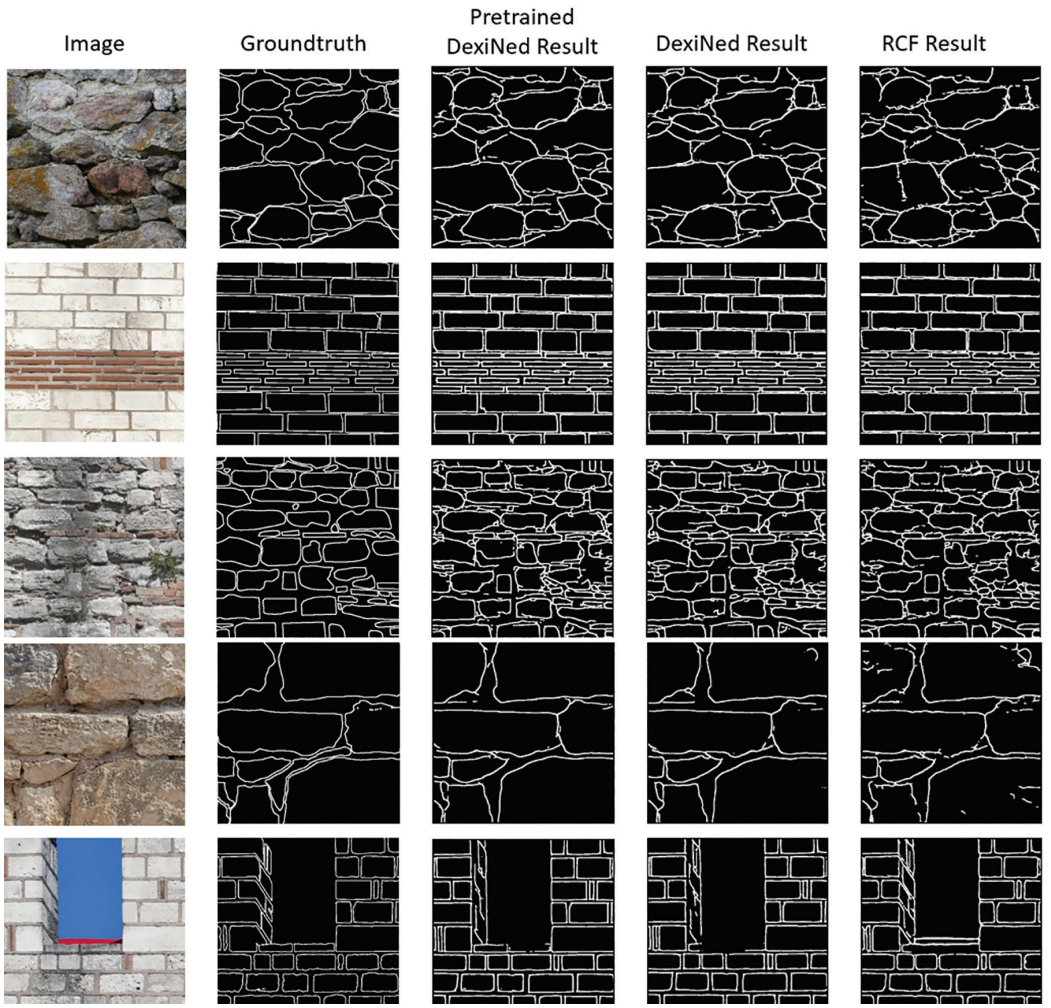



Figure 7. The resulting images of our data set.

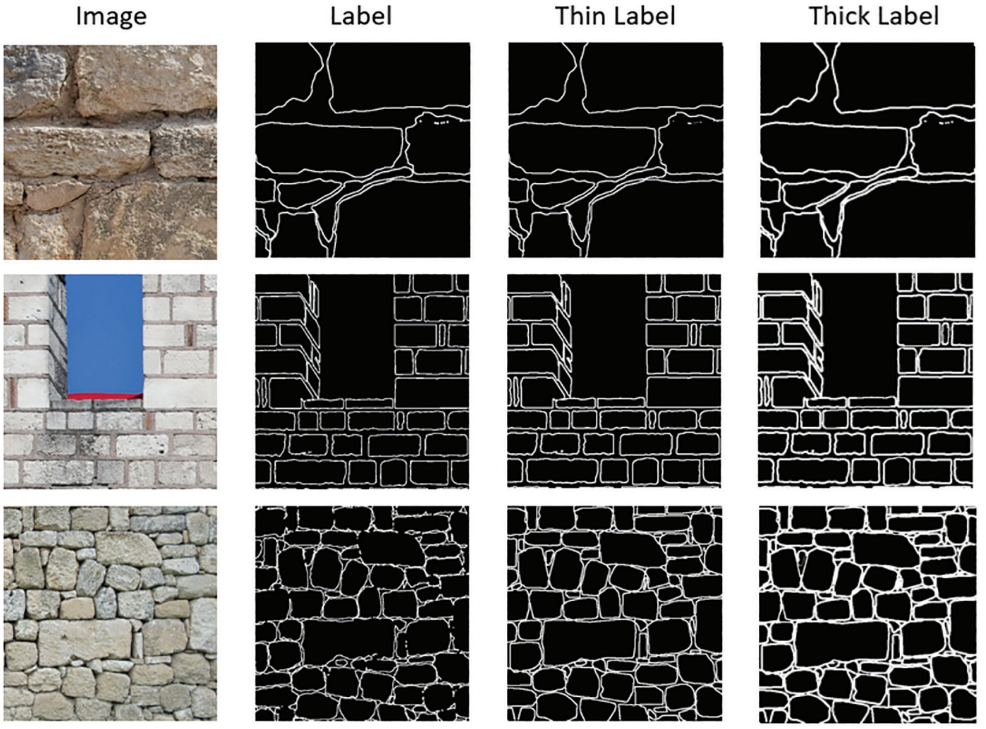



Figure 8. Visualization of raw, thin and thick labels.

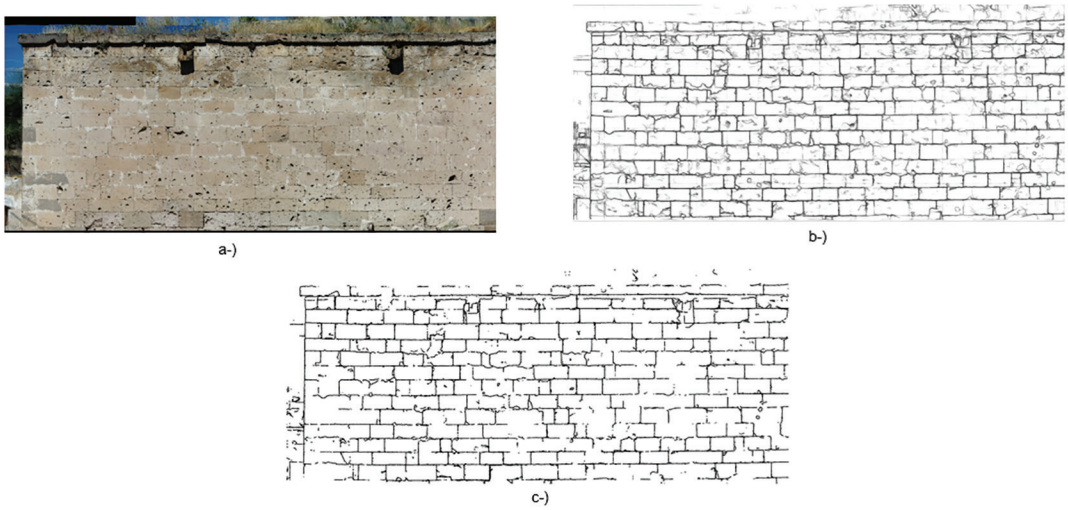



Figure 9. (a) Orthoimage. (b) Prediction. (c) Postprocessing result.

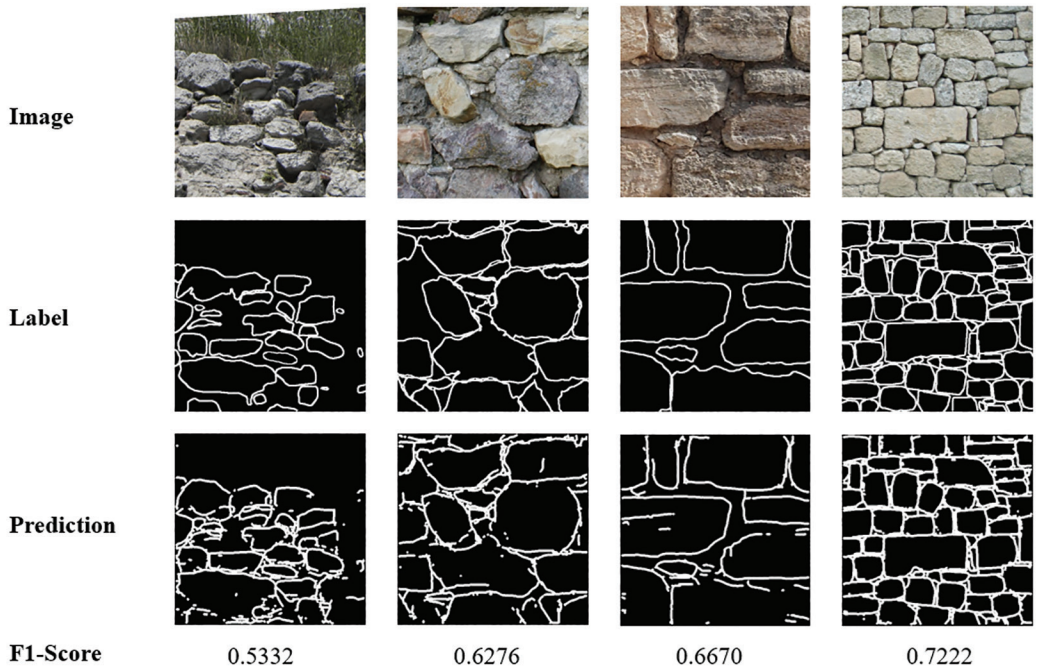



Figure 10. Effect of structural features or material characteristics on accuracy.

monument, which dates back to circa the 12th century, is located in Bahceli village in the Cappadocia region of Türkiye. The prediction and postprocessing results are given in Figure 9b and 9c, respectively. It can be seen that the DL architecture was able to extract very compact edges from the facade.

On the other hand, there are some limitations and drawbacks of the proposed networks. Background and vegetation seem to affect extraction performance, as can be seen in the first column of Figure 10. Additionally, structural features or material characteristics of the stones decreased accuracy, which is visible in the second column of Figure 10. However, higher F1 scores are obtained for objects with topologically homogeneous borders (the last two columns of Figure 10).

As the number of layers and blocks increases in the architecture of both networks, morphological detail is lost. Both networks do not generate binary images as edge and non-edge. This requires postprocessing as we proposed. However, since the gray values of extracted edges may be similar with the background, it can lead to the loss of some details during postprocessing.

## Conclusions

CH reflects the architecture, social structure, development, and lifestyle of a society. As they are vulnerable to disasters, climate change, tourism, and human actions, it is vital to detect and monitor damages for sustainable management and preservation of cultural monuments. Accurate, efficient, and scalable CH management systems with new technologies are required to avoid natural and anthropogenic negative impacts or reduce their effects. In order to hand down the historical and natural heritage to future generations and to protect the world heritage and to overcome the need to develop systems for the detection of damage and the sustainable monitoring of the historical and natural heritage and with new technologies, the proposed DL-based approaches generated encouraging results.

Within the scope of the study, a labeled data set was generated specific to CH contour and crack details using images from different historical structures in order to automatically extract contours and cracks in historical buildings.

Although various technologies are used for CH studies, the use of DL in the field is still limited. One of the most important scientific contributions of the study is the use, for the first time, of novel architectures for contour and crack extraction for CH monitoring and documentation. This study presents a novel use-case example of DL for CH documentation and monitoring. It is observed that the developed architectures can also be used for CH in Türkiye and all over the world for sustainable detection and monitoring.

Conventional techniques are time consuming and may be applicable to large heritage or archaeological sites. UAVs can be a great cost-effective data source for CH. Another significant outcome of the study is that a combination of DL and UAVs can provide fast, accurate, and low-cost information by integration of DL. The proposed method can enable the generation of prediction models and strategies for climate-neutral CH and short-, mid-, and long-term risk assessment. Our solution can be used as a tool to generate strategies that can be identified for security, stabilization measures, and damage risk assessments, including short-, medium-, and long-term risk estimation and to eliminate the vulnerability of materials against climate change and natural disasters.

According to the literature review in the "Introduction" section, several solutions are proposed for edge detection from images that are focused mostly on the BSDS500 data set (Arbelaez *et al.* 2011) consisting of various scene images. The obtained F1 scores in the literature vary between approximately 60% and 80%. Since our study focuses on contour and crack extraction from complex CH structure images, our results cannot be compared with the mentioned related studies due to the content of the target objects. Therefore, the proposed solution is a pioneer study tailored to our data set specifically for CH documentation and monitoring.

The DexiNed and RCF architectures have been used for the first time in the literature for this specific task. Additionally, the proposed study aims to create infrastructure for innovative digital technologies and to integrate them to ensure the sustainable use and protection of CH from natural disasters and climate change impacts, in line with UNESCO's recommendations and with future research studies in the field of climate change and CH (Sabbioni *et al.* 2008).

Adaptation of CH to climate change is one of the new and important challenges, and research on this subject is limited (Sesana *et al.* 2018). The need to develop preparedness for climate change and natural disasters, national and regional adaptation strategies, and intersectoral risk management plans has become a strategic target. In 2015, the World Heritage Committee supported the United Nations Framework Convention on Climate Change to prevent the effects of climate change on heritage. These initiatives also include UNESCO's advisory bodies: the International Center for the Conservation and Restoration of Cultural Property, the International Union for Conservation of Nature, and the International Council for Monuments and Sites, the last of which established a Climate Change and Heritage Working Group in 2017 to protect CH from the effects of climate change. In this context, one of the most important steps is the Adaptation to Climate Change strategy of the European Commission as part of the European Green Deal Commitment in 2021. Within the scope of strengthening the resilience of CH, the Paris Agreement recognizes adaptation as a global challenge and goal and adopts the 2030 agenda for sustainable development. In addition, Targets 11.4 and 13 of the Sustainable Development Goals aim to strengthen efforts to protect and preserve the world's cultural and natural assets and to take urgent action to combat climate change and its effects. Therefore, the proposed study can aid the monitoring process of CH protection against climate change, natural disasters, and anthropogenic effects.

In future studies, we aim to enlarge our data set and generate improved versions of the DexiNed and RCF architectures by structural alterations to obtain more accurate results specifically for CH.

## Acknowledgments

This research was funded by the Scientific and Technological Research Council of Türkiye (TUBITAK) 1001 program under project no. 122Y017.

## References

Arbelaez, P., M. Maire, C. Fowlkes and J. Malik. 2011. Counter detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(5):898–916.

Bakirman, T., B. Bayram, B. Akpinar, M. F. Karabulut, O. C. Bayrak, A. Yigitoglu and D. Z. Seker. 2020. Implementation of ultra-light UAV systems for cultural heritage documentation. *Journal of Cultural Heritage* 44:174–184.

Bazazian, D. and M. E. Pares. 2021. EDC-Net: Edge detection capsule network for 3D point clouds. *Applied Sciences* 11(4):1883.

BIPED (Barcelona Images for Perceptual Edge Detection). 2020. <https://xavysp.github.io/MBIPED> Accessed 15 June 2022.

Canny, J. 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(6):679–698.

Chen, X., Q. Lin, S. Kim, J. G. Carbonell and E. P. Xing. 2012. Smoothing proximal gradient method for general structured sparse regression. *Annals of Applied Statistics* 6(2):719–752.

Deng, J., W. Dong, R. Socher, L. Li, K.Li and L. Fei. 2009. ImageNet: A large-scale hierarchical image database. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, held in Miami, FL, USA, 20-25 June 2009.

Deng, R., C. Shen, S. Liu, H. Wang and X. Liu. 2018. Learning to predict crisp boundaries. Pages 570–586 in *Computer Vision—ECCV*, held in Munich, Germany, 8-14 September 2018.

Dollar, P., Z. Tu and S. Belongie. 2006. Supervised learning of edges and object boundaries. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, held in New York, NY, USA, 17-22 June 2006.

Fiorucci, M., M. Khoroshiltseva, M. Pontil, A. Traviglia, A. D. Bue and S. James. 2020. Machine learning for cultural heritage: A survey. *Pattern Recognition Letters* 133:102–108.

Ganin, Y. and V. Lempitsky. 2014. N4-fields: Neural network nearest neighbor fields for image transforms. In *Asian Conference on Computer Vision*, held in Singapore, Singapore, 1-5 November 2014.

He, J., S. Zhang, M. Yang, Y. Shan and T. Huang. 2019. BDCN: Bi-directional cascade network for perceptual edge detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, held in Long Beach, CA, USA, 15-20 June 2019.

Hu, Z., M. Zhen, X. Bai, H. Fu and C. Tai. 2020. JSENet: Joint semantic segmentation and edge detection network for 3D point clouds. In *European Conference on Computer Vision (ECCV)*, held in Glasgow, Scotland, 23-28 August 2020.

Isola, P., D. Zoran, D. Krishnan and E. H. Adelson. 2014. Crisp boundary detection using pointwise mutual information. Pages 799–814 in *European Conference on Computer Vision*, held in Zurich, Switzerland, 6-12 September 2014.

Jing, J., S. Liu, G. Wang, W. Zhang and C. Sun. 2022. Recent advances on image edge detection: A comprehensive review. *Neurocomputing* 503:259–271.

Kelm, A. P., V. S. Rao and U. Zölzel. 2019. Object contour and edge detection with RefineContourNet. In *Computer Analysis of Images and Patterns, 18th International Conference (CAIP)*, held in Salerno, Italy, 3-5 September 2019.

Kivinen, J. J., C.K.I. Williams and N. Heess. 2014. Visual boundary prediction: A deep neural prediction network and quality dissection. *Proceedings of Machine Learning Research* 33:512–521.

Liu, Y., M. M. Cheng, D. P. Fan, L. Zhang, J. W. Bian and D. Tao. 2022. Semantic edge detection with diverse deep supervision. *International Journal of Computer Vision* 130:179–198.

Liu, Y., N. Cheng, L. Zhang and X. Bai. 2017. Richer convolutional features for edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41(8):1939–1946.

Liu, Y. and M. S. Lew. 2016. Learning relaxed deep supervision for better edge detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, held in Las Vegas, NV, USA, 27-30 June 2016.

Markham, A., E. Osipova, K. Lafrenz Samuels, and A. Caldas. 2016. *World Heritage and Tourism in a Changing Climate*. Paris: UNESCO.

Marr, D. and E. Hildreth. 1980. Theory of edge detection. *Proceedings of the Royal Society of London Series B* 207:187–217.

Martin, D. R., C. C. Fowlkes and J. Malik. 2004. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(5):530–549.

McIlhagga, W. 2011. The Canny edge detector revisited. *International Journal of Computer Vision* 91:251–261.

Pu, M., Y. Huang, Q. Guan and H. Ling. 2021. RINDNet: Edge detection for discontinuity in reflectance, illumination, normal and depth. Pages 6859–6868 in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, held in Montreal, QC, Canada, 10-17 October 2021.

Pu, M., Y. Huang, Y. Lui, Q. Guan and H. Ling. 2022. EDTER: Edge detection with transformer. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, held in New Orleans, LA, USA, 18-24 June 2022.

Ren, X. 2008. Multi-scale improves boundary detection in natural images. In *Computer Vision – European Conference on Computer Vision (ECCV)*, held in Marseille, France, 12-18 October 2008.

Sabbioni, C., M. Cassar, P. Brimblecombe and R. A. Lefevre. 2008. *Vulnerability of Cultural Heritage to Climate Change*. Strasbourg: Council of Europe.

Sesana, E., A. S. Gagnon, C. Bertolin and J. Hughes. 2018. Adapting cultural heritage to climate change risks: Perspectives of cultural heritage experts in Europe. *Geosciences* 8(8):305.

Sesana, E., A. S. Gagnon, C. Ciantelli, J. Cassar and J. J. Hughes. 2021. Climate change impacts on cultural heritage: A literature review. *WIREs Climate Change* 12(4). https://doi.org/10.1002/wcc.710.

Shustip, P. F., S. Ubaru, V. Kalantzis and H. Avron. 2022. PCENet: High dimensional surrogate modeling for learning uncertainty. *arXiv preprint:* 2202.05063 .

Sobel, E. 1970. *Camera Models and Machine Perception*. Ph.D. dissertation, Stanford University.

Soria, X., G. Pomboza-Junez and A. G. Sappa. 2022. LDC: Lightweight dense CNN for edge detection. *IEEE Access* 10: 68281–68290.

Soria, X., E. Riba and A. Sappa. 2020. Dense extreme inception network: Towards a robust CNN model for edge detection, In *IEEE Winter Conference on Applications of Computer Vision (WACV '20)*, held in Snowmass Village, CO, USA, 1-5 March 2020.

Su, Z., W. Liu, Z. Yu, D. Hu, Q.Liao, Q. Tian, M. Pietikäinen and L. Liu. 2021. Pixel difference networks for efficient edge detection. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, held in Montreal, QC, Canada, 10-17 October 2021.

Torre, V. and T. A. Poggio. 1986. On edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(2):147–163.

Wang, X., Y. Xu, K. Xu, A. Tagliasacchi, B. Zhou, A. Mahdavi-Amiri and Z. Hao. 2020. PIE-NET: Parametric inference of point cloud edges. In 34th Conference on Neural Information Processing Systems (NeurIPS 2020), held in Vancouver, BC, Canada, 6-12 December 2020.

Wibisono, J. K. and H. Hang. 2021. Fined: Fast inference network for edge detection. In *IEEE International Conference on Multimedia and Expo (ICME)*, held in Shenzhen, China, 5-9 July 2021.

World Tourism Organization. 2020. *Supporting Jobs and Economies Through Travel and Tourism —A Call for Action to Mitigate the Socio-Economic Impact of COVID-19 and Accelerate Recovery*. Madrid: UNWTO.

Xie, S. and Z. Tu. 2015. Holistically-nested edge detection. In *IEEE International Conference on Computer Vision*, held in Santiago, Chile, 7-13 December 2015.

Yang, J., B. Price, S. Cohen, H. Lee and M. Yang. 2016. Object contour detection with a fully convolutional encoder-decoder network. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, held in Las Vegas, NV, USA, 27-30 June 2016.

Yang, K., S. Gao, C. Li and Y. Li. 2013. Efficient color boundary detection with color-opponent mechanisms. In 2013 IEEE Conference on Computer Vision and Pattern Recognition, held in Portland, Ore., 23–28 June 2013.

Yu, L., X. Li, C. Fu, D. Cohen-Or and P. Heng. 2018. EC-Net: An edge-aware point set consolidation network. Pages 398–414 in *European Conference on Computer Vision (ECCV)*, held in Munich, Germany, 8-14 September 2018.

Yu, Z., W. Liu, Y. Zou, C. Feng, S. Ramalingam, B.V.K.V. Kumar and J. Kautz. 2018, Simultaneous edge alignment and learning. Pages 400–417 in *European Conference on Computer Vision*, held in Munich, Germany, 8-14 September 2018.

## IN-PRESS ARTICLES

Jian Wu, Shifeng Fu, Peng Chen, Qinghui Chen, Xiang Pan. Validation of Island 3D-mapping based on UAV spatial point cloud optimization: A case study in Dongluo Island of China.

Dan Li, Hanjie Wu, Yujian Wang, Xiaojun Li, Fanqiang Kong, and Qiang Wang. Lightweight Parallel Octave Convolutional Neural Network for Hyperspectral Image Classification.

Ruifeng Ma, Xuming Ge, Qing Zhu, Xin Jia, Min Chen, Liu Tao. Model-driven precise degradation analysis method of highway marking using Mobile Laser Scanning point clouds.

Shuanggen Jin, Ayman M.Elameen, Daniel Olago. Identification of drought events in major basins of Africa based on weighted water storage deficit index from GRACE measurements.

Bo Xu, Daiwei Zhang, Han Hu, Qing Zhu, Qiang Wang, Xuming Ge, Min Chen, Yan Zhou. Spherical Hough Transform for Robust Line Detection toward a 2D-3D Integrated Mobile Mapping System.

Oussama Mezouar, Fatiha Meskine, Issam Boukerch. Automatic Satellite Images Ortho-rectification using K-means based Cascaded Meta-heuristic Algorithm.

Xinyan Pang, Na Ren, Changqing Zhu, Shuitao Guo, Ying Xiong. Blind and Robust Watermarking Algorithm for Remote Sensing Images Resistant to Geometric Attacks.

# WHO'S WHO IN ASPRS

Founded in 1934, the American Society for Photogrammetry and Remote Sensing (ASPRS) is a scientific association serving thousands of professional members around the world. Our mission is to advance knowledge and improve understanding of mapping sciences to promote the responsible applications of photogrammetry, remote sensing, geographic information systems (GIS) and supporting technologies.

## BOARD OF DIRECTORS

### BOARD OFFICERS

**President**
Lorraine B. Amenda, PLS, CP
Towill, Inc

**President-Elect**
Bandana Kar
Oak Ridge National Lab

**Vice President**
Amr Abd-Elrahman
University of Florida

**Past President**
Christopher Parrish, Ph.D
Oregon State University

**Treasurer**
John McCombs
NOAA

**Secretary**
Harold Rempel
ESP Associates, Inc.

### COUNCIL OFFICERS

ASPRS has six councils. To learn more, visit https://www.asprs.org/Councils.html.

**Sustaining Members Council**
Chair: Ryan Bowe
Deputy Chair: Melissa Martin

**Technical Division Directors Council**
Chair: Hope Morgan
Deputy Chair:

**Standing Committee Chairs Council**
Chair:
Deputy Chair:

**Early-Career Professionals Council**
Chair: Youssef Kaddoura
Deputy Chair:

**Region Officers Council**
Chair: Demetrio Zourarakis
Deputy Chair: Jason Krueger

**Student Advisory Council**
Chair: Oscar Duran
Deputy Chair:

### TECHNICAL DIVISION OFFICERS

ASPRS has seven professional divisions. To learn more, visit https://www.asprs.org/Divisions.html.

**Geographic Information Systems Division**
Director: Denise Theunissen
Assistant Director: Jin Lee

**Lidar Division**
Director: Ajit Sampath
Assistant Director: Mat Bethel

**Photogrammetric Applications Division**
Director: Ben Wilkinson
Assistant Director: Hank Theiss

**Primary Data Acquisition Division**
Director: Srini Dharmapuri
Assistant Director: Ravi Soneja

**Professional Practice Division**
Director: Hope Morgan
Assistant Director: Matt Elious

**Remote Sensing Applications Division**
Director: Tao Liu
Assistant Director: Indu Jeyachandran

**Unmanned Autonomous Systems (UAS)**
Director: Jacob Lopez
Assistant Director: Bahram Salehi

### REGION PRESIDENTS

ASPRS has 13 regions to serve the United States. To learn more, visit https://www.asprs.org/regions.html.

**Alaska Region**

**Cascadia Region**
Jimmy Schulz

**Eastern Great Lakes Region**
Craig Fry

**Florida Region**
Matt LaLuzerne

**Gulf South**
Cody Condron

**Heartland Region**
Whit Lynn

**Mid-South Region**
David Hughes

**North Atlantic Region**
Kurt Lutz

**Northeast Region**

**Pacific Southwest Region**
Omar Mora

**Potomac Region**
Jason Brown

**Rocky Mountain Region**
Trent Casi

**Western Great Lakes Region**
Adam Smith

# Validation of Island 3D-Mapping Based on UAV Spatial Point Cloud Optimization: A Case Study in Dongluo Island of China

Jian Wu, Shifeng Fu, Peng Chen, Qinghui Chen, and Xiang Pan

## Abstract

*The unmanned aerial vehicle (UAV) remote sensing is of small volume, low cost, fine timeliness, and high spatial resolution, and has the special advantage on island surveying. Focus on the inaccurate elevation of non-ground point cloud without lidar device, this study explored a methodology for island three-dimensional (3D) mapping and modelling based on spatial point clouds optimization with a K-Nearest Neighbors Adaptive Inverse Distance Weighted (K-AIDW) interpolation algorithm. By classifying the UAV point clouds into ground, vegatetation, and structure, the K-AIDW algorithm was applied to optimize the elevations of non-ground point clouds (vegetation and structure) to recalculate Z values.The aerophotogrammetry result was generated based on the optimized spatial point clouds. Finally, the 3D model of Dongluo Island was reconstructed and rendered in Metashape. The accuracy evaluation result shows that the max-errors of ground control points (–0.0154 in X, 0.0305 in Y, and 0.0133 in Z) and the checkpoints (–0.091 in X, –0.176 in Y, and 0.338 in Z) can meet the error-tolerance requirements of the corresponding terrain on the 1:500 scale set by the national standard of GB/T 23236-2009 in China. It is found that the K-AIDW algorithm displayed the best Z accuracy (root-mean-square error of 0.2538) compared with IDW (0.3668) and no-optimized (1.6012), proving it is an effective methodology for improving 3D-modelling accuracy of island.*

## Introduction

As a big maritime country with vast sea waters and numerous islands, China has over 11 000 islands accounting for 0.8% of China's land area (MNR of the P.R.C. 2018). As one of the critical land resources, islands are pivotal for protecting the marine environment, maintaining ecological balance, strengthening the marine economy, expanding the development space, and defending national rights and interests. Notably, most islands are uninhabited. Islands are generally far away from the mainland and are affected by inconvenient transportation, landing difficulties, steep terrain, dense vegetation, and poor communication signals, all of which result from the island's unique geographical location. Hence, the island's basic geographic information cannot be efficiently collected through traditional methods such as real-time kinematic (RTK) and total station surveys (Yuan *et al.* 2018).

## UAV Remote Sensing

Unmanned aerial vehicle remote sensing (UAVRS), involving a platform of small volume, low cost, good timeliness, and high spatial resolution, is specifically advantageous in island monitoring and information acquisition (Wu *et al.* 2019). Therefore, it has become the leading low-altitude remote sensing platform (Li *et al.* 2016). Spatial remote sensing information is obtained in an automatic, intelligent, professional, and rapid way by combining several technologies such as UAV, telemetry and telecontrol, communications, positioning, and orientation system (POS), a global positioning system (GPS), differential positioning, and remote sensing (RS). Furthermore, UAVRS also involves application technology, real-time processing, modelling, and analysis, providing an advanced aerial RS technology solution (Li *et al.* 2014). Recently, UAVRS has played an irreplaceable role in ecological environment monitoring (Messinger *et al.* 2016; Hodgson *et al.* 2016), vegetation information extraction (Dong *et al.* 2019; Mafanya *et al.* 2017), emergency disaster (Li *et al.* 2008), glacier monitoring (Bhardwaj *et al.* 2016), and other fields. UAVRS applications are quite broad, ranging from military to commercial and civil industries, which have attracted much research interest (Zhang *et al.* 2018a). However, due to the unique environment of the island, which is relatively independent and fragile, some disadvantages can be anticipated, including: strong winds, poor takeoff and landing conditions, poor signal, and irregular shorelines, all of which can create safety risks and technical difficulties in UAV application on the island relying on the development of photogrammetry as well as UAV remote sensing in the future. Consequently, careful planning is required, looking for periods of less wind, usually in the early morning. When wind speed exceeds tolerance requirements (usually Grade 6), it is unsafe for UAV flight. For another, large water bodies such as sea area along the shoreline can hinder the application of matching techniques. In these cases, masking techniques are usually applied to avoid these areas being used in point matching (Gonçalves *et al.* 2015).

## Previous UAV Research of 3D Mapping

The photogrammetric potential of UAVRS has been evaluated recently in several studies (Zarco-Tejada *et al.* 2014; O'Driscoll 2018 ; Gevaert *et al.* 2017; Siebert and Teizer 2014; Gonçalves *et al.* 2015), which is of great importance to various applications including topographic modelling, mapping, engineering, construction environmental monitoring, agriculture, etc. (Ajayi *et al.* 2018). Three-dimensional (3D) visualization modelling, as an indispensable technical asset and a component of the digital land and earth, is playing an essential role in island monitoring and information management. Much research has been conducted in UAV photogrammetry for 3D mapping and modelling. For example, Siebert and Teizer (2014) evaluated a UAV system that was built to rapidly and autonomously acquire mobile 3D mapping data. Hudzietz and Saripalli (2011) successfully used structure from motion (SfM) techniques for the reconstruction of aerial imagery from landscapes. Jizhou *et al.* (2004) designed and implemented an algorithm for 3D reconstruction of city buildings from multiple images using a single UAV. Different photogrammetry software packages optimize and improve the aerotriangulation algorithm uniquely, aiming to improve speed and accuracy. Some are free and open-source, such as Bundler and CMVS (Furukawa and Ponce 2010) or Apero and Mic-Mac (Deseilligny and

Third Institute of Oceanography, Ministry of Nature Resources, Xiamen 361005, China (fushifeng@tio.org.cn).

*Contributed by Dorota Iwaszczuk, August 12, 2022 (sent for review October 25, 2022; reviewed by Han Hu, Zhehui Yang).*

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 3, March 2023, pp. 173–182.
0099-1112/22/173–182

doi: 10.14358/PERS.22-00109R2

Clery 2011) and some are commercial. Typically, the core algorithm of a commercial software is proprietary and classified.

Precisely modelling can intuitively display geological disasters on islands, collect island vegetation information, quantitatively estimate vegetation parameters, and monitor the island shoreline erosion along with the status of its development and use. Until now, UAVRS have provided an alternative for the remote monitoring of island from lower altitudes unaffected by cloud coverage. To date, several studies have focused on island topographic information extracting and monitoring (Yuan *et al.* 2018; Zhang *et al.* 2018b). However, references to the use of UAVs for island 3D mapping and precisely-modelling are still scarce in the scientific literature. In view of studies have been mentioned above, research generally focuses more on the 3D reconstruction algorithm optimization, visual effect, and processing efficiency of UAVRS, with less concern on validation of modelling accuracy. Nevertheless, the problem of the "non-ground" point cloud elevation calculated by the aerotriangulation algorithm is the surface elevation rather than the terrain height, which will reduce the digital mapping accuracy of Digital Elevation Model (DEM) to some extent.

To this point, building on experience from previous research, the present study explored an island 3D-mapping and modelling methodology with spatial point clouds optimization extracted from UAV. A K-Nearest Neighbors (K-NN) adaptive inverse distance weighted (Yan *et al.* 2020) interpolation algorithm was applied to recalculate the "non-ground" point cloud elevation to improve the 3D mapping accuracy. The experimental results demonstrated that this methodology can obviously improve the elevation accuracy of "non-ground" point cloud and reconstruct the 3D-visual model for Chengzhou Island, which is an effective technique for the rapid and precise monitoring of islands.

## Methods

### Study Area

This study was carried out in Dongluo Island which is situated at the East Sea area of Niutou Bay, Songxia Town, Changle City, Fujian Province (see Figure 1). It is located at 25°45.8' N, 119°40.3' E, and is about 3.5 km away far from the mainland. As the largest island in Dongluo Archipelago, Dongluo Island is surrounded by the islands Xiaoluo, Xiaolun, Dalun, Dalun North, Dalun South, and Dongyin, which is uninhabited and is the size of about 1.26 km long (north to south) and 0.47 km wide (east to west), with a coastline of about 5.6 km and an area of about 0.63 km2. The terrain of Dongluo Island is flat as most slope is under 20 degrees. The altitude increases from the northeast to the southwest ranging from 0.88 m to 64.09 m. The Dongluo Island is the subtropical monsoon zone with relatively high temperature and abundant rainfall throughout the year. The mean annual temperature is 19.3 °C and the average annual maximal and minimal temperature is 38.7 °C and –1.2 °C, respectively. Rain events concentrate seasonally from May to September. The prevailing wind direction of this area is NNE–ENE, with annually average wind speed of 4.1 m/s.

### Methodological Framework

The core operating process of UAV aerophotogrammetry lies in the analytical aerial triangulation technology, called aerotriangulation for short, which originates from computer vision research. Specifically, relying on the stereophotogrammetry principle, the same feature points are extracted from overlapping photos, the photo's exterior orientation elements are solved, and the coordinates of unknown points are calculated using ground control points (GCPs) coordinates to reconstruct the 3D geographical space (Zhang *et al.* 2018b).

In this current study, Metashape from Agisoft was used for 3D-modelling. Based on multi-view 3D reconstruction technology, it is essentially automated commercial software and the workflow algorithms are not fully documented. According to the relevant literature (Gonçalves *et al.* 2015; Agisoft 2021), it adopts an algorithm called the SfM, which differs from the traditional aerial survey software. The Scale Invariant Feature Transform (SIFT) operators are used to extract the same feature points from each photo in SfM, and the corresponding descriptors are obtained. Then, the descriptors are matched using the random sample and consensus (RANSAC) algorithm, aiming to limit gross errors and mismatching. Finally, A camera model specifies the transformation from point coordinates in the local camera coordinate

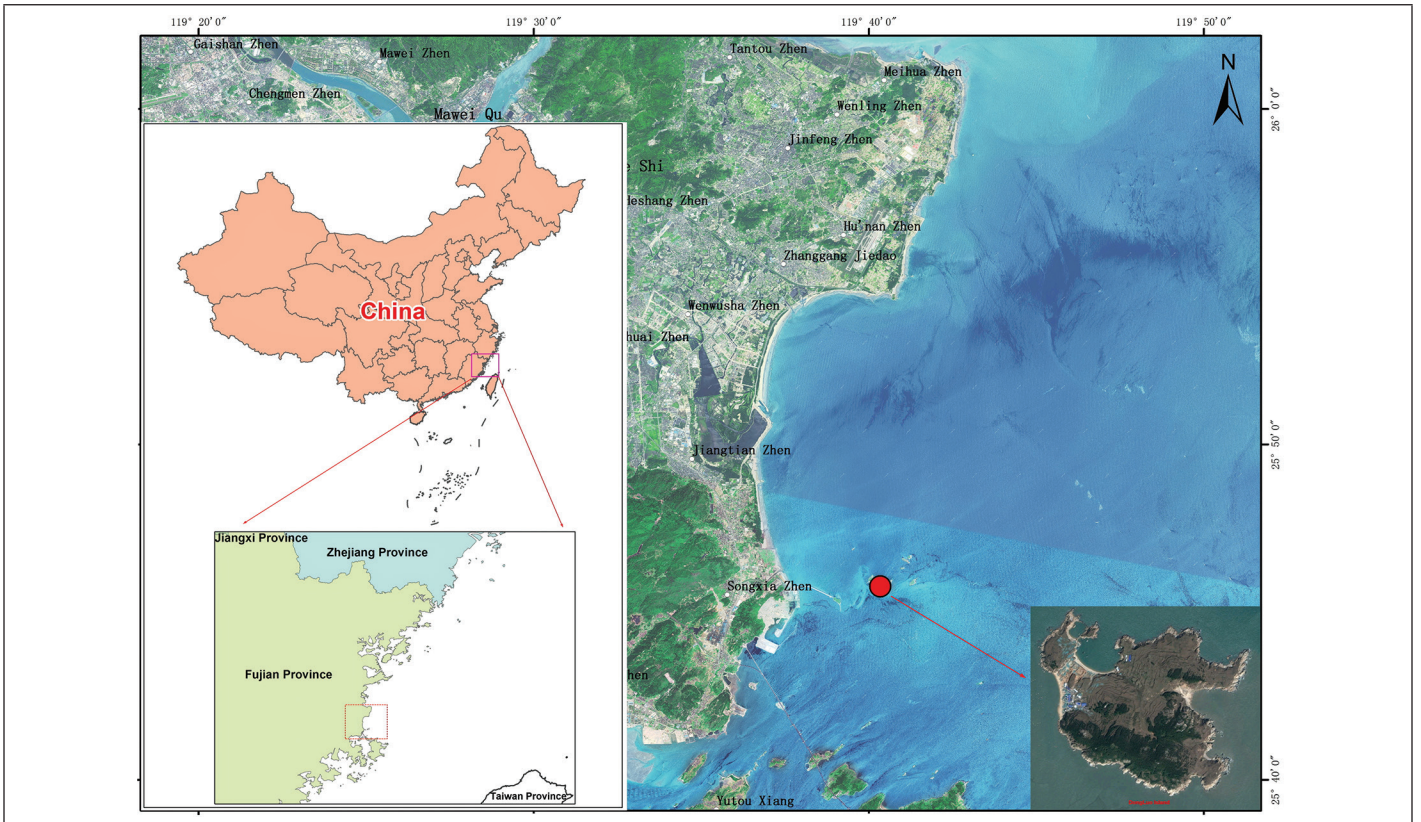


Figure 1. The location of Dongluo Island in China.

system ($x$, $y$) to the pixel coordinates in the image frame ($x'$, $y'$). The mathematical formulation is given as (Agisoft 2021):

$$x' = x\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + k_4 r^8\right) + P_1\left(r^2 + 2x^2\right) + 2P_2 xy \quad (1)$$

$$y' = y\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + k_4 r^8\right) + P_2\left(r^2 + 2y^2\right) + 2P_1 xy \quad (2)$$

where $r = \sqrt{x^2 + y^2}$. Then the pixel projected coordinates ($col$, $row$) of the point will be:

$$col = 0.5w + C_x + \mathrm{x}'f + \mathrm{x}'B_1 + y'B_2 \quad (3)$$

$$row = 0.5h + C_y + y'f \quad (4)$$

where ($col$, $row$) are the projected point coordinates in the image coordinate system, ($x$, $y$) are point coordinates in the local camera coordinate system, ($x'$, $y'$) are point coordinates in the image coordinate system, $f$ is the focal length, ($C_x$, $C_y$) is principal point offset, ($k_1$, $k_2$, $k_3$, $k_4$) is radial distortion coefficients (dimensionless), ($P_1$, $P_2$) is tangential distortion coefficients (dimensionless), ($B_1$, $B_2$) is affinity and non-orthogonality (skew) coefficients (in pixels), and ($w$, $h$) is image width and height (in pixels).

This objective of this study is aimed to: (i) evaluate the potential of UAVRS to island 3D mapping and (ii) improve the altitude accuracy of 3D by "non-ground" point cloud. In order to achieve the precise digital terrain model (DTM) of "non-ground" point cloud such as buildings, vegetation, and other structures instead of digital surface model (DSM), a K-NN adaptive inverse distance weighted (K-AIDW) interpolation algorithm was applied to the spatial point cloud optimization. This algorithm can take both spatial correlation and heterogeneity into account simultaneously and improve the interpolation accuracy without the needs of parameters input for users. Firstly, the sample data is divided into three classes (high, medium, and low) according to the statistical characteristics, the formula equation is as follows:

$$T_i = \begin{cases} high, & Z_i > H_i \\ medium, & L_i \le Z_i \le H_i \\ low, & Z_i < L_i \end{cases} \quad (5)$$

where $H_i$, $L_i$ is the classification threshold of each point. According to the classification result, the different weight adjustment coefficient of $\alpha_i$ is computed as:

$$\alpha_i = (k_i / w)/D_i^2 \quad (6)$$

where $\alpha_i$ is the weight adjustment coefficient, $w$ is the total number of the first order neighboring sample points, $D_i$ is the Euclidean distance between unknown point and sample point, and $k_i$ is the number of each class. Then the K-AIDW interpolation algorithm is construed as:

$$z(x) = \frac{\sum_{i=1}^{k_1} \frac{k_1}{d_i^n} z_i + \sum_{j=1}^{k_2} \frac{k_2}{d_j^n} z_j + \sum_{k=1}^{k_3} \frac{k_3}{d_k^n} z_k}{\sum_{i=1}^{k_1} \frac{k_1}{d_i^n} + \sum_{j=1}^{k_2} \frac{k_2}{d_j^n} + \sum_{k=1}^{k_3} \frac{k_3}{d_k^n}} \quad (7)$$

where $z(x)$ is the point calculated by the K-AIDW algorithm, $z_i$, $z_j$, $z_k$ is the attribution of each class point, and $d_i$, $d_j$, $d_k$ respectively, denotes the Euclidean distance of each class point.

The methodology of this study can be summarized as: data acquisition, spatial point cloud processing, accuracy validation, and 3D mapping and modeling, with the corresponding technical flowchart illustrated in Figure 2. GCPs and UAV aerial surveys were carried out. Flight POS data and aerial images were obtained for input in Trimble Business Center (TBC) for aerotriangulation. Once the accurate image orientation was calculated, dense point clouds were obtained and were classified as vegetation, structure, ground, and noise. Spatial point cloud optimization to enhance elevation of non-ground points (vegetation and structure) was conducted using the K-AIDW algorithm. Z accuracy assessment of 12 checkpoints was compared with the conventional IDW algorithm and no optimization. Subsequently, digital maps of digital orthophoto map (DOM), DTM, and DSM were generated and collected for 3D reconstruction and modeling in Metashape.

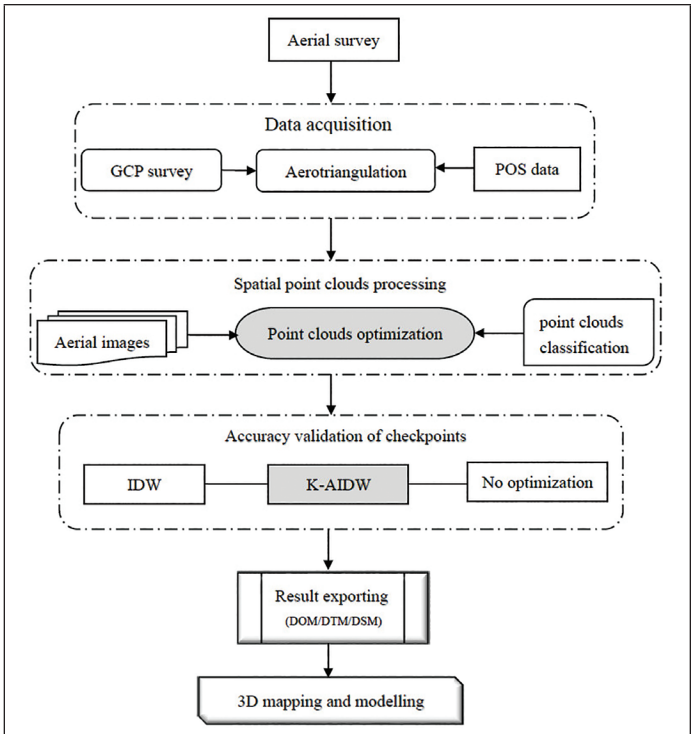



Figure 2. The flowchart of the proposed methodology. IDW = inverse distance weighted; K-AIDW = K-NN adaptive inverse distance weighted; 3D = three-dimensional.

## Data Acquisition

The UAV platform involved is Trimble's UX5 fixed-wing UAV. It weighs 2.9 kg, has a flight altitude of 75–750 m with the Ground Sample Distance (GSD) of 1.0–10.5 cm, has a maximum cruising speed of 80 km/h, and its endurance is about 40 minutes in a single flight. The UAV can fly stably up to a wind scale of six. The catapult-assisted take-off and belly-landing mode are adopted, and the landing area is 50×50m. Thus, this UAV is especially suitable for flight monitoring in remote medium/small islands due to its features. The UAV is equipped with a SONY NEX-5T miniature SLR camera calibrated by the manufacturer, with the main parameters presented in Table 1.

Table 1. The parameters of NEX-5T.

| Parameters | Features | Parameters | Features |
|---|---|---|---|
| Pixel | 16 MP (4912 × 3264) | Focal length | 15 mm |
| Sensor | APS-C CMOS (23.5 × 15.6mm) | Aperture | 4.5–22 |
| Shutter | 1/4000 ~ 30 S | FOV | 110.4° |
| Weight | 218 g | — | — |

The flight was performed on the morning of 18 November 2017 at the east sea area of Songxia Town of Fujian Province, P.R.C. with the wind level was 2–3, covering the Dongluo Island and its surrounding sea areas. The survey area is about 1.15 km$^2$ requiring 22 flight strips and the flight altitude was 180 m during about 30 minutes and 627 photos were obtained with a forward and side overlap of 80% (see Figure 3). Eight GCPs were set in this study, as shown in Figure 4.

## Data Processing

The software INPHO® affiliated to Trimble company, developed by Prof. Ackermann, is a famous digital photogrammetry system in Europe, which has a robust automatic point matching and high-precision point prediction ability (Zhang *et al.* 2016) and integrates many proceeding functions such as aerotriangulation, image matching and correction, DTM/DSM extraction, orthophoto mosaicking, and so on. Trimble Business Centre (TBC), integrating most of INPHO functions into its photogrammetry sub-module, is the supporting software for UX5 data processing. In this study, TBC is applied for aerotriangulation and Metashape is used for 3D-visual modelling.

- Project setup. Firstly, China Geodetic Coordinate System 2000 (CGS2000) with 3 degree was defined as the project coordinate system. The POS data as well as the GCPS were imported into TBC to generate the flight information (Figure 5). There are 22 strips for the arial triangulation and the area has a planimetric extent of about 1660 × 1411 m.
- Relative adjustment of the survey stations. The bundle adjustment was applied (Wang *et al.* 2017) and the tie points were automatically calculated by combining the feature-based matching and least-squares matching algorithm. As is shown in Figure 6, there are 23 335 tie points in the project, and the point size and color reflects the number of images containing the point.
- Absolute adjustment of GCPs. The artificial signals were placed in the survey area and the locations were shown in Figure 4. The GCPs are provide to the image orientation and accurate georeferencing. The accurate coordinate measurements were performed with a Trimble SPS985 global navigation satellite system (GNSS) receiver which is capable of supporting existing and planned GNSS satellite signals, including GPS, *GLONASS*, *Galileo*, *Quasi Zenith Satellite System*, and *Compass* that can operate in RTK mode using a data connection to a permanent GNSS network. The China Geodetic Coordinate System 2000 (CGCS2000) is the unified coordinate system used in P.R.C. The GCPS survey of this study was carried out in a sunny day with fine network and satellite signal. Normally, the position accuracy is very high, especially in an unobstructed environment that more than 15 satellites were used in each GCPs determination. The predicted accuracy was always better than 0.8 cm in horizontal and better than 1.5 cm in vertical which were confirmed with repeatedly tests.

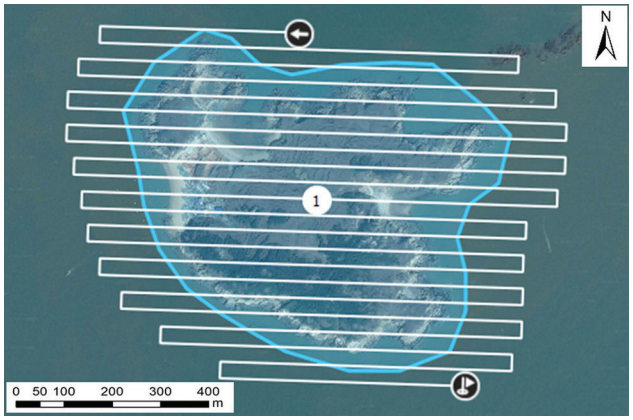


Figure 3. The flight overview.

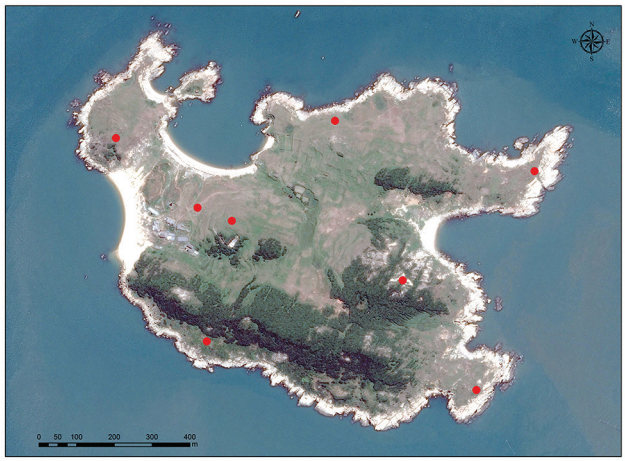


Figure 4. The distribution map of Ground Control Points.

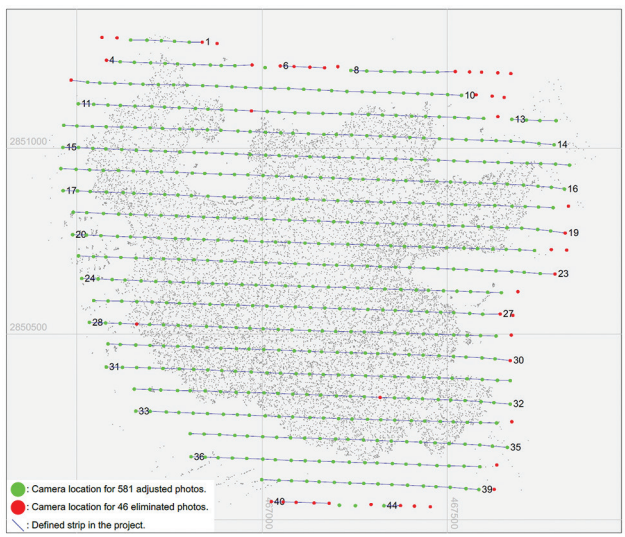


Figure 5. The strips and survey stations overview.

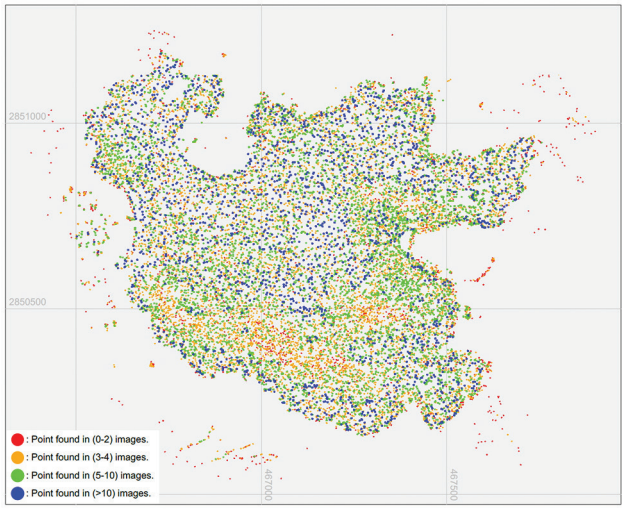


Figure 6. The tie points distribution.

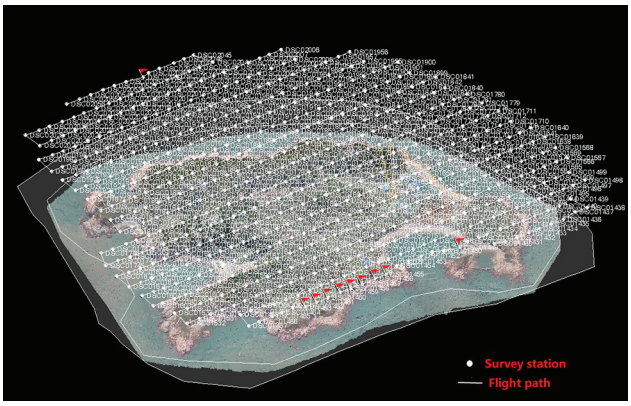


Figure 7. The spatial point clouds of Dongluo Island.

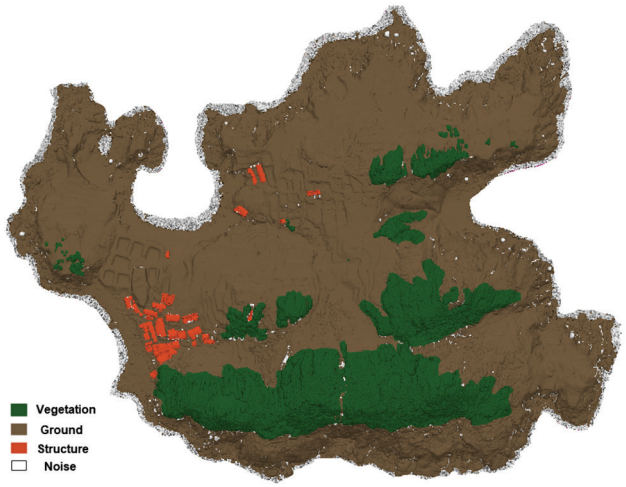


Figure 8. The classification of spatial point clouds.

- Spatial point cloud generation. The spatial point clouds were generated based on the absolute adjustment result (see Figure 7).
- Spatial point cloud optimization. The spatial point clouds were imported and edited in UASMaster module of TBC software to classify into four types as shown in Figure 8 through the spatial filtering tool. Then the K-AIDW algorithm, adopted to recalculate the elevation of non-ground points, such as vegetation and structure based on the surrounding ground points, that retains specific terrain details better reflecting the terrain changes so as to improve DEM accurate .The results are illustrated in Figure 9 and Figure 10.
- The final digital image results are generated based on the optimized spatial point cloud data (Figure 11) such as DOM and DTM.

## Results and Discussion

### 3D Modelling

The aerotriangulation algorithm mainly comprises the following steps (Zhang *et al.* 2018a): (a) Camera calibration to determine the interior orientation elements; searching/matching feature points; (b) Camera point adjustment based on feature points and POS data to solve the

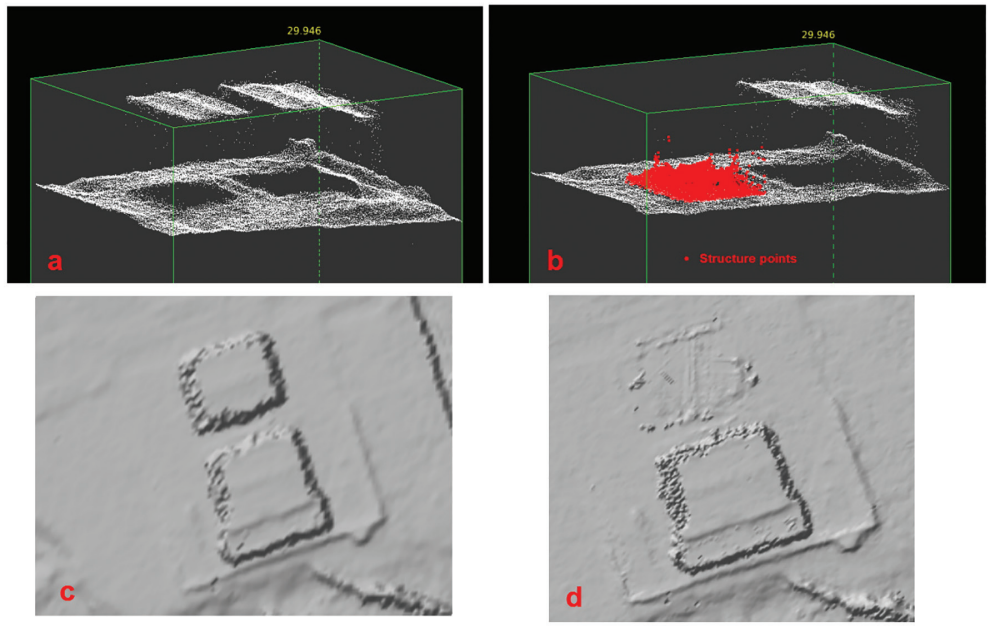


Figure 9. The optimization of structure point clouds: (a) (before) and (b) (after) in Z profile, (c) (before) and (d) (after) in XY profile.

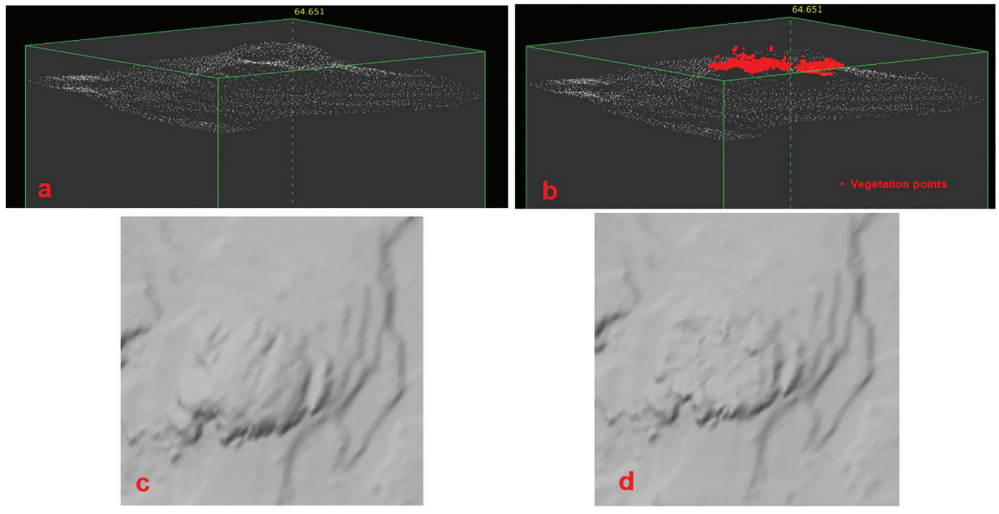


Figure 10. The optimization of vegetation point clouds: (a) (before) and (b) (after) in Z profile, (c) (before) and (d) (after) in XY profile.

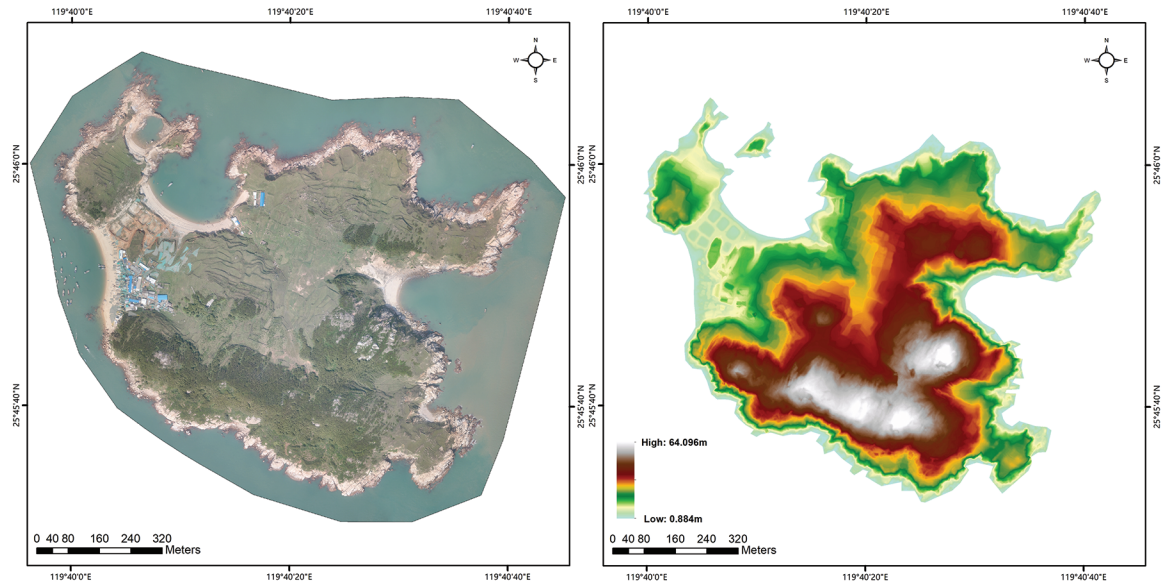



Figure 11. The digital orthophoto map (Ground Sample Distance ((GSD) = 5 cm) and digital terrain model maps.

Figure 12. A real-time rendering of three-dimensional visualization model.

exterior orientation elements and restore the photo position and attitude; (c) control point adjustment to generate spatial point clouds; and (d) generation of the DSM/DTM as well as the DOM through point clouds.

Metashape is a 3D modelling software that exploits computer vision techniques, developed by the Russian Agisoft Company, formerly known as Photoscan. This software generates a natural and exquisite 3D model affording real visual effects through digital 3D reconstruction and texture mapping.We used Metashape for 3D modelling of island.

There are several step-wise procedures as follows: (1) data import; (2) align photos; (3) geo-referencing; (4) build dense clouds; (5) build geometry; (6) build texture; (7) render 3D model; (8) export result. Each step gave several possibilities to adjust parameters, which have influence on the accuracy and structure of the results and the processing time. As to this research, the aerotriangulation result was directly imported into Metashape to build a 3D visualization scene. Then, the adaptive texture was generated and 3D texture mapping was carried out in the UV mode and a refined 3D model were constructed and exported. The 3D model of Dongluo Island is rendered in real-time, with the final effects illustrated in Figure 12.

## Accuracy Evaluation

The accuracy of the proposed research is verified by the mean error (M), standard deviation (SD), and root-mean-square error (RMSE), which is given by:

$$M = \Sigma_{i=1}^{n} \Delta_i / n \quad (8)$$

$$SD = \Sigma_{i=1}^{n} \Delta^2 i / n \quad (9)$$

$$RMSE = \sqrt{\Sigma_{i=1}^{n} \Delta^2 i / n} \quad (10)$$

where $\Delta_i$ represents the difference between the calculated and the measured value, and $n$ is the number of measurements.

### Aerotriangulation Accuracy

According to the adjustment report, the aerotriangulation's sigma naught in this regional network is 0.66. Table 2 and Figure 13 provide the terrain points' overall accuracy and the projection center's SD per station, respectively.

The adjustment report for the entire project involving all survey stations is generated. As is shown in Figure 13, the SD of each station on the island is very small, while stations with a large SD are concentrated at the edge of the survey area located above the sea area around Dongluo Island. The SD values of these stations are high due to objective factors such as the low photo overlap rate and the inaccurate extraction of water tie points. The spatial point clouds corresponding to these stations are eliminated during the post-processing spatial point cloud optimization to improve the aerotriangulation accuracy.

### Control Point Accuracy

The position of eight GCPs are shown in Figure 4 and accuracy of GCP is evaluated, with M values presented in Table 3.

### Checkpoint Accuracy

According to the Chinese technical standard of CH/T 9008.1-2010, the "Digital Products of Fundamental Geographic Information 1:500, 1:1000, and 1:2000 Digital Line Graphs" divides the topography into four types: flat land (slope <2°), hilly land (2° ≤ slope < 6°), mountainous land (6° ≤ slope < 25°) and alpine land (slope ≥25°).

In Figure 14b, we selected a total of 12 checkpoints for accuracy verification including four ground checkpoints (JC05, JC06, JC07, JC09), four vegetation checkpoints (JC02, JC03, JC11, JC12), and four structure checkpoints (JC01, JC04, JC08, JC10). As to Figure 14a, the checkpoints were classified into three types with the slope map based on the National Standard of CH/T 9008.1-2010, which were hilly (JC05, JC09, JC10), mountainous (JC03, JC04, JC06, JC07, JC08, JC12), and alpine (JC01, JC02, JC11). The checkpoints information

Table 2. The standard deviation of terrain points.

| *X* | *Y* | *Z* | TOTAL |
|---|---|---|---|
| 0.0415 | 0.0520 | 0.1247 | 0.1414 |

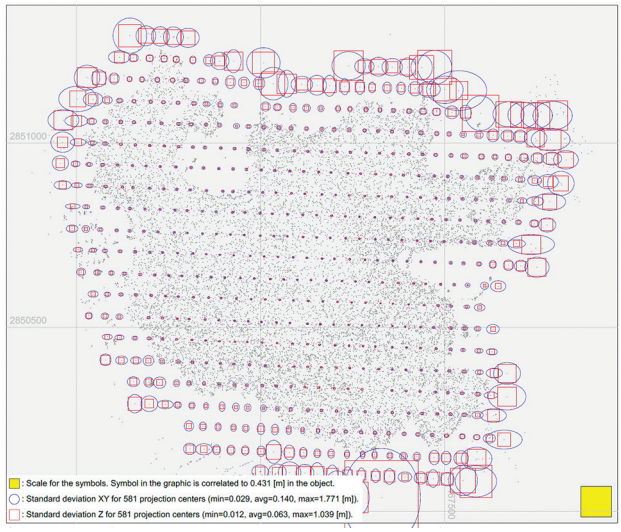



Figure 13. The standard deviation map of projection centers.

Table 3. The mean errors of ground control points.

| ID | *X*/m | *Y*/m | *Z*/m |
|---|---|---|---|
| KZ01 | 0.0070 | 0.0305 | –0.0071 |
| KZ02 | 0.0004 | –0.0044 | –0.0007 |
| KZ03 | –0.0035 | –0.0012 | –0.0027 |
| KZ04 | –0.0002 | –0.0005 | 0.0017 |
| KZ05 | 0.0072 | –0.0015 | –0.0005 |
| KZ06 | –0.0016 | 0.0008 | 0.0016 |
| KZ07 | 0.0062 | –0.0043 | –0.0054 |
| KZ08 | –0.0154 | –0.0193 | 0.0133 |
| **RMSE** | **0.0070** | **0.0130** | **0.0058** |

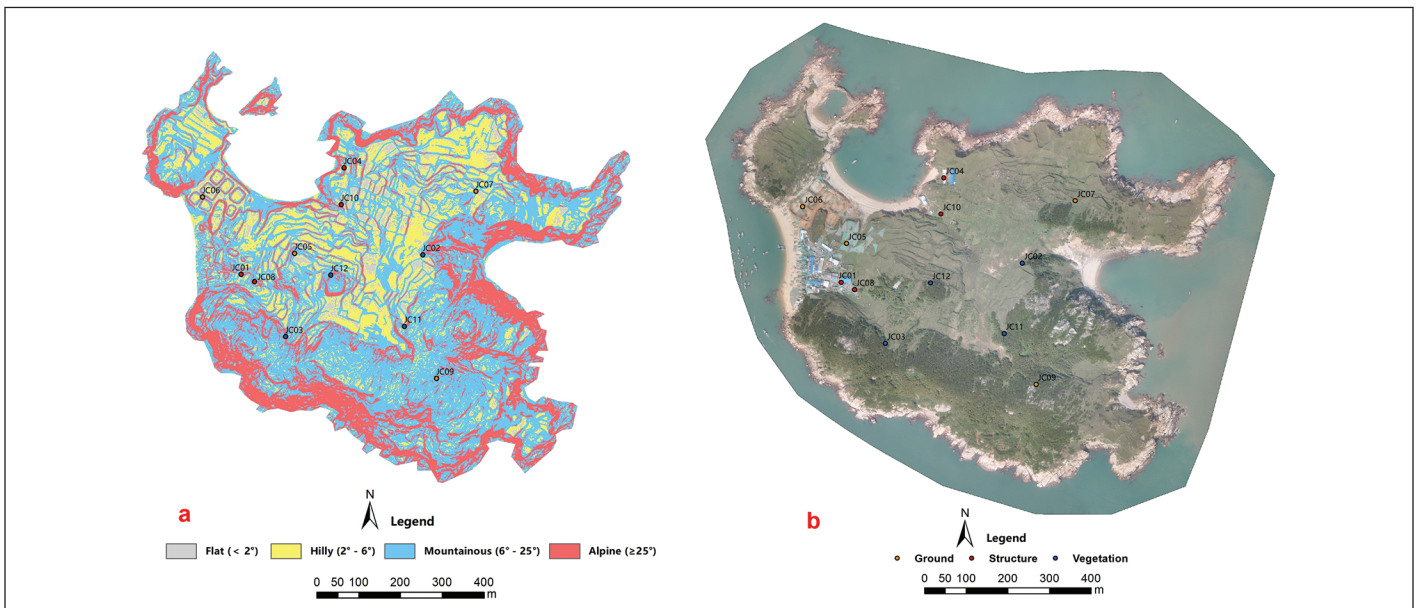



Figure 14. The distribution of checkpoints: (a) locations in slope map; (b) locations in digital orthophoto map.

summary and accuary assessment as well as the mean-error distribution of K-AIDW were illustrated in Table 4 and Figure 15, respectively.

*Error Comparison of the Z Values*

Z-error comparison of the checkpoints was performed to quantitatively assess and validate the K-AIDW algorithm. The brush tool of UASMaster, based on the conventional IDW algorithm, can interpolate and optimize point cloud elevation to edit the terrain. This study used the checkpoints interpolated by this tool in UASMaster to explain the improvement of K-AIDW. Checkpoints directly exported from the aerotriangulation results were as well applied to verify the validity. Table 4 and Figure 16 illustrate the comparison results and error distribution.

Specifically, checkpoints of JC05, JC06, JC07, and JC09 were ground points that did not require recalculation by algorithms, whose mean error as well as error coefficient were consistent that were not introduced for comparison. Compared to non-optimized, it revealed that Z-error of JC04 was enhanced the most after optimization using K-AIDW, with the mean error decreasing from 2.4 to –0.285 and the error coefficient from 21.28% to 2.53%. Moreover, the JC01 was also notably improved, with the mean error reducing from 2.15 to 0.216 and the error coefficient from 18.12% to 1.82%, both of which are structures. Figure 16 reveals that the error coefficient of each checkpoint is under 5% after optimized by K-AIDW, indicating a satisfactory accuracy of the Z-values, and the structure checkpoints (JC01, JC04, JC08, and JC10) are better improved with the error coefficient than the optimized vegetation checkpoints (JC02, JC03, JC11, and JC12). Meanwhile, all checkpoints were improved obviously. With the conventional IDW algorithm, the Z-errors optimized with the K-AIDW algorithm were generally more accurate, except for JC08, whose error coefficient was 2.06% when compared with IDW of 1.88%. With vegetation checkpoints (JC02, JC03, JC11, and JC12), the error coefficient of K-AIDW ranged from 0.73% to 0.96% indicating a better result compared with IDW of 1.02% to 1.76%. It was found that mean errors of IDW were all positive (0.396 to 0.621), reflecting the concern of less optimization of vegetation points, which were all positive from 0.396 to 0.621, decreasing the Z accuracy when compared to K-AIDW (–0.315 to 0.338). Referring to the structures (JC01, JC04, JC08, and JC10), the K-AIDW and IDW algorithms displayed similar precision in which average error coefficients were 2.045% and 2.69%, respectively. The K-AIDW RMSE (0.2538) was less than with IDW (0.3668) as shown in Table 4, demonstrating better optimization and accuracy with the K-AIDW algorithm.

## Analysis and Discussion

Until now, there have been several national standards and technical specifications of aerial photogrammetry in China, such as "Specifications for Aerotriangulation of Digital Aerophotogrammetry" (GB/T 23236-2009), "Digital Products of Fundamental Geographic Information 1:500, 1:1000, and 1:2000 Digital Elevation Model" (CH/T 9008.2-2010), "Digital Products of Fundamental Geographic Information 1:500, 1:1000, and 1:2000 Digital Orthophoto Maps" (CH/T 9008.3-2010), and "Specifications for Office Operation of Low-altitude Digital Aerophotogrammetry" (CH/Z 3003-2010). By comparing the different references, this study used the one with the highest accuracy requirements of GB/T 23236-2009 for accuracy evaluation and verification, as presented in Table 5.

Considering the ground control points, the max-errors of XY and Z of 0.0305 m and 0.0133 m, respectively, were much smaller than the required values on the 1:500 scale for any land type. Therefore, GCP accuracy requirements on the 1:500 scale were met.

As shown in Table 4, the 12 checkpoints had been classified into three landform types, alpine, mountainous, and hilly. Alpine land (JC01, JC02, and JC11), max-errors of XY and Z of –0.061 m and –0.315 m, respectively, were much smaller than the required values on the 1:500 scale (0.35 m in XY, 0.6 m in Z), indicating checkpoint accuracy was met. With regard to the mountainous land (JC03, JC04, JC06, JC07, JC08, and JC12), max-error of XY and Z of –0.176 m and 0.338 m, respectively, was smaller than the required values on the 1:500 scale

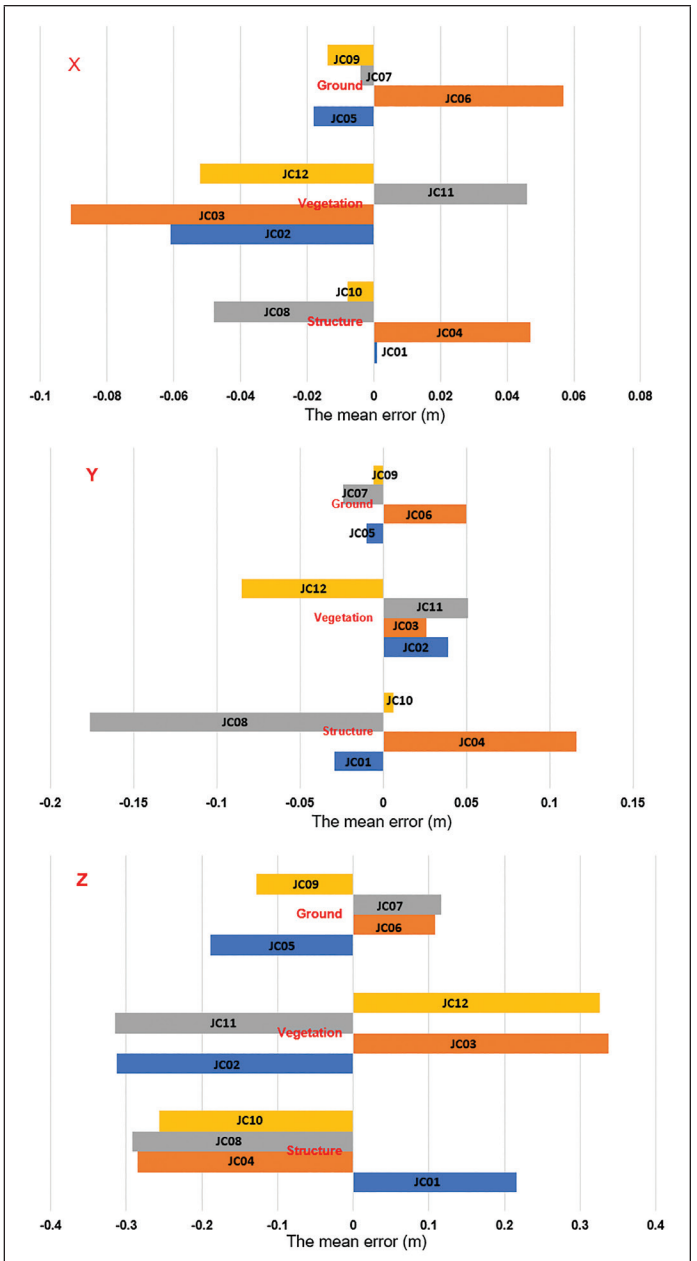



Figure 15. The mean error of K-NN adaptive inverse distance weighted in XYZ profiles.

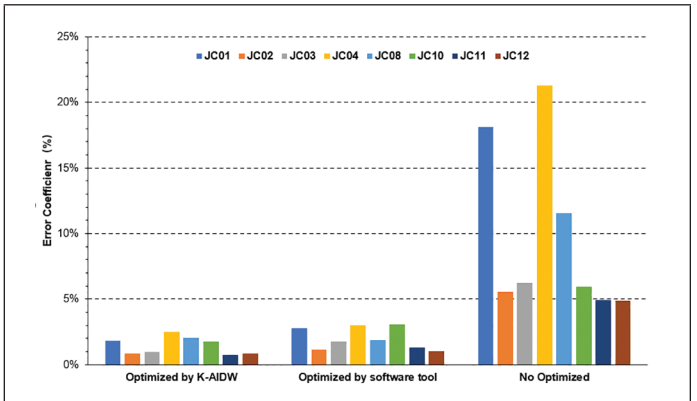



Figure 16. The Z-error comparison of the checkpoints (only non-ground checkpoints).

Table 4. The checkpoints summary and error analysis.

| | | | *X*/m | *Y*/m | *Z*/m | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Optimized by K-AIDW | | Optimized by software tool | | No Optimized | |
| ID | Type | Landform | mean errors | mean errors | mean errors | error coefficient (%)** | mean errors | error coefficient (%)** | mean errors | error coefficient (%)** |
| JC01 | Structure | Alpine | 0.001 | –0.029 | 0.216 | 1.82 | –0.332 | 2.8 | 2.15 | 18.12 |
| JC02 | Vegetation | Alpine | –0.061 | 0.039 | –0.312 | 0.87 | 0.415 | 1.17 | 1.98 | 5.57 |
| JC03 | Vegetation | Mountainous | –0.091 | 0.026 | 0.338 | 0.96 | 0.621 | 1.76 | 2.20 | 6.22 |
| JC04 | Structure | Mountainous | 0.047 | 0.116 | –0.285 | 2.53 | 0.341 | 3.02 | 2.40 | 21.28 |
| JC05 | Ground | Hilly | –0.018 | –0.01 | –0.189 | 0.52 | –0.189 | 0.52 | –0.189 | 0.52 |
| JC06 | Ground | Mountainous | 0.057 | 0.05 | 0.108 | 1.15 | 0.108 | 1.15 | 0.108 | 1.15 |
| JC07 | Ground | Mountainous | –0.004 | –0.024 | 0.116 | 0.33 | 0.116 | 0.33 | 0.116 | 0.33 |
| JC08 | Structure | Mountainous | –0.048 | –0.176 | –0.292 | 2.06 | –0.266 | 1.88 | 1.63 | 11.53 |
| JC09 | Ground | Hilly | –0.014 | –0.006 | –0.128 | 0.25 | –0.128 | 0.25 | –0.128 | 0.25 |
| JC10 | Structure | Hilly | –0.008 | 0.006 | –0.256 | 1.77 | –0.442 | 3.06 | 0.86 | 5.95 |
| JC11 | Vegetation | Alpine | 0.046 | 0.051 | –0.315 | 0.73 | 0.575 | 1.32 | 2.14 | 4.93 |
| JC12 | Vegetation | Mountainous | –0.052 | –0.085 | 0.326 | 0.84 | 0.396 | 1.02 | 1.89 | 4.86 |
| **RMSE** | **-** | **-** | **0.0458** | **0.0710** | **0.2538** | **-** | **0.3668** | **-** | **1.6012** | **-** |

** Error Coefficient = | (Calculated Value-Measured Value)/Measured Value | × 100%; K-AIDW = K-NN adaptive inverse distance weighted.

Table 5. The tolerance errors of different scales in Specifications for Aerotriangulation of Digital Aerophotogrammetry (GB/T 23236-2009).

| | | Tolerance Errors of Plane | | | | Tolerance Errors of Altitude | | | |
|---|---|---|---|---|---|---|---|---|---|
| SCALE | TYPE | Flat | Hilly | Mountainous | Alpine | Flat | Hilly | Mountainous | Alpine |
| 1:500 | ground control points | 0.13 | 0.13 | 0.2 | 0.2 | 0.11 | 0.20 | 0.26 | 0.4 |
| | checkpoints | 0.175 | 0.175 | 0.35 | 0.35 | 0.15 | 0.28 | 0.4 | 0.6 |
| | common points | 0.35 | 0.35 | 0.55 | 0.55 | 0.3 | 0.56 | 0.7 | 1.0 |
| 1:1000 | ground control points | 0.3 | 0.3 | 0.4 | 0.4 | 0.20 | 0.26 | 0.4 | 0.75 |
| | checkpoints | 0.5 | 0.5 | 0.7 | 0.7 | 0.28 | 0.4 | 0.6 | 1.2 |
| | common points | 0.8 | 0.8 | 1.1 | 1.1 | 0.56 | 0.7 | 1.0 | 2.0 |

(0.35 m in XY, 0.4 m in Z), confirming they met accuracy requirements. For hilly regions (JC05, JC09, and JC10), it was found that XY max-error of –0.018 m and Z max-error of –0.256 m were smaller than the requirements on the 1:500 scale (0.175 m in XY, 0.28 m in Z), which fulfilled accuracy requirements. Consequently, accuracy validation of classified checkpoints based on terrain met the requirements for 1:500 scales of national standard (GB/T 23236-2009) in China.

By analyzing the M values as illustrated in Figure 15, the structure checkpoints are the most accurate than the other two types of checkpoints with the X errors between 0.001 (JC10) to 0.048 (JC04), while the vegetation is the lowest with the X errors between 0.046 (JC11) to 0.091 (JC03). As to the Y error, the ground checkpoints are the most precise with the M errors between 0.006 (JC09) to 0.05 (JC06) than the two other types of checkpoints, while the structure is the lowest with the errors between 0.006 (JC10) to 0.176 (JC08).

Focusing on comparison of checkpoints in Table 4, it was concluded that the K-AIDW algorithm had the best accuracy with minimal RMSE of 0.2538, while the IDW result was 0.3668, and maximal RMSE of 1.6012 in non-optimized. As analyzed above, the phenomenon of less optimization of vegetation with the IDW algorithm led to the large positive value of Z-error, decreasing Z accuracy. In structure checkpoints, the two algorithms had a similar Z-error, except in JC08, the Z-error of IDW was less than that of K-AIDW, indicating a better Z accuracy. In comparison with the non-optimized checkpoints, the Z-error coefficient of the K-AIDW algorithm had been remarkably improved, particularly in structure checkpoints, which were more precise than vegetation checkpoints. Results for Z-errors of checkpoints found that ground checkpoint errors were between 0.1 and 0.2 m, structure checkpoint errors were between 0.2 to 0.3 m, and vegetation checkpoint errors were between 0.3 to 0.4 m. Among the checkpoint locations shown in Figure 14, JC02 and JC03, both vegetation checkpoints had the most significant Z-errors of –0.312 and 0.338, respectively. JC06 and JC07, both ground checkpoints, had the lowest Z-errors of 0.108 and 0.116, respectively.

As elevations of vegetation checkpoints could meet mapping requirements for the 1:500 scale of National Standard (GB/T 23236-2009) after recalculation and optimization by the K-AIDW algorithm, revealing elevation accuracy of these points have been greatly improved. However, the Z accuracies of vegetation checkpoints were still lower than the structure and ground checkpoints due to the influence of objective factors such as low signal and position accuracy in the grove and sampling points layout subject to the natural conditions on the island. In brief, the accuracies of non-ground points were notably raised, indicating that Z-values of non-ground points could be advanced with the K-AIDW algorithm of the developed optimization procedure. Additionally, the overall checkpoints' accuracy in XY is higher than in the Z direction. In terms of Z profile, except for the vegetation checkpoints, the other checkpoint ‘s Z accuracy is less than 0.3m, indicating the high modelling accuracy.

## Conclusions

Considering Dongluo Island as the study area, this study developed an island 3D mapping and modelling method using UAV spatial point cloud optimization. Precisely, UAV aerotriangulation was calculated in TBC. Then the point clouds were imported and edited in UASMaster. By eliminating the noise and false points, the remaining point clouds were classified into ground, structure, and vegetation by spatial filtering tools. After that, the "non-ground" points (structure and vegetation) were optimized and recalulated with K-AIDW algorithm to obtain the true terrain height and the digital mapping outcomes (DOM, DSM, and DTM) were relied on the optimized spatial point clouds. Finally, the aerotriangulation

result was imported to Metashape to rebulid the 3D-visual model of Dongluo Island, enhancing a better 3D-visual effect.

In this study, the sigma naught of aerotriangulation was 0.66, less than one pixel, while average standard deviations of terrain points were 0.0415 m (X), 0.0520 m (Y), and 0.1247 m (Z). The max-errors of GCPs were –0.0154 m in X, 0.0305 m in Y, and 0.0133 m in Z. For the check-points, the accuracy in XY was higher than that in Z, and the max-errors were –0.091 m in X, –0.176 m in Y, and 0.338 m in Z, which fulfilled the error-tolerance requirements of their corresponding landforms on the 1:500 scale as set by the national standard of "Specifications for Aerotriangulation of Digital Aerophotogrammetry" (GB/T 23236-2009) in China. Moreover, according to the error coefficient, Z accuracies of non-ground (vegetation and structure) point clouds were notably enhanced after optimization with the proposed K-AIDW algorithm as compared with non-optimization. The K-AIDW algorithm produced a better effect in vegetation optimization compared with the conventional IDW algorithm. Consequently, based on accurate POS data calculation, GCPs measurement combined with spatial point clouds optimization, low-altitude UAVRS met island 3D mapping and modelling requirements on the scale of 1:500 and the K-AIDW algorithm proved to be an effective and feasible approach for interpolation and optimization of the elevation of non-ground point clouds from UAV.

Currently, the UAVRS technology is developing at a rapid pace. Further research shall focus on accurately restoring the aerial attitudes of photos and correcting POS data using machine learning,, post-processing kinematic technology, and software intelligent optimization. New methods such as oblique photogrammetry and lidar can be adopted to extract high-density and high-precision spatial point clouds and achieve more terrain features. Such approaches will reduce the manual processing workload, improve working efficiency and quality, so as to increase the 3D mapping and modelling accuracy.

## Acknowledgments

This work was supported by the China-ASEAN Maritime Cooperation fund “Marine Protected Areas Network in China-ASEAN Countries”, the Scientific Research Foundation of Third Institute of Oceanography, MNR (NO. 2016017), and the Natural Science Foundation of Fujian Province (Grant number 2016J05101).

## References

Agisoft. 2021. *Agisoft Metashape User Manual (Professional Edition), Version 1.7*. St. Petersburg, Russia: Agisoft.

Agisoft. 2022. Agisoft Helpdesk Portal. <https://agisoft.freshdesk.com/support/home> Accessed 30 June 2022.

Ajayi, O. G., M. Palmer and A.A. Salubi. 2018. Modelling farmland topography for suitable site selection of dam construction using unmanned aerial vehicle (UAV) photogrammetry. *Remote Sensing Applications: Society and Environment* 11:220–230.

Bhardwaj, A., L. Sam, F. J. Martin-Torres and R. Kumar. 2016. UAVs as remote sensing platform in glaciology: Present applications and future prospects. *Remote Sensing of Environment* 175:196–204.

Bing, L. and H. Yuhong. 2017. Species classification using unmanned aerial vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland. *ISPRS Journal of Photogrammetry and Remote Sensing* 128:73–85.

Deseilligny, M. P. and I. Clery. 2011. Apero an open source bundle adjustment software for automatic calibration and orientation of set of images. Pages 269–276 in *Proceedings of ISPRS International Workshop on 3D Virtual Reconstruction and Visualisation of Complex Architectures*, held in Trento, Italy, 2–4 March 2011.

Dong X. Y., J. Li and H. Y. Chen. 2019. Extraction of individual tree information based on remote sensing images from an unmanned aerial vehicle. *Journal of Remote Sensing* 6:1268–1280.

Furukawa, Y. and J. Ponce. 2010. Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(8):1362–1376.

Gevaert, C. M., C. Persello and R. Sliuzas. 2017. Informal settlement classification using point-cloud and image-based features from UAV data. *ISPRS Journal of Photogrammetry and Remote Sensing* 125:225–236.

Gonçalves, J. A. and R. Henriques. 2015. UAV photogrammetry for topographic monitoring of coastal areas. *ISPRS Journal of Photogrammetry and Remote Sensing* 104:101–111.

Hodgson, J. C., S. M. Baylis and R. Mott. 2016. Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports* 6:2257.

Hudzietz, B. P. and S. Saripalli. 2011. An experimental evaluation of 3d terrain mapping with an autonomous helicopter. In *Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g)*, held in Zurich, Switzerland, September 2011. Edited by H. Eisenbeiss, M. Kunz, and H. Ingensand.

Jizhou, W., L. Zongjian and L. Chengming. 2004. Reconstruction of buildings from a single UAV image. *Proceedings International Society for Photogrammetry and Remote Sensing Congress* 6:100–103.

Li, D. R. 2016. The “Internet Plus M” space-based information services. *Journal of Remote Sensing* 20(5):708–715.

Li, D. R. and M. Li. 2014. Research advance and application prospect of unmanned aerial vehicle remote sensing system. *Geomatics and Information Science of Wuhan University* 39(5):505–513.

Li, D. R., X. L. Chen and X. B. Cai. 2008. Spatial information techniques in rapid response to Wenchuan earthquake. *Journal of Remote Sensing* 12(6):841–851.

Mafanya, M., P. Tsele and J. Botai. 2017. Evaluating pixel and object based image classification techniques for mapping plant invasions from UAV derived aerial imagery: Harrisia pomanensis as a case study. *ISPRS Journal of Photogrammetry and Remote Sensing* 129:1–11.

Messinger, M., G. P. Asner and M. Silman. 2016. Rapid assessments of amazon forest structure and biomass using small unmanned aerial systems. *Remote Sensing* 8(8):615.

Ministry of Natural Resources of the P.R.C (MNR of P.R.C.). 2018. *Bulletin of Island Statistics Survey in 2017 of China*. Governmental Bulletin, MNR, China.

O’Driscoll, J. 2018. Landscape applications of photogrammetry using unmanned aerial vehicles. *Journal of Archaeological Science: Reports* 22: 32-44.

Siebert, S. and J. Teizer. 2014. Mobile 3D mapping for surveying earthwork projects using an unmanned aerial vehicle (UAV) system. *Automation in Construction* 41:1–14.

Wang, X. J., H. Shan and Y. Zhang. 2017. Quick processing of emergency surveying and mapping data from UAV based on INPHO. *Standardization of Surveying and Mapping* 33(2):36–38.

Wu, J., P. Chen, S. F. Fu, Q. H. Chen and X. Pan. 2019. Study on cooperative regression of island leaf area index (LAI) based on UAV multiple remote sensing data. *Journal of Ocean Technology* 38(6):1–8.

Yan, J. B., X. Q. Duan and W. W. Zheng. 2020. An adaptive IDW algorithm involving spatial heterogeneity. *Geomatics and Information Science of Wuhan University* 45(1):97–104.

Yuan, L., Q. D. Liu and Y. G. Sun. 2018. Application of UAV image in the island topographic map of 1:500. *Hydrographic Surveying and Charting* 38(1):43–46.

Zarco-Tejada, P. J., R. Diaz-Varela and V. Angileri. 2014. Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. *European Journal of Agronomy* 55:89–99.

Zhang, C., S. Yang and C. Zhao. 2018a. Topographic data accuracy verification of small consumer UAV. *Journal of Remote Sensing* 1:185–195.

Zhang, K., H. Zhang and X. Lu. 2018b. Application of unmanned aerial vehicle aerial photogrammetry in island. *Bulletin of Surveying and Mapping* 9:87–90.

Zhang, X. P. and N. Yu. 2016. Highly-efficient methods for aerial triangulation based on INPHO. *Geospatial Information* 14(10):75–77.

# MCAFNet: Multi-Channel Attention Fusion Network-Based CNN for Remote Sensing Scene Classification

Jingming Xia, Yao Zhou, Ling Tan, and Yue Ding

## Abstract

*Remote sensing scene images are characterized by intra-class diversity and inter-class similarity. When recognizing remote sensing images, traditional image classification algorithms based on deep learning only extract the global features of scene images, ignoring the important role of local key features in classification, which limits the ability of feature expression and restricts the improvement of classification accuracy. Therefore, this paper presents a multi-channel attention fusion network (MCAFNet). First, three channels are used to extract the features of the image. The channel "spatial attention module" is added after the maximum pooling layer of two channels to get the global and local key features of the image. The other channel uses the original model to extract the deep features of the image. Second, features extracted from different channels are effectively fused by the fusion module. Finally, an adaptive weight loss function is designed to automatically adjust the losses in different types of loss functions. Three challenging data sets, UC Merced Land-Use Dataset (UCM), Aerial Image Dataset (AID), and Northwestern Polytechnic University Dataset (NWPU), are selected for the experiment. Experimental results show that our algorithm can effectively recognize scenes and obtain competitive classification results.*

## Introduction

With the advancement of remote sensing technology in recent years, the resolution of remote sensing images obtained from remote sensing satellites has steadily increased. However, the features of high-resolution images contain a great deal of redundant information, which leads to the problems of intra-class differences and inter-class similarity of remote sensing images, and thus brings great challenges to the accurate classification of remote sensing scene images. Intra-class diversity manifests itself, for instance, in the port scene as differences in size, color, shape, and background of different ports, which may lead to misclassification. The inter-class similarity indicates that different categories of scenes share similar texture characteristics. For instance, sparse residential area, medium residential area, and dense residential area are frequently difficult to classify. Consequently, it is easy to misclassify these images during the process of scene classification.

In recent years, the classification of remote sensing images has made extensive use of deep learning. Convolutional Neural Network (CNN) is an important deep learning technology. As a powerful image processing tool, CNN has achieved excellent application performance in variety of fields (Guan *et al.* 2022; Jiang *et al.* 2022; Krizhevsky *et al.* 2012; Long *et al.* 2015; Ren *et al.* 2015), especially in image classification (He *et al.* 2016; Huang *et al.* 2017; Zheng *et al.* 2021). The convolutional layer is the central component of a convolutional neural network (CNN), which extracts feature information by fusing spatial and channel information of local receptive fields. In the fields of image classification and object detection, models such as Visual Geometry Group Network (VGGNet) (Simonyan and Zisserman 2015), DenseNet (Huang *et al.* 2017), GoogleNet (Szegedy *et al.* 2015) and MobileNetV2 (Sandler *et al.* 2018) have achieved good performance over the years. Penatti *et al.* (2015) applied CNN to the classification of remote sensing image scene and extracted high-level visual features using CNN's end-to-end feature learning, demonstrating the network's strong generalization ability. However, this is only a preliminary application of CNN to remote sensing image classification and does not make good use of all the capabilities of CNN. Therefore, two-stage deep feature fusion convolutional neural network was proposed by Liu *et al.* (2017). The network is capable of adaptively fusing the feature information of the intermediate layer and the fully connected layer, making full use of the abundant information in the shallow layer, and effectively enhancing the classification accuracy. However, a large number of parameters are generated during training, in order to reduce the number of parameters and improve the model's accuracy. Zeng *et al.* (2018) proposed an end-to-end convolutional neural network that combines global contextual features and local object features, which achieves good accuracy in scene classification and effectively reduces network complexity.

However, although single-channel CNN can be used as a basic model for classification, their ability to extract features is limited by a certain depth. This prompts us to use multi-channel structure to extract features from images and fuse the features of different channels to improve the feature representation. Numerous studies have demonstrated that the multi-channel structure enables the network to extract and fuse different feature maps on multiple channels, thereby enhancing the performance of remote sensing image classification. Shi *et al.* (2020) proposed a bilinear-based method for feature extraction that combines depthwise separable convolution and standard convolution to extract image features and fuses the features extracted by the two channels. The complexity of the model is greatly reduced. To improve the classification accuracy of the model, Ma *et al.* (2020) designed a two-branch interactive spatial channel collaborative attention augmentation network (SCCA-net) for multi-resolution classification to improves the classification accuracy of the model. The sampling angle is adapted adaptively based on the texture distribution of homogeneous regions in order to capture neighborhood information that is more advantageous for classification. Shi *et al.* (2021) therefore proposed a convolutional neural network based on an attention mechanism and a multi-branch structure, which combined the attention mechanism with hybrid convolutions and attained good classification performance. Li *et al.* (2021) proposed an attention-based convolutional neural network for multi-layer feature aggregation due to the large number of parameters that are generated during training. In order to improve scene classification performance, the network aggregates valuable information from multi-layer convolutional features based on the semantic regions it has

Jingming Xia, Yao Zhou, and Yue Ding are with the School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, China (yaozhou@nuist.edu.cn).

Ling Tan is with the School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China.

*Contributed by Dongdong Wang, September 26, 2022 (sent for review October 25, 2022).*

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 3, March 2023, pp. 183–192.
0099-1112/22/183–192

doi: 10.14358/PERS.22-00121R2

learned. Although the above methods use multi-branch structure to extract image features, their fusion method is only achieved by channel splicing or numerical addition, and the extracted features cannot be well fused.

The problem of intra-class diversity and inter-class similarity in remote sensing image classification requires attention to the global and local key features of the image. Numerous studies have shown that adding an attention module to a convolutional neural network can improve the network's performance. This strategy enables the convolutional neural network to focus on the target region and obtain more detailed information regarding the target, while ignoring irrelevant information. By establishing interdependencies between channels, Hu *et al.* (2018) designed squeeze and excitation networks (SENet) to improve the model's feature response. The squeeze operation enables the network to squeeze features spatially and obtain a global receptive field through a global average pooling operation. The excitation operation is performed through two consecutive fully connected layers. SENet is a separate substructure that can be embedded in various network structures. However, embedding inevitably adds some parameters and computation in some classification networks. Therefore, Wang *et al.* (2020) proposed a lightweight efficient channel attention (ECA) module to replace the fully connected computation in SENet, using one-dimensional convolution to learn the relationship between local channels. Global average pooling to aggregate each channel significantly reduces the number of parameters. To enable the classification network to achieve higher performance after adding the attention module, Woo *et al.* (2018) developed a convolutional block attention module (CBAM). By providing an intermediate feature map, the attention weights are sequentially inferred along the two dimensions of space and channel, and then multiplied with the original feature map to adaptively adjust the features, thereby improving the performance of the model. Wang *et al.* (2018) developed an attention mechanism for CNN and introduced it into scene classification in order to highlight critical information in remote sensing images and discard irrelevant information. This method reduces the number of learned parameters by reducing high-level semantic and spatial features. Sun *et al.* (2022) proposed an encoder-decoder network for remote sensing image classification that fuses spatial attention and spectral channel attention. The network encodes spatial and spectral multi-channel contextual information and achieves good classification results. The intra-class diversity problem refers to the frequent misclassification of the same scene type in remote sensing image classification, due to differences in size, background, etc. The aforementioned studies demonstrate that the attention mechanism can alleviate this problem by increasing the network's focus on the region of interest while ignoring background features. However, the above work does not focus on the location of global and local key features in the image at the same time, so the accurate classification cannot be achieved.

To overcome these difficulties, a high-performance multi-channel attention fusion network (MCAFNet) is designed for remote sensing image scene classification. The main contributions are as follows.

(1) Aiming at the problem that the feature extraction ability of single channel is insufficient and the extracted features cannot be reasonably fused, a multi-channel extraction structure is designed based on VGG16 to obtain more effective features. The feature dimension and feature redundancy are reduced by reducing the convolution kernel. At the same time, through the feature fusion algorithm, the features extracted from different channels are effectively fused by strengthening the weight of the extracted feature map.
(2) Aiming at the problem of inter-class similarity and intra-class diversity in remote sensing images, this paper proposes a global-local attention module, which uses channel attention and spatial attention to pay attention to the global features and local key features of images.
(3) In actual classification, some data sets have class imbalance problems. This paper designs an improved adaptive weight loss function, increased tuning parameters can make the function more flexible.

## Method

### Overall Framework

VGG16 (Simonyan and Zisserman 2015) is a deep learning network model. VGG16 contains five convolutional modules and three fully connected layers. After each convolution module is a max-pooling layer, the number of convolution channels starts from 64 in the first layer, and doubles after each max-pooling layer until 512. Each convolution module contains multiple convolutional layers with 3 × 3 convolution kernels.

Based on VGG16, we design a new convolutional neural network for image classification. The network adopts a multi-channel structure and adds an attention mechanism module to two of the channels, and then uses fusion method to fuse the features extracted by the multichannel. Its original structure is shown in Figure 1.

Firstly, feature extraction is performed on the image using multi-channel structure. The first channel can make better use of the original features, the second channel can be more sensitive to image changes, and the third channel is the original VGG16 network, which can extract image details.

Secondly, the attention module is added after the first and third maximum pooling layers of VGG16, which solves the problem of information loss due to network depth when extracting image features. Attention module includes channel attention module and spatial attention module. The former obtains the global features of the image, which enhances the effective information in the channel and suppresses the invalid information. The latter is used to extract the features of the area of interest; that is, the local key features in the task body related to the classification of remote sensing images, so as to improve the recognition accuracy of the model for similar scenes. Adding an attention module after the first maximum pooling layer allows the underlying feature mapping to contain more image information. An attention module is added after the third maximum pooling layer to enhance the ability to represent intermediate features.

Finally, if stitching fusion is used directly to fuse features extracted from multiple channels, the result must be highly redundant. Therefore, we halve the number of convolution cores in the model and use the

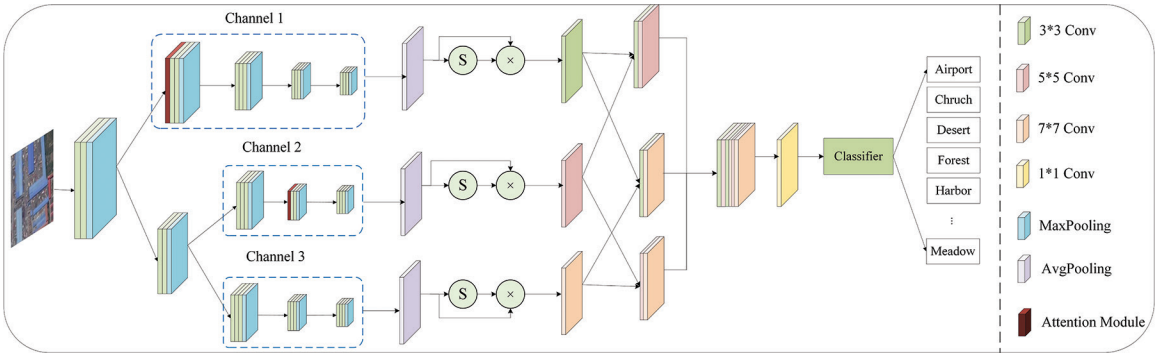



Figure 1. The overall framework of multi-channel attention fusion network (MCAFNet).

fusion module to enhance the features extracted from the three channels by weighting them, then fuse them. After the fusion module, the feature fusion results are sent to the classifier to get the classification results.

## Attention Module

### *Channel Attention Module*

Channel attention is to focus on the importance of each feature channel of the input image, and then enhance or suppress different channels for different tasks. As shown in Figure 2, for an input feature $F$ of size $H\times W\times C$, the global average pooling and global maximum pooling are first used for the feature map $F$ to aggregate the global information of the feature map and obtain the channel descriptions of two 1×1×C, the average pooling feature $F^c_{avg}$ and the maximum pooling feature . Then the features $F^c_{avg}$ and $F^c_{max}$ output from the two channels are add fused to obtain more discriminative features. Then a multi-layer perceptron with hidden layers is added to obtain the weighted channels of each feature for learning the final channel attention feature map $M_c(F)$. To reduce the parameter overhead, the size of the hidden layer is set to $C/r$, where $r$ is the compression ratio. Finally, the input feature map and the channel weights are multiplied to enhance the effective features. The weight coefficients are obtained by the sigmoid activation function. The equation of Channel Attention Module (CAM) is as follows.

$$\begin{aligned} M_c(F) &= \sigma(\sigma(MLP(Maxpooling(F) + Avgpooling(F))) \times F) \\ &= \sigma(\sigma(W_1(W_0(F^c_{max} + F^c_{avg})))) \end{aligned} \quad (1)$$

where $\sigma$ is the sigmoid activation function, $W_0 \in R^{(C/r)\times C}$, $W_1 \in R^{C\times(C/r)}$, $W_0$ and $W_1$ are the weights of the Multi Layer Perception (MLP), and $F$ is the input feature map.

### *Spatial Attention Module*

The spatial attention module finds the most important part of the network for processing, which can effectively improve the feature representation. As shown in Figure 3, to focus on semantically similar features in the spatial neighborhood, for the input features of size $H\times W\times C$, first, local average pooling and local maximum pooling are applied along the respective channels to obtain the feature maps $F_{LAP}$ and $F_{LMP}$, respectively, and obtain the neighborhood information of . Then, the channel information of the feature graph is aggregated using global maximum pooling and global average pooling on the two channels, respectively. Then they are concatenated separately to obtain a better feature representation capability. To mitigate the feature loss due to multiple pooling operations, the features are further learned using 3×3 convolution, respectively. Then the respective weights are multiplied with $F_{LAP}$ and $F_{LMP}$, respectively, to obtain feature maps $F'_{LAP}$ and $F'_{LMP}$, enhance the effective features. Then concatenation is performed to obtain a better feature representation. Finally, upsampling is performed by deconvolution to obtain a spatially attentive feature map $M_s(F)$ with the same scale as the original input image or feature map. The equation of Spatial Attention Module (SAM) is as follows:

$$F'_{LMP} = f^{3\times3}[Maxpooling(F_{LMP}); Avgpooling(F_{LMP})] \quad (2)$$

$$F'_{LAP} = f^{3\times3}[Maxpooling(F_{LAP}); Avgpooling(F_{LAP})] \quad (3)$$

$$M_s(F) = \sigma(f_d^{1\times1}[F'_{LMP}; F'_{LAP}]) \quad (4)$$

where $\sigma$ is the sigmoid activation function, $f^{3\times3}$ represents 3×3 convolution, $f_d^{1\times1}$ represents 1×1 deconvolution, and $F$ is the input feature map.

### *Combine Channel Attention Module and Spatial Attention Module*

Given a feature map, attention complementation is performed in parallel using the attention module CAM and SAM calculations. The former determines the importance of each feature channel, and the latter finds the key features in the network. As shown in Figure 4, the overall computation process can be summarized as follows:

$$F' = \sigma(SAM(CAM(F)\otimes F)\otimes F)\otimes F \quad (5)$$

where $\sigma$ is the sigmoid activation function, $SAM$ represents local spatial attention, and $CAM$ represents the global channel attention.

### *Loss Function*

In order to automatically adjust the weight of the ground object loss, this paper improves an adaptive weight algorithm based on the cross-entropy loss function to improve the loss evaluation ability of the

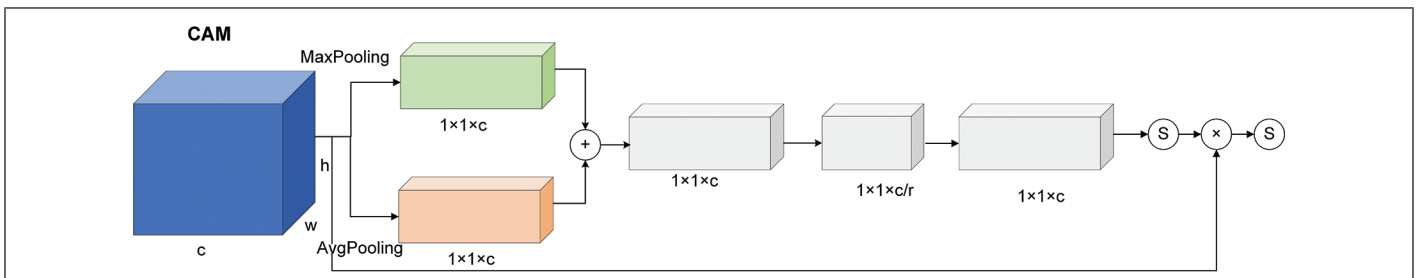



Figure 2. Channel Attention Module.

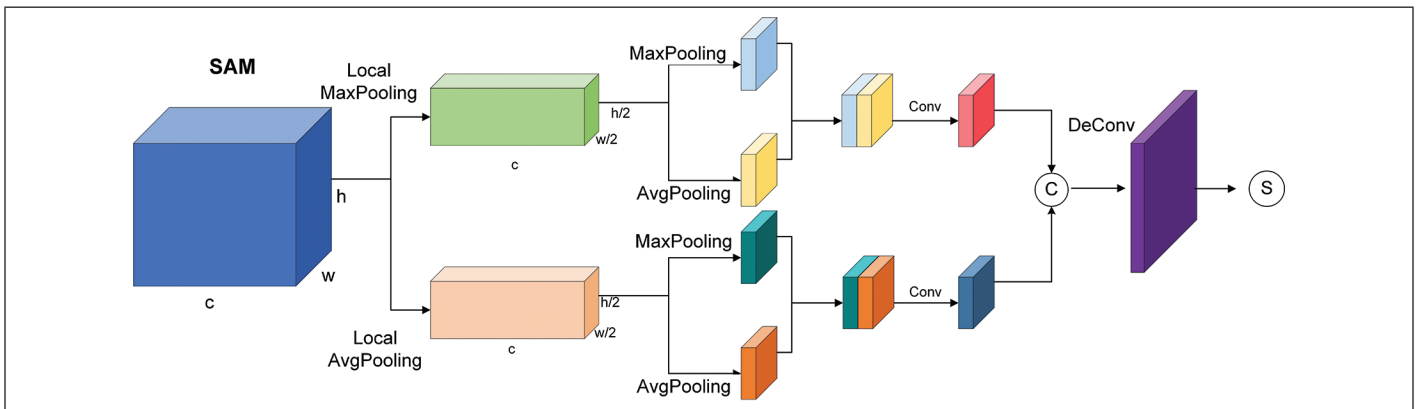



Figure 3. Spatial Attention Module.

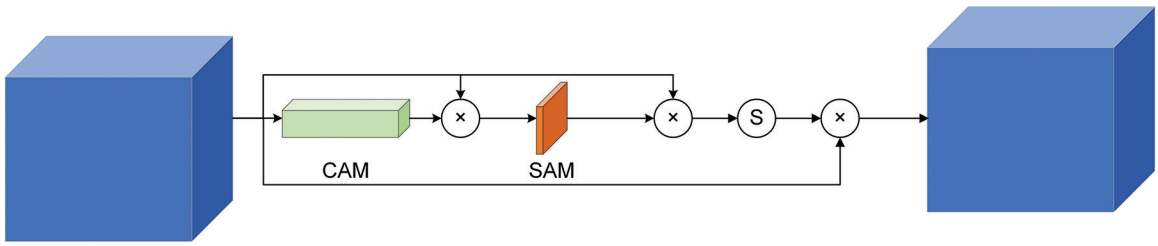


Figure 4. Combine Channel Attention Module (CAM) and Spatial Attention Module (SAM).

model. Cross-entropy was mainly used to evaluate the difference information between two probability distributions, which can be understood as the difference between the information entropy during fitting and the information entropy of the sample annotation, which is expressed as:

$$Loss = -\frac{1}{n}(y_i(ln_a) + (1-y_i)ln(1-a)) \tag{6}$$

$$\frac{\partial Loss}{\partial w_j} = \frac{1}{n}\sum_x x_j(\sigma(z)-y)$$
$$\frac{\partial Loss}{\partial b} = \frac{1}{n}\sum_x (\sigma(z)-y) \tag{7}$$

where

$$z = wx + b \tag{8}$$

where *Loss* represents the loss value, *x* represents the sample, *y* represents the actual value, *a* represents the output value, and *n* represents the number of samples. Then the parameter $\sigma(z) - y$ indicates the distance between the output value and the actual expected value. When the error is larger, the corresponding gradient is larger, and the parameters *w* and *b* are adjusted faster, which can speed up the training speed. High-resolution images are characterized by intra-class diversity and inter-class similarity. To solve this problem, Paszke *et al.* (2016) designed an adaptive weight calculation method. In this paper, the method is improved by adding adjustment parameters to make the adaptive weight function more flexible. The improved class weight function is defined as:

$$w_{class} = \frac{1}{\ln(\alpha + \beta P_{class})} \tag{9}$$

where $w_{class}$ represents the weight of different categories, $P_{class}$ represents the proportion of samples of this class, and $\alpha$ and $\beta$ are two hyperparameters used to adjust the shape and range of the weight function, During the experiment, $\alpha$ is set to 1.1 and $\beta$ is set to 0.35.

## Experimental Analysis and Discussion

### Data Sets

This paper selects three data sets (UC Merced Land-Use Dataset (UCM) (Yang and Newsam 2012), Aerial Image Dataset (AID) (Xia *et al.* 2017), and Northwestern Polytechnic University Dataset (NWPU) (Cheng *et al.* 2017)) to train MCAFNet and conduct comparative image classification experiments in order to validate the performance of the algorithm.

The UCM data set contains a total of 2100 scene images of 21 categories, 100 images per category, the size of the images is 256 × 256 pixels, and the spatial resolution is 0.3 m. The AID data set contains a total of 10 000 scene images of 30 categories, 220–420 images per category, the size of the images is 600 × 600 pixels, and the spatial resolution is 0.5–8 m. The NWPU data set contains 31,500 images divided into 45 categories, with each category contains 700 scene images with an image size of 256 × 256 pixels. Image resolution between 0.2 m and 30 m.

### Experimental Setup

The experiments were carried out under the Ubuntu Pytorch framework, running on an NVIDIA 2080Ti with 11 GB of video memory. Adam was used as the optimizer in the training process. All models were trained for 100 epochs, the batch size was set to 64, and the initial learning rate was 0.001. To improve the reliability of classification results, all experiments were repeated 10 times and three data sets were randomly divided. For UCM data set, 50% and 80% training samples are used in the experiment. For the AID data set, 20% and 50% training samples were used in the experiment. For the NWPU data set, 10% and 20% training samples are used in the experiment.

### Evaluation Metrics

(1) Overall Accuracy (OA): Refers to the ratio between the number of correctly classified samples and the number of all samples.
(2) Standard Deviation (STD): It is a reflection of the difference between all values and their mean values. The performance of the model varies with each training session. Therefore, the reliability of the model requires multiple experiments to calculate the standard deviation.
(3) Confusion Matrix (CM): Confusion matrix is a two-dimensional table, which is used to analyze the classification error and confusion degree between classes and visualize the performance of the algorithm.

### Compared to State-of-the-Art Methods

In this section, in order to evaluate the performance of our method, we compared it with some advanced methods in recent years on three remote sensing scene image data sets, namely UCM, AID, and NWPU, as shown in Table 1 to Table 3.

Table 1 shows the classification performance comparison between our method and advanced method on UCM data set. The UCM data set was divided into 80% training ratios. After analyzing the existing advanced methods, it is found that the classification accuracy will reach 98% in 2018 and 99% in 2022, and the classification accuracy will gradually become saturated. Compared with 99%, our method achieves a higher accuracy of 99.14%, which is 0.14% higher. This fully shows that our method can distinguish the importance of different features, improve the performance of features, and thus improve the classification accuracy.

Table 1. Overall accuracies (%) of different methods with 80% training ratio in the UCM data set.

| Methods | Training Ratio (80%) |
|---|---|
| VGG-VD-16 (Xia *et al.* 2017) | 95.21 |
| MCNN (Liu *et al.* 2017) | 96.66 |
| Two-Stream (Yu and Liu 2018) | 98.02 ±1.03 |
| MDFR (Zhang *et al.* 2019) | 98.02 ± 0.51 |
| conv5-MSP5-FV (Zheng *et al.* 2019) | 98.60 |
| Attention GANs (Yu *et al.* 2020) | 97.69 ± 0.69 |
| SAFF (Cao *et al.* 2021) | 97.02 ± 0.78 |
| Attention CNN + H-GCN (Gao *et al.* 2021) | 99.00 ± 0.43 |
| Ours | 99.14 ± 0.27 |

Table 2 shows the classification performance comparison between our method and advanced method on AID data set. The AID data set was divided into 20% and 50% training ratios. It can be seen from the table that our method achieves 93.72% and 96.06% classification accuracy, respectively. Compared with the advanced methods in recent years. Our method has obtained significant classification results, and the classification accuracy has been improved by 0.66% and 0.28%, respectively, higher than previous advanced methods.

Table 3 shows the classification performance comparison between our method and advanced method on NWPU data set. NWPU data set is the data set with the largest data volume and category number in our experiment. The AID data set was divided into 10% and 20% training ratios. It can be seen from the table that our method achieves the classification accuracy of 91.97% and 93.86%, respectively. At the training rate of 10%, the overall accuracy is 1.22% and 0.58% higher than the previous advanced methods of Combined CNN and GCN and Attention CNN + H – GCN, respectively. At the training rate of 20%, the overall accuracy is 0.99% and 0.24% higher than that of the previous advanced methods of Combined CNN and GCN and Attention CNN + H – GCN, respectively.

### Reports of Confusion Matrix

In order to understand the effect of our method, the confusion matrix is drawn to discuss the advantages and disadvantages of the proposed method. In order to show the recognition effect in more detail, the confusion matrix of recognition results on the three data sets of UCM, AID, and NWPU is drawn, as shown in Figure 5–Figure 9. In this confusion matrix, the values on the diagonal are all correct prediction results, and the remaining values are all wrong prediction results caused by model misjudgment. Each row of the matrix represents the real category, while each column of the matrix represents the prediction label of the model.

Figure 5 shows the confusion matrix with a training ratio of 80% on the UCM data set. It can be seen from the figure that most of the 21 scene categories have achieved 100% classification accuracy, and several categories with poor effects have also achieved more than 90% accuracy. Medium dense residential areas are wrongly classified as dense residential areas. This is because the shape and texture features of the two categories are too similar. This makes it difficult to distinguish models effectively, but our model is still valid. For example, highways, intersections, overpasses, and runways are roads. We succeeded in classifying them correctly.

Table 2. Overall accuracies (%) of different methods with 20% and 50% training ratio in the AID data set.

| Methods | Training Ratio (20%) | Training Ratio (50%) |
|---|---|---|
| VGG-VD-16 (Xia *et al.* 2017) | 86.59 ± 0.29 | 89.64 ± 0.36 |
| MCNN (Liu *et al.* 2017) | — | 91.80 ± 0.22 |
| Two-Stream (Yu and Liu 2018) | — | 94.65 ± 0.33 |
| MDFR (Zhang *et al.* 2019) | 90.62 ± 0.27 | 93.37 ± 0.29 |
| conv5-MSP5-FV (Zheng *et al.* 2019) | — | 93.90 |
| Attention GANs (Yu *et al.* 2020) | 92.30 ± 0.24 | 94.93 ±0.21 |
| SAFF (Cao *et al.* 2021) | 90.25 ± 0.29 | 93.83 ± 0.28 |
| Attention CNN + H-GCN (Gao *et al.* 2021) | 93.06 ± 0.26 | 95.78 ± 0.37 |
| Ours | 93.72 ± 0.28 | 96.06 ± 0.29 |

Table 3. Overall accuracies (%) of different methods with 10% and 20% training ratio in the NWPU data set.

| Methods | Training Ratio (10%) | Training Ratio (20%) |
|---|---|---|
| VGG-VD-16 (Xia *et al.* 2017) | 87.15 ± 0.45 | 90.36 ± 0.18 |
| Two-Stream (Yu and Liu 2018) | — | 83.16 ± 0.18 |
| MDFR (Zhang *et al.* 2019) | 83.37 ± 0.26 | 86.89 ± 0.17 |
| Attention GANs (Yu *et al.* 2020) | 86.11 ± 0.22 | 89.44 ± 0.18 |
| SCCov (He *et al.* 2020) | 89.30 ± 0.35 | 92.10 ± 0.25 |
| SAFF (Cao *et al.* 2021) | 84.38 ± 0.19 | 87.86 ± 0.14 |
| Combined CNN and GCN (Liang *et al.* 2021) | 90.75 ± 0.21 | 92.87 ± 0.13 |
| Attention CNN + H-GCN (Gao *et al.* 2021) | 91.39 ± 0.19 | 93.62 ± 0.28 |
| Ours | 91.97 ± 0.24 | 93.86 ± 0.17 |

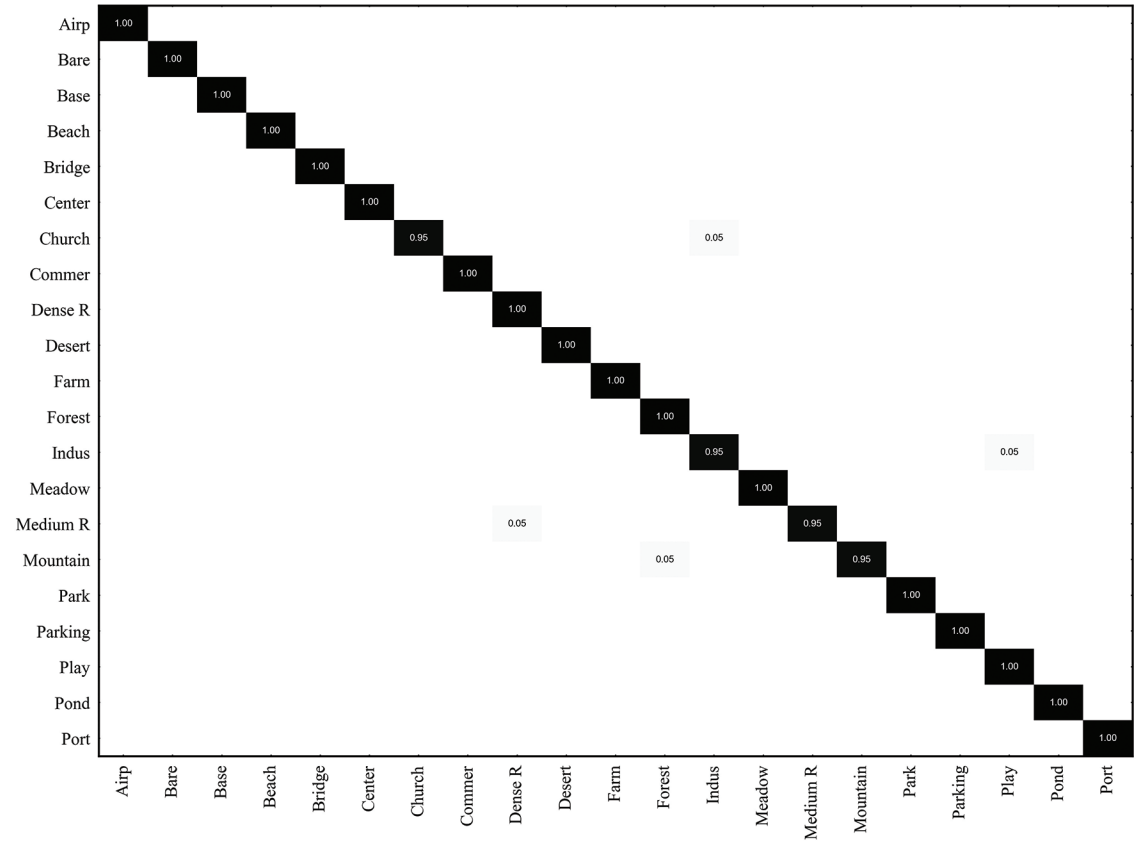



Figure 5. Confusion matrix of our method on UCM data set under the training ratio of 80%.

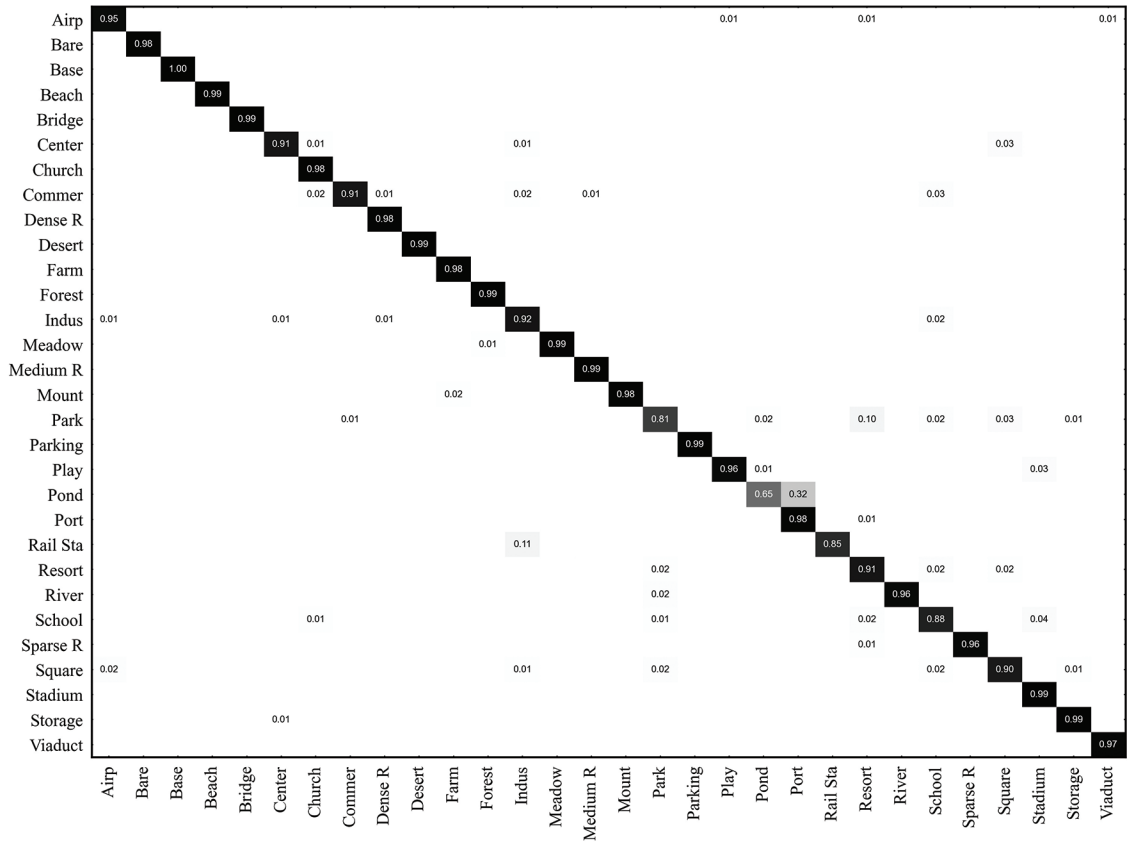



Figure 6. Confusion matrix of our method on AID data set under the training ratio of 20%.

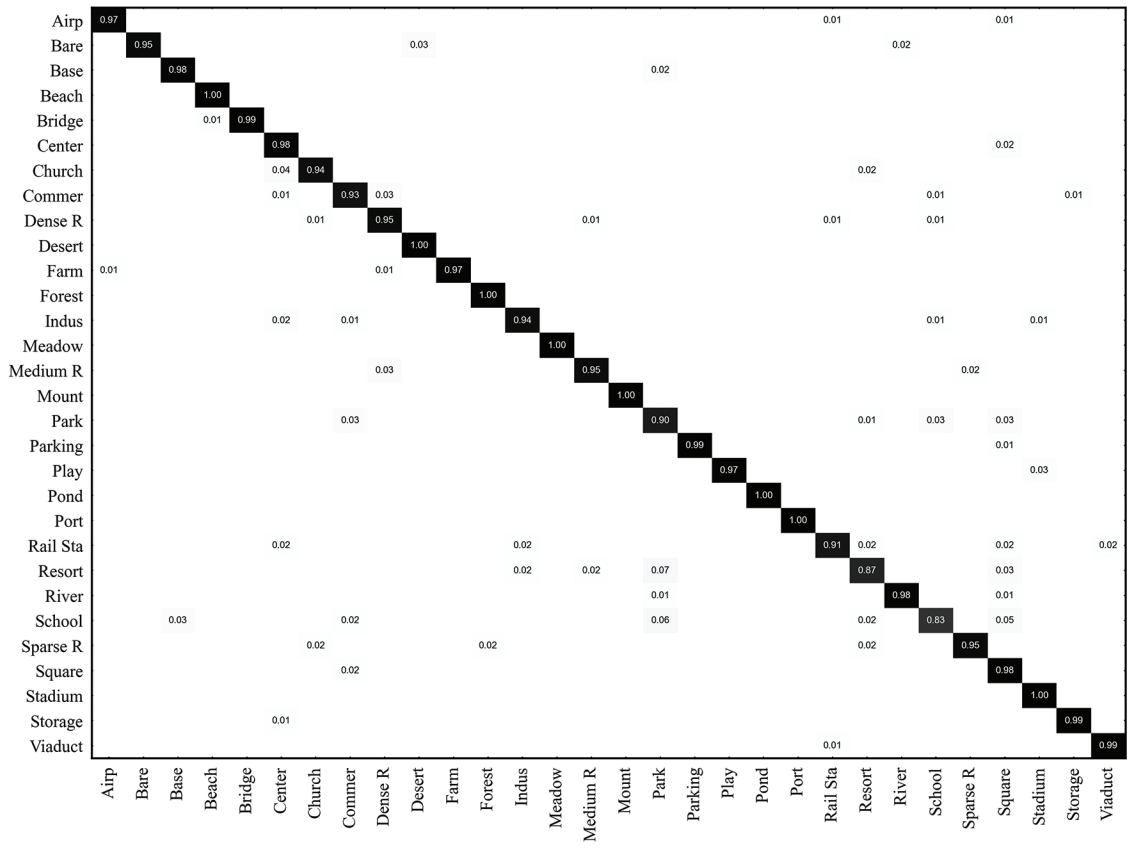



Figure 7. Confusion matrix of our method on AID data set under the training ratio of 50%.

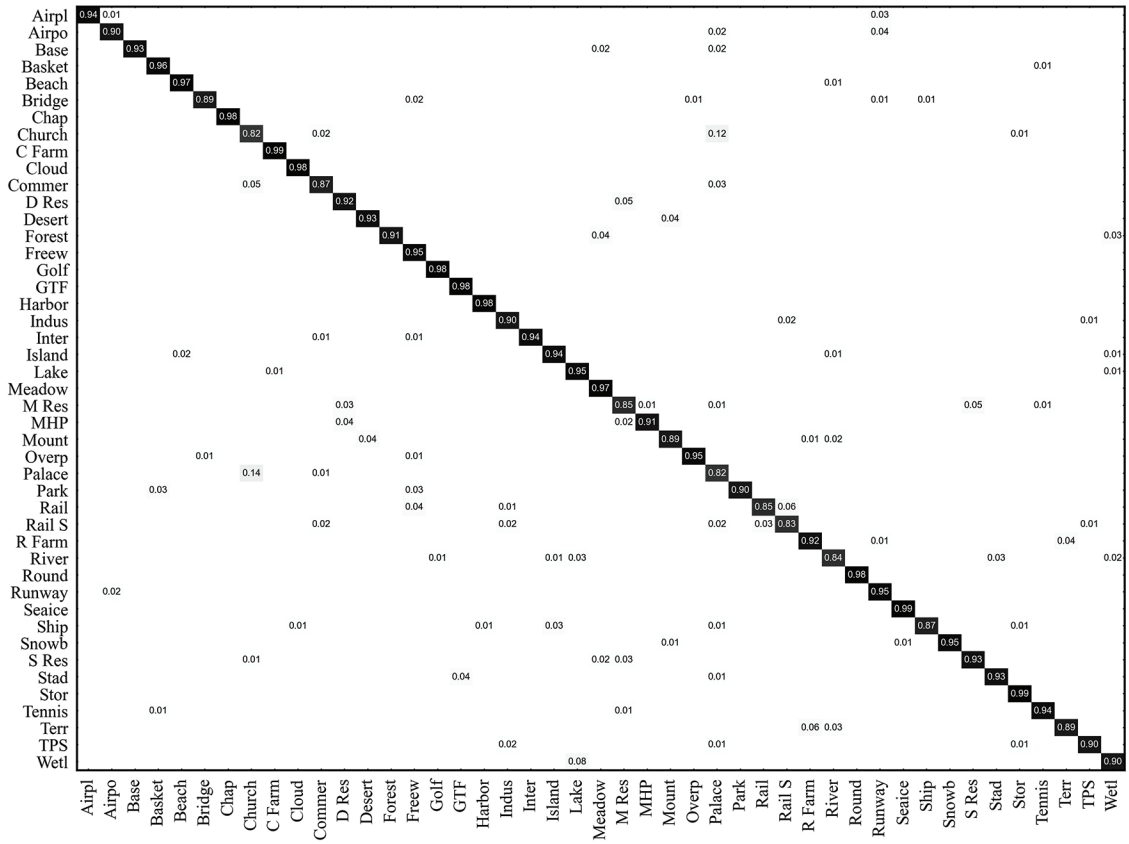



Figure 8. Confusion matrix of our method on NWPU data set under the training ratio of 10%.

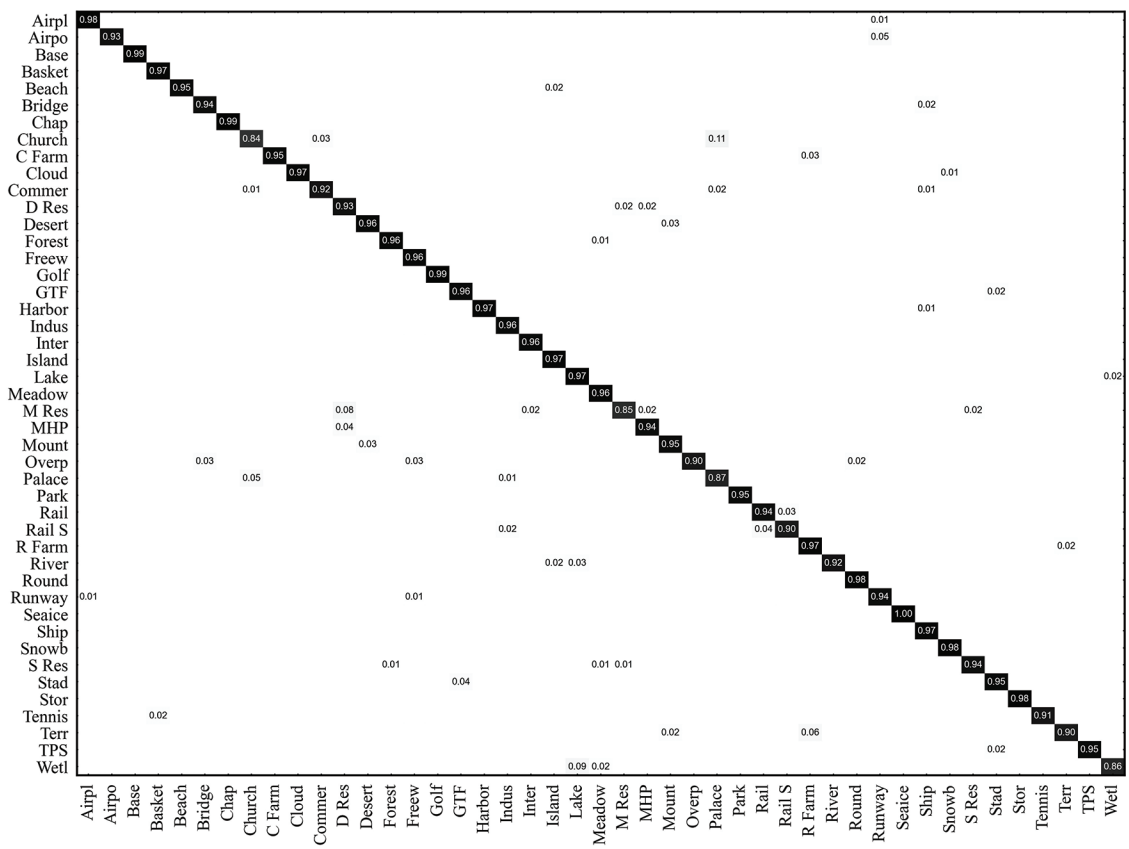



Figure 9. Confusion matrix of our method on NWPU data set under the training ratio of 20%.

Figure 6 and Figure 7 show the confusion matrix when the training rate of AID data set is 20% and 50%, respectively. In the confusion matrix with a training rate of 20%, the worst effect is Pond, which has only achieved 65% accuracy, 32% of it is incorrectly identified as port, because the two categories are highly similar in some areas. In the confusion matrix with a training rate of 50%, only the classification accuracy of three categories is lower than 90%, and good results have been achieved.

The confusion matrix of NWPU data set is shown in Figure 8 and Figure 9 shows that when the training rate is 10% and 20%, the classification accuracy of 35 and 40 categories is higher than 90%. Under the two different training rates, the accuracy rate of medium dense residential areas is low, and the training accuracy is kept at 85%, which shows that with the increase of training data, the accuracy rate has not improved, and the model has not learned from the data. This is due to the extremely similar architectural style and layout between categories.

## Ablation Study

### *Attention Position Ablation Study*

The location of attention modules has an important impact on the performance of the network. In this paper, attention modules are added after the different maximum pooling layers of VGG16, and the best location for adding modules is selected through experiments. VGG16 has five maximum pooling layers. The model for adding attention modules after the first maximum pooling layer is VGG-AM1, and the model for adding attention modules after the second maximum pooling layer is VGG-AM2, and so on.

In this paper, five network structures from VGG-AM1 to VGG-AM5 are used for experiments on three data sets, namely UCM, AID, and NWPU. The experimental results are shown in Table 4. VGG-AM1 achieves the best precision on the UCM data set. On the AID data set, VGG-AM2 has the best recognition accuracy; on the NWPU data set, VGG-AM3 has the best recognition accuracy. CNN extracts the depth of image features, but the feature mapping after depth extraction becomes abstract. Adding the attention module in the shallow layer is beneficial to feature mapping to obtain better feature representation. The results show that the attention module can enhance the features of more interesting visual areas in similar scenes, thus improving the recognition accuracy.

Figure 10 shows the accuracy curve of VGG-AM3 on the AID data set and the NWPU data set. It can be seen from the figure that the recognition accuracy is relatively stable. Although the accuracy of VGG-AM3 is not the highest, after multi-layer feature extraction, the model filters the redundant information of the image, and the obtained feature map excludes more interference information, indicating that VGG-AM3 has a better network structure.

## Multi-Channel Fusion Ablation Study

To obtain a more comprehensive feature map, this paper proposes a three-channel feature extraction model.

Table 4. Experimental results of attention module.

| | UCM (%) | AID (%) | | NWPU (%) | |
|---|---|---|---|---|---|
| **Methods** | **80** | **20** | **50** | **10** | **20** |
| VGG-AM1 | 91.69 | 89.37 | 91.66 | 85.42 | 87.35 |
| VGG-AM2 | 91.61 | 89.42 | 92.65 | 86.16 | 87.09 |
| VGG-AM3 | 91.53 | 88.76 | 91.25 | 86.32 | 87.13 |
| VGG-AM4 | 91.38 | 88.63 | 90.41 | 84.24 | 86.69 |
| VGG-AM5 | 90.48 | 87.95 | 90.62 | 84.19 | 86.27 |

Table 5. Experimental results of multi-channel network.

| | UCM (%) | AID (%) | | NWPU (%) | |
|---|---|---|---|---|---|
| **Models** | **80** | **20** | **20** | **10** | **20** |
| VGG-AM1 | 91.69 | 89.37 | 91.66 | 85.42 | 87.35 |
| VGG-AM2 | 91.61 | 89.42 | 92.65 | 86.16 | 87.09 |
| Multichannel1 | 97.15 | 91.68 | 93.47 | 89.36 | 91.34 |
| Multichannel2 | 98.03 | 92.18 | 94.96 | 90.58 | 91.83 |

In the experiments on NWPU, VGG-AM1 achieved the best results in five classification comparisons, and VGG-AM3 achieved better accuracy in five classification comparisons. Therefore, VGG-AM1 and VGG-AM3 are selected as two channels of feature extraction structure in this paper. The first channel is VGG-AM1; the second channel is VGG-AM3; and the third channel is VGG16. There is no attention module in this channel, and the features obtained are more abstract, which can fully retain the texture and detail information of the image. The multi-channel mechanism gives consideration to both the deep details of the image and the more interesting information in the vision.

This paper proposes two multi-channel feature fusion methods: Multichannel1 and Multichannel2. Multichannel1 fuses the features of the final three channels through add, and then transfers them to the classifier. Multichannel2 splices the features of the final three channels through concatenation, and then transfers them to the classifier.

In order to verify the performance of Multichannel1 and Multichannel2, the recognition results of the multi-channel network are compared with those of the best VGG-AM1 and VGG-AM2. The results show that Multichannel2 has the best recognition effect on the three data sets, as shown in Table 5.

The recognition accuracy of Multichannel2 on NWPU data set is 91.83%, which is 4.48% higher than VGG-AM1; The recognition accuracy of Multichannel2 on AID data set is 94.96%, which is 2.31% higher than VGG-AM2. Therefore, this paper selects Multichannel 2 as the feature fusion method of multi-channel modules.

Figure 11 shows the decline chart of precision and loss function of experimental results under AID data set. In Figure 7a, the accuracy curve of Multichannel2 tends to be stable at about time period = 70, and the loss curve is more stable than VGG-AM1. Although the recognition accuracy of Multichannel2 and VGG-AM2 is only 0.05% different, the value of VGG-AM2 is oscillatory.

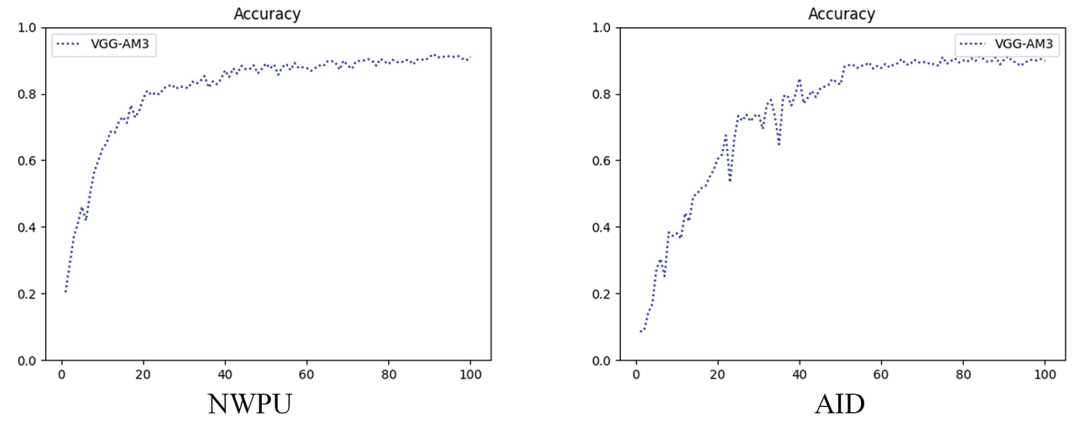



Figure 10. Accuracy curve of VGG-AM3.

The experimental results show that the multi-channel feature extraction structure has stable performance and better recognition accuracy than the single channel pattern. However, because the network structure based on Multichannel2 fusion method is more complex than VGG-AM, the iteration speed is reduced in the training process.

Table 6. Loss function ablation experiment results.

| | UCM (%) | AID (%) | | NWPU (%) | |
|---|---|---|---|---|---|
| **Loss** | **80** | **20** | **50** | **10** | **20** |
| Mean-Squared Error Loss | 96.56 | 91.68 | 94.37 | 87.25 | 90.32 |
| Mean-Absolute Error Loss | 97.35 | 91.14 | 93.71 | 89.46 | 90.61 |
| Cross Entropy Loss | 98.03 | 92.18 | 94.96 | 90.58 | 91.83 |
| MCAFNet | 99.14 | 93.72 | 96.06 | 91.97 | 93.86 |

MCAFNet = multi-channel attention fusion network.

### Loss Function Ablation Study

Compared with the Multichannel2 using the traditional cross entropy loss function, MCAFNet uses the improved loss function and achieves better results. The results are shown in Table 6.

### Visualization

To better illustrate the impact of the proposed algorithm on remote sensing scene classification, and to make the performance of the proposed MCAFNet method more intuitive, we visualize the remote sensing images in the data set using the Grad-CAM method, as shown in Figure 12. It can be seen from the figure that the recognition accuracy of the visualization results without the attention module in the main part of the recognition task is low, and our method can give good consideration to both the global features and local key features of the image, so we get a good visualization effect.

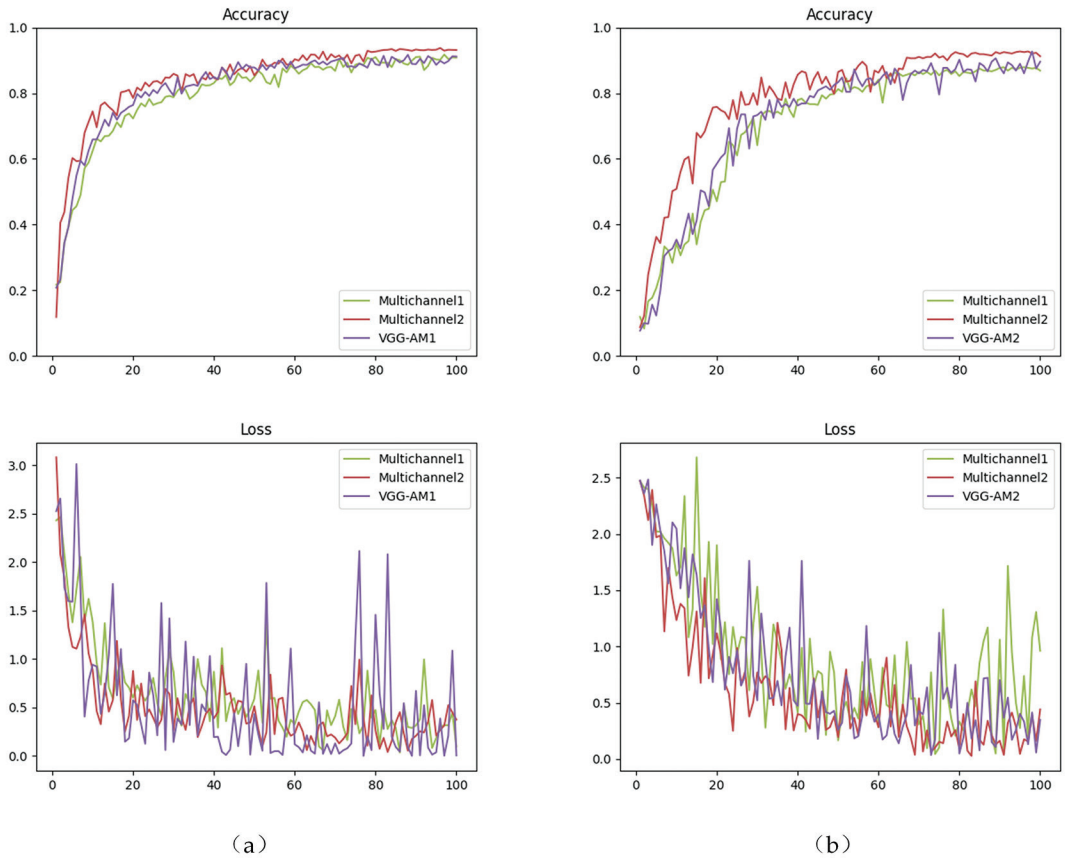



Figure 11. Recognition accuracy of three network structures.

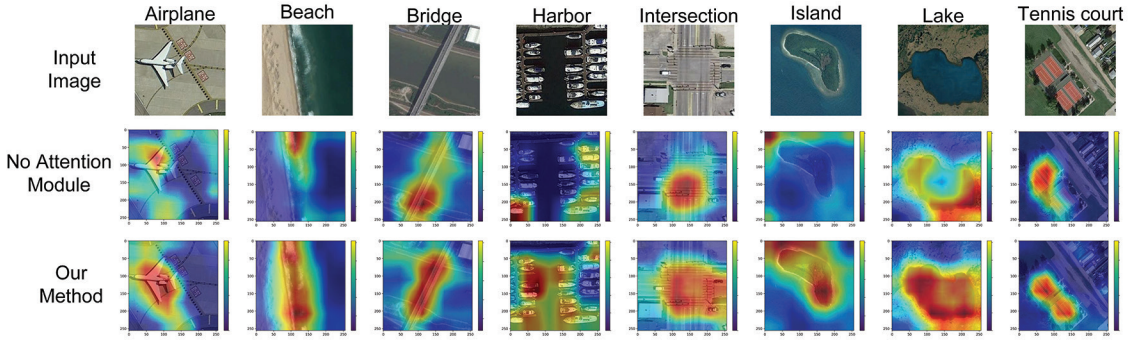



Figure 12. Grad-CAM visualization results of some remote sensing image samples.

## Conclusion

For the intra-class diversity and inter-class similarity problems in urban remote sensing scenes, a CNN-based MCAFNet is proposed in this paper. The network is based on VGG16, and the attention module is added after different pooling layers to obtain global and local key features of images, and a multi-channel feature extraction and fusion structure is used to obtain effective image features. Meanwhile, the recognition difficulties caused by category imbalance are mitigated by adaptive loss functions. The experimental results show that our method can effectively recognize complex and similar scenes in urban remote sensing images.

## Acknowledgments

We acknowledge the support from Jiangsu Province Industry-Academia-Research Fund (Grant: BY2022459). The remote sensing image classification data sets used in this study can be obtained from the following websites: https://hyper.ai/datasets/5431, https://hyper.ai/datasets/5446, https://hyper.ai/datasets/5449.

## References

Cao, R., L. Fang, T. Lu and N. He. 2021. Self-attention-based deep feature fusion for remote sensing scene classification. *Journal of IEEE Geoscience and Remote Sensing Letters* 18(1):43–47.

Cheng, G., J. Han and X. Lu. 2017. Remote sensing image scene classification: Benchmark and state of the art. *Journal of Institute of Electrical and Electronics Engineers* 105(10):1865–1883.

Gao, Y., J. Shi, J. Li and R. Wang. 2021. Remote sensing scene classification based on high-order graph convolutional network. *Journal of European Journal of Remote Sensing* 54(1):141–155.

Guan, Z., X. Miao, Y. Mu, Q. Sun, Q. Ye and D. Gao. 2022. Forest fire segmentation from aerial imagery data using an improved instance segmentation model. *Journal of Remote Sensing* 14(13):31–59.

He, K., X. Zhang, S. Ren and J. Sun. 2016. Computer vision and pattern recognition. Deep residual learning for image recognition. Pages 770–778 in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, held in Las Vegas Nev.: 27–30 June 2016.

He, N., L. Fang, S. Li, J. Plaza and A. Plaza. 2020. Skip-connected covariance network for remote sensing scene classification. *Journal of IEEE Transactions on Neural Networks and Learning Systems* 31(5):1461–1474.

Hu, J., L. Shen and G. Sun. 2018. Computer vision and pattern recognition. Squeeze-and-excitation networks. Pages 7132–7141 in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, held in Salt Lake City, Utah, 18–23 June 2018.

Huang, G., Z. Liu, L. Van Der Maaten and K. Q. Weinberger. 2017. Densely connected convolutional networks. Pages 4700–4708 in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, held in Honolulu, Hawaii, 9–12 November 2017.

Huang, Z., Z. Pan and B. Lei. 2017. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Journal of Remote Sensing* 9(9):17–28.

Jiang, M., L. Xu and D. A. Clausi. 2022. Sea ice–water classification of radarsat-2 imagery based on residual neural networks (RESNET) with regional pooling. *Journal of Remote Sensing* 14(13):63–79.

Krizhevsky, A., I. Sutskever and G. E. Hinton. 2012. Imagenet classification with deep convolutional neural networks. *Journal of Advances in Neural Information Processing Systems* 25(1):84–90.

Li, M., L. Lei, Y. Tang, Y. Sun and G. Kuang. 2021. An attention-guided multilayer feature aggregation network for remote sensing image scene classification. *Journal of Remote Sensing* 13(16):31–47.

Liang, J., Y. Deng and D. Zeng. 2021. A deep neural network combined CNN and GCN for remote sensing scene classification. *Journal of Selected Topics in Applied Earth Observations Remote Sensing* 13(2):4325–4338.

Liu, Y., Y. Liu and L. Ding. 2017. Scene classification based on two-stage deep feature fusion. *Journal of IEEE Geoscience and Remote Sensing Letters* 15(2):183–186.

Long, J., E. Shelhamer and T. Darrell. 2015. Fully convolutional networks for semantic segmentation. *Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(4):640–651.

Ma, W., J. Zhao, H. Zhu, J. Shen, L. Jiao, Y. Wu and B. Hou. 2020. A spatial-channel collaborative attention network for enhancement of multiresolution classification. *Journal of Remote Sensing* 13(1):19–36.

Paszke, A., A. Chaurasia, S. Kim and E. Culurciello. 2016. International conference on learning representations. Enet: A deep neural network architecture for real-time semantic segmentation. Pages 1–10 in *Proceedings of the International Conference on Learning Representations*, held in Toulon Var, France, 24–26 April 2016.

Penatti, O. A., K. Nogueira and J. A. Dos Santos. 2015. Computer vision and pattern recognition workshops. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains. Pages 44–51 in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, held in Boston, Mass., 7–12 June 2016.

Ren, S., K. He, R. Girshick and J. Sun. 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. *Journal of Advances in Neural Information Processing Systems* 28(1):91–99.

Sandler, M., A. Howard, M. Zhu, A. Zhmoginov and L. Chen. 2018. Computer vision and pattern recognition. Mobilenetv2: Inverted residuals and linear bottlenecks. In 4510–4520 in *Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, Utah, 18–22 June 2018.

Shi, C., T. Wang and L. Wang. 2020. Branch feature fusion convolution network for remote sensing scene classification. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13(4):5194–5210.

Shi, C., X. Zhao and L. Wang. 2021. A multi-branch feature fusion strategy based on an attention mechanism for remote sensing image scene classification. *Journal of Remote Sensing* 13(10):19–34.

Simonyan, K. and A. Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. *Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(6):379–490.

Sun, J., J. Zhang, X. Gao, M. Wang, D. Ou, X. Wu and D. Zhang. 2022. Fusing spatial attention with spectral-channel attention mechanism for hyperspectral image classification via encoder–decoder networks. *Journal of Remote Sensing* 14(9):19–31.

Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich. 2015. Conference on computer vision and pattern recognition. Going deeper with convolutions. Pages 1–9 in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, held in Boston, Mass., 7–12 June 2015.

Wang, Q., B. Wu, P. Zhu, P. Li, W. Zuo and Q. Hu. 2020. Conference on computer vision pattern recognition. Eca-net: Efficient channel attention for deep convolutional neural networks. Pages 11531–11539 in *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition*, held in Seattle, Wash., 16–21 June 2020.

Wang, Q., S. Liu, J. Chanussot and X. Li. 2018. Scene classification with recurrent attention of VHR remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 57(2):1155–1167.

Woo, S., J. Park, J.-Y. Lee and I. S. Kweon. 2018. CBAM: Convolutional block attention module. *Journal of European Computer Vision*:3–19.

Xia, G.-S., J. Hu, F. Hu, B. Shi, X. Bai, Y. Zhong, L. Zhang and X. Lu. 2017. Aid: A benchmark data set for performance evaluation of aerial scene classification. *Journal of IEEE Geoscience and Remote Sensing Letters* 55(7):3965–3981.

Yang, Y. and S. Newsam. 2012. Geographic image retrieval using local invariant features. *IEEE Transactions on Geoscience and Remote Sensing*, 51(2):818–832.

Yu, Y. and F. Liu. 2018. A two-stream deep fusion framework for high-resolution aerial scene classification. *Journal of Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2018/8639367.

Yu, Y., X. Li and F. Liu. 2020. Attention GANS: Unsupervised deep feature learning for aerial scene classification. *Journal of IEEE Transactions on Geoscience and Remote Sensing* 58(1):519–531.

Zeng, D., S. Chen, B. Chen and S. Li. 2018. Improving remote sensing scene classification by integrating global-context and local-object features. *Journal of Remote Sensing* 10(5):7–21.

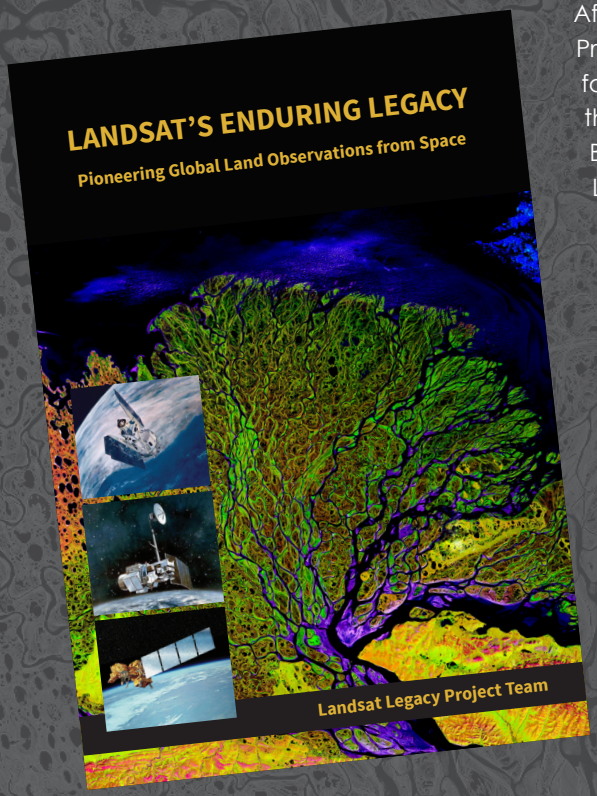
Zhang, J., M. Zhang, L. Shi, W. Yan and B. Pan. 2019. A multi-scale approach for remote sensing scene classification based on feature maps selection and region representation. *Journal of Remote Sensing* 11(21):2504.

Zheng, X., T. Gong, X. Li and X. Lu. 2021. Generalized scene classification from small-scale datasets with multitask learning. *Journal of IEEE Transactions on Geoscience and Remote Sensing* 60(1):1–11.

Zheng, X., Y. Yuan and X. Lu. 2019. A deep scene representation for aerial scene classification. *Journal of IEEE Transactions on Geoscience and Remote Sensing* 57(7):4799–4809.

# Landsat's Enduring Legacy

## Pioneering Global Land Observations from Space




After more than 15 years of research and writing, the Landsat Legacy Project Team published, in collaboration with the American Society for Photogrammetry and Remote Sensing (ASPRS), a seminal work on the nearly half-century of monitoring the Earth's lands with Landsat. Born of technologies that evolved from the Second World War, Landsat not only pioneered global land monitoring but in the process drove innovation in digital imaging technologies and encouraged development of global imagery archives. Access to this imagery led to early breakthroughs in natural resources assessments, particularly for agriculture, forestry, and geology. The technical Landsat remote sensing revolution was not simple or straightforward. Early conflicts between civilian and defense satellite remote sensing users gave way to disagreements over whether the Landsat system should be a public service or a private enterprise. The failed attempts to privatize Landsat nearly led to its demise. Only the combined engagement of civilian and defense organizations ultimately saved this pioneer satellite land monitoring program. With the emergence of 21st century Earth system science research, the full value of the Landsat concept and its continuous 45-year global archive has been recognized and embraced. Discussion of Landsat's future continues but its heritage will not be forgotten.

The pioneering satellite system's vital history is captured in this notable volume on Landsat's Enduring Legacy.

**Landsat Legacy Project Team**

Samuel N. Goward
Darrel L. Williams
Terry Arvidson
Laura E. P. Rocchio
James R. Irons
Carol A. Russell
Shaida S. Johnston

**Landsat's Enduring Legacy**

Hardback. 2017, ISBN 1-57083-101-7

| | |
|---|---|
| Member/Non-member | $48* |
| Student Member | $36* |

* Plus shipping

**Order online at**

**www.asprs.org/landsat**

asprs THE IMAGING & GEOSPATIAL INFORMATION SOCIETY