WHITTLES PUBLISHING’S STABLE OF CLASSIC GEOMATICS BOOKS INCLUDES THE THREE WINNERS OF THE PRESTIGIOUS KARL KRAUS MEDAL AWARDED BY ISPRS

- Close Range Photogrammetry
  - Thomas Luhmann
  - Bernd Hofmann
  - Stephen Rolfe
  - Ian Hay

- Unmanned Vehicle Systems for Geomatics Toward Robotic Mapping
  - Costas Armenakis
  - Petros Patias

- 3D/4D City Modelling: From Sensors to Application
  - Stefan Hinz, Andreas Brain, and Martin Wehrmann
  - ISBN: 978-184995-390-0

- High Resolution Optical Satellite Imagery
  - ISBN: 978-184995-475-4

- Univariate Statistical Analysis: Theory and Practice

- Airborne and Terrestrial Laser Scanning
  - ISBN: 978-1804445-87-6

- Stable of Classic Geomatics Books

BROWSE OUR WEBSITE TO SEE OUR FULL RANGE OF ACCLAIMED GEOMATICS BOOKS

WWW.WHITTLES PUBLISHING.COM

OUR LIST CONTINUES TO EXPAND WITH THESE NEW TITLES

WHITTLES PUBLISHING OFFERS A 20% DISCOUNT TO ASPRS MEMBERS USING CODE WPASA1PR52
ASPRS Aerial Data Catalog
“THE SOURCE FOR FINDING AERIAL COLLECTIONS”

HTTP://DPAC.ASPRS.ORG

The ASPRS Aerial Data Catalog is a tool allowing owners of aerial photography to list details and contact information about individual collections.

By providing this free and open metadata catalog with no commercial interests, the Data Preservation and Archiving Committee (DPAC) aims to provide a definitive metadata resource for all users in the geospatial community to locate previously unknown imagery.

DPAC hopes this Catalog will contribute to the protection and preservation of aerial photography around the world!

ASPRS Members: We Need Your Help!
There are three ways to get involved

1. USE
Use the catalog to browse over 5,000 entries from all 50 states and many countries. Millions of frames from as early as 1924!

2. SUPPLY
Caretakers of collections, with or without metadata, should contact DPAC to add their datasets to the catalog free of charge!

3. TELL
Spread the word about the catalog! New users and data collections are key to making this a useful tool for the community!

For More Details or To Get Involved Contact:

DAVID RUZ • druz@quantumspatial.com • 510-834-2001 OR DAVID DAVIS • dday@kasurveys.com • 215-677-3119
Teledyne Optech, a Teledyne Technologies company and global leader in advanced lidar sensors has extended its bestselling Galaxy lineup to include the CM2000, a new sensor specifically designed for corridor mapping.

With a true measuring rate of up to 2 million points per second, the Galaxy CM2000 delivers precise detail of fine corridor elements such as electric wires and conductors, distribution power poles, railway signs, cellular tower antennas, as well the ability to detect fine changes in the ground over time for pipeline monitoring.

The CM2000’s adjustable field of view, provides users with the flexibility to narrow the field of view to the exact width of their corridor and thus concentrate the laser measurements on their precise target. This combination of an adjustable field of view, and a measurement rate of 2 million points per second, will deliver data resolution that allows for advanced analytics, insights, and decision making.

The Galaxy CM2000 has the smallest laser footprint in the market allowing for complete detection of towers, transmission and distribution wires and attachment points on power poles. In addition, the CM2000 has built in roll compensation which corrects for turbulence on the aircraft, maintaining a constant swath width on the ground.

“The true advantage of the Galaxy CM2000 is its ability to improve the mapping of the infrastructures that we all rely on every day – including the electric grid we count on for power, or the roads and rail we depend on for safe travel. This is accomplished by providing an astounding level of detail via true 2 million points per second straight to the ground and a small laser footprint that allows for the modelling of complex targets like electric towers and distribution wires,” commented Malek Singer, Airborne Product Manager at Teledyne Optech.

For more information, visit https://www.teledyneoptech.com/en/home/.

Teledyne Optech, a Teledyne Technologies company and global leader in advanced lidar sensors has extended its bestselling Galaxy lineup to include the CM2000, a new sensor specifically designed for corridor mapping.

With a true measuring rate of up to 2 million points per second, the Galaxy CM2000 delivers precise detail of fine corridor elements such as electric wires and conductors, distribution power poles, railway signs, cellular tower antennas, as well the ability to detect fine changes in the ground over time for pipeline monitoring.

The CM2000’s adjustable field of view, provides users with the flexibility to narrow the field of view to the exact width of their corridor and thus concentrate the laser measurements on their precise target. This combination of an adjustable field of view, and a measurement rate of 2 million points per second, will deliver data resolution that allows for advanced analytics, insights, and decision making.

The Galaxy CM2000 has the smallest laser footprint in the market allowing for complete detection of towers, transmission and distribution wires and attachment points on power poles. In addition, the CM2000 has built in roll compensation which corrects for turbulence on the aircraft, maintaining a constant swath width on the ground.

“The true advantage of the Galaxy CM2000 is its ability to improve the mapping of the infrastructures that we all rely on every day – including the electric grid we count on for power, or the roads and rail we depend on for safe travel. This is accomplished by providing an astounding level of detail via true 2 million points per second straight to the ground and a small laser footprint that allows for the modelling of complex targets like electric towers and distribution wires,” commented Malek Singer, Airborne Product Manager at Teledyne Optech.

For more information, visit https://www.teledyneoptech.com/en/home/.

**ANNOUNCEMENTS**

The International LIDAR Mapping Forum (ILMF) and Geo Week hosted their 3rd annual 2020 Lidar Leader Awards presentation where the board recognized GeoCue Group’s True View® 410 3D Imaging System (3DIS™) as the 2020 Outstanding Innovation in Lidar winner.

Introduced to the market in June 2019, the True View 410 LIDAR/camera fusion sensor is the first drone-carried data collection system designed from the ground-up to match high-quality photogrammetric imagery with lidar point clouds. Featuring Quanergy’s M8 Ultra lidar sensor, dual GeoCue photogrammetric cameras and an Applanix Position and Orientation System (POS), the True View 410 produces high-accuracy 3D colorized lidar point clouds within minutes of landing. The sensor has a 120-degree fused field of view and can be carried by any drone capable of managing a payload of 2.25 kg.

The True View 410 system provides all the technology needed to realize the entire 3D Imaging workflow from start to finish; sensor, Applanix positioning software and GeoCue’s True View EVO comprehensive postprocessing/analytic software. Rather than spending hours waiting on image reconstruction software to create a 3D point cloud, the True View 410 yields high accuracy products such as colorized bare earth point cloud, volumetrics, elevation models and other deliverable products in short order.

The award also cited GeoCue’s novel “Hardware-as-a-Service” subscription model that allows customers to subscribe to a True View 410 3DIS for periods as short as one month. This novel business model provides customers a very low risk approach to trying drone 3D imaging technology as well as providing a very cost-effective way to deal with business fluctuations.

“We are extremely pleased to be recognized by the ILMF for this innovation award,” said Lewis Graham, GeoCue Group’s President and CTO. My engineering teams worked long hours to bring this product to market; it is very rewarding to see their efforts recognized with this honor. It was a true team effort with terrific collaborative support from the engineers at Applanix and Quanergy. I am still amazed to see the stunning, 3D colorized point clouds produced by this sensor.”

To learn more, visit www.geocue.com.

Waller, Todd & Sadler, a Woolpert Company, provided integrated design services for the renovation of Café 1201 at Old Dominion University. Located inside Webb University Center, the facility offers students easy access to food options throughout the day at a site that serves as a social hub for the university community.

Woolpert Architect and Practice Leader Stelios Xystros said the project transformed a “tired, all-you-care-to-eat dining facility into a bright and modern dining center.” The project also added a catering kitchen and a smaller residential dining concept, Ruby’s Café, named after longtime employee Ruby Milteer.

For more information, visit woolpert.com.

**CALENDAR**

- 9-13 November 2020, URISA GIS Leadership Academy, St. Petersburg, Florida. For more information, visit https://www.urisa.org/education-events/urisa-gis-leadership-academy/.


- 7-11 June 2021, URISA GIS Leadership Academy, Minneapolis, Minnesota. For more information, visit https://www.urisa.org/education-events/urisa-gis-leadership-academy/.
541 Heliport Detection Using Artificial Neural Networks
Emre Başeski

Automatic image exploitation is a critical technology for quick content analysis of high-resolution remote sensing images. The presence of a heliport on an image usually implies an important facility, such as military facilities. Therefore, detection of heliports can reveal critical information about the content of an image. In this article, two learning-based algorithms are presented that make use of artificial neural networks to detect H-shaped, light-colored heliports.

547 Semi-Automatic Building Extraction from WorldView-2 Imagery Using Taguchi Optimization
Hasan Tonbul and Taskin Kavzoglu

Due to the complex spectral and spatial structures of remotely sensed images, the delineation of land use/land cover classes using conventional approaches is a challenging task. This article tackles the problem of seeking optimal parameters of multi-resolution segmentation for a classification task using WorldView-2 imagery.

557 Precise Extraction of Citrus Fruit Trees from a Digital Surface Model Using a Unified Strategy: Detection, Delineation, and Clustering
Ali Öğün Ok and Aslı Özdarıcı-Ok

In this study, we present an original unified strategy for the precise extraction of individual citrus fruit trees from single digital surface model (DSM) input data. A probabilistic method combining the circular shape information with the knowledge of the local maxima in the DSM has been used for the detection of the candidate trees.

571 Performance Analysis of Advanced Decision Forest Algorithms in Hyperspectral Image Classification
Ismail Çolakçı and Omer Hatib Ertekin

In this study, the performances of random forest (RF), rotation forest (RoF), and canonical correlation forest (CCF) algorithms were compared and analyzed for classification of hyperspectral imagery.

581 Analyzing the Contribution of Training Algorithms on Deep Neural Networks for Hyperspectral Image Classification
Mehmet Akif Günen, Umit Haluk Atasever, and Erkan Beytdok

Autoencoder (AE)-based deep neural networks learn complex problems by generating feature-space conjugates of input data. The learning success of an AE is too sensitive for a training algorithm. The problem of hyperspectral image (HSI) classification by using spectral features of pixels is a highly complex problem due to its multi-dimensional and excessive data nature. In this paper, the contribution of three gradient-based training algorithms (i.e., scaled conjugate gradient (SCG), gradient descent (GD), and resilient backpropagation algorithms (RP)) on the solution of the HSI classification problem by using AE was analyzed.

527 GIS Tips & Tricks—Need to Generate Multiple Profiles Quickly; Another Approach!
By Al Karlin, Ph.D, CMS-L, GISP

535 In Memoriam—Charles E. “Chuck” Olson
By Raechel Portelli, Colin Brooks, Nancy French, Kathleen Bergen, Laura Bourgeau-Chavez, and Marguerite Madden
Jezero is a 45-kilometer (28-mile) wide ancient impact crater located in the north-west corner of a larger impact basin on Mars—essentially an impact crater within an impact crater. It is noteworthy because it once contained a lake, as evidenced by delta deposits. Previously, scientists discovered carbonate minerals throughout the crater. Using data taken by NASA’s Mars Reconnaissance Orbiter (MRO), Horgan and her team recently discovered evidence that some of these carbonate minerals may have formed in the lake.

“Carbonates are important because they are really good at trapping anything that existed within that environment, such as microbes, organics, or certain textures that provide evidence of past microbial life,” said Brad Garzynski, a graduate student at Purdue who works with Horgan. “But before we go to Jezero, it is really important to gain context on how these carbonates form on Earth in order to focus our search for signs for life.”

It just so happens that Lake Salda is the only known lake on Earth that contains the carbonates and depositional features (deltas) similar to those found at Jezero Crater. The black and white inset image shows Jezero Crater as observed by MRO’s Context Camera. Spectral data showed signatures of carbonates on the western edge of the crater, which scientists believe to be the shoreline and beaches of an ancient lake. The carbonates are also present in the delta, which is the planned site of the Perseverance landing.

The background image shows Lake Salda on June 8, 2020, as observed by the Operational Land Imager (OLI) on Landsat 8. The lake contains alluvial fans full of rock deposits eroded and washed down from the surrounding bedrock (similar to the delta in Jezero). By studying how material is deposited in Lake Salda, the team can learn more about the various depositional processes at Lake Jezero.

The white shoreline around Lake Salda is comprised of sands and gravels that are dominated by hydromagnesite, which is similar to the carbonate minerals detected at Jezero. Horgan explained that the hydromagnesite sediments along Lake Salda’s shoreline are thought to have eroded from large mounds called “microbialites”—rocks formed with the help of microbes. In Lake Salda, they formed from microbial mats that lived just beneath the surface of the water near the shoreline. As the microbialites grew, they incorporated carbonate materials and created large terrace islands. Visit https://landsat.visibleearth.nasa.gov/view.php?id=147041 to see an image of the terrace island in Lake Salda.

PHOTOGRAHMETRIC ENGINEERING & REMOTE SENSING

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, and Remote Sensing Applications. Additional information on the functioning of the Technical Divisions and the Society can be found in the Yearbook issue of PE&RS.

Correspondence relating to all business and editorial matters pertaining to this and other Society publications should be directed to the American Society for Photogrammetry and Remote Sensing, 425 Barlow Place, Suite 210, Bethesda, Maryland 20814-2144, including inquiries, memberships, subscriptions, changes in address, manuscripts for publication, advertising, back issues, and publications. The telephone number of the Society Headquarters is 301-495-0296; the fax number is 225-408-4422; web address is www.asprs.org.

PE&RS. PEARS (ISSN0099-1112) is published monthly by the American Society for Photogrammetry and Remote Sensing, 425 Barlow Place, Suite 210, Bethesda, Maryland 20814-2144. Periodicals postage paid at Bethesda, Maryland and at additional mailing offices.

SUBSCRIPTION. For the 2020 subscription year, ASPRS is offering two options to our PE&RS subscribers—a e-Subscription and the print edition. E-subscribers can plus-up their subscriptions with printed copies for a small additional charge. Print and Electronic subscriptions are on a calendar-year basis that runs from January through December. We recommend that customers who choose both e-Subscription and print (e-Subscription + Print) renew on a calendar-year basis.

The rate of the e-Subscription (digital) Site License Only for USA and Non-USA is $100.00 USD. e-Subscription (digital) Site License Only for Canada* is $109.00 USD. Site License Only for Canada* is $109.00 USD. e-Subscription (digital) Plus Print for USA is $1365.00 USD. e-Subscription (digital) Plus Print for Canada* is $1424.00 USD.

The rate of the e-Subscription (digital) Site License Only for USA and Non-USA is $1395.00 USD. Printed-Subscription Only for USA is $1065.00 USD. Printed-Subscription Only for Canada* is $1124.00 USD. Printed-Subscription Only for Canada* is $1124.00 USD.

*Note: e-Subscription/ Printed-Subscription Ony-Subscription Plus Print for Canada include 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.

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. Or you can receive the print copy of PE&RS, which is available to all members types for an additional fee of $60.00 USD for shipping USA, $65.00 USD for Canada, or $75.00 USD for international shipping. Dues for ASPRS Members outside of the U.S. will now be the same as for members residing in the U.S. Annual dues for Regular members (Active Member) is $150.00 USD; for Student members $50.00 USD for USA and Canada $64.00 USD, $60.00 USD for Other Foreign members. A tax of 5% for Canada’s Goods and Services Tax (GST #135123065) is applied to all members residing in Canada.

COPYRIGHT 2020. Copyright by the American Society for Photogrammetry and Remote Sensing. Reproduction of this issue or any part thereof (except short quotations for use in preparing technical and scientific papers) may be made only after obtaining the specific approval of the Managing Editor. The Society is not responsible for any statements made or opinions expressed in technical papers, advertisements, or other portions of this publication. Printed in the United States of America.

PERMISSION TO PHOTOCOPY. The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner’s consent that copies of the article may be made for personal or internal use or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per copy fee of 3.00 USD through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, Massachusetts 01923, for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale.

Be a part of ASPRS Social Media:
facebook.com/ASPRS.org
linkedin.com/groups/274512/profile
twitter.com/ASPRSorg
youtube.com/user/ASPRS

SEPTEMBER 2020
Integration of Varied Spatial Resolution Data

Remote sensing has and continues to change extremely rapidly. Those changes have included varied platforms, new sensors, and improved data processing methods. With the current availability of many imagery types, this column urges renewed educational emphasis on the integration of varied spatial resolution data for improved spatial analysis.

The changes of platforms and sensors is evident in the resolutions of imagery available to scientists and decision makers. Spectral resolution has expanded by the acquisition of hyperspectral imagery with some sensors providing 512 bands. Temporal resolution has always been frequent with meteorological sensors but is now high with fine spatial resolution systems. The small-sat constellation of the commercial company Planet is able to acquire global imagery daily with very fine spatial resolution. Radiometric resolution has increased from the 6 and 7 bits of early Landsats to 12 and 16 bit imagery today.

One of the major changes in resolution has been the availability of fine spatial resolution imagery from satellites. Initially the community thought that this imagery would replace aerial photography but the high initial image costs did not make that viable. However, there is now fine spatial resolution satellite imagery available at little or no cost.

However, as with aerial photography, fine spatial resolution satellite imagery has generally small footprints. The imagery is often 20 km or less per side making it very difficult to acquire and accurately extract information over large areas. Most watersheds and local governmental administrative units would require hundreds of images.

Analysis of medium and coarse spatial resolution satellite imagery (10 m pixels and greater) is an effective way to assess regional, continental or global phenomena because of its synoptic coverage, frequency of image acquisition, large footprint and often inexpensive cost. The disadvantage of medium spatial resolution imagery is the lack of detail. Biases often occur because of the large pixel size of the data. Less common surface features are often underestimated, causing a negative bias, and more common components can be overestimated, causing a positive bias.

The purpose of this column is to establish interest among remote sensing educators in the integration of imagery at different spatial resolutions to provide improved spatial statistics for multiple applications. Many remote sensing scientists believe that the most important information that they can provide are accurate maps. However, the reality often is that many decisions are made based upon statistical analysis such as the extent and rate of wetland loss, deforestation or urban expansion.

Regression Estimation

Regression estimation is a statistical sampling technique to combine the synoptic coverage of a coarse spatial resolution sensor with the improved detail of a finer spatial resolution sensor (Gallego, 2004). This technique requires the analyst to map phenomena first using coarse spatial resolution imagery. Next, fine spatial resolution images are acquired for samples within the study area. Phenomena are likewise mapped with the fine spatial resolution data. A regression analysis or other statistical procedure is then performed to determine a correlation or relationship between the two sets of data. If a good correlation exists, the more accurate, finer spatially detailed imagery can be used to calibrate the coarse spatial imagery using a correction factor (Nelson, 1989). In this manner, regression estimation provides more accurate statistical information than only using coarse imagery.

This column was prompted in part based upon a review of 12 textbooks in remote sensing to determine if they referenced regression estimation or similar approaches for surface inventories. Interestingly, only two had any reference to the method and they were quite dated, thus the concern that the procedure is not commonly included in curriculums.

Applications

There are numerous examples of regression estimation. They generally employ imagery of different spatial resolutions and footprints but the method can also use field data rather than fine spatial resolution imagery. It is also possible to
have several stages of regression estimation using different spatial resolution sensors (Koeln and Kollasch, 2000). For example, a product at 1 m could be used to adjust 5 m SPOT imagery as a first stage. The corrected SPOT could then be used to adjust 30 m Landsat, and it in turn to correct MODIS 250 m imagery. This is often referred to as nested sampling or Nested Area Frame Sampling and is a module in at least one of the standard image processing systems.

One study using regression estimation was to locate open water in North Dakota suitable for waterfowl. The original Landsat estimates of the area of ponds were only about 70% of the actual open water compared to validation data. The regression estimates using aerial photography samples increased those estimates to within 8% of the actual extent.

Regression estimation was performed over seven land covers with an emphasis on forests in Brazil using aerial photographs and Landsat images. The regression estimation procedure was able to provide accurate results in a time-efficient and cost-effective manner with better results than from Landsat independently. Landsat imagery (30 m) in conjunction with AVHRR (1 km) images accurately estimated Canadian burned forest areas to calculate carbon storage (Fraser, 2004).

A study to ascertain the most accurate method to monitor biomass burning in Central Africa determined the best approach would utilize data from both fine and coarse spatial resolution sensors and a regression estimator strategy. Improved walrus counts were obtained using this strategy for a rookery in the North Pacific Ocean (Barber et al., 1991).

Regression estimation has also been employed for mapping agriculture. One study used field data with Landsat to accurately determine the amount of winter rice area in Bangladesh (Haack and Rafter, 2010). Figure 1 illustrates the relationship between the two data sets. In Tanzania, crop statistics were accurately determined via this approach and similarly aerial photography and Landsat correctly estimated the amount of wheat in a region in Brazil.

**Summary**

There is a record of successful applications of the regression estimation strategy to improve the spatial statistics of a variety of surface features. Unfortunately, this method does not seem to be widely used, understood or even included in current remote sensing textbooks. Given the increased availability of fine spatial resolution satellite-based remote sensing data at minimal or no cost, this column encourages wider introduction of this very effective technique in university curriculum.

---

**References**


**Author**

Dr. Barry Haack is a Professor of Geographic and Cartographic Sciences at George Mason University in Fairfax, Virginia, USA and is an ASPRS Fellow.
Need to Generate Multiple Profiles Quickly; Another Approach!

One of the recurring themes in this column is that when using GIS software, there are always multiple ways to accomplish the same task. Last month, I demonstrated a tip using the GlobalMapper Path Profile tools to create multiple perpendicular profiles along a user-defined path. [Quick tip: When using GlobalMapper, if you have a user defined line or line shapefile already in the project, the sequence of, (1) selecting that line, right-clicking on it, (2) choosing “Analysis/Measurement” and then the (3) PATH PROFILE option, will invoke the same menu(s) to construct perpendicular profiles as demonstrated last month.]

The GlobalMapper Path Profile tool works against rasters (DEMs, DTMs, etc.) and lidar data, but it requires a GlobalMapper Lidar License to work against a lidar point cloud. As an alternative to the GlobalMapper Lidar License, there are similar tools, deeply hidden in “LP360 for ArcGIS”, the “Export Profile by Selected Features” and “Export Profile by Selected Graphic” tools to accomplish these tasks.

The tools are included with all levels of LP360 for ArcGIS but are not found on any of the default toolbars. To access the Export Profile tools, an ArcGIS-user with a LP360 for ArcGIS license activated can add the tools to any toolbar through the ArcGIS user interface by:

1. Double-clicking anywhere on the gray-space of the UI (or Customize | Customize Mode… from the Standard Toolbar) to open the Customize dialog,
2. Choose the Commands Tab and scroll down to LP360 (left window)
3. Scroll to the “Export Profiles by Selected Features” (or “Export Profiles by Selected Graphic”) in the right window, and
4. Drag the selected tool onto any available toolbar on the UI, Figure 1.
5. Dismiss the Customize dialog by pressing “Close”

Because you can only drag one tool at a time, if you want to use both tools, you will need to drag them independently. [TIP: You can change the icon for the tool. When you drag the default icon to the toolbar before you dismiss the Customize dialog, right-click and select the “Change Button Image” option. A menu of optional button images will appear. Clicking on one will change the tool’s icon.] You will also notice that the tools are inactive (grayed out) until you have a point cloud loaded (through LP360 for ArcGIS) AND line feature(s) (or line graphics) selected.

With a LP360 LAS_Layer loaded and a 2D- or 3D- line feature (shapefile or feature class) selected the “Export Profile by Selected Feature” tool becomes active, as in Figure 2.
In Figure 2, I placed the tools on the LP360 Main Toolbar and with a line feature selected (cyan line using the Esri “selection” tool), the “Export Profile by Selected Feature” is active (red circled tool).

Clicking on the “Export Profile by Selected Feature” tool, invokes the Profile Export Settings dialog for you to specify a shapefile to output the 3D profiles into, the Profile Length (in map units) and the Step Amount (distance between profiles in map units), along with a CheckBox that provides the option of exporting the elevation along the profiles only at the beginning, end and where it crosses the selected feature.) The default is to construct a 3D vertex at each place the profile crosses the internally-constructed TIN edge. The tool will produce a 3D-line shapefile; each record representing a profile with the vertices encoded in the line.

![Profile Export Settings](image)

For this example, I designated an output file, Figure 3, with profiles to be 75.00m in length (37.5m to either side of the selected line), and spaced every 50.00m along the selected line. As with all ArcGIS tools, if multiple objects are selected, the tool will process each object. Pressing “OK” will start the tool processing.

When the processing is complete, LP360 for ArcGIS will report to the screen, Figure 4, and clicking on “Yes” will add the new shapefile to your map. You may need to adjust the symbology, line thickness and color to suit your map preferences. I increased the line thickness and made the line color red in the examples in Figure 5.

In Figure 6, opening the Attribute Table (left table) for the output shapefile will show an individual 3D record (Polyline ZM) for each profile generated, and upon opening an Esri editing session, selecting an individual profile line and choosing the “Edit Sketch Properties” tool in ArcGIS, will reveal the selected profile's elevations (right table). The selected profile line (highlighted in the left table) is located approximately midway between the tables on the map. The Edit Sketch Properties table (on right) shows the X, Y and Z (elevation) values for the vertices along the selected profile.

Of course, the profiles are also viewable in the LP360 3D window as 3D features in the point cloud in Figure 7.

Finally, just a few last tips:

1. If you want ground (ASPRS Class 2) profiles (the usual case), make sure that you apply a filter to your point cloud. The Export Profile tools respect applied filters, so if no filters are applied, the elevations along the profile will be extracted from the unfiltered (entire) point cloud, so elevations in trees and above the surface will be included,
Borrowing inspiration from a well-known agency motto, a better title for Dr. Dwivedi’s book might have possibly been “Remote Sensing for a Sustainable Land.” At a time in which concern about monitoring the nature, intensity, extent, and rates of land and soil changes reaches fever pitches, this book brings to the reader, a fusion of foundational knowledge of remote sensing and the opportunities this science and allied technologies present for successful application to the problem.

This book is a timely effort in the sense that it resonates intimately with the Sustainable Development Goals (SDGs) just fleshed out by the United Nations for their 2030 Agenda (https://sustainabledevelopment.un.org/topics/desertificationlanddegradationanddrought). In them desertification, drought, and land degradation processes are made an explicit target for remediation, prevention, and conservation efforts.

The unspoken but compelling premise of the book is that planetary health is synonymous with land health, which is in turn inextricably linked to soil health, even though these types of “health” are hard to measure and quantify. Thus, the terms “land” and “soil” are used interchangeably in the context of this book, both serving as substrate to a good portion of life on Earth. Dr. Dwivedi sets out to explore examples of the ways in which Earth-observing, remote sensing science and technologies can be applied to the problems of monitoring, detecting, modeling, mapping, and remediating negative changes in land and soils.

While the slant towards soil science and technology will satisfy readers with a background related to this specific domain, the book provides a bridge to a geospatial community by providing specific examples of sensor data exploitation and utilization. The perspective of both domains is thus enriched with a broader view and greater appeal by showcasing a diversity of remote sensing methods used to measure, monitor, and assess numerical indicators of land and soil health, for management purposes.

To illustrate the methodologies employed, the author chose for monitoring a sample of negative land and soil changes of interest, such as erosion, acidification, alkalinization and salinization. Desertification, vegetation degradation, and land loss to urbanization and industrial use are mentioned in the section on fundamentals of land degradation (Chapter 5). It would have been useful if the author had included examples of the effects of invasive species, and of the effects on shoreline changes due to sea/ocean rise, and due to the increasing frequency of violent storms.

Not unlike the structure followed by Dr. Dwivedi’s previous book (Remote Sensing of Soils, 2017. Springer-Verlag, GmbH, Germany), the first four chapters take up 148 pages or 40% of the book and are devoted to providing a solid exposition of the fundamentals of remote sensing (i.e. Chapter 1—An Introduction to Geospatial Technology; Chapter 2—Passive Remote Sensing; Chapter 3—Active Remote Sensing; Chapter 4—Digital Image Processing). Chapters 5 through 10 cover the remaining 230 pages and they guide the reader through a sampler of land degradation processes (i.e. Chapter 5—An Introduction to Land Degradation; Chapter 6—Water Erosion; Chapter 7—Wind Erosion; Chapter 8—Soil Salinization and Alkalinization; Chapter 9—Soil Acidification; Chapter 10—Waterlogging; Chapter 11—Land Degradation due to Mining, Aquaculture, and Shifting Cultivation; Chapter 12—Drought). The effects of land use (i.e. mining, aquaculture and shifting cultivation) on land health are exemplified in Chapter 11.

The role played by global databases (i.e. big data) and the
potentials for data mining using artificial intelligence, and machine and deep learning is addressed in Chapter 13—Land Degradation Information Systems.

Each chapter has a section on the basics of the land degradation process and an ulterior section on the role of remote- and sometimes proximal sensing, and on data exploitation workflows to generate numerical inputs ultimately translated into physical and chemical indicators - for use with descriptive and predictive biophysical-mathematical models specific to each process.

The book is well-written, with copious references offered at the end of each chapter. Rarely, a typo is found (e.g. Figure 11.6). For the benefit of readers, the term “annexure” should have been replaced with the more widely used term “annex”. Surprisingly, some duplicated paragraphs occur (e.g. § 8.7.2.1.1.8 “Error Assessment” and § 9.5.2.1.7 “Accuracy Estimation”). The same information appears on Chapter 4 in Section 4.8. Some map figures appear distorted on Chapter 12.

This work provides a clear exposition on land and soil degradation processes occurring in select geographies for which remotely sensed data are ultimately used as inputs to numerical and stochastic models for mapping, describing and predicting the current extent and severity and the response to remediation, mitigation and conservation efforts.

This valuable book serves its best purpose when used as a primer in remote sensing for specialists in land and soil sciences, as it emphasizes how sensor data are generated, what types of data are generated and how measurements and information are derived for use in their domain. If the book were to be applied to the opposite crowd, it would introduce the Earth-observing remote sensing specialist to specific applications of his/her craft to the perennial but very timely and momentous topics of land and soil conservation and remediation.

2. Although the generated profile shapefile is in the coordinate system of the ArcGIS Map document, the exported shapefile will not have an associated .PRJ file. It is a good idea to use ArcGIS to add the projection file, and

3. Drawing a line graphic with the ArcGIS line graphic tool, will activate the “Export Profile by Selected Graphic” tool. Clicking on the “Export Profile by Selected Graphic” tool with the line graphic selected will invoke the same dialogs as the “Export Profile by Selected Feature” tool and perform the same tasks.

And there you have it.

Al Karlin, Ph.D., CMS-L, GISP is with Dewberry’s geospatial and technical services group in Tampa, Florida. As a senior GIS Geospatial Scientist, he works with all aspects of lidar, remote sensing, photogrammetry, and GIS-related projects.
pronounced “kir-ih-bahss,” the islands were originally settled by Austronesians thousands of years ago. Around the 14th century AD, the islands were invaded by Fijians and Tongans. The first recorded European encounter with Kiribati was by the Spanish explorer Quiros in 1606. By 1820, all of the islands had been charted. At that time, the Russian hydrographer A. I. Krusenstern gave the group the name Gilbert Islands. Until 1870, many British and American whaling vessels sought sperm whales in Gilbertese waters. Starting in 1850, trading vessels passed through seeking at first coconut oil and later copra. In the 1860s, slave ships known as “blackbirders” carried off islanders to work in plantations in Peru, Fiji, Tahiti, Hawaii, and Australia. The male population was decimated, and European disease (measles) took a further large toll on the “I’Kiribati” people. The Ellice group (now Tuvalu – PE&RS, December 2001) and the Gilbert Islands became a British Protectorate in 1892. In 1975, the Ellice Islands seceded from the colony and became the independent nation of Tuvalu. On 12 July 1979, Kiribati obtained its own independence from the United Kingdom and became a republic within the Commonwealth.

Kiribati consists of a three-island group: the Gilbert Islands, the Line Islands, and the Phoenix Islands. These groups of islands straddle the equator and the International Date Line, about half way between Hawaii and Australia. In 1995, The Republic of Kiribati proclaimed that all of its territory lies in the same time zone as the Gilbert Islands group (GMT + 12). The total land area equals 717 sq km, about four times the size of Washington, D.C. The country is composed of mostly low-lying coral atolls surrounded by extensive reefs.

The lowest point is the Pacific Ocean, the highest point is on Banaba Island (81 m). Twenty of the 33 islands are inhabited, and the capital is Tarawa.

According to Russell Fox of the Ordnance Survey, “We revised our Tarawa 1:50,000 map in 1997 and produced new 1:25,000 photomapping of the Line and Phoenix Islands in 1995/96 using Australian Army aerial photography, but we did not do any geodetic survey work in that decade. I believe that a GPS survey, aerial photography and new mapping (on WGS84) project was planned for Tarawa for 1998/99 under Australian aid. The following notes summarise what we know about Kiribati.

**Background**

The Republic of Kiribati (pronounced “kiribass”) comprises the Gilbert, Phoenix and Line island groups in the central Pacific Ocean. Total land area 717 sq km, total sea area 5.2 million sq km, greatest extent 4000 km W-E by 2000 km N-S. The modern republic has its genesis in a British protectorate proclaimed in 1892, which became the Gilbert and Eli-
ce Islands Colony in 1916. Independence was granted to the Gilbert, Phoenix and Line Islands in 1979 as the Republic of Kiribati. The Ellice Islands had previously seceded, in 1975, as the Republic of Tuvalu. The International Date Line passes through Kiribati, but the Government of Kiribati has (legitimately) legislated that the Line and Phoenix Islands will observe the same date as the Gilbert Islands.

**Survey History**

The Royal Navy carried out hydrographic surveys, based on astro fixes, from the early nineteenth century onward. Little survey and mapping work was done by the British colonial authorities before WW2. Japanese military occupation of the Gilbert Islands in 1941 lead to extensive aerial photography and mapping by the US armed forces, who drove out the enemy in fierce fighting. Post-WW2: Aerial photography: Gilbert Islands RNZAF 1962/63; USN 1964; Fiji Lands, Mines and Survey Department 1968/69. The 1968/69 cover was used subsequently by the Directorate of Overseas Surveys (DOS) for mapping. Line Islands RAF 1950s. Phoenix Islands RNZAF 1962/63. All Kiribati, 1984/85, by the Royal Australian Survey Corps, as part of Operation Anon, one of the Australian Army’s Pacific Doppler campaigns.

**Survey**

DOS control surveys 1967-73, by Tellurometer traversing, covered all the islands in the Gilbert group, with each atoll on its own astro datum and local Transverse Mercator grid. International Spheroid (DOS’s default spheroid for the Pacific) was used. Christmas Island was surveyed by the RNZN, British Military Survey and contract staff between 1941 and the 1960s, and by the Kiribati Survey Department in 1979/80. DOS then brought those surveys together, computing on Christmas Island 1967 Astro Datum, International Spheroid, Kirimiti Local [TM] Grid. The first modern survey work in the Phoenix and Line Islands was the precise Doppler campaign of 1984/85. Mapping: DOS produced 1:25,000 (a few sheets were done at 1:12,500 & 1:10,000), photomap series of all the islands between 1972 and 1996. The larger atolls, Tarawa and Kirimiti (Christmas Island), were also mapped at 1:50 000 on single sheets. Generally two editions were produced, one showing the local TM grid, and the other (for military use) UTM grid. Early editions of the Christmas Island sheet showed UTM Grid only. Southern Tarawa was mapped at 1:2500 and 1:1250.” All of these classical Datums are referenced to the International 1924 ellipsoid where a = 6,378,388 km, and 1/f = 297.

Starting with specifics for the Gilbert Islands Group, for the Abaiaang Datum of 1962, the Datum origin is at the first-order station Flagstaff of Government Station, also called HMS Cook Astro “H” where \( \Phi_0 = 01° 49’ 25.029˝ N \) and \( \Lambda_0 = 173° 01’ 25.830˝ \) East of Greenwich. According to a British Admiralty Report of Survey file, “The Government flagstaff on a coral rock plinth in the centre of the Government Station – Taburao.” The local grid is based on the Transverse Mercator projection with the central meridian, \( \lambda_0 = 172° 55’ E \), the False Easting = 20 km, and the False Northing = zero. The Scale Factor at Origin is unity (\( m_0 = 1.0 \)). From Abaiaang Datum of 1962 to WGS 84 Datum, the converted Doppler solution is \( \Delta X = +254.8 \), \( Y = –322.4 \), and \( \Delta Z = –270.0 \). The Operation ANON 1984-85 solution is \( \Delta X = +254.3 \), \( \Delta Y = –323.4 \), and \( \Delta Z = –275.6 \). For the Abemama Datum of 1944, the Datum origin is at Signal Station, southwestern tip of Steve Island (Station 'Flag'), where \( \Phi_0 = 00° 27’ 36˝ N \) and \( \Lambda_0 = 173° 49’ 11” \) East of Greenwich. For the Abemama Datum of 1959, the Datum origin is at Cook Astro Point where \( \Phi_0 = 00° 24’ 19.02” N \), \( \Lambda_0 = 173° 55’ 36.57” \) East of Greenwich, and \( H_0 = 2.14 \). The local grid is based on the Transverse Mercator projection with the central meridian, \( \lambda_0 = 173° 51’ E \), the False Easting = 20 km, and the False Northing = 100 km. The Scale Factor at Origin is unity (\( m_0 = 1.0 \)). From Abemama Datum of 1959 to WGS 84 Datum, the converted Doppler solution is \( \Delta X = +289.4 \), \( \Delta Y = +656.2 \), and \( \Delta Z = +303.4 \). Note that the UTM coordinates of local traverses depend on the position of Observation Spot that was connected by traverse to C where \( \phi_0 = 00° 24’ 29.02” N \) and \( \lambda_0 = 173° 55’ 36.57” E \).

For the Arorae Datum of 1965, the Datum origin is at Arorae Astro Observation Spot where \( \Phi_0 = 02° 38’ 36.7” S \) and \( \Lambda_0 = 176° 49’ 33.3” \) East of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, \( \lambda_0 = 176° 49’ E \), the False Easting = 10 km, and the False Northing = 500 km. The Scale Factor at Origin is unity (\( m_0 = 1.0 \)). From Arorae Datum of 1965 to WGS 84 Datum, the converted Doppler solution is \( \Delta X = +221.4 \), \( \Delta Y = –34.4 \), and \( \Delta Z = –21.6 \).

For the Beru Datum of 1970, the Datum origin is at third-order station BRZ 19 where \( \Phi_0 = 01° 19’ 29.9632” S, \Lambda_0 = 157° 59’ 16.9134” \) East of Greenwich, and \( H_0 = 1.73 \). The local grid is based on the Transverse Mercator projection with the central meridian, \( \lambda_0 = 157° 59’ E \), the False Easting = 10 km, and the False Northing = 300 km. The Scale Factor at Origin is unity (\( m_0 = 1.0 \)). From Beru Datum of 1970 to WGS 84 Datum, the converted Doppler solution is \( \Delta X = +179.9 \), \( \Delta Y = –595.3 \), and \( \Delta Z = +6.96 \). The Operation ANON 1984-85 solution is \( \Delta X = +181.3 \), \( \Delta Y = –585.6 \), and \( \Delta Z = –7.2 \). For the Butaritari Datum of 1965, the Datum origin is at third-order station BTZ 26 where \( \Phi_0 = 03° 15’ 40.629” N, \Lambda_0 = 172° 41’ 45.8381” \) East of Greenwich, and \( H_0 = 1.87 \). The local grid is based on the Transverse Mercator projection with the central meridian, \( \lambda_0 = 172° 50’ E \), the False Easting = 20 km, and the False Northing = zero. The Scale Factor at Origin is unity (\( m_0 = 1.0 \)). From Butaritari Datum of 1965 to WGS 84 Datum, the converted Doppler solution is \( \Delta X = +253.8 \), \( \Delta Y = +6.1 \), and \( \Delta Z = +528.2 \). The Operation ANON 1984-85 solution is: \( \Delta X = +254.2 \), \( \Delta Y = +3.2 \), and \( \Delta Z = +544.2 \).

For the Kuria Datum of 1962, the Datum origin is at HMS Cook Astro, Kuria 1962 where \( \Phi_0 = 00° 13’ 00.4” N \) and \( \Lambda_0 = 173° 23’ 06.8” \) East of Greenwich. The local Kuria and Arana-
ka Grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 173^\circ 30'\ E$, the False Easting = 30 km, and the False Northing = 100 km. The Scale Factor at Origin is unity ($m_o = 1.0$). From Kuria Datum of 1962 to WGS 84 Datum, the converted Doppler solution is based on a fourth parameter where a Z-axis rotation is performed first (algebraically added to the longitude), where $R_z = 102.84765'$ and then the standard three-parameter shift is performed where $\Delta X = +219.1 \ m$, $\Delta Y = -24.9 \ m$, and $\Delta Z = +137.0 \ m$. The Operation ANON 1984-85 solution is $R_z = 102.804'$, $\Delta X = +218.6 \ m$, $\Delta Y = -24.8 \ m$, and $\Delta Z = +140.1 \ m$.

For the Little Makin Datum of 1972, the Datum origin is at station Bikati Astro where $\Phi_0 = 03^\circ 16'\ 19.90''\ N$ and $\lambda_o = 172^\circ 40'\ 36.21''\ E$ of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 172^\circ 50'\ E$, the False Easting = 20 km, and the False Northing = zero. The Scale Factor at Origin is unity ($m_o = 1.0$). From Little Makin Datum of 1972 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +243.4 \ m$, $\Delta Y = +221.1 \ m$, and $\Delta Z = -104.1 \ m$. The Operation ANON 1984-85 solution is $\Delta X = +239.5 \ m$, $\Delta Y = +189.9 \ m$, and $\Delta Z = -121.6 \ m$.

For the Maiana Datum of 1965, the Datum origin is at Maiana Astro 1965. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 173^\circ 02'\ E$, the False Easting = 20 km, and the False Northing = zero. The Scale Factor at Origin is unity ($m_o = 1.0$). From Maiana Datum of 1965 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +215.5 \ m$, $\Delta Y = -27.9 \ m$, and $\Delta Z = -159.7 \ m$.

For the Marakei Datum of 1969, the Datum origin is at SW Point A where $\Phi_0 = 01^\circ 55'\ 58''\ S$ and $\lambda_o = 173^\circ 15'\ 22''\ E$ of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 173^\circ 16'\ E$, the False Easting = 10 km, and the False Northing = zero. The Scale Factor at Origin is unity ($m_o = 1.0$). From Marakei Datum of 1969 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +188.1 \ m$, $\Delta Y = -237.6 \ m$, and $\Delta Z = -185.6 \ m$. The Operation ANON 1984-85 solution is $\Delta X = +188.7 \ m$, $\Delta Y = -229.9 \ m$, and $\Delta Z = -189.3 \ m$.

For the Nikunau Datum of 1965, the Datum origin is at third-order station NKZ 7 where $\Phi_0 = 01^\circ 23'\ 28.9196''\ S$ and $\lambda_o = 172^\circ 28'\ 46.4327''\ E$ of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 172^\circ 27'\ E$, the False Easting = 10 km, and the False Northing = 300 km. The Scale Factor at Origin is unity ($m_o = 1.0$). From Nikunau Datum of 1965 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +229.8 \ m$, $\Delta Y = -38.4 \ m$, and $\Delta Z = -311.7 \ m$. The Operation ANON 1984-85 solution is $\Delta X = +230.3 \ m$, $\Delta Y = -17.7 \ m$, and $\Delta Z = -315.3 \ m$.

For the Nikunau Datum of 1959, the Datum origin is at Government Flagstaff where $\Phi_0 = 01^\circ 20'\ 44.37''\ S$ and $\lambda_o = 176^\circ 26'\ 31.36''\ E$ of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 176^\circ 27'\ E$, the False Easting = 10 km, and the False Northing = 300 km. The Scale Factor at Origin is unity ($m_o = 1.0$). From Nikunau Datum of 1959 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +221.9 \ m$, $\Delta Y = -99.5 \ m$, and $\Delta Z = -926.8 \ m$. The Operation ANON 1984-85 solution is $\Delta X = +221.2 \ m$, $\Delta Y = -96.3 \ m$, and $\Delta Z = -835.1 \ m$. Note that the UTM coordinates of local traverses in 1963 of the Nououiti Survey depend on the position A being $\phi_o = 00^\circ 40'\ 16.4''\ S$ and $\lambda_o = 174^\circ 27'\ 28''\ E$.

For the Onotoa Datum of 1970, the Datum origin is at first-order station ONZ 7. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 175^\circ 33'\ E$, the False Easting = 10 km, and the False Northing = 300 km. The Scale Factor at Origin is unity ($m_o = 1.0$). From Onotoa Datum of 1970 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +244.4 \ m$, $\Delta Y = +197.9 \ m$, and $\Delta Z = -243.1 \ m$. The Operation ANON 1984-85 solution is $\Delta X = +245.2 \ m$, $\Delta Y = +203.7 \ m$, and $\Delta Z = -248.4 \ m$.

For the Tabiteuea Datum of 1959, the Datum origin is at third-order station TBZ 1 Astro where $\Phi_0 = 01^\circ 28'\ 05.6''\ S$ and $\lambda_o = 175^\circ 03'\ 15.0''\ E$ of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 175^\circ 33'\ E$, the False Easting = 30 km, and the False Northing = 300 km. The Scale Factor at Origin is unity ($m_o = 1.0$). From Tabiteuea Datum of 1959 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +250.8 \ m$, $\Delta Y = +235.6 \ m$, and $\Delta Z = -377.2 \ m$.

For the Tamana Datum of 1962, the Datum origin is at Astro Observation Spot where $\Phi_0 = 02^\circ 30'\ 09.0''\ S$ and $\lambda_o = 175^\circ 58'\ 45.8''\ E$ of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 174^\circ 53'\ E$, the False Easting = 30 km, and the False Northing = 400 km. The Scale Factor at Origin is unity ($m_o = 1.0$). From Tamana Datum of 1962 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +202.3 \ m$, $\Delta Y = -181.6 \ m$, and $\Delta Z = +128.2 \ m$. The Operation ANON 1984-85 solution is $\Delta X = +206.8 \ m$, $\Delta Y = -159.8 \ m$, and $\Delta Z = +95.7 \ m$.

For the Tarawa Datum of 1966, the Datum origin is at first-order station Tarawa SECOR AMS 1966 where $\Phi_0 = 01^\circ 21'\ 42.13''\ N$ and $\lambda_o = 172^\circ 55'\ 47.27''\ E$ of Greenwich. The local grid is based on the Transverse Mercator projection with the central meridian, $\lambda_o = 173^\circ 02'\ E$, the False Easting = 20 km, and the False Northing = zero. The Scale Factor at Origin is unity ($m_o = 1.0$). From Tarawa Datum of 1966 to WGS 84 Datum, the converted Doppler solution is $\Delta X = +176.0 \ m$, $\Delta Y = -421.8 \ m$, and $\Delta Z = +282.3 \ m$. The Operation ANON 1984-85 solution is $\Delta X = +184.0 \ m$, $\Delta Y = -356.3 \ m$, and $\Delta Z = +287.5 \ m$. For the Betio Anchorage Survey of 1959, note that the UTM...
coordinates of local traverses depend on the position C being \( \phi_0 = 01^\circ 19'\ 42.98''\ N \) and \( \lambda_0 = 172^\circ \ 58'\ 31.747''\ E \).

For the Line Island group, according to John W. Hager, the origin is at Station Beacon on Kiritimati (Christmas Island) where \( \phi_0 = 01^\circ \ 59'\ 08''\ S \) and \( \lambda_0 = 157^\circ \ 29'\ 00''\ W \) West of Greenwich. Established by N. J. Till, hydrographic surveyor, M.N.Z.I.S. in April-September 1941. English Harbor Observation Spot on Tabuaerau (Fanning Island) is the Datum origin where \( \phi_0 = 03^\circ \ 16'\ 50''\ W \) West of Greenwich. The position is described by N. J. Till, hydrographic surveyor, M.N.Z.I.S. in April-September 1941. English Harbor Observation Spot on Tabuaerau (Fanning Island) is the Datum origin where \( \phi_0 = 03^\circ \ 51'\ 23''\ S \) and \( \lambda_0 = 159^\circ \ 21'\ 50''\ W \) West of Greenwich.

For the Phoenix Island Group, Birnie Island Astro is where \( \phi_0 = 03^\circ \ 35'\ 07.875''\ S \) and \( \lambda_0 = 171^\circ \ 31'\ 03.194''\ W \) West of Greenwich. Kanton 1939 Datum at the American Eclipse Expedition Pier, USS Bushnel where \( \phi_0 = 02^\circ \ 49'\ 07.2''\ S \) and \( \lambda_0 = 157^\circ \ 53.5''\ W \) West of Greenwich. The position is described by monolith on the western side of the island. Kanton 1963 is where \( \phi_0 = 02^\circ \ 47'\ 20.99''\ W \pm 0.3'' \), \( \lambda_0 = 171^\circ \ 39'\ 49.00''\ W \pm 0.3'' \) West of Greenwich, and the reference azimuth is \( \alpha_0 = 75^\circ \ 15'\ 19.15''\ W \pm 0.15'' \) to CAN AZI from south (USC\&GS second order).

Canton Astro 1966 at Canton SECOR Astro is where \( \phi_0 = 02^\circ \ 46'\ 28.99''\ S \pm 0.04''\ W \) and \( \lambda_0 = 171^\circ \ 42'\ 53.5''\ W \pm 0.05'' \) West, and the azimuth is \( \alpha_0 = 385^\circ \ 51'\ 02'\ 65''\ \pm 0.11''\ (SIC) \) from SECOR RM-1 to TBM #3 (SECOR AZ MK) from South, referenced to the International 1909 ellipsoid. Enderbury Island Astro is where \( \phi_0 = 03^\circ \ 08'\ 30.140''\ S \), \( \lambda_0 = 174^\circ \ 32'\ 27.71''\ W \) West of Greenwich, and the reference azimuth is \( \alpha_0 = 252^\circ \ 25'\ 24.41'' \) to station Line. The Hull Island Astro Datum is where \( \phi_0 = 04^\circ \ 29'\ 15.263''\ S \), \( \lambda_0 = 172^\circ \ 10'\ 15.188''\ W \) West of Greenwich, and the reference azimuth is \( \alpha_0 = 001^\circ \ 50'\ 22.20'' \) to station Base. The McKean Island Astro is where \( \phi_0 = 03^\circ \ 35'\ 51.375''\ S \), \( \lambda_0 = 174^\circ \ 07'\ 37.522'' \) West of Greenwich, and the reference azimuth is \( \alpha_0 = 005^\circ \ 04'\ 59.26'' \) to station North. The Phoenix Island Astro is where \( \phi_0 = 03^\circ \ 43'\ 13.375''\ S \), \( \lambda_0 = 170^\circ \ 42'\ 56.004'' \) West of Greenwich, and the reference azimuth is \( \alpha_0 = 309^\circ \ 23'\ 37.76'' \) to station South. Sydney Island Astro is where \( \phi_0 = 04^\circ \ 26'\ 57.975''\ S \), \( \lambda_0 = 171^\circ \ 15'\ 43.885'' \) West of Greenwich, and the reference azimuth is \( \alpha_0 = 009^\circ \ 45'\ 57.97'' \) to station Nee. For both the Line Islands Group and the Phoenix Islands Group, the ellipsoid of reference is the Clark 1866 unless otherwise noted. Thanks again go to Russell Fox of the U.K. Ordnance Survey; John W. Hager, retired from AMS/DMA/NIMA; and Richard W. Stevenson, head of the Reference and Bibliography Section, and Gary Fitzpatrick, senior reference librarian, both of the Library of Congress; and David Llewellyn, senior draftsman, Lands and Survey Division, Bairiki, Tarawa, Republic of Kiribati.

### UPDATE

A number of geodetic surveys have been performed on various islands by Geoscience Australia:


The contents of this column reflect the views of the author, who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the American Society for Photogrammetry and Remote Sensing and/or the Louisiana State University Center for GeoInformatics (C\(^G\)). This column was previously published in *PE&RS*.
Dr. Charles E. Olson Jr. was a pioneering member of ASPRS from 1956 until he passed away on June 28, 2020. Chuck dedicated his life’s work to the interpretation of aerial photographs, and his love of education shined through in all of his endeavors, from being a U.S. Navy air photo instructor, working with Elderwise and with youth via Michigan Envirothon, and as a tutor through the Ann Arbor Rotary Service.

Professionally, Chuck’s love of forestry and air photo interpretation started early at the School of Forestry at the University of Michigan, where he received his Bachelor of Science in Forestry degree in 1952, followed with a Master’s of Forestry in Forest Management degree in 1953 from the University of Minnesota. It was in that same year that Chuck began working as an instructor of air photo interpretation and photogrammetry for the U.S. Navy, retiring at the rank of Captain from the Naval Reserves in 1987. The Commanding Officer of the Atlantic Intelligence Center once introduced Dr. Olson as “The best photo interpreter in the U.S. Navy.” During active duty, his expertise in oblique photo metrics and radar image interpretation was regularly requested for complex assignments. For over 30 years in the Naval Reserves, he taught recruits the science and the art of image interpretation. He received a Ph.D. in Forestry from the University of Michigan in 1969.

Although offered full-time positions in remote sensing by the Navy and the CIA, Dr. Olson decided to further his education and pursue an academic career in natural resources management. In 1956, he began working with the University of Illinois Forestry Department, teaching photo-interpretation and pursuing a second Master’s of Forestry in Photo-Interpretation and Photogrammetry. It would be a combination of his education and his service that would result in one of his most famous works, the seminal Photogrammetric Engineering publication entitled, “Elements of Image Interpretation” (Olson 1960). This paper brought together disparate references of image understanding into a holistic framework that continues to serve as a fundamental component of remote sensing education today. Indeed, the Elements of Image Interpretation (i.e., tone, shape, size, pattern, association, shadow, texture, and resolution) still form the basis of many GIS and image processing algorithms from feature extraction to object-based image analysis and machine/deep learning.

In 1962, Dr. Olson presented his concept of the sources and characteristics of energy signals recorded in remotely sensed images as the Energy Flow Profile (EFP) to the Second Remote Sensing of Environment Symposium held at the Institute of Science and Technology at the University of Michigan, Ann Arbor. Building on Dr. Steven Spurr’s invention of the parallax wedge in 1945, Dr. Olson also developed the Michigan Parallax Wedge. This all-purpose tool provided the necessary measurement graticules to perform complex photogrammetric operations and was used by generations of aerial photograph practitioners. These early achievements in photo interpretation and remote sensing were standard approaches in practice and education for decades, and they remain influential in the geospatial field to this day.

From 1963 to 1969, Olson held a joint appointment at the University of Michigan School of Natural Resources and the Infrared Physics Lab of Willow Run Laboratories, the predecessor to the Environmental Research Institute of Michigan (ERIM). After completing his Ph.D., he spent most of his career working at the University of Michigan where he eventually became Dean of the School of Natural Resources. In his 35 years at Michigan, Chuck had one of the highest rates of successful mentoring in the School’s history, graduating over sixty M.S. and Ph.D. students and influencing countless undergraduate and graduate students. They took his courses in Map and Image Interpretation, Remote Sensing of Environment, and Digital Image Processing.

Chuck was an active member of ASPRS locally in Michigan through the Eastern Great Lakes Region, nationally and internationally teaching air photo interpretation and inspiring numerous students to careers in remote sensing. He was a Remote Sensing Instructor at the Remote Sensing Center for East Africa, Nairobi, Kenya in 1981. Even after retiring as Emeritus Professor from the University of Michigan in 1999, Chuck was a notable voice among his colleagues and friends at these meetings, often bringing his wife, Connie, along to scout out the nearest Carnegie Libraries.
Chuck served as the ASPRS National Director from the Eastern Great Lakes Region from 2002 until 2008. In 2005, Brain Huberty recalls that Chuck came up with the brilliant idea to have a joint interregional meeting on the Lake Michigan Car Ferry between Manitowoc, Wisconsin, and Ludington, Michigan. The Eastern Great Lakes Region members got on the ferry early in the morning on August 19 and had their meeting going across to Manitowoc, WI. Western Great Lakes members got on the ferry around noon for a joint meeting in a small theater on the ferry heading back to Ludington.

His service to ASPRS earned Chuck several Society awards, including the 2011 Outstanding Workshop Instructor Award for his excellent conference workshops on image understanding. He received three Presidential Citations for Meritorious Service and the Ford Bartlett Award. Chuck was honored as an ASPRS Fellow in 1998, and in 2009, as an Honorary Member of ASPRS, the highest award a member can receive in recognition of distinguished lifetime service to ASPRS and advancing the science and use of geospatial information sciences. Chuck also began the ASPRS Oral History Project and completed 56 interviews, several of which were the basis for Reflection of the Past series in PE&RS.

A gentleman and outgoing member, he made everyone feel welcome at ASPRS, always eager to talk about the finer points of aerial photos. He leaves a long-lasting impact on each colleague who has been fortunate to experience his wisdom, intelligence, and practical knowledge. He will be remembered as a giant on whose shoulders we all stand within the remote sensing community.

### Selected Publications by Charles Olson

- **Elements of Photographic Interpretation Common to Several Sensors**
  
  C. E. Olson, Jr.
  
  *Photogrammetric Engineering*, 1960

- **Seasonal Trends in Light Reflectance from Tree Foliage**
  
  C. E. Olson, Jr.
  
  *International Archives of Photogrammetry*, 1963

- **Infrared Sensors and their Potential Contributions to Forestry**
  
  C. E. Olson, Jr.
  
  *Papers of the Michigan Academy of Science, Arts and Letters*, 1965

- **Accuracy of Land-use Interpretation from Infrared Imagery in the to 5.5 Micron Band**
  
  C. E. Olson, Jr.
  
  *Annals of the American Association of Geographers*, 1967

- **Infrared Scanning Techniques for Big Game Censusing**
  
  G. W. Croon, D. R. McCullough, C. E. Olson, Jr., and L. M. Queal
  
  *Journal of Wildlife Management*, 1968

- **Multispectral Sensing of Forest Stress**
  
  W. G. Rohde and C. E. Olson, Jr.
  

- **Assessing Spruce Budworm Damage with Small-Format Aerial Photographs**
  
  J. McCarthy, C. E. Olson, Jr., and J. A. Witter.
  

- **Use of Geographic Information Systems to Develop Habitat Suitability Models**
  
  M. L. Donovan, D. L. Rabe, and C. E. Olson, Jr.,
  

- **Minimizing Classification Errors Arising from Skewed Distributions of Satellite-Observed Brightness Values**
  
  C. E. Olson, Jr.
  

- **Is 80% Accuracy Good Enough?**
  
  C. E. Olson, Jr.
  
  In, *Proceedings of the ASPRS SemiAnnual Meeting*, 2008

- **The Fallacy of Normality in Remotely Sensed Data**
  
  C. E. Olson, Jr.
  
  In, *Proceedings of the ASPRS Annual Convention*, 2009
Some selected memories from ASPRS members include the following.

Chuck’s Image Interpretation Elements became the basis for my Ph.D. work, and friendship with Chuck emerged after I coincidently sent him an email profusely praising his contributions to the field on his birthday in 2012. His encouragement, generosity, and thoughtful feedback continue to serve as the core of my research in the history of the remote sensing field, and the human factors that underpin current remote sensing practices. There is not a student of my own who does not engage with Chuck’s work and who does not grow to appreciate what he has done for the discipline.

~ Dr. Raechel Portelli, Assistant Professor, Michigan State University


Sorry to hear this very sad news. He impressed me by attending every ASPRS annual conference, attending sessions and stopping at booths – a truly amazing and respected gentleman.

He will be missed.

~ Jie Shan, Purdue University

As you know, Chuck was a giant in our field and an amazing gentleman. I got the pleasure of having his support early in my career and then working with him later on. He will be greatly missed.

~ Russ Congalton, University of New Hampshire, ASPRS Past President

For a long time, it seemed to me that he was always present at every convention, regardless of location or size, keen to discuss fine points of aerial photos . . .

~ James Campbell, Virginia Polytechnic Institute and State University

How lucky we were all to experience a great mentor. Hopefully more will follow because of the experience.

~ Janet Degner, University of Florida

This was extremely sad news this morning. Chuck made everyone feel welcome and also an important part of ASPRS.

~ Lorraine Amenda, Towill, Inc., ASPRS Board of Directors— Secretary

To the remembrances and tributes, I would like to add “...gentleman... “

~ Alan M Mikuni, GeoWing Mapping, Inc., ASPRS Past President

Wow. I am heart broken. I will miss his constant and reassuring presence.

~ Kass Green, Kass Green & Associates, ASPRS Past President

I first met Chuck at the U of M in 1965, as an undergraduate. What a loss!

~ Clifford Greve, Chair of the ASPRS Foundation Board of Trustees, and ASPRS Past President

Chuck will be missed by all of us.

~ Qassim Abdullah, Woolpert, Inc.

Chuck was a highly dedicated advocate of all things’ ASPRS’! And he educated a lot more that hopefully follow his footsteps regarding volunteering to help ASPRS sustain itself.

~ Terry Keating, ASPRS Past President

Thanks for passing along this sad news.

~ Thomas Lillesand, ASPRS Past President

Truly sad news.

~ Jeff Lovin, Woolpert, Inc., ASPRS President

ASPRS would like to thank Raechel Portelli, Colin Brooks, Nancy French, Kathleen Bergen, Laura Bourgeau-Chavez, and Marguerite Madden for writing this memorial to Chuck Olsen.
ANNOUNCING NEW MEMBERS & COMMITTEE CHAIRS OF OUR STUDENT ADVISORY COUNCIL (SAC) TEAM

The ASPRS SAC announces the election of four new members/Committee Chairs to its team: Madison Fung, Lauren McKinney-Wise, Terra McKee, and Evan Vega.

We are delighted to welcome these four wonderful geospatial leaders who have a lot to share with SAC. They possess unique backgrounds along with diverse experience at the ASPRS Chapter level, which makes them a great asset in strengthening ASPRS, an important role our student members play.

The new leadership and existing members of SAC will continue working on SAC growth. They play a critical role in providing the resources and expertise needed for SAC to effectively contribute to ASPRS through student involvement and in providing opportunities for students.

Since its inception in 2006, SAC has offered lots of great leadership opportunities for young geospatial students to build confidence, step outside their comfort zones, and learn, and grow from their experiences. In 2020 we have been expanding our social media presence towards building an engaged student community.

We would like to thank our ex-Chair, Victoria Scholl, for successfully leading the team to achieve last year’s goals. We truly value her involvement this year and her unlimited support by providing insight, guidance, and dedicated service.
Karen Schuckman Named as Lidar Leader

CONGRATULATIONS TO OUR GEO-RENAISSANCE WOMAN, KAREN SCHUCKMAN!

International Lidar Mapping Forum (ILMF) and LIDAR Magazine recently announced the recipients of the third annual Lidar Leader Awards during an award ceremony held virtually. ILMF – in cooperation with LIDAR Magazine – designed this unique program to recognize excellence in five distinct categories: Outstanding Personal Achievement in Lidar, Outstanding Team Achievement in Lidar, Outstanding Enterprise Achievement in Lidar, Outstanding Innovation in Lidar, and Outstanding University Achievement in Lidar.

Karen Schuckman of Pennsylvania State University and ASPRS was one of two recipients of the Outstanding Personal Achievement in Lidar category. For two decades, Mrs. Schuckman has made numerous noteworthy contributions and has sustained forward-thinking leadership in advancing lidar technologies in the mainstream geospatial/mapping community. She has been one of the most influential professionals in the lidar field, leading the first statewide lidar mapping program, helping form the ASPRS Lidar Division, coordinating standards development, and teaching in the Penn State Online Geospatial Program.

Lisa Murray, Group Director at Diversified Communications stated of the awards, “This is the third year for the Lidar Leader Awards. Over 80 submissions were received highlighting excellent work in the industry. We would like to thank all the nominees and nominators, our esteemed Advisory Board, as well as our partners at LIDAR Magazine for helping us recognize great achievements in the field of lidar technology. Especially during this time of physical distancing, these awards are a great way to bring the community together to help celebrate each other.”

Dr. A. Stewart Walker, Managing Editor at Lidar Magazine commented, “As always, it is a privilege to participate in the Lidar Leader Awards. The high quality of nominations has been maintained and again we have worthy winners. It is especially important, in these difficult times when face-to-face events are not possible, to recognize excellence and the virtual ceremony proved a popular way to do so.”

JOIN ASPRS IN WELCOMING TWO NEW SUSTAINING MEMBER COMPANIES TO OUR FAMILY

Green Grid Inc. (GGI) is a climate-focused geospatial information and operational technology (IT/OT) and risk management service business serving energy, utility and regulatory customers for operational safety and efficiency improvement. Their field services digitalization and automation solutions are beneficial to the following use cases.

- Utilities Infrastructure and Vegetation Management
- Climate Change & Wildfire Risk Mitigation
- Public Safety & Hazard Mitigation
- Natural Resource Management
- Architectural, Engineering & Construction Virtual Inspection
- Precision Agriculture

Their technology toolbox includes AI, Smart Sensors, IIOT, Bigdata, GIS, Cloud & Mobile, lidar, Photogrammetry, M/HS Imagery and more tools and workflows.

“There are approximately 640,000 miles of electric transmission and 5,500,000 miles of distribution powerlines nationally. Some of these network assets have exceeded their safe operational life limit. Knowing where they are spatially and how they are doing temporally is essential for safe and reliable electric grid operations,” said Chinmoy Saha, CEO of Green Grid.

For more information about Green Grid, Inc. visit www.greengridinc.com/.

FR Aleman is an award-winning consulting firm with expertise in engineering, geospatial/surveying and mapping, and 3D subsurface utility engineering. For the past three decades, their dedicated professionals have been providing professional engineering services and innovative solutions through technology to public and private sector clients throughout the State of Florida, the United States, the Caribbean, and Central and South America.

Their team is composed of professionals and technicians who have a longstanding commitment to excellence and innovation working with integrity, initiative, and ingenuity with unsurpassed performance with over a century of combined experience.

FR Aleman understands that no one technology can derive a solution for every project. That is why they continually invest in new technologies, state-of-the-art equipment, and incorporate a variety of cloud-based proprietary applications, artificial intelligence, and augmented reality for safe, programmatic, and integrated approaches.

FR Aleman is certified by the State of Florida as a Woman-Owned Minority Business Enterprise (W/MBE) and by certain entities as a Disadvantaged Business Enterprise (DBE) and Small Business Enterprise (SBE).

For more information on FR Aleman, visit https://fr-aleman.com/.

NEW ASPRS MEMBERS

ASPRS would like to welcome the following new members!

At Large
Daniel Hrouda
Hector R. Rivera
Hari Rajan Saravanan

Florida
Nyakeh B. Aruna
Jennifer Larsen
Greg Scott

Mid South
Jay Q. Drake, PLS
Karina Juarez
Cathy McPherson
David Weld
Casey Aaron Johnston

North Atlantic
Joseph E. Romano, PLS

Pacific Southwest
Dylan M. Ayers

Potomac
Matthew W. Huchla

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

Your Path To Success In The Geospatial Community
Heliport Detection Using Artificial Neural Networks

Emre Başeski

Abstract
Automatic image exploitation is a critical technology for quick content analysis of high-resolution remote sensing images. The presence of a heliport in an image usually implies an important facility, such as military facilities. Therefore, detection of heliports can reveal critical information about the content of an image. In this article, two learning-based algorithms are presented that make use of artificial neural networks to detect H-shaped, light-colored heliports. The first algorithm is based on shape analysis of the heliport candidate segments using classical artificial neural networks. The second algorithm uses deep-learning techniques. While deep learning can solve difficult problems successfully, classical-learning approaches can be tuned easily to obtain fast and reasonable results. Therefore, although the main objective of this article is heliport detection, it also compares a deep-learning based approach with a classical learning-based approach and discusses advantages and disadvantages of both techniques.

Introduction
The increased resolution and the amount of commercially available remote sensing data are a new challenge for image analysis. With the increase in resolution, even a small portion of an image can contain details that are difficult for a human operator to analyze. In addition, the increasing number of imaging sensors produce a massive amount of imagery data that is impossible to analyze without intelligent image-processing algorithms. Therefore, automatic image-exploitation algorithms are very important for analyzing the content of this huge data.

In this work, the problem of fully automated heliport detection is discussed and two different approaches are presented. The presence of a heliport in an image often points to an important facility, such as government buildings or military facilities. The heliports that are studied in this work are H-shaped, light-colored structures with different sizes and fonts. Previous work on heliport detection was done by Başeski (2018), with an algorithm analyzing the shape of binary segments.

Automatic detection of artificial structures in remote sensing images is an important area of research. Classical image-processing techniques and learning-based approaches are used to solve different problems. Extensive information on different approaches to artificial-object detection is given by Cheng and Han (2016) and Blaschke (2010). Classical techniques (e.g., Mueller, Segl and Kaufmann 2004) often use a combination of region- and edge-based techniques to identify large and potentially artificial areas, whereas more recent studies tend to use learning-based approaches (e.g., Ingлада 2007; Han et al. 2015; G.-S. Xia et al. 2017; Zhu et al. 2017).

Recent developments in deep learning have led to significant breakthroughs in the field of image processing. Success in image classification has been carried to object detection through networks that can also locate objects in the image. Networks such as Fast R-CNN (Girshick 2015), YOLO (Redmon et al. 2016), and SSD (Liu et al. 2016) have achieved tremendous performance improvements over conventional methods in object detection.

The performance problem of these networks on relatively small objects is solved by Faster R-CNN (Ren et al. 2017).

There seems to have been a significant increase in the use of deep-learning techniques in recent years by the remote sensing community. For instance, F. Xia and Li (2018) conducted a study on airport detection in remote sensing data via the SSD algorithm. SSD is a network that has difficulty detecting small objects due to the nature of its anchor-selection mechanism. The contribution of S. Chen, Zhan, and Zhang (2018) improved the classical SSD algorithm for detecting relatively small objects in satellite images. Ying et al. (2018) present a survey on image classification for remote sensing. Deep-learning architectures such as U-Net (Ronneberger, Fischer and Brox 2015) and DeepPUNet (Li et al. 2018) have also been shown to produce quite successful results in the land use classification problem.

In this work, two different algorithms for heliport detection are presented. The first algorithm is based on classical artificial neural networks. The second is an algorithm developed with deep-learning techniques. The technique based on classical neural networks is mainly based on the analysis of the shape of the heliport candidate segments. The deep-learning technique is based on a Faster R-CNN (Ren et al. 2017) architecture. Although deep learning seems to solve fairly difficult problems successfully, classical-learning approaches can be tuned easily to obtain fast and reasonable results. Therefore, a side objective of this article is to compare a deep learning-based approach with a classical learning-based approach and discuss the advantages and disadvantages of both techniques. Also, analysis of factors such as shape size, deformation of the shape, and effect of contrast on heliport detection is discussed.

Algorithm
In this section, two different algorithms that find heliports in satellite and aerial images are discussed. The first algorithm is based on classical artificial neural networks. The second algorithm exploits deep-learning techniques. The classical neural network is trained with a data set used for character recognition. It performs shape analysis on potential heliport candidate segments. In the deep learning-based algorithm, a convolutional artificial neural network trained with a large number of heliport images is used to find the location of heliports.

Automated detection of heliports is a challenging problem due to variable shape size, changing background, pale edges, and surrounding shapes. Also, the color of runways fades over time, and a heliport may have different color shades. Figure 1 presents some sample heliport images to highlight the difficulties of the problem.

The evaluation of the presented methods was performed on 32 images containing 66 heliports. The images were collected from Google Maps and selected from different regions of the world to include different features (different backgrounds, image content, vegetation, etc.). By using the same
data set, the presented methods are also compared with the algorithm of Başeski (2018).

The method of Başeski (2018) relies on classical image-processing techniques. The algorithm iteratively thresholds the image and performs shape analysis on binary segments using symmetry, connectivity, and cleanliness. The shape-analysis algorithm is fuzzy enough to handle shape distortions and fading color but it is not as flexible as the deep-learning algorithm presented in this work for partially occluded heliports. The algorithm was able to accurately locate 60 heliports out of 66 in the experimental set, could not find six, and produced one false positive. Therefore, 90.9% recall and 98.36% precision values were obtained. The main advantage of the algorithm is that it can be developed with a small number of sample images, since it is not learning based. Also, its precision value is very high. The most important disadvantage is that it is not able to find deformed or partially occluded heliports.

The Algorithm Based on Classical Neural Networks
The classical neural-network approach follows a similar strategy as that of Başeski (2018), except for the shape analysis. While that method used some geometrical properties of the binary segments, the method presented in this work uses a shallow network to perform shape analysis.

Trying to find a single threshold for an image that can reveal the binary shape of heliports is an ill-posed problem. Instead of searching for the ideal threshold, the algorithm iteratively binarizes the image and performs shape analysis on the connected components. The flow diagram of the algorithm is presented in Figure 2. The algorithm starts with the initial threshold calculation. Then the image is binarized and the connected components are calculated. The shape-analysis step determines whether a connected component looks like an H or not. Once all the connected components are analyzed, the threshold value is incremented by a certain integer and the whole process is repeated until the threshold value reaches 255.

In order to determine the initial threshold value, high contrast between the heliport and its surrounding pixels needs to be created. The process starts by reducing the brightness of the image \( I \) by subtracting \( m \) (Equation 1a), the value that divides the histogram of \( I \) into two equal parts. After the brightness is reduced, the contrast of the segment is boosted by Equation 1b. The initial threshold value is then calculated from the resultant image \( I' \) using Otsu’s (1979) thresholding algorithm. Once a binary image is obtained, the connected components are calculated for further shape analysis.

\[
I'' = I - m \quad (1a)
\]

\[
I' = I' \ast a - m \quad (1b)
\]

\[
a = \frac{255 + m}{\max(I')} \quad (1c)
\]

The shape analysis is performed by a neural network that is designed to classify numbers and letters. The network is a three-layer artificial neural network (Figure 3). The fundamental function of the neural network is to decide whether a shape of 28×28 pixels resembles the letter H.

The neural network was trained by using the EMNIST letter database (Cohen et al. 2017). The database contains 697 932 characters for training and 116 323 characters for the test. These characters consist of numbers and upper- and lowercase letters (62 classes in total) in the English alphabet. There are 3152 H characters in the training set and 521 in the test set. Since the EMNIST database consists of 28×28-pixel images, the first layer of our neural network contains 28 × 28 = 784 neurons. A single hidden layer of 100 neurons was used in the study. The output layer consists of 62 neurons.

For a segment similar to the H shape, output neuron 17 produces a high value, while the other neurons produce high values for non-H segments. In the training of the network, the learning rate was chosen as 0.1, the batch size was 25, and the cost function was selected as cross entropy. Weights and biases are calculated within 150 epochs.

In order to determine whether a segment is similar to the H shape, the segment is first aligned in such a way that the centroid is at the origin of the coordinate system, the long side is parallel to the y-axis, and the short side is parallel to the x-axis. The resulting image is rescaled to 28×28 pixels and pushed into the neural network for the shape analysis.

In Figure 4, two sample detection results are presented. While the one on the left is a rather large and clean heliport, the one on the right side has a faint color. The algorithm can successfully detect both heliports without any false positive.

Figure 5 displays two results for large residential regions with a complex background. Although the background is large and contains complex patterns, the algorithm still can detect the heliports without any false positive.

The algorithm can also handle multiple heliports in a single image (Figure 6). Since the shape-analysis part of the algorithm handles each segment separately, multiple heliports with different dimensions, colors, and distortion can be handled successfully.

Industrial areas usually pose a great challenge for automated image-exploitation algorithms, because of the complex nature of the image content. As shown in Figure 7, the algorithm can handle such complex environments with very high precision.
In Figure 8, four different scenarios are presented where the algorithm fails to detect the heliports. In the top left, the algorithm is able to find the heliport on the right but unable to detect the one on the left. Although both heliports are extremely deformed by the mark on top of the H, the thresholding process could separate the H shape for the one on the right. The top right heliport is extremely small and its shape is highly deformed. The bottom right heliport is partially occluded by a line passing through the middle of the shape. The one undetected heliport on the bottom left has a very faint color, and the initial threshold was unable to create a segment that looked like an H.

The algorithm was able to correctly detect 57 of the 66 heliports in the test set, was unable to find nine of them, and produced four false alarms. Therefore, 86.36% recall and 93.44% precision values were obtained.

The main advantage of the algorithm is that the data needed to train the network are from a different research area. Since the training data were labeled and ready for training, the creation and training of the neural network required much less effort than performing these procedures for deep learning. In addition, without the need to develop special algorithms for shape analysis, a general method that can successfully solve the problem has been introduced. However, the overall method shows worse performance than the problem-specific algorithm of Başeski (2018). There was an increase in the frequency of misclassification and a decrease in precision. It should be noted that shape deformation and extra lines on the shape are a significant problem for this algorithm.

The Algorithm Based on Deep Learning

In this study, Faster R-CNN (X. Chen and Mulam 2017) was used as the deep-learning architecture. Faster R-CNN consists of three basic networks. The first, the head portion, uses the first few layers of a previously trained network, such as VGG16 or ResNet-50, to find useful features. The first few layers of the network learn to recognize common features, such as edges and color segments, with good distinguishing capabilities for different problems. The next layers, however, learn higher-level features. The second network, the region proposal network, creates nonbackground objects and creates candidate image fragments for the classification network. The classification network, which is the last basic element, classifies the candidates proposed by the region proposal network. In this way, the same network can find both objects in the image and the locations of these objects.

One hundred fifty-two images were used for training the network. A detailed description of the data set is given by Başeski (2019). The images are 0.5-m-resolution Google Earth images. The minimum width and height of heliport bounding boxes are 10 pixels, the average width is 49 pixels, and the average height is 48 pixels. The dataset directory and labeling structure is compatible with Pascal VOC. It contains heliports with different sizes, colors, and backgrounds. (The source code for the PyTorch implementation of Faster R-CNN can be found at https://github.com/TAMU-VITA/dehaze/tree/master/code/iodh/tf-faster-rcnn.) For the training, a pretrained VGG16 network was used as the feature-extraction network. A momentum optimizer was used as the optimizer, and the
batch size was chosen as 10. Of the 152 images, 69 were used for training, 56 were used for validation, and 27 were used for testing. The model was tested with images that were not part of the training process. The 70 000-iteration training process took about seven hours on a single Tesla K40c graphics processing unit. With the model created, it takes approximately 0.7 second to analyze an image of 640×640 pixels.

The main advantage of the deep learning-based algorithm is that it can successfully find heliports that are partially covered or have significantly deformed parts. The algorithm was able to find 63 of the 66 heliports in the experiment set correctly, was unable to find three of them, and produced three false positives. Therefore, 95.45% recall and 95.45% precision values were obtained. In Figure 9, two sample detection results are presented. The image on the left side contains a heliport that cannot be detected by classical algorithms. Even though the superfluous structure of the shape was misleading for shape analysis by classical algorithms, the deep learning-based algorithm successfully found the heliport.

In Figure 10, sample detection results are shown for a case where the contrast is poor and the leftmost heliport is more deformed than the others.

One of the factors affecting the success of the algorithm is that the images are 8-bit, and the parameters used in the conversion from 16-bit to 8-bit can decrease the contrast between the heliports and their surrounding pixels. In order to avoid problems that may occur due to low contrast, each image has been given as the input to the network three times after increasing the contrast slightly. The contrast was increased by multiplying color values by 1.2 and 1.4.

In Figure 11, the effect of a hand-drawn line on a heliport is analyzed. Looking carefully at the result boxes, it can be seen that the drawn line changes the result, but the algorithm continues to find the heliport successfully. For shape-based algorithms, these lines radically change the binary segment shape. Therefore, classical algorithms cannot classify the shape as an H.

It is quite possible to encounter heliports that have lines on top, especially at civil airports. Heliports that are on taxiways or parking areas usually have this kind of line passing through. Therefore, it is important for an automatic heliport-detection algorithm to be able to work with heliports that have distracting markings on top. In Figure 12, a sample heliport-detection result is shown where the heliport is on a runway and there is an external line passing through the H shape. Note that both classical methods fail to find this heliport, because of the line passing through.

In Figure 13, another scenario is shown, where the H shape is partially occluded by helicopters. Since heliports are landing zones for helicopters, this is not an unusual scenario, and it is important for the algorithm to be able to handle such cases. As with other partially covered heliports, only the deep learning-based method can perform successfully.

In Figure 14, a sample result is shown for two deformed heliports. The heliport on the left side is extremely deformed, and even though almost half of the shape is missing, the algorithm can successfully perform the detection.

In Figure 15, two sample results are displayed for a complicated background. For both cases, the algorithm is able to detect the heliports.

Figure 16 shows a scenario in which the algorithm is unable to detect one of the heliports (top left). Note that the color of the missing heliport is extremely faint, and the height-to-length ratio is larger than for the overall heliports.

In Figure 17, the algorithm detects an extra heliport (top right corner). The main reason for this is that the structure is visually similar to a deformed heliport with a frame around it. Note that flexibility, which allows the algorithm to work in case of deformation and faint color, may in some cases produce false positives.
The algorithm is able to work robustly on both clean-shaped and deformed heliports. Although the numerical performances of the algorithms appear close to each other, the experiments were done with a limited number of images that contain few occluded and extremely deformed heliports. Therefore, the addition of difficult scenarios to the test set will convert the balance in favor of the deep learning-based algorithm.

The most obvious advantage of this algorithm is that it can produce stable results even on very difficult heliports and does not require special code writing for the problem. In addition, model performance can be further increased over time with extra training data.

The most important challenge in the deep learning-based technique is the need to collect a large amount of data for the preparation of training set. Although the 152 images that were used for training here are a small set for deep learning, the relatively clear shape of the heliports allowed successful training with such a small data set.

In studies that include the use of deep learning in remote sensing, the data-collection phase is one of the main challenges. Not only collection of the data but also labeling of structures to be searched in the image requires great effort. In addition, the training process requires special computers with powerful graphics processing units. A small disadvantage of the algorithm is that it finds the bounding rectangle of the heliport, not the sensitive segmentation of the H shape.

Figure 13. Heliports that are partially covered by helicopters.

Figure 14. Sample detection result for two deformed heliports.

Figure 15. Sample results for complicated background.

Figure 16. A sample scenario where the algorithm fails to detect one of the heliports (topmost one).

Figure 17. A false-positive scenario (top right corner).
Conclusion and Discussion
In this study, two algorithms are presented which find heliports in satellite and aerial images. The first algorithm is based on classical artificial neural networks. The second algorithm is completely developed by deep-learning techniques. The classical neural network is trained with a data set used for character recognition, and its main objective is shape analysis. In the deep learning-based algorithm, a convolutional neural network is used to find the location of the heliport as the network output.

Experiments to measure the performance of algorithms were performed on images collected from Google Maps. Images were selected from different regions of the world to include different features (different background, different image content, different vegetation). The algorithms presented in this work are also compared with a reference algorithm (Başeski 2018) on the same data set. In Table 1, algorithm performances are given for the test set. The reference algorithm works with very high precision for the relatively clean runways; however, its performance starts degrading if there is partial occlusion or some parts of the figure are missing. The reference algorithm and the method based on classical neural networks both analyze the shape of candidate segments. Although there is a slight loss of performance compared to the reference method, training of the classical-learning method is possible with reference data from another problem, without developing a specific algorithm for the problem. The deep learning-based algorithm works stably with complex background images, faint color, and partial occlusions. Although external segments passing through the heliport, such as runway lines or helicopters, cause performance loss for the reference algorithm and the classical neural-network approach, the deep learning-based approach handles such deformations with a very slight drop in precision. The main disadvantage of the deep learning-based approach is that the process of collecting and labeling the training data is time consuming and difficult.

Table 1. Summary of algorithm performances.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Başeski (2018)</td>
<td>98.36</td>
<td>90.9</td>
</tr>
<tr>
<td>Classical neural network</td>
<td>93.44</td>
<td>86.36</td>
</tr>
<tr>
<td>Deep learning-based</td>
<td>95.45</td>
<td>95.45</td>
</tr>
</tbody>
</table>

Although the numerical performances of the algorithms appear close to each other, the experiments were done with a limited number of images that contain few occluded and extremely deformed heliports. Therefore, the addition of difficult scenarios to the test set will convert the balance in favor of the deep learning-based algorithm.

Acknowledgments
This work was supported by Havelsan.

References


Semi-Automatic Building Extraction from WorldView-2 Imagery Using Taguchi Optimization

Hasan Tonbul and Taskin Kavzoglu

Abstract
Due to the complex spectral and spatial structures of remotely sensed images, the delineation of land use/land cover classes using conventional approaches is a challenging task. This article tackles the problem of seeking optimal parameters of multi-resolution segmentation for a classification task using WorldView-2 imagery. Taguchi optimization was applied to search optimal parameters using the plateau objective function (POF) and quality rate (Qr) as fitness criteria. Analysis of variance was also used to estimate the contributions of the parameters for POF and Qr, separately. The scale parameter was the most effective one, with contribution levels of 87.45% and 56.87% for POF and Qr, respectively. Linear regression and support-vector regression methods were used to predict the results of the experiment. Test results revealed that Taguchi optimization was more effective than linear regression and support-vector regression for predicting POF and Qr values.

Introduction
Producing a land use/land cover (LU/LC) map significantly facilitates the dissemination of high-resolution images. However, high intraclass spectral heterogeneity and interclass spectral similarity of objects in high-resolution images are an important problem for classification (Kavzoglu 2017). Therefore, traditional pixel-based approaches assuming that each land cover class has a distinctive spectral signature fail to achieve a high degree of classification accuracy (Johnson and Jozdani 2018). Accordingly, as an alternative to traditional pixel-based image-analysis techniques, object-based image analysis (OBIA)—introduced by Blaschke (2010)—has emerged in the analysis of remotely sensed imagery.

Image segmentation is the first and most critical stage of the OBIA approach. Segmentation can be described as the process of creating homogeneous image objects using the color, texture, and shape properties of the images. It is usually managed by three parameters: scale, shape, and compactness. The process of defining the optimum combination of these parameters is very challenging and time consuming (Kavzoglu, Yildiz Erdemir and Tonbul 2017; Kavzoglu and Tonbul 2018). Other parameters including band weighting and spectral indices are all left in the hands of the user. In addition, the subjective choices of the analyst in the selection of these parameters and spectral, spatial, contextual, and textural features undermine the reliability, repeatability, and effectiveness of the segmentation process. The multi-resolution segmentation (MRS) algorithm, introduced by Baatz and Schäpe (2000), is the most popular and widely used segmentation algorithm, available in the eCognition® software package (Trimble Geospatial Imaging). In the literature, these parameters, particularly shape and compactness, are set as default parameters or with a trial-and-error strategy (e.g., Halabisky, Moskal and Hall 2011; Lowe and Guo 2011; Kavzoglu and Yildiz 2014; Chen et al. 2015; Kavzoglu, Colkesen and Yomralioglu 2015). It is obvious that this is an open question that needs to be investigated to push OBIA to a better and more popular position. Therefore, using an optimization technique can be effective for decreasing the time and cost of the trial-and-error method (Cheng et al. 2014; Gibril, Shafri and Hamedianfar 2017).

Some approaches exist in the literature for assessing the quality of image segmentation, which can be divided into two categories: supervised and unsupervised (Zhang 1996; Zhang et al. 2008). The purpose of supervised approaches is to define the most appropriate segmentation by evaluating the overlap of ground-reference polygons with the created image objects. Supervised evaluation techniques are based on modeling the geometric or arithmetic relationship between reference objects and corresponding image objects (Liu et al. 2012; Yang et al. 2015; Su and Zhang 2017). The existing supervised segmentation metrics, such as area fit index (Lucieer and Stein 2012), quality rate (Winter 2000), Euclidean distance (Liu et al. 2012), and shape index (Neubert, Herold and Meinel 2006), have been applied and evaluated independently in many studies (Clinton et al. 2010; Zhang et al. 2015; Novelli et al. 2017). On the other hand, unsupervised methods are applied without using reference polygons for comparison. Unsupervised methods basically aim to define segmentation parameters that maximize the average intrasegment homogeneity and intersegment heterogeneity of segmentation results (Johnson and Xie 2011; Johnson and Jozdani 2018). An unsupervised objective function proposed by Espindola et al. (2006) integrates the weighted variance and spatial autocorrelation (Moran’s index) of the image objects to estimate segmentation quality. Furthermore, Drögut, Tiede, and Levick (2010) have proposed a tool called the estimation of scale parameter tool, which evaluates variations in local variance depending on scale values to determine optimum parameters for MRS. There is a plethora of literature on the use of unsupervised quality-assessment methods for segmentation (Johnson and Xie 2011; Zhang, Xiao and Feng 2012; Kavzoglu et al. 2017).

In this study, the statistical optimization technique developed by Taguchi (1986) is used to optimize MRS parameters for semi-automated mapping of buildings. The use of the Taguchi optimization technique is gaining prominence in remote sensing applications, and the technique has lately been applied in various studies. For example, Moosavi, Talebi, and Shirmohammadi (2014) use Taguchi optimization for the analysis of landslide inventory maps using advanced pixel-based and object-oriented approaches. In that study, Taguchi optimization was used to optimize the parameters of artificial neural network and support vector machine classifiers together with the MRS parameters. Idrees and Pradhan (2016) studied Taguchi optimization and the objective function to optimize segmentation parameters for automatic cave-bird detection from terrestrial laser-scanning intensity images, and concluded that Taguchi optimization is useful for distinguishing...
years, birds from cave walls with high accuracy. Idrees et al. (2016) used the Taguchi orthogonal array to optimize MRS segmentation parameters for detecting and counting cave-roosting birds. They report that all MRS parameters, not only the scale parameter, collectively influence the quality of segmentation. Gibril et al. (2017) investigated the optimization of segmentation parameters by using the Taguchi technique in heterogeneous urban areas to semi-automatically map asbestos cement roofs. The results reveal that the MRS-based Taguchi technique performed a good delineation in the boundaries of roofing materials and other impervious surfaces. Zare, Behnia, and Gabriels (2019) applied the Taguchi technique to optimize the parameters of support vector machine classification by using kappa coefficients, concluding that optimized support vector machine classification by the Taguchi method can be efficiently used to produce LULC change maps.

Although many studies have applied Taguchi optimization, few of them use statistical analysis for the assessment of MRS parameters with this optimization method (Luo, Li and Liu 2017). The purpose of this article is to investigate the effectiveness of the Taguchi optimization method for obtaining optimum MRS parameters based on two widely used segmentation quality metrics: the unsupervised plateau objective function (POF) and the supervised quality rate (Qr). We implemented an L25 (53) orthogonal array to minimize the number of experiments. The MRS parameters of scale, shape, and compactness were considered as test values in Taguchi optimization. Analysis of variance was used to determine the statistical significance of MRS parameters, and the reliability of the Taguchi technique was verified by a confirmation test. The ultimate goal was to semi-automatically extract and map buildings through OBIA-based image classification. A robust machine-learning algorithm, namely random forest, was selected to perform object-based image classification for LULC classes considering the segments created by the MRS algorithm for the selected WorldView-2 imagery.

Study Site and Data Set
In this study, a multispectral WorldView-2 high-resolution satellite image containing eight spectral bands (0.45–1.05 µm) at 2-m spatial resolution obtained on March 21, 2012, was used. The study site is located in the city of San Clemente, California, USA. Figure 1 shows the study site (4000×4000 pixels), covering major impervious surfaces of building rooftops, asphalt roads, bare soil areas, and forested lands. In fact, six LULC classes—asphalt road, bare soil, forest, white roof, red roof, and concrete—covering the bulk of the test site were determined and identified in this study.

The image is freely available and was downloaded from DigitalGlobe (https://www.digitalglobe.com/resources/product-samples). The panchromatic and multispectral bands of the WorldView-2 image were resampled using Gram–Schmidt panchromatic sharpening, and subsequent analyses were performed on the 0.5-m pan-sharpened imagery.

Methodology
The Taguchi statistical optimization technique was used to optimize MRS parameters and minimize the number of experiments to extract buildings from a WorldView-2 image using object-based image analysis. The main processing steps of the method are given as a flowchart in Figure 2. In the first stage, preprocessing and pan-sharpening operations were applied to the WorldView-2 image. Object-based image classification was performed in two steps. Segmentation was the first process, which determines the image-object size. In the second stage, the Taguchi optimization technique was used to set the optimum parameter combination, which was used in the classification of the WorldView-2 imagery (stage 3). In the final stage, accuracy was assessed using the confusion matrix.

Multi-Resolution Image Segmentation (MRS)
The MRS algorithm is a region-based algorithm based on local homogeneity criteria. It starts with a single pixel and collects...
pixels of different shapes, sizes, and properties in the form of image objects until it reaches a user-defined homogeneity level or threshold (Baatz and Schäpe 2000). Thus, the maximum heterogeneity allowed for the created image objects is determined. The MRS process has three main parameters: scale, shape, and compactness. The scale parameter is considered as the most effective factor controlling the average image-object size (Kavzoglu et al. 2017). It determines the maximum heterogeneity allowed for generated image objects. As the scale value increases, the image-object size also grows. The shape parameter affects the separation of LULC classes depending on the color and texture information, and the compactness parameter helps to define the image-object boundaries as sharper or softer (Kavzoglu et al. 2017). Values of shape and compactness parameters range between 0 and 1.

### Taguchi Optimization Technique

Taguchi optimization (Taguchi 1986) is fundamentally based on the concept of the quality loss function. It has been mostly used to optimize design parameters and minimize testing time and experimental costs for engineering research and applications (Wang and Huang 2007). The use of a Taguchi orthogonal array reduces the number of possible combinations of experiments to a limited number and makes the experimental design easier and more consistent. Taguchi uses $L_a(b^c)$ notation for naming the orthogonal arrays, where $a$ is the total number of experimental runs, $b$ is the number of levels for each factor, and $c$ is the number of variables used in the experiment. It is applied to examine the effect on segmentation accuracy not only of each parameter but also of each pair of parameters, as well as to optimize the parameters by considering the interaction effect. The main steps of Taguchi optimization can be given in order as follows (Chen et al. 1996):

1. Detection of factors and interactions
2. Determination of the levels of each factor
3. Selection of an appropriate orthogonal matrix
4. Transfer of factors and interactions to columns of orthogonal matrices
5. Conducting experiments
6. Analysis of data and determination of optimal levels
7. Validation tests

Taguchi optimization utilizes a loss function to analyze the deviation between the experimental values and the desired values. This function is further relocated into a signal-to-noise ratio (SNR; Mandal et al. 2011; Kvak 2014), a logarithmic function of the desired outcome of the Taguchi method that provides an optimization measure with constraints and variables that should be minimized or maximized. There are three SNR functions with specific characteristics: smaller-the-best, larger-the-best, and nominal-the-best. The objective of this article is to determine the maximum POF and Qr. Therefore, the larger-the-better characteristic was selected:

$$\text{SNR} = -10\log_{10}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{1}{y_i}\right),$$  \hspace{2cm} (1)

where $y_i$ is the observed data in the experiment and $n$ is the number of experiments (Mandal et al. 2011).

### Design of Experiment

The Taguchi technique was utilized to design experiments for three parameters (scale, shape, compactness) with five defined levels. Based on the level combinations in orthogonal arrays, POF and Qr values were calculated for each experimental design. The objective function $F$ measures the within-segment variance $V$ and the between-segments autocorrelation index (Moran’s $I$ index) (Espindola et al. 2006):

$$V = \frac{\sum_{i=1}^n a_i \cdot v_i}{\sum_{i=1}^n a_i},$$  \hspace{2cm} (2)

where $v_i$ denotes the variance of segment $i$ and $a_i$ shows the area of this segment. Moran’s $I$ index is calculated as

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (\bar{y}_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (\bar{y}_i - \bar{y})^2 \left( \sum_{i=1}^n \sum_{j=1}^n w_{ij} \right)},$$  \hspace{2cm} (3)

where $n$ represents the total number of segments; $w_{ij}$ is the spatial proximity measure, assessing the spatial adjacency of regions $R_i$ and $R_j$; $\bar{y}_i$ is the mean spectral value of region $R_i$; and $\bar{y}$ denotes the mean spectral value of the image. The value of Moran’s $I$ index varies between $-1$ (dispersed) and $+1$ (clustered). A value of 0 indicates perfect randomness (i.e., no correlation). The presented objective function incorporates the variance and the autocorrelation in the objective function.
\[ F_{\text{plateau}} = F(v) + F(I). \]  

(4)

where \(F(v)\) and \(F(I)\) are normalized functions, expressed as

\[ F(x) = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}. \]  

(5)

The POF metric is a combination of intrasegment variance and spatial autocorrelation (Martha et al. 2011). According to Johnson and Xie (2011), the POF metric is a good indicator in the assessment of segmentation quality. The second measure used in this study is Qr, proposed by Winter (2000). Many studies mention that Qr is highly effective in segmentation quality-assessment studies (e.g., Liu et al. 2012; Belgiu and Drăguţ 2014; Kazvoglu and Tonbul 2018). The Qr metric measures the overlap between reference objects (polygons) corresponding to segments and is calculated as

\[ Qr = \frac{A_{(o)} \bigcap A_{(g)}}{A_{(o)} \bigcup A_{(g)}} \]  

(6)

where \(A_{(o)}\) shows the total area of reference objects and \(A_{(g)}\) illustrates the total area of corresponding created segments. Values of Qr vary in the range of \([0, 1]\); for perfect segmentation, it should be 1. In order to investigate the performance of the Qr metric, 30 manually digitized reference polygons were identified, including different types of buildings.

**Results**

In this study, the statistical Taguchi optimization technique was used to find optimal MRS parameters (i.e., scale, shape, compactness). The MRS parameters determined at five levels are given in Table 1. In order to determine the optimum MRS parameters in a wider range, the \(L_{16}(5^3)\) orthogonal array (three factors each, with five levels, and a total of 25 experiments), which is also frequently used in the literature, was preferred. It should be emphasized that with the use of Taguchi optimization, the number of possible combinations is reduced from 125 (i.e., \(5^3\)) to 25, and then the parameter optimization process is initiated. The levels defined in Table 1 are determined to prevent over- and undersegmentation according to the size of the objects in the image. In other words, scale values below 10 were exposed to oversegmentation and those above 30 were exposed to undersegmentation, considering LULC characteristics in the study area.

The optimization of the MRS parameters was evaluated separately using POF and Qr values. The larger-the-better performance characteristic for POF and Qr was implemented to acquire the optimal MRS parameters. Consequently, the SNR should be maximized. The best value for each parameter was estimated according to the largest SNR obtained at all levels for that parameter. Segmentation results for the 25 combinations of MRS parameters were estimated by calculating their POF, Qr, and SNR values, as shown in Table 2.

The effect of each parameter (scale, shape, compactness) on the POF and Qr was analyzed using the SNR response table (Table 3). This table, estimated using Taguchi optimization, shows the optimum controlling factor levels for POF and Qr. The best level for each controlling factor was determined by considering the highest SNR levels of the controlling factor. It should be noted that the bold values in Table 3 represent the optimum MRS parameter levels for POF and Qr. Delta values express the difference between the highest and lowest average response values for each factor, and rank is generated from the delta values.

The SNR values for controlling factors based on POF and Qr are also given in Figure 3 for better visual assessment of changes in mean SNR values and observation of the optimal shape, scale, and compactness values.

Optimum MRS parameters for maximizing POF and Qr can be straightforwardly observed in these graphs. The optimum levels for scale, shape, and compactness were determined as 10–0.1–0.7 for POF (Figure 3a) and 10–0.1–0.9 for Qr (Figure 3b). An important result is that the same scale and shape parameters were estimated with both POF and Qr. However, the compactness parameter was calculated at different values, indicating a slight difference, particularly considering that the compactness parameter is regarded as the least effective one.

<table>
<thead>
<tr>
<th>Level</th>
<th>Scale</th>
<th>Shape</th>
<th>Compactness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>0.9</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 1. Levels designed for segmentation parameters used in optimization process.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Scale</th>
<th>Shape</th>
<th>Compactness</th>
<th>POF</th>
<th>SNR for POF</th>
<th>Qr</th>
<th>SNR for Qr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>0.1</td>
<td>0.1</td>
<td>1.000</td>
<td>0.000</td>
<td>0.910</td>
<td>-0.817</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>0.3</td>
<td>0.3</td>
<td>0.980</td>
<td>-0.178</td>
<td>0.905</td>
<td>-0.871</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>0.5</td>
<td>0.5</td>
<td>0.956</td>
<td>-0.387</td>
<td>0.898</td>
<td>-0.933</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>0.7</td>
<td>0.7</td>
<td>0.933</td>
<td>-0.605</td>
<td>0.896</td>
<td>-0.955</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>0.9</td>
<td>0.9</td>
<td>0.897</td>
<td>-0.944</td>
<td>0.879</td>
<td>-1.124</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>0.1</td>
<td>0.3</td>
<td>0.914</td>
<td>-0.779</td>
<td>0.895</td>
<td>-0.959</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>0.3</td>
<td>0.5</td>
<td>0.911</td>
<td>-0.810</td>
<td>0.892</td>
<td>-0.990</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>0.5</td>
<td>0.7</td>
<td>0.872</td>
<td>-1.189</td>
<td>0.883</td>
<td>-1.076</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>0.7</td>
<td>0.9</td>
<td>0.867</td>
<td>-1.235</td>
<td>0.877</td>
<td>-1.141</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>0.9</td>
<td>0.1</td>
<td>0.844</td>
<td>-1.472</td>
<td>0.843</td>
<td>-1.485</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>0.1</td>
<td>0.5</td>
<td>0.839</td>
<td>-1.526</td>
<td>0.873</td>
<td>-1.185</td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>0.3</td>
<td>0.7</td>
<td>0.848</td>
<td>-1.430</td>
<td>0.879</td>
<td>-1.129</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>0.5</td>
<td>0.9</td>
<td>0.839</td>
<td>-1.526</td>
<td>0.874</td>
<td>-1.172</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
<td>0.7</td>
<td>0.1</td>
<td>0.775</td>
<td>-2.217</td>
<td>0.850</td>
<td>-1.416</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
<td>0.9</td>
<td>0.3</td>
<td>0.829</td>
<td>-1.625</td>
<td>0.827</td>
<td>-1.649</td>
</tr>
<tr>
<td>16</td>
<td>25</td>
<td>0.1</td>
<td>0.7</td>
<td>0.776</td>
<td>-2.202</td>
<td>0.851</td>
<td>-1.403</td>
</tr>
<tr>
<td>17</td>
<td>25</td>
<td>0.3</td>
<td>0.9</td>
<td>0.787</td>
<td>-2.082</td>
<td>0.856</td>
<td>-1.356</td>
</tr>
<tr>
<td>18</td>
<td>25</td>
<td>0.5</td>
<td>0.1</td>
<td>0.730</td>
<td>-2.734</td>
<td>0.842</td>
<td>-1.496</td>
</tr>
<tr>
<td>19</td>
<td>25</td>
<td>0.7</td>
<td>0.3</td>
<td>0.716</td>
<td>-2.903</td>
<td>0.838</td>
<td>-1.536</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>0.9</td>
<td>0.5</td>
<td>0.720</td>
<td>-2.853</td>
<td>0.787</td>
<td>-2.077</td>
</tr>
<tr>
<td>21</td>
<td>30</td>
<td>0.1</td>
<td>0.7</td>
<td>0.730</td>
<td>-2.728</td>
<td>0.845</td>
<td>-1.464</td>
</tr>
<tr>
<td>22</td>
<td>30</td>
<td>0.3</td>
<td>0.9</td>
<td>0.727</td>
<td>-2.775</td>
<td>0.840</td>
<td>-1.516</td>
</tr>
<tr>
<td>23</td>
<td>30</td>
<td>0.5</td>
<td>0.3</td>
<td>0.734</td>
<td>-2.682</td>
<td>0.846</td>
<td>-1.449</td>
</tr>
<tr>
<td>24</td>
<td>30</td>
<td>0.7</td>
<td>0.5</td>
<td>0.692</td>
<td>-3.194</td>
<td>0.818</td>
<td>-1.742</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>0.9</td>
<td>0.7</td>
<td>0.795</td>
<td>-1.993</td>
<td>0.748</td>
<td>-2.521</td>
</tr>
</tbody>
</table>

Table 2. Signal-to-noise-ratio (SNR) values estimated from plateau objective function (POF) and quality rate (Qr).

<table>
<thead>
<tr>
<th>Level</th>
<th>Scale</th>
<th>Shape</th>
<th>Compactness</th>
<th>POF</th>
<th>SNR for POF</th>
<th>Qr</th>
<th>SNR for Qr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.423</td>
<td>-1.447</td>
<td>-1.840</td>
<td>-0.940</td>
<td>-1.166</td>
<td>-1.346</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1.097</td>
<td>-1.455</td>
<td>-1.634</td>
<td>-1.130</td>
<td>-1.171</td>
<td>-1.293</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1.665</td>
<td>-1.704</td>
<td>-1.754</td>
<td>-1.308</td>
<td>-1.225</td>
<td>-1.385</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-2.355</td>
<td>-2.031</td>
<td>-1.484</td>
<td>-1.574</td>
<td>-1.356</td>
<td>-1.415</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-2.874</td>
<td>-1.771</td>
<td>-1.703</td>
<td>-1.738</td>
<td>-1.772</td>
<td>-1.251</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Signal-to-noise-ratio response for controlling factors using plateau objective function (POF) and quality rate (Qr).

| Boldface indicates the highest value for a parameter. |
Analysis of variance (ANOVA) was implemented to determine the statistical significance of the MRS parameters. ANOVA is a statistical method used to specify the discrete interactions of all controlling factors in the test design (Kivak 2014). In this study, ANOVA was used to investigate the contributions of the scale, shape, and compactness parameters on POF and Qr. The ANOVA results are presented in Table 4.

The percentage contributions of scale, shape, and compactness were computed as 87.45%, 5.67%, and 1.72%, respectively, for POF. It is clear that the scale parameter is by far the most predominant parameter according to the estimated POF values. The percentage contributions of scale, shape, and compactness were calculated as 56.87%, 35.12%, and 2.42%, respectively, for Qr. The percentage error is considerably low for both POF (5.16%) and Qr (5.59%). It is obvious that the scale parameter is the most important parameter for both POF and Qr. The shape parameter is the second most effective parameter.

Regression analyses are used to determine and analyze the relationship between a dependent variable and multiple independent variables. In this study, POF and Qr values were considered dependent variables, and MRS parameters independent variables. Regression analysis was used to estimate predictive equations for POF and Qr. The predictive equations were performed for linear regression (LR) and support vector regression (SVR). The predictive equations acquired by the LR model for POF and Qr are

\[
P_{OF} = 1.0807 - 0.011418 \times \text{Scale} - 0.0618 \times \text{Shape} + 0.0141 \times \text{Compactness} \\
Q_{r} = 0.9697 - 0.003991 \times \text{Scale} - 0.0672 \times \text{Shape} + 0.0036 \times \text{Compactness}.\]

The LR model computed \( R^2 \) values of 88.73% and 84.93%, respectively, for POF and Qr. Figure 4 shows the assessment of estimated test results and predicted values with the LR model.

Predictive equations of POF and Qr were also established using the SVR model:

![Figure 3. The effect of multi-resolution segmentation parameters on signal-to-noise-ratio response characteristics using (a) plateau objective function (POF) and (b) quality rate (Qr).](image_url)

![Table 4. Analysis-of-variance results for controlling factors using plateau objective function (POF) and quality rate (Qr).](table_url)
The SVR model calculated $R^2$ values of 88.02% and 83.05%, respectively, for POF and Qr. Figure 5 shows the relationship between estimated test results and values predicted by the SVR model. As shown in Figures 4 and 5, there is good adherence between the models and the experimental data for POF and Qr.

In Taguchi optimization, a confirmation test is applied to eliminate concerns about the selection of controlling parameters or assumptions of responses (Co 2008; Mandal et al. 2011). In this study, confirmation tests were implemented for optimum level and random level of controlling factors. Optimal POF ($POF_{opt}$) and Qr ($Qr_{opt}$) values were evaluated as

$$POF_{opt} = (Sc_1 - Av_{POF}) + (Sh_1 - Av_{POF}) + (C_4 - Av_{POF}) + Av_{POF}$$

(11)

$$Qr_{opt} = (Sc_1 - Av_{Qr}) + (Sh_1 - Av_{Qr}) + (C_5 - Av_{Qr}) + Av_{Qr},$$

(12)

where $Sc_i$ indicates the first scale level; $Sh_i$ shows the first shape level; $C_i$ and $C_j$ show the fourth and fifth compactness levels; and $Av_{POF}$ and $Av_{Qr}$ indicate the average values of all estimated POF and Qr values (Table 5; highest values in bold). Afterward, $POF_{opt}$ and $Qr_{opt}$ were estimated, using Taguchi optimization, as 0.993 and 0.922, respectively (Table 6).

Table 6 illustrates the comparison of the test results and the values predicted by Taguchi optimization and the LR and SVR regression equations (Equations 7–10).
It is clear that minimum errors were calculated for Taguchi optimization, and it achieved higher accuracy compared to the LR and SVR models. The maximum percentages of error calculated with SVR model prediction for POF and Qr were estimated as 4.10% and 4.92%, respectively. Predicted and experimental values are also quite close to each other for Taguchi optimization. According to Yıldırım, Kivak, and Erzincanlı (2019), the estimated error should be less than 20% to obtain reliable statistical analysis. The results presented in Table 6 clearly indicate that the error values estimated by Taguchi and the LR and SVR models are well below that acceptable limit.

In order to analyze the effect of parameter selection on classification accuracy, object-based image classification was performed on the segmented images. The classification process was carried out using the optimum MRS parameters obtained from the Taguchi optimization technique for POF and Qr values. The random-forest classifier, which is a popular ensemble method, was used in the classification process. In total, 43 spectral and geometric features (i.e., mean, minimum, maximum, standard deviation, ratio values of objects at all bands, normalized difference vegetation index for two near-infrared bands, and area-of-pixel values) were considered in the classification process. It should be noted that the weights of all bands were set to 1. In order to evaluate the accuracy of classification, assessment was performed using confusion matrices considering validation data sets consisting of 10 000 randomly selected pixels per class (Table 7). The results of the two classification performances are very close: Overall accuracies of 91.56% (k = 0.90) and 90.77% (k = 0.89) were achieved using POF and Qr values, respectively, based on optimal MRS parameters. The accuracy differences between the two approaches were less than 1%. Image objects of the “red roof” class were well separated from other LULC classes, giving over 94% producer and user accuracies for both approaches. The highest user accuracy was estimated for the “white roof” class, at 99.40%, using the segments constructed by considering Qr values. The worst user accuracy (78.38%) was calculated for the “concrete” class considering Qr values.

The classified images produced using optimum MRS parameters determined by the Taguchi optimization technique based on POF and Qr values are presented in Figure 6. Some conclusions can be easily drawn from the thematic maps even after a visual assessment. First, it can be seen that Figure 6a and 6b shows great resemblance between the two thematic maps. However, there are some parts of the maps that were classified differently—particularly, some spectrally similar classes (e.g., concrete and asphalt road) were misclassified in some parts of the thematic LULC maps. Results also show that optimization of MRS parameters based on the Taguchi optimization technique provided good separation of buildings and other impervious surface boundaries.

**Discussion**

Given that the accuracy of OBIA depends on the quality of segmentation, robust Taguchi optimization was performed in this article to optimize MRS parameters in a heterogeneous urban area to extract buildings more accurately. In the literature, the

<table>
<thead>
<tr>
<th>Land Use/Land Cover Class</th>
<th>POF Producer</th>
<th>POF User</th>
<th>Qr Producer</th>
<th>Qr User</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asphalt road</td>
<td>0.88</td>
<td>0.87</td>
<td>0.85</td>
<td>0.87</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.85</td>
<td>0.98</td>
<td>0.83</td>
<td>0.97</td>
</tr>
<tr>
<td>Concrete</td>
<td>0.92</td>
<td>0.80</td>
<td>0.92</td>
<td>0.78</td>
</tr>
<tr>
<td>Forest</td>
<td>0.99</td>
<td>0.90</td>
<td>0.98</td>
<td>0.89</td>
</tr>
<tr>
<td>Red roof</td>
<td>0.96</td>
<td>0.97</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td>White roof</td>
<td>0.89</td>
<td>0.99</td>
<td>0.90</td>
<td>0.99</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
<td>91.56</td>
<td>90.77</td>
<td>k = 0.90</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Figure 6. Classification results for image objects generated by segmentation parameters determined by Taguchi optimization considering (a) plateau objective function (POF) and (b) quality rate (Qr).
The determination of optimum MRS parameters has usually been carried out by determining a single parameter, whereas Taguchi optimization stands out to the fore with optimum value determination for all MRS parameters. Furthermore, Taguchi optimization is easy to apply and has proven to be effective in determining the appropriate segmentation parameters of MRS (i.e., scale, shape, and compactness) with a limited number of trials.

In previous studies, the segmentation-quality metric of POF has been generally used to determine optimum MRS parameters using Taguchi optimization (Moosavi et al. 2014; Idrees and Pradhan 2016; Idrees et al. 2016; Gibril et al. 2017). The main difference that distinguishes this study from others is the analysis of supervised (Qr) and unsupervised (POF) segmentation quality metrics in determining the optimum MRS parameters. Taguchi optimization based on SNR function systematically analyzed the set of input variables (i.e., POF and Qr) to obtain optimum factor levels for scale, shape, and compactness parameters. The ANOVA test conducted to determine the importance of the MRS parameters showed that scale was the most important parameter for both POF and Qr. In addition to Taguchi optimization for confirmation tests of control factors, optimal and random levels were performed for SVR and LR equations. While regression results were found to be acceptable, Taguchi optimization yielded better results compared to regression (Table 6).

**Conclusion**

The Taguchi optimization technique is an effective approach for parameter evaluation and optimization. In this article, the problem of optimum parameter selection for the MRS algorithm was studied using Taguchi optimization. The segmentation process was implemented on a WorldView-2 satellite image using the MRS algorithm embedded in eCognition Developer software. In order to evaluate the quality of segmentation for different scale, shape, and compactness parameters, supervised (Qr) and unsupervised (POF) quality metrics were used as fitness criteria for optimization. The experimental results were then assessed using ANOVA. The following conclusions can be drawn from our results:

- The optimal values of the controlling factors (i.e., scale, shape, and compactness) for maximizing POF and Qr using the SNR were estimated using Taguchi optimization. The optimum setting for MRS parameters using POF and Qr values, respectively, were estimated as 10–0.1–0.7 and 10–0.1–0.9 for scale, shape, and compactness. The optimum values estimated for scale and shape were the same for both POF and Qr; however, the compactness parameter was slightly different (0.7 for POF and 0.9 for Qr).
- According to the ANOVA, scale was the most significant parameter for POF and Qr, with respective contributions of 87.45% and 56.87%. Shape was the second most effective parameter, and compactness had a very limited effect on the creation of image objects.
- The Taguchi confirmation test revealed a very good relationship, with high correlation coefficients (POF$_{opt}$ = 0.993 and Qr$_{opt}$ = 0.922) between the estimated and predicted values. Compared to the correlation results of LR and SVR, Taguchi optimization definitely produced better results.
- Predicted values of POF and Qr using LR and SVR were also close to the experimental values, in that error values estimated by Taguchi, LR, and SVR models were all less than 5%. This means that the regression models all performed well in predicting optimal MRS parameters.
- The overall accuracy obtained using the optimum MRS parameters was 91.56% for POF and 90.77% for Qr. Overall accuracy higher than 90% indicates accurate classification with well-estimated segmentation parameters using Taguchi optimization.

The results of this research support the idea that Taguchi optimization is an effective technique for optimizing MRS parameters. Therefore, it can be concluded that Taguchi optimization is a good alternative to other regression methods, in this case LR and SVR. However, the reliability of Taguchi optimization for the determination of optimal parameters should be investigated in future studies by using different data sets with different spectral and spatial characteristics. Further research may also focus on analyzing the effect of other segmentation-quality metrics together with the parameterization of machine-learning algorithms used in remote sensing studies.

**Acknowledgments**

The WorldView-2 image data used in this study were obtained from DigitalGlobe as an online product sample, which is only authorized for research and educational purposes.

**References**


Photogrammetric Engineering and Remote Sensing (PE&RS) is seeking submissions for a special issue on Urban Remote Sensing.

The formulation of the 17 Sustainable Development Goals (SDGs) is a major leap towards humankind’s quest for sustainability. In recent decades, global urban areas have been rapidly expanding, especially in developing countries. The prospect is that the urbanization rate will reach 60% by 2030. Urban expansion will inevitably increase vulnerability to natural hazards, natural vegetation cover decline and arable land loss, urban heat islands, air pollution, hydrological circle alteration and biotic homogenization. Since urban ecosystems are strongly influenced by anthropogenic activities, a considerable amount of research has been conducted all around the world to understand the spatial patterns, driving forces and the ecological and social consequences of urbanization. It is not only crucial for characterizing the ecological consequences of urbanization but also for developing effective economic, social and environmental policies in order to mitigate its adverse impacts.

Remote sensing has been widely used for investigating urban environment and the associated drivers during the urbanization process, as it can quickly and frequently monitor large area surface change with lower cost, compared to filed survey or in situ measurements. Digital archives of remotely sensed data provide an excellent opportunity to study historical urban changes and to relate their spatio-temporal patterns to environmental and human factors. With the rapid development of Earth observation techniques, it has become convenient to obtain a large number of remotely-sensed imagery over a certain area at different times, from hundreds of Earth observation platforms. However, this brings challenges to researchers to timely process the remote sensing big data as well as to rapidly transfer the data into information and knowledge.

Considering this, this special issue of PE&RS is aimed at reporting novel studies on exploiting remote sensing big data to monitor and improve urban environment, and showing the potential of remote sensing in developing sustainable cities, including but not limited to:

- Urban remote sensing big data
- Remote sensing information interpretation
- Urban expansion, dynamics and associated environment consequences
- Remote sensing of urban water quality
- Remote sensing of urban thermal environment
- Remote sensing of urban geological environment
- Urban sustainability assessment
- Urban sustainable development
- Urban Spatiotemporal analysis
- Urban Sustainability Indicators
- Urban environmental Monitoring

Papers must be original contributions, not previously published or submitted to other journals. Submissions based on previous published or submitted conference papers may be considered provided they are considerably improved and extended. Papers must follow the instructions for authors at http://asprs-pers.edmgr.com/.

**Important Dates**
- July 1, 2020   Submission system opening
- October 31, 2020   Submission system closing
- Planned publication date: Dec. 2020

**Guest Editors**

**Zhenfeng Shao, Wuhan University, China**
Prof. Zhenfeng Shao, Professor at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. His research interests include urban remote sensing. He is now an associate editor of Email: shaozhenfeng@whu.edu.cn.

**Orhan Altan, Istanbul Technical University, Turkey**
Prof. Orhan ALTAN, Professor at the Department of Geomatics Engineering, Istanbul Technical University, Turkey. He is Past President and Honorary Member of ISPRS, Honorary Member of Science Academy. He has published more than 200 scientific papers in scientific journals and conferences, and editor or co-editor of more than 20 international books. Email: oaltan@itu.edu.tr

**J.L.van Genderen, University of Twente, Netherlands**
Professor J.L. van Genderen, Prof. at the Department of Earth Observation Science, Faculty of Geo-information Science and Earth Observation(ITC),University of Twente, Netherlands, he is an associate Editor of Geo-Spatial Information Science. His research interests include photogrammetry and remote sensing, urban remote sensing and computer vision. Email: Genderen@alumni.itc.nl.
Precise Extraction of Citrus Fruit Trees from a Digital Surface Model Using a Unified Strategy: Detection, Delineation, and Clustering

Ali Ozgun Ok and Asli Ozdarici-Ok

Abstract
In this study, we present an original unified strategy for the precise extraction of individual citrus fruit trees from single digital surface model (DSM) input data. A probabilistic method combining the circular shape information with the knowledge of the local maxima in the DSM has been used for the detection of the candidate trees. An active contour is applied within each detected region to extract the borders of the objects. Thereafter, all extracted objects are seamlessly divided into clusters considering a new feature data set formed by (1) the properties of trees, (2) planting parameters, and (3) neighborhood relations. This original clustering stage has led to two new contributions: (1) particular objects or clustered structures having distinctive characters and relationships other than the citrus objects can be identified and eliminated, and (2) the information revealed by clustering can be used to recover missing citrus objects within and/or nearby each cluster. The main finding of this research is that a successful clustering can provide valuable input for identifying incorrect and missing information in terms of citrus tree extraction. The proposed strategy is validated in eight test sites selected from the northern part of Mersin province of Turkey. The results achieved are also compared with the state-of-the-art methods developed for tree extraction, and the success of the proposed unified strategy is clearly highlighted.

Introduction
Citrus orchards play an important role in fruit production in the world. As reported by the Food and Agriculture Organization (2017), in excess of 124 million tons of citrus products were cultivated in 2016. Today, various types of citrus trees (e.g., lemon, orange, grapefruit, tangerine, etc.) are cultivated on millions of acres in more than 50 countries. These statistics clearly indicate that the development of efficient strategies supplying reliable and up-to-date yield information is essential for citrus products worldwide.

Remote sensing combined with image processing knowledge plays a critical part in extracting individual trees and their coverages (Ozdarici-Ok 2015). The techniques developed so far mainly prefer to handle mainly three-dimensional point clouds either from LiDAR or dense image matching, particularly to generate a digital surface model (DSM) that is then converted to a normalized form (nDSM) or a crown height model (CHM) (Zhen et al. 2016). It is now a standard way to produce high-quality point clouds with very high levels of detail from unmanned aerial vehicle (UAV) image data sets with high overlaps. In addition, the use of UAVs for civil purposes is rapidly increasing since it offers an initial investment cost at affordable levels. Therefore, we favor a dense surface model generated from a UAV platform as a single input source.

In this study, we present a new unified strategy for the extraction of individual citrus fruit trees (Figure 1). A probabilistic method (Ok and Ozdarici-Ok 2018a) combining the circular shape of trees with the knowledge of local maxima (LM) in a DSM has been used for the detection stage. An active contour method with a negative contraction bias was applied considering each object region to extract the boundaries of the detected objects (Ok and Ozdarici-Ok 2018b). Thereafter, all (tree) objects are seamlessly divided into clusters considering (1) the properties of trees, (2) planting parameters, and (3) neighborhood relations. An agglomerative hierarchical

Figure 1. The proposed unified strategy.
clustering method was used during the clustering process. In this way, first, particular objects or clustered structures having distinctive characters and structural relationships other than the citrus objects were identified and eliminated. Second, the information revealed by clustering was used to recover missing citrus objects within and/or nearby each cluster.

Eight test sites were selected from the northern part of Mersin province containing one of the most productive citrus orchards of Turkey to evaluate the performance of the proposed approach. The overall pixel- and object-based F1-scores computed were 91.7% and 94.7%, respectively. The results were also compared with the results of four state-of-the-art methods proposed in the literature, and the assessments have shown that the developed approach can produce quite successful results by extracting the citrus trees from a single DSM.

Our newly proposed contributions in this research are provided below:

- This study is the first attempt proposing a unified strategy that mutually integrates the detection, delineation, and clustering of citrus trees. To our knowledge, previous efforts on this topic merely focused on the two primary stages, detection and/or delineation, and such a demanding clustering process has not been developed until this time.
- The original clustering stage involves a five-dimensional feature data set particularly designed to represent the (1) properties of trees, (2) planting parameters, and (3) neighborhood relations. To our knowledge, such a feature data set for the clustering of (citrus) trees in orchards has not been developed previously.
- We propose two new expansions of clustering: (1) objects revealing distinctive characters and structural relationships other than the citrus objects are identified and eliminated, and (2) information revealed by clustering is adapted to recover missing citrus objects within and/or nearby each cluster. In this way, we greatly benefit from the clustering results, and prove that a successful clustering can provide valuable input for identifying incorrect and missing information in terms of citrus tree extraction.

The remainder of this article is organized as follows. In the section “Related Work,” we summarize the previous work performed in this context. The section “The Proposed Unified Framework” provides the details of the proposed unified framework. The experiments related to our approach and the comparisons to the state of the art are presented in this section “Data set, Evaluation, and Selection of Parameters.” Finally, in the section “Results and Discussion,” we give concluding remarks and make suggestions for possible future works.

Related Work

In this study, we deeply investigate the previous attempts involving a DSM→DSM→CHM as an input data considering the topic tree extraction. For further information related to other approaches dealing with different data sets, refer to review articles (Fassnacht et al. 2016; Zhen et al. 2016) and/or comparative works conducted (Kaartinen et al. 2012; Vauhkonen et al. 2012; Jakubowski et al. 2013; Wallace et al. 2014; Dalponte et al. 2015b; Eysn et al. 2015). In this section, we first present the main strategy followed by each approach and then discuss its limitation(s).

Most of the previous approaches in the context of tree extraction benefit from a generic input—→DSM or, specifically, CHM—in which a pixel contains an elevation value above the ground to help the extraction of trees, such as in a forest area. Hyyppä et al. (2001a) presented the first unified attempt to derive single-tree information from laser scanner-derived CHM, and Hyyppä et al. (2001b) demonstrated for the first time that the tree heights extracted from an airborne laser scanning (ALS) campaign can successfully be used to initialize a segmentation process. Their strategy was based on two fundamental steps after a prefiltering step: seed point extraction and seeded region growing. However, the performance of their approach deteriorates in dense regions where single trees become hard to identify. In a different study, LM filtering was proposed to identify tree locations from a LiDAR-based CHM (Popescu and Wynne 2004). The algorithm was proposed to run with both circular and square window LM filters, with and without data fusion with optical data. In addition, to differentiate between deciduous and pine trees in the absence of optical information, the filter size was calculated based on the relationship between height and crown size. Nevertheless, the filter window size must be carefully selected to achieve successful results in the LM technique. Thereafter, spatial wavelet analysis technique (Falkowski et al. 2006) was proposed to better estimate the two critical parameters (tree height and crown diameter) based on the tree locations found using the LM technique proposed in Popescu and Wynne (2004).

In a different study, Koch et al. (2006) used a simple LM method considering the four-connective neighborhood of a pixel. Starting from the pixels labeled as LM, regions are formed and finally segmented. However, their automatic segmentation method underestimates the tree number significantly, especially for deciduous trees. In a similar context, a search for LM in a canopy maxima model with variable window sizes (Chen et al. 2006) was found to be a better method of detecting individual treetops in savanna woodland. In addition, marker-controlled watershed segmentation was applied to isolate individual trees. However, that method requires field data to train the parameters and is not valid for other types of trees in a forest.

Wolf and Heipke (2007) adapted both color-infrared aerial imagery and a DSM to the extraction and delineation of single trees. Their approach was based on a geometric and radiometric model of a tree and segmentation at multiple scales of DSM followed by scale selection and a refinement step using active contours. The GSD of the input DSM was found to be the main limiting factor of their approach, whereas the trees exhibiting more than one large crown and dense environments having partial occlusions or noncircular crown shapes constitute the other important factors influencing their results. The extended maxima transformation of CHM was proposed to detect treetops in Kwak et al. (2007) followed by a watershed segmentation for the delineation step. However, the maxima transform loses efficiency if small conical swellings exist on a crown surface in which different parameters are required to maximize the performance of different types of trees.

Multiple stereo aerial images and ALS data were used to generate a CHM in Hirschmugl et al. (2007). Thereafter, aerial images and CHM were used to delineate trees with a rule-based seeded region growing. Their results confirm that dominant trees are more likely to be detected than trees of intermediate or lower layers, and manual processing is suggested to remove superfluous seeds. A new feature set extraction approach for tree species classification was presented in Puttonen et al. (2010). The approach was based on dividing a single tree crown into illuminated and shaded parts with the help of a CHM and then using averaged color values in digital aerial images as classification features. However, the improvements of the approach are computed to be in the order of a few percentage points in overall classification accuracy.

Prior knowledge was introduced into single-tree extraction in Heinzel et al. (2011). An optical image is classified with a pixel-based strategy (mathematical morphology based on iterative granulometry) whose number of classes was defined
manually, and the classification results were used as prior information for the delineation of the trees with the approach in Koch et al. (2006), where CHMs were used as input. Nonetheless, problems are observed in cases where no clear morphological boundary exists, like overtopped trees or highly interlocked crowns. An adaptive method based on a Poisson forest stand model applied to a CHM generated from laser scanning data was proposed in Ene et al. (2012). In their work, two Poisson methods were developed based on the assumption that the tree locations are generated by a homogeneous spatial Poisson process. However, the key factors controlling the number of trees detected are stated to be the GSD of the CHM and the size of the low-pass filter performed. In addition, the tree heights were systematically underestimated. LM and marker-controlled watershed segmentation over CHM generated from a synthetic LiDAR data set was revisited in Wang et al. (2013). They found a consistent error toward the underestimation of crown diameters due to the nature of LiDAR methods. However, the approach is not tested with real data sets, and their synthetic data rely on multiple assumptions, including the regular distribution of trees in the stand, symmetric shape of crown, and flat terrain.

A scale-adaptive region-growing method based on DSM was proposed in Palenichka et al. (2013). In that study, LM of the multi-scale, multi-component isotropic attention operator was used to extract interest points, and a scale-adaptive region-growing algorithm was developed for the pixelwise segmentation of forest areas. Finally, the segmented forest areas are analyzed at low scale range, which corresponds to tree crown sizes. However, the approach is tested with a single test site, and it is not known whether the parameters selected for that site are transferrable across different types of trees or sites. A variable-area local maxima (VLM) algorithm that incorporates predictions of the metabolic scaling theory for processing CHMs was proposed in Swetnam and Falk (2014). A k-nearest neighbor search was adapted as an allometric rule of the VLM to filter out spurious LMs and to better identify neighboring trees. The authors stated that the accuracy of the VLM varied by forest type and by cover percentage. In addition, representing each tree with a circle may underestimate the area computed for individual trees in a forest.

A supervised strategy to delineate trees from LiDAR data was proposed by Liu et al. (2015). In their approach, a fishing net dragging method based on watershed segmentation was proposed for crown boundary refinement, and random forest classification was then used to separate individual trees and their branches. The CHM and two pseudowaveforms were used as features in the classification. Because the approach is still based on the watershed segmentation, it carries the same limitations of their predecessors, such as not working well over areas where the valley shape between trees is not “V” or “U” shaped or resulting in merged tree crowns.

Another study proposed a two-step marker-controlled region-growing algorithm for individual tree crown delineation by combining CHM and orthoimagery (Zhen et al. 2014). The results confirm that additional information from orthoimagery reduces errors occurring due to small trees; however, the selection of height threshold is critical to balance commission and omission errors. In a different work, an agent-based region-growing approach was proposed by Zhen et al. (2015) for tree detection using LiDAR-based CHM considering both growth and competition mechanisms. As the marker-controlled region growing was involved in their work, the success of the agent-based approach still depends largely on the selection of treetops and the thresholds in the growth procedure.

In a comparative study, Dalponte et al. (2015b) proposed a tree delineation approach fixing LMs within a CHM as treetops and using a decision tree method to grow individual crowns around the LMs. Their comparative results showed that the delineation methods tested provided very different results, and the best performance was influenced by the accuracy statistic computed. Matsuki et al. (2015) combined CHM with hyperspectral data during the classification of 16 tree species over the Tama Forest Science Garden in Tokyo, Japan. The tree crown features (height, size, and shape) were extracted using the region growing based on LMs detected from CHM and fused with spectral features during support vector machine (SVM) classification. The approach suffers from the same problems using the LM-based strategy; that is, small trees are overlooked, and multiple LMs are detected over a single large tree. In Heenkenda et al. (2015), mangrove treetops were detected using LMs of a DSM, and the delineation of trees was done using a region-growing method that was tested on different data layer combinations from high-resolution optical images. Nevertheless, the authors argued that the imagery must be separated into homogeneous species stands for the best results, and the validation of the detected locations of treetops is recommended before applying the region-growing algorithm.

A semisupervised SVM classifier was presented in Dalponte et al. (2014, 2015a) to extract individual trees from hyperspectral and laser scanning data. In their approach, the automated delineation of the trees was performed using the approach in Ene et al. (2012), and they stated that the semisupervised classification methods can be very useful in the framework of tree species classification. However, the tasks required for the integrated classification, such as coregistration of ALS and hyperspectral data, and relating the tree crowns detected with the field measured trees used in the training phase may have negative effects on the classification results if not properly performed.

A hierarchical approach to the three-dimensional segmentation of single trees in a forest area using LiDAR data was presented in Paris et al. (2016). That approach combines the raster CHM and LiDAR point cloud to better identify and delineate the tree crowns, and it was shown that the detection rate of the tree crowns was improved with respect to the state of the art. However, that approach still requires the original LiDAR point cloud as an input other than the CHM. A raster slicing method applied to CHM with user-defined data ranges (slice heights) was presented in Hadaš and Estornell (2016). The optimal slice height was chosen as the one where the standard deviation of the centroid distances between the detected and the reference data was the smallest. However, the approach was not tested with trees in a dense environment. An orthophoto and aDSM were jointly used in Bulatov et al. (2016) to detect tree crowns. Their approach benefits from a classification result and includes several steps, including the correction of DTM and the detection and classification of individual trees. However, the approach involves empirical thresholds that need to be rigorously tested in different environments. CHMs derived from ALS data and UAV images for the extraction of individual trees were compared in Li et al. (2016). Based on the results, the UAV-based CHMs, generated using the ALS-derived DTM, can achieve better performance in delineating individual tree attributes than ALS data alone. Mohan et al. (2017) tested automatic individual tree detection using an LM-based algorithm on UAV-derived CHMs. They evaluated the impacts of two window sizes on the performance of the LM algorithm and concluded that the algorithm tested needs to be continuously improved to adapt different scenarios of tree canopies. In a more recent work, Koc-San et al. (2018) proposed an approach based on a circular Hough transform to extract citrus trees using UAV images and DSMS. However, their approach fails for the cases where the shadows of the trees are not visible in dense environments of citrus trees. Another
more recent work conducted by Selim et al. (2018) revisited
semiautomated object-based segmentation and classification
for the extraction of citrus trees from UAV images and a DSM.
However, their semiautomated approach requires significant
user-based input, and counting the trees in dense areas is
found to be inaccurate.

**The Proposed Unified Framework**

The first step is formed by a detection phase that combines
the probabilistic information of LM of DSM together with
the circular shape structure of trees. The second step is the extraction of boundaries of the detected tree objects. Once the coverage of each tree object has been identified, they are separated into clusters, depending on certain key parameters (i.e., tree characteristics, planting parameters, and neighboring relations). We benefit from the clustering results in two ways. First, particular objects or clustered structures other than the citrus objects are identified and eliminated. Second, missing citrus objects within and/or nearby each cluster are recovered.

**Detection of Candidate Tree Locations**

The state-of-the-art approach used for the detection phase was developed in Ok and Ozdarici-Ok (2018a). Therefore, in this section, we only briefly go over the key parts of that approach.

The detection process starts with the calculation of probabilistic LM regions with the multi-level extended maxima transform, which benefits from an array of height thresholds (hi, i = 1, 2, . . . , n) to compute the probabilistic LM regions:

\[
P_{LM} = \left( \frac{\sum_{i=1}^{n} \text{regionalmax}_{i} \left[ \text{rec}(J, I) \right]}{n} \right)^{\frac{1}{2}}
\]  

(1)

In Equation 1, I and J denote the input DSM and the difference of threshold values (hi) from the DSM (DSM − hi), respectively, and the term rec(I, J) defines a grayscale reconstruction. The term regionalmax in refers to the collection of the regional maximum areas in the reconstructed DSM, while the term γ is a user-defined sensitivity parameter. This probabilistic information is adapted during the generation of the combined orientation symmetry image \(O^r\), which is based on the given bound (m) of radial symmetry for a given radius r:

\[
O^{r} = \sum_{i=r}^{m} \left( \frac{(O_{i})^{r}}{\max \{(O_{i})^{r}\}} \right) \cdot P_{LM} \cdot G_{\sigma}
\]  

(2)

In Equation 2, each level of radial symmetry is defined by \(a_{r}\), while \(O_{i}\) is the orientation image of the input DSM for each radius, \(R = \{r_{\min}, r_{\min+1}, \ldots, r_{\min}\}\). The term \(G_{\sigma}\) defines a two-di-

The delineation phase starts with the development of a reliable influence region of each candidate tree location detected in the previous section. For that purpose, initial boundaries of the influence regions have been identified by applying watershed transformation in view of the detected candidate locations. Thereafter, all irrelevant region boundary-

\[
\arg \min_{\lambda, \mu, \rho} \left\{ \int_{\text{inside}} I(p) - \mu \right\} + \frac{\rho}{\lambda^{2}} \left\{ \int_{\text{outside}} I(p) - \mu \right\} + \frac{\lambda}{\rho} \cdot \text{Len}(C) + \frac{\mu}{\lambda} \cdot \text{Area}(C)
\]  

In Equation 3, C defines the boundary of a closed contour, and \(\lambda, \mu, \rho\) are the average of pixel values inside and outside of C, respectively. The operators Len and Area represent the length of the contour and the area inside the contour, respectively. The parameters \(\lambda\) and \(\mu\) weight the inside and outside terms, respectively. The term \(\rho\) determines the level of smoothness of the contour, the term \(\sigma\) is the contraction bias that controls the tendency of the contour to grow outwards or shrink inward depending on the sign of the term (positive or negative), and I and p denote the image (in our case DSM) and their pixels, respectively.

This delineation process was followed by two postprocessing steps to avoid potential false alarms: (1) the elevation values under each delineated region were statistically evalu-

In this study, we propose a five-dimensional feature data set \(X = [X_1, X_2, \ldots, X_K]\) to be used in the clustering process of the K number of (tree) objects delineated in the previous section. The first feature set \(X = [r_1^1, r_1^2, \ldots, r_1^k; k = 1, 2, \ldots, K]\) is composed of the radii of the circles computed with the help of the number of pixels within each object delineated:

![Figure 2. Objects (trees) detected for test site 1. (a) UAV image; (b) DSM, the only input to the method (white = relatively high elevation, black = relatively low elevation); and (c) center positions of the detected objects (trees).](image-url)
In this study, we implemented agglomerative hierarchical clustering to reveal different clusters/group structures in a data set. In this way, we anticipate that the objects with relatively different radii (small vs. large) are assigned to different clusters.

The second feature set \( \mathbf{X}_k = \mathbf{d}^2_{ik} = [d^2_{i1}, d^2_{i2}, \ldots, d^2_{ik}]^T; k = 1, 2, \ldots, K \) considers the closest two-dimensional distances \( d^2_{ik} \) by considering the boundaries of the objects delineated:

\[
d^2_{ik} = \min_{\mathbf{m}_i} \left( d\left( \mathbf{m}_i, \mathbf{b}_j \right) \right)
\]

In Equation 5, \( \mathbf{b}_j \) is the \( j \)-th boundary pixel of an object \( C \) and \( \mathbf{m}_i \) indicates the \( i \)-th boundary pixel of any object \( M \). The two-dimensional proximity in image space, \( d(.) \), is computed by the Euclidean distance. We expect that this feature set will help to split up objects having different forms of canopy closures (dense/relatively close distance vs. sparse/relatively long distance).

The third feature set \( \mathbf{X}_c = \mathbf{d}^2_{ik} = [d^2_{i1}, d^2_{i2}, \ldots, d^2_{ik}]^T; k = 1, 2, \ldots, K \) is formed by taking into account the center locations of the objects delineated:

\[
d^2_{ik} = \min_{\mathbf{m}_i} \left( d\left( \mathbf{m}_i, \mathbf{c}_k \right) \right)
\]

In Equation 6, \( \mathbf{m}_i \) denotes all delineated objects except for \( C \) and \( \mathbf{c}_k \) indicates the \( k \)-th center location of an object \( C \). Once again, the two-dimensional proximity in image space, \( d(.) \), is computed by the Euclidean distance. We expect that this feature set will highlight the differences between objects having different planting patterns (frequent vs. regular).

The last two feature sets \( \mathbf{X}_x \) and \( \mathbf{X}_y \) define the positions of the center location of each object \( \mathbf{c}_k = [x_k, y_k] \) with respect to the image space (i.e., according to the top left corner):

\[
\mathbf{X}_x = \mathbf{x}^T_{\mathbf{c}} = [x^c, y^c, \ldots, x^c_K]^T
\]

\[
\mathbf{X}_y = \mathbf{y}^T_{\mathbf{c}} = [y^c, y^c, \ldots, y^c_K]^T
\]

In this way, the objects presenting the same feature information in \( \mathbf{X}_x, \mathbf{X}_y, \) and \( \mathbf{X}_c \) but located relatively far away from each other in the image space can be assigned to different clusters.

The clustering process refers to numerical methods applied to reveal different clusters/group structures in a data set. In this study, we implemented agglomerative hierarchical clustering during the clustering process for two reasons: (1) we believe that the feature data set we propose can be partitioned in a manner that would be relevant for interpretation as a concept of hierarchy, and (2) we would like to have meaningful partitions as close as a human operator.

Agglomerative hierarchical clustering starts with the calculation of the distances between each of the objects in a data set. In this way, similarities or discrepancies between each object are revealed. In this context, various distance measures can be used (e.g., Euclidean, standardized Euclidean, Manhattan, Mahalanobis, cosine, and Pearson correlation) (Duda et al. 2000). After the first step, the clustering starts by taking into account the distances; that is, each object is joined to the closest object in the feature set with the shortest distance (thus the most similar object). In the next steps, pairs of clusters are merged or agglomerated since each step after the first step may contain more than one object in a cluster. In this sense, the distances to be used in the computation must include the cluster similarities or differences, and different approaches for linking the clusters are proposed, such as single, average, complete, weighted average, and ward hierarchical (Duda et al. 2000). Finally, clustering is finalized once all data sets are merged under a single cluster. The output of the clustering of test image 1 is illustrated in Figure 4 (the information of the related parameters is provided in the section “Data set, Evaluation, and Selection of Parameters”).

**Promoting the Clustering Results**

The clustering results can provide valuable input for identifying incorrect and missing information in terms of citrus tree extraction. In this respect, we propose two new expansions of clustering: (1) objects revealing distinctive characters and structural relationships other than the citrus objects are identified and eliminated, and (2) information revealed by clustering is adapted to recover missing citrus objects within and/or nearby each cluster.

**Eliminating Improper Objects**

As can be seen from Figure 4c, although the clustering result can be regarded as successful in a general sense, there are still certain cases where individual objects are improperly assigned to a cluster, such as objects whose computed radii (i.e., using Equation 4) or canopy closure forms (i.e., using Equation 5) are significantly different from their surrounding objects (even if they are sewn in accordance with a planting row). Thus, a new split–merge–clean strategy that takes into account the planting characteristics of trees is proposed to mitigate these errors.

In the initial splitting stage, each cluster is handled separately to check whether all objects in a cluster satisfy a certain object-to-object minimum distance (i.e., \( \tau_{\text{min}} \)), computed as in Equation 5. If the distance of any object in a cluster is found to exceed \( \tau_{\text{min}} \), it is excluded from the cluster and allocated in a new one. In this way, we ensure that all object-to-object distances in each cluster are consistent and representative (Figure 5b). In the merging stage, starting from the cluster with the least number of objects, the labels of the objects in a close neighborhood are examined in an iterative manner.

---

**Figure 3.** The extracted tree objects to test site 1 (a) Input image, DSM, and input center of objects; (b) the delineated tree objects (white color); and (c) updated center locations of the delineated tree objects.
For that purpose, a range search method detecting all neighbors of objects (by considering their center locations) within a specified maximum distance (i.e., $\tau_D^{\min}$) is used. Thereafter, each object label is set to the majority cluster label observed in the neighborhood if at least two neighbors of each object have the same cluster label (except for ones having the same cluster label of the object). During this iterative process, the labels of all objects satisfying the above criterion are updated consistently, and the iterations are continued until no change is observed for all clusters, preserving a certain minimum number of objects (i.e., $\tau_N$). This merging stage stipulates a smoothing effect on all object labels available in the entire image and aims to minimize the number of objects erroneously assigned to a label within or nearby each cluster (Figure 5c). In the final cleaning stage, first a polygon representing the boundaries of each cluster is formed, and we check whether an intersection exists between any of the polygons generated. If an intersection is found, the iterative process described above is applied for all objects causing the intersection, and in this way, we confirm that the bounding polygon of a cluster is computed as distinctly as possible from the bounding polygons of other clusters (Figure 5d).
Once cluster labels of all objects are properly assigned, we compute the average ($\mu_1^C$ and $\mu_2^C$) and standard deviation ($\sigma_1^C$ and $\sigma_2^C$) of the two feature sets separately for each cluster: (1) the radius of the circle ($r_1^C$ in Equation 4) and (2) the minimum distances for center locations ($d_{min}^C$ in Equation 6). Finally, we identify and subsequently eliminate the cluster(s) belonging to noncitrus objects if a cluster comprises a limited number of objects (i.e., $< n_0$) or the computed average $\mu_1^C$ is smaller than the minimum distance (i.e., $r_{th}^C$).

**Recovering Missing Objects**

The detection strategy used in the section “Detection of Candidate Tree Locations” uniquely combines probabilistic LM information with orientation-based radial symmetry. In this way, most of the citrus trees are detected successfully without having a strict limitation, such as planting pattern and orientation, texture, shape, and elevation. However, some of the trees (mostly the newly planted nonbearing types) may still provide critically weak evidence for both orientation symmetry and LM information due to their irregular (small) shapes and therefore are missed by the detection step. However, such trees, especially young ones, are important to monitoring and recording the trees in newly planted orchards. Therefore, in this section, a new method to recover such missing information with the help of clustering information is proposed.

In the first stage, we convert the probabilistic LM information in Equation 1 into simple binary form ($p_{LM}>0$). This binary image holds all LM regions available in the input DSM along with many false alarms, as expected. Thereafter, we remove all LM regions that correspond to the objects of clusters labeled as accepted, as discussed in the previous section. For all remaining ones (Figure 6a), we label connected components (CC) using eight-neighborhood connectivity (total of $Q$ components), and for each component ($q = 1, 2, \ldots, Q$), we compute the two feature sets: (1) the radius of the circle ($r_{CC}^q$) and (2) the minimum distances for center locations ($d_{CC}^q$). The computation of the former is straightforward for each component, as given in Equation 4; however, the latter one for each component is computed only with respect to the objects’ center locations of accepted clusters (similar to the one presented in Equation 6) (Figure 6b).

In the second stage, we initiate an iterative framework to accept or reject the components in LM binary image within or nearby each cluster. To do that, we first collect all components which are found to be within the range of a minimum ($r_{LM}^{min}$) and maximum ($r_{LM}^{max}$) distance thresholds (i.e., $r_{LM}^{min} < d_{CC}^q < r_{LM}^{max}$). Thereafter, we apply two different inlier tests depending on the component’s location: (1) if a component is located within or on the boundaries of any accepted cluster in image space, we validate only its center position ($\mu_1^C - 3\sigma_1^C < d_{CC}^q < \mu_2^C + 3\sigma_2^C$) to decide whether it is an inlier for the cluster under consideration, and (2) if a component is located outside the boundaries of all accepted clusters in image space, we validate both its center position ($\mu_1^C - \sigma_1^C < d_{CC}^q < \mu_2^C + \sigma_2^C$) and the radius ($r_{CC}^q - 3\sigma_1 < r_{CC}^q < r_{CC}^q + 3\sigma_1$) for all clusters nearby the component. Note that, in the first test, we only expect an agreement for the center position of each component regarding the cluster parameters, as we tolerate three standard deviations from the mean of the cluster. However, during the second test, we not only expect a good arrangement for the center position of each component (just one standard deviation) but also check the radius compatibility of each component during comparison (an inlier returns for components whose radii are found to be less three standard deviations from the mean radius of the cluster). Finally, if a component is labeled as an inlier for any of the accepted clusters (Figure 6c), we delineate its boundaries (as in the section “Delineation of Candidate Citrus Trees”) and update the cluster information. The iterations continue until no new component can be assigned to any of the clusters. Ultimately, the objects in all clusters are recognized as citrus trees.

**Data Set, Evaluation, and Selection of Parameters**

Eight test images (Figure 7) covering the most productive citrus orchards from the northern part of Mersin province located on the Mediterranean coast of Turkey has been selected for this study. The input DSMs ($GSD \approx 3.5$ cm) were generated using the Semi-Global Matching (SGM) approach available in Pix4D software (Pix4D 2015) from overlapping (80% forward and 60% side overlaps) UAV images collected with a Smartplane (Table 1). The SGM approach produced point clouds of an average density of 40 to 50 points per m$^2$, and the point clouds were then used to generate the DSMs of the test images. We did not perform (manual or automatic) blunder check for either the point clouds or the DSMs, and therefore the DSMs may contain errors. We reduced the GSD of the test DSMs by a factor of two ($\approx 7$ cm) to avoid unnecessary detail in test DSMs and to facilitate processing.

The reference images with the boundaries of citrus trees were produced manually, and three quality measures are used to evaluate the pixel-based performance of the proposed strategy:

\[
\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (9)
\]

\[
\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (10)
\]

\[
F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)
\]

where TP are true positives, FP are false positives, and FN are false negatives. The operator $\| \|$ denotes the number of pixels assigned to each category, and the $F_1$-score reflects the overall performance. The object-based performance of the proposed approach is also evaluated with the same measures given in Equation 11. An output object was labeled as TP if at least 60% of the object is overlapped with the reference data, as FP if the object does not coincide with any part of the

---

**Figure 6.** Recovering missing objects for test image 1. (a) All LM regions remaining (in black), (b) all center locations (white and black signs correspond to LMs and already accepted ones, respectively), and (c) center locations of components labeled as an inlier (in white).
reference data, and as FN if the output object corresponds to reference data with a limited amount of overlap (<60%). Note that all objects that are partially visible at the boundaries of the test images are considered as well for the evaluation.

Like any object extraction pipeline, our strategy also requires a number of parameters to be set by the user. A large number of experiments have been carried out using different scenarios for all parameters involved in the sections "Detection of Candidate Tree Locations" and "Delineation of Candidate Citrus Trees," and optimum parameter values have been reached by investigating the effect of each parameter considering the precision, recall, and F1-scores (Ok and Ozdarici-Ok 2018a, 2018b). Therefore, in this study, we focus only the parameters required for the parts newly proposed in this research (the sections "Clustering of Trees" and "Promoting the Clustering Results").

The cophenetic correlation coefficient expresses how successfully the hierarchical tree constructed represents the original distances (or dissimilarities) between the observations used to construct the tree. The cophenetic correlation coefficients computed for test image 1 are given in Table 2, and note that similar results are also computed for all other test images. As shown in Table 2, the application of the cophenetic correlation coefficient revealed a combination of

---

### Table 1. Technical specifications of the UAV used.

<table>
<thead>
<tr>
<th>Platform (Type)</th>
<th>Aircraft (Smart Plane)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand</td>
<td>SmartOne-C</td>
</tr>
<tr>
<td>Maximum speed</td>
<td>70 km/hr</td>
</tr>
<tr>
<td>ø/wingspan</td>
<td>120 cm</td>
</tr>
<tr>
<td>Vehicle mass</td>
<td>1.2 kg, approx. (including camera and battery)</td>
</tr>
<tr>
<td>Propulsion</td>
<td>200-W electrical engine</td>
</tr>
<tr>
<td>Battery</td>
<td>11.1-V LiPo cell</td>
</tr>
<tr>
<td>Maximum payload mass</td>
<td>0.6 kg</td>
</tr>
<tr>
<td>Flight time</td>
<td>Up to 90 min (depends on payload and wind)</td>
</tr>
<tr>
<td>Wind tolerance</td>
<td>13 m/s</td>
</tr>
</tbody>
</table>

---

### Table 2. Cophenetic correlation coefficients computed for test image 1. (The most suitable methods are those that produce scores close to 1.)

<table>
<thead>
<tr>
<th>Distance Measures</th>
<th>Linkage Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single</td>
</tr>
<tr>
<td>Euclidean</td>
<td>0.24649</td>
</tr>
<tr>
<td>Standardized Euclidean</td>
<td>0.69170</td>
</tr>
<tr>
<td>Manhattan</td>
<td>0.26300</td>
</tr>
<tr>
<td>Hamming</td>
<td>0.10655</td>
</tr>
<tr>
<td>Cosine</td>
<td>0.32568</td>
</tr>
<tr>
<td>Pearson correlation</td>
<td>0.26508</td>
</tr>
</tbody>
</table>
two suitable algorithms for our feature data set: standardized Euclidean as the distance measure and average as the linkage method; thus, we implemented these methods during the construction of our hierarchical tree structure (Figure 8).

In hierarchical clustering, the consistency (or inconsistency) of each link in a hierarchical cluster tree can be investigated. This can be performed by comparing the height of a link in a cluster hierarchy with the average height of links below it, and if inconsistent links are found in the hierarchy, this may indicate the border of a natural division in a data set. Unfortunately, during our evaluations, we could not find expressive links characterizing inconsistencies for our entire test image data set (presumably because the test images are chosen to represent a diverse planting pattern of orchards composed of citrus trees with different characteristics and ages). Therefore, we perform the clustering through the predefined number of clusters, and Figure 9a, b presents the performance results of our strategy regarding the numbers of clusters. Our tests performed with different numbers of clusters reveal that this parameter can be set across a range of values (5 to 15) without large impact on the overall \(F_1\)-score performance (Figure 9a, b). It is also noteworthy to mention that the results without applying the clustering (i.e., the number of clusters is set to 1) produced recall around 50% for both the pixel- and the object-based evaluations, thus proving the importance of clustering stage developed. For that parameter, we fixed the number clusters to 10 for all experiments (see Figure 8), as it provided slightly better performance for both the overall pixel- and object-based \(F_1\)-scores. We also stress that if comparatively large regions in size are processed (compared to the ones tested in this research), this parameter and the related experiments should be reconsidered to achieve stable results.

For the two parameters, minimum \(\tau_D^{\text{min}}\) and maximum \(\tau_D^{\text{max}}\) distances, even though both parameters can be intuitively set by considering the planting requirements of (citrus) trees, we tested our strategy using different distance configurations (Figure 9c–f). In most cases, planting at 3.5×3.5-m intervals is done in the first planting stage when planting a citrus orchard. However, as shown in Figure 9d, setting \(\tau_D^{\text{min}}\) to exactly 3.5 m reduced the overall object-based recall performance dramatically, and this is due mostly to the slightly imprecise extraction of positions of (citrus) objects in image space. Accordingly, we get the best overall performance when \(\tau_D^{\text{min}}\) is set to 3 m. Considering the \(\tau_D^{\text{max}}\), our tests (Figure 9e, f) revealed that setting a neighborhood of 8.5 m around each (citrus) object provides the best balance between the overall pixel- and object-based results.

We also evaluated the effect of the parameter \(\tau_N\), the minimum number of objects allowed to form a cluster in image space (Figure 9g, h). We find that the overall performances across a range of the parameter \(\tau_N\) are similar. However, we also observe that the performances drop slightly when \(\tau_N\) increases, and this fact is due mainly to the small-sized citrus clusters formed close to the borders of the test images, which are eventually eliminated as \(\tau_N\) increases.

Results and Discussion
We illustrate the extraction results of our approach in Figure 10. The results of the proposed strategy as well as the results of the state-of-the-art approaches are presented in Figure 11. In addition to visual illustration, all numerical results of the proposed approach are listed in Tables 3 and 4. For a numerical assessment for all test images, we also present the results of our previous work in Ok and Ozdarici-Ok (2018b), providing an opportunity to carry out a fair quantitative comparison.

For this study, the pixel-based recall had an overall performance of 91.3% with a range of 83.3% to 94.0%, the overall
performance of pixel-based precision was 92.2% with a range of 69.7% to 98.5%, and the pixel-based $F_1$-score was 91.7% and varied from 75.9% to 94.5% (Table 3). In view of the object-based evaluation, among the 6254 reference citrus tree objects, 5716 (91.4%) objects were detected (given an object overlap threshold of 60%). In total, we missed 538 trees and falsely detected just 98 objects; thus, our approach provided highly robust and reliable results in case of an automated strategy.

With the help of newly proposed approaches in this study, we have the unique capability to filter out (groups of) irrelevant objects and at the same time recover some of the missing ones. In this way, our approach balances the measures precision and recall well and provides the best $F_1$-scores in most of the test images. For our approach, test image 7 has the
lowest performance (Tables 3 and 4). That area is dominated largely by a newly planted citrus orchard composed of young nonbearing trees with an average height of 1 m. Although we achieved a successful object-based precision score (96.3%) thanks to our unified framework, the pixel-based precision score was slightly less than 70%. This is due mainly to the delineation step in which the boundaries of some of the very small trees are overextracted, causing false alarms (Figure 11). If we omit the results of test image 7, all computed $F_s$-scores for our approach, for both pixel- and object-based evaluation, exceed the 91% bound, proving the success of the proposed unified framework.

The comparison of the state-of-the-art approaches also highlights the superior performance of the unified strategy. Note that we used nDSMs of the test images to satisfy the input requirements for three state-of-the-art approaches (i.e., Popescu and Wynne 2004; Swetnam and Falk 2014; Dalponte et al. 2015b) and run those approaches with the parameter configurations that provided the best performance. Not surprisingly, test image 7, wherein we computed the worst results for our approach, clearly outperforms the state of the art, especially considering the object-based scenario. For that image, we secure at least a 24.5% overall $F_s$-score performance improvement compared to the results of the earlier methods. In contrast, our approach provided relatively poor results for test image 4 during the comparative assessment. This is due to two reasons: (1) the citrus objects being planted have the least regularity compared to the other test images, and (2) the image is fully comprised of objects that are nothing but citrus; thus, the earlier methods seem to produce better results in such circumstances. More specifically, an issue still generating false positives for our approach is observed for different type of objects that both (1) support circular evidence and (2) form a group (Figure 10, last row, last column, upper left). For such cases, we still have no apparent solution. In addition, we did not evaluate the center locations found for individual trees due to missing observations in field, and we expect an increase of potential errors with increasing density of citrus trees. In spite of these challenging issues, we believe that the results of the proposed approach are indeed quite promising for the extraction of the citrus trees in an automated manner.

The implementation and processing were performed on a computer with an Intel i7 processor with 2.40 GHz and 16 GB RAM through MATLAB programming language, except for the approach developed by Dalponte et al. (2015b), which is available as an R package. The number of pixels in each test image is provided in Figure 7. We use the built-in parallel processing (with four cores) available in MATLAB to speed up the processing for our approach; nevertheless, our approach performed three times slower than the state of the art, requiring a total of 45 minutes to process the entire data set. However, note that the previous methods tested require an nDSM input; thus, a substantial amount of time should also be added to their processing times in that respect for a fair comparison.

### Conclusions

In this study, a new unified strategy consisting of detection, delineation, and clustering stages for the extraction of citrus fruit trees is presented. With the help of this newly proposed strategy, we have the capability to accurately recover individual citrus trees, their boundaries, and the clusters they form even in dense environments with a single input data (i.e., DSM).

#### Table 3. Pixel-based comparison of the results of the state of the art and the proposed strategy.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>$F_s$</td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>1</td>
<td>70.9</td>
<td>92.2</td>
<td>80.2</td>
<td>70.6</td>
<td>78.1</td>
</tr>
<tr>
<td>2</td>
<td>73.5</td>
<td>89.4</td>
<td>80.6</td>
<td>75.6</td>
<td>61.7</td>
</tr>
<tr>
<td>3</td>
<td>75.2</td>
<td>94.3</td>
<td>83.7</td>
<td>71.8</td>
<td>66.3</td>
</tr>
<tr>
<td>4</td>
<td>97.3</td>
<td>95.7</td>
<td>96.5</td>
<td>89.8</td>
<td>40.5</td>
</tr>
<tr>
<td>5</td>
<td>57.1</td>
<td>97.0</td>
<td>71.9</td>
<td>72.5</td>
<td>63.3</td>
</tr>
<tr>
<td>6</td>
<td>87.4</td>
<td>95.6</td>
<td>91.3</td>
<td>90.6</td>
<td>49.4</td>
</tr>
<tr>
<td>7</td>
<td>68.9</td>
<td>77.2</td>
<td>72.8</td>
<td>57.1</td>
<td>41.0</td>
</tr>
<tr>
<td>8</td>
<td>76.4</td>
<td>95.2</td>
<td>84.8</td>
<td>85.8</td>
<td>57.8</td>
</tr>
<tr>
<td>Overall</td>
<td>77.3</td>
<td>94.4</td>
<td>85.0</td>
<td>79.7</td>
<td>56.6</td>
</tr>
</tbody>
</table>

#### Table 4. Object-based comparison of the results of the state of the art and the proposed strategy.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Rec</td>
<td>$F_s$</td>
<td>Pre</td>
<td>Rec</td>
</tr>
<tr>
<td>1</td>
<td>62.0</td>
<td>97.8</td>
<td>75.9</td>
<td>89.3</td>
<td>81.3</td>
</tr>
<tr>
<td>2</td>
<td>71.4</td>
<td>85.2</td>
<td>77.7</td>
<td>78.3</td>
<td>42.2</td>
</tr>
<tr>
<td>3</td>
<td>81.9</td>
<td>96.4</td>
<td>88.5</td>
<td>87.6</td>
<td>56.1</td>
</tr>
<tr>
<td>4</td>
<td>89.2</td>
<td>98.5</td>
<td>93.7</td>
<td>97.2</td>
<td>29.8</td>
</tr>
<tr>
<td>5</td>
<td>53.1</td>
<td>99.1</td>
<td>69.1</td>
<td>53.5</td>
<td>59.5</td>
</tr>
<tr>
<td>6</td>
<td>77.7</td>
<td>98.0</td>
<td>86.7</td>
<td>86.5</td>
<td>37.8</td>
</tr>
<tr>
<td>7</td>
<td>85.7</td>
<td>52.9</td>
<td>65.4</td>
<td>85.5</td>
<td>19.3</td>
</tr>
<tr>
<td>8</td>
<td>72.2</td>
<td>100.0</td>
<td>83.9</td>
<td>73.9</td>
<td>50.0</td>
</tr>
<tr>
<td>Overall</td>
<td>73.7</td>
<td>91.9</td>
<td>81.8</td>
<td>81.1</td>
<td>46.6</td>
</tr>
</tbody>
</table>
Our strategy benefits from the LM information and near-circular shapes of citrus trees for detection, and the delineation of citrus trees is performed independently in influence regions with active contours. Finally, the results are promoted using the clustering strategy proposed, considering the characteristics of trees, planting parameters, and neighboring relations. In this way, the objects other than citrus trees (other tree species, man-made, etc.) are eliminated as a cluster basis, and errors caused by different types of objects are minimized.

The proposed strategy is tested using eight images covering citrus trees with different planting patterns, orientations, textures, shapes, and elevations. For all test data, the overall pixel-based and object-based $F_1$-scores were computed as 91.7% and 94.7%, respectively. The results achieved are also compared with the state-of-the-art methods developed for tree extraction, and the success of the proposed unified strategy is clearly demonstrated.

In this study, manually delineated tree crowns were used as reference data, and such data may also include subjective errors. In addition, the highest peak location of a tree might not correctly represent the trunk’s location; therefore, detailed fieldwork must be performed to collect the correct trunk locations to perform a reliable comparison in that respect.

Currently, we have the possibility to identify and separate different groups of citrus trees in orchards. Therefore, an interesting next step would be the accurate identification of the boundaries of orchards, which might be an input to an automated orchard boundary-updating system. In this context, our approach can be expanded to other orchards that involve types of trees other than citrus as long as the shape characteristics of the trees in the DSM bear a resemblance to the citrus trees. In addition, graph-based approaches where the smoothness effect can be directly imposed during clustering might be another interesting research direction. We also plan to test the performance of an approach in different (large) regions with varying topography, especially other than the Mediterranean type, and errors caused by different types of objects are minimized.

The authors are also gratefully to thank Dr. Michele Dalponte and Dr. Tyson L. Swetnam for their constructive comments. The authors would like to thank Dr. Michele Dalponte and Dr. Tyson L. Swetnam for the help and for graciously sharing the codes (http://mparkan.github.io/Digital-Forestry-Toolbox) of their approaches. The authors are also grateful to two anonymous reviewers and the associate editor for their constructive comments.

References


<table>
<thead>
<tr>
<th>SUSTAINING MEMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerial Services, Inc.</td>
</tr>
<tr>
<td>Cedar Falls, Iowa</td>
</tr>
<tr>
<td><a href="http://www.AerialServicesInc.com">www.AerialServicesInc.com</a></td>
</tr>
<tr>
<td>Member Since: 5/2001</td>
</tr>
<tr>
<td>AIRBUS DEFENSE AND SPACE</td>
</tr>
<tr>
<td>Chantilly, Virginia</td>
</tr>
<tr>
<td><a href="http://www.intelligence-airbusds.com">www.intelligence-airbusds.com</a></td>
</tr>
<tr>
<td>Member Since: 11/2018</td>
</tr>
<tr>
<td>Ayres Associates</td>
</tr>
<tr>
<td>Madison, Wisconsin</td>
</tr>
<tr>
<td><a href="http://www.AyresAssociates.com">www.AyresAssociates.com</a></td>
</tr>
<tr>
<td>Member Since: 1/1953</td>
</tr>
<tr>
<td>CARDINAL SYSTEMS, LLC</td>
</tr>
<tr>
<td>Flagler Beach, Florida</td>
</tr>
<tr>
<td><a href="http://www.cardinalsystems.net">www.cardinalsystems.net</a></td>
</tr>
<tr>
<td>Member Since: 1/2001</td>
</tr>
<tr>
<td>Diewberry</td>
</tr>
<tr>
<td>Fairfax, Virginia</td>
</tr>
<tr>
<td><a href="http://www.diewberry.com">www.diewberry.com</a></td>
</tr>
<tr>
<td>Member Since: 1/1985</td>
</tr>
<tr>
<td>Miami, Florida</td>
</tr>
<tr>
<td><a href="https://fr-aleman.com/">https://fr-aleman.com/</a></td>
</tr>
<tr>
<td>Member Since: 7/2020</td>
</tr>
<tr>
<td>MEMBERSHIP BENEFITS</td>
</tr>
<tr>
<td>Benefits of an ASPRS Membership</td>
</tr>
<tr>
<td>- Complimentary and discounted Employee Membership*</td>
</tr>
<tr>
<td>- E-mail blast to full ASPRS membership per year*</td>
</tr>
<tr>
<td>- Professional Certification Application fee discount for any employee</td>
</tr>
<tr>
<td>- Discounts on PE&amp;RS Classified Ad</td>
</tr>
<tr>
<td>- Member price for PE&amp;RS publications</td>
</tr>
<tr>
<td>- Discount on group registration to ASPRS virtual conferences</td>
</tr>
<tr>
<td>- Sustaining Member company listing in ASPRS directory/website</td>
</tr>
<tr>
<td>- Hot link to company website from Sustaining Member company listing page on ASPRS website</td>
</tr>
<tr>
<td>- Press Release Priority Listing in PE&amp;RS Industry News</td>
</tr>
<tr>
<td>- Priority publishing of Highlight Articles in PE&amp;RS plus, 20% discount off cover fee, one per year</td>
</tr>
<tr>
<td>- Discount on PE&amp;RS full-page advertisement</td>
</tr>
<tr>
<td>- Exhibit discounts at ASPRS sponsored conferences (exception ASPRS/ILMF)</td>
</tr>
<tr>
<td>- Free training webinar registrations per year*</td>
</tr>
<tr>
<td>- Discount on additional training webinar registrations for employees</td>
</tr>
<tr>
<td>- Discount for each new SMC member brought on board (Discount for first year only)</td>
</tr>
<tr>
<td>*quantity depends on membership level</td>
</tr>
<tr>
<td>Sanborn Map Company</td>
</tr>
<tr>
<td>Colorado Springs, Colorado</td>
</tr>
<tr>
<td><a href="http://www.sanborn.com">www.sanborn.com</a></td>
</tr>
<tr>
<td>Member Since: 10/1984</td>
</tr>
<tr>
<td>T3 Global Strategies, Inc.</td>
</tr>
<tr>
<td>Bridgeville, Pennsylvania</td>
</tr>
<tr>
<td><a href="https://t3gs.com/">https://t3gs.com/</a></td>
</tr>
<tr>
<td>Member Since: 6/2020</td>
</tr>
<tr>
<td>Towill, Inc.</td>
</tr>
<tr>
<td>San Francisco, California</td>
</tr>
<tr>
<td>Member Since: 1/2020</td>
</tr>
<tr>
<td>Woolpert LLP</td>
</tr>
<tr>
<td>Dayton, Ohio</td>
</tr>
<tr>
<td>Member Since: 1/1985</td>
</tr>
</tbody>
</table>
Performance Analysis of Advanced Decision Forest Algorithms in Hyperspectral Image Classification

Ismail Colkesen and Omer Habib Ertekin

Abstract
In this study, the performances of random forest (RF), rotation forest (RoF), and canonical correlation forest (CCF) algorithms were compared and analyzed for classification of hyperspectral imagery. For this purpose, the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Indian Pine (IP), the Reflective Optics System Imaging Spectrometer University of Pavia, and the AVIRIS Kennedy Space Center (KSC) data sets were used as main data sources. In addition to the confusion matrix–derived accuracy measures (overall accuracy, kappa coefficient, F-scores), the performances of the algorithms were analyzed in detail considering three diversity measures (Q statistics, correlations, and interrater agreements) and a kappa-error diagram. Results showed that the highest classification accuracies (87% for IP, 94% for PU, and 93% for KSC data sets) were achieved with the use of CCF algorithm, and improvements in classification accuracy were statistically significant compared to RF and RoF. Based on the diversity measures and the kappa-error diagram, individual learners in the CCF ensemble were found to be more diverse and accurate.

Introduction
Having accurate, reliable, and timely information about spatial distributions of the natural and artificial objects on the Earth’s surface plays a major role in many global and local scale applications. Remote sensing technologies have been used as an important tool for gathering such valuable information, and their primary products, satellite imageries, have been considered as a main data source in many studies. Producing thematic maps representing the land cover and land use (LULC) types of the Earth’s surface by means of image classification is a widely used technique for extracting meaningful information from the remotely sensed imagery. The reliability of thematic maps derived by image classification at the point of representation of the natural and artificial surface features is critical for the success of many studies requiring LULC information. Thematic map accuracy is strongly correlated to the success of the classification process, and it varies depending mainly on the LULC types in the study area, the characteristics of the used images, the scale of the study, and the method employed to perform the classification task (Lu and Weng 2007). To date, many classification techniques and methods have been developed and used in applications in order to perform the classification task (Tso and Mather 2009; Li et al. 2014).

Innovations in remote sensing technologies and sensor systems have allowed satellite images to develop, especially in terms of spatial and spectral resolution. Hyperspectral images consisting of hundreds of narrow spectral bands are one of the main products of the remote sensing technologies that make it possible to gather a high level of spectral information from the particular area of the Earth’s surface. Although the supervised classification technique is one of the most common tools for the analysis of hyperspectral imagery and producing of LULC map, the classification of hyperspectral imagery is still a challenging issue due mainly to the curse of dimensionality and the limited number of training samples (Ghamisi et al. 2014).

The number of training samples required to accurately determine class boundaries is generally considered to be a function of the number of spectral bands (Foody et al. 2006; Thenkabail et al. 2014). Also, higher spectral resolution allows discriminating different materials, resulting in a larger number of classes to be classified. However, it is not always possible to collect sufficient number of samples for each LULC class to find effective class boundaries using traditional parametric classification algorithm (e.g., naive Bayes and maximum likelihood). In order to overcome these problems and to conduct supervised classification task, the use of nonparametric classifiers or machine learning algorithms (e.g., support vector machines and neural networks) have recently become a vibrant research topic in the remote sensing area (Kavzoglu and Colkesen 2009; Ghamisi et al. 2016; Maxwell et al. 2018).

In recent years, there has been renewed interest in the use of ensemble learning algorithms for the classification of satellite imagery (Dietterich 2000; Rokach 2010; Jurek et al. 2014; Gislason et al. 2006; Kavzoglu and Colkesen 2013; Colkesen and Kavzoglu 2017). The main idea behind the ensemble learning is to construct a set of multiple classifiers and then make a classification decision on the test samples by aggregating their individual predictions (Kuncheva 2014). Within the ensemble learning frameworks, decision tree–based multiple classifier systems, such as random forest (RF) and rotation forest (RoF), have received increased attention in hyperspectral image classification studies due to their ability to improve the prediction performance of a weak classifier (i.e., decision tree) as well as handling high-dimensional data classification problems. For example, Chan and Paelinckx (2008) evaluated the classification performance of an RF and decision tree–based AdaBoost algorithm using Airborne HyMap hyperspectral image having 126 spectral bands distributed between 0.4 and 2.5 µm. Performances of the algorithms were also compared with an artificial neural network (ANN) classifier. The results of the study indicated that tree-based ensemble models showed similar classification performances, but both outperformed the ANN classifier in terms of overall accuracy (OA). Xia et al. (2014) conducted a comparative study on the classification performances of decision tree–based bagging, AdaBoost, RF, RoF ensemble models, and support vector machines (SVM) using three well-known hyperspectral data sets. Results showed that the decision tree–based RoF algorithm using principal component analysis produced higher classification accuracy compared with other methods in terms of overall accuracies. More recently, a novel ensemble learning method, called canonical correlation forest (CCF), based on...
the use of canonical correlation analysis (CCA) for hyperplane splitting, was introduced by Rainforth and Wood (2015). In the past few years, CCF has had limited use for the classification of satellite imagery. For example, Colkesen and Kavzoglu (2017) investigated the use of the CCF algorithm for classifying Sentinel-2 and Landsat-8 multispectral imagery. In addition, RF and RoF algorithms were evaluated for comparative analysis, and it was observed that the CCF algorithm outperformed RF and RoF. Xia et al. (2017) explored the performance of CCF for the classification of six well-known hyperspectral images using three spatial classification strategies (MPr, EMAPs, and EICA-RGF). Hu et al. (2018) applied the CCF algorithm to Sentinel-1 dual-Pol SAR data and compared its performance to the SVM classifier. Although the above-mentioned comparative studies underlined the effectiveness of CCF, it is essential to verify the effectiveness of the algorithm for various data sets and different classification problems.

The main purpose of this study was to analyze and compare the performances of decision tree–based RF, RoF, and CCF classifier ensemble models in hyperspectral image classification. For this purpose, three hyperspectral data sets recorded by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and the Reflective Optics System Imaging Spectrometer (ROSSIS) were used to test classification performances of the ensemble learning models. Performance analyses were conducted based on confusion matrix–derived accuracy measures (i.e., OA, kappa coefficient, and F-score). A non-parametric McNemar test was also used to conduct a pairwise comparison of the performance of the ensemble algorithms. For further analysis of the performances of RF, RoF, and CCF, three diversity measures including pairwise (i.e., Q statistic and correlations) and nonpairwise (i.e., interrater agreement) measures suggested by Kuncheva and Whitaker (2003) were estimated. Moreover, kappa-error diagrams were used to analyze the relationship between diversity and individual accuracy of ensemble learners.

**Test Site and Data**

In this study, the performance of decision tree–based RF, RoF, and CCF ensemble learning algorithms were evaluated for the classification of hyperspectral images. For this purpose, three popular benchmark hyperspectral data sets recorded by AVIRIS and ROSSIS sensors were used. The first hyperspectral data set was acquired by AVIRIS sensor over the Indian Pine site in northwestern Indiana (USA). The spatial resolution is 20 m, and the size of images is 145×145 pixels. The data set originally has 220 spectral bands in wavelengths ranging from 0.4 to 2.5 µm. After removing the 20 water absorption bands (104–108, 150–163), the remaining 200 spectral bands were used in this study. Ground-truth data containing a total of 10,366 labeled pixels and 16 land cover classes are available for the data set. A false color composite of the Indian Pine image with bands of 47, 23, and 13 and the corresponding ground-truth data with class names are given in Figure 1.

The second hyperspectral data set was collected by ROSSIS sensor over the University of Pavia, Italy. The image size is 610×340 pixels with a spatial resolution of 1.3 m. After removing the 12 noise bands, the remaining 103 spectral bands with wavelengths ranging from 0.43 to 0.86 µm were used in this study. The available ground-truth data set of the University of Pavia containing a total of 42,776 labeled pixels is composed of nine land cover classes. A false color composite of the University of Pavia image with bands of 103, 56, and 31 and the corresponding ground-truth data with class names are given in Figure 2.

The third hyperspectral data set was acquired by the AVIRIS sensor over the Kennedy Space Center (KSC), Florida (USA). The image size is 512×614 pixels with a spatial resolution of 18 m. After removing water absorption and noisy bands, the remaining 176 spectral bands with wavelengths ranging from 0.4 to 2.5 µm were used in this study. The available ground-truth data set of the KSC containing a total of 5,211 labeled pixels is composed of 13 land cover classes. A false color composite of the KSC image with bands of 40, 25, and 15 and the corresponding ground-truth data with class names are given in Figure 3.

**Ensemble Learning Methods**

Ensemble learning methods, also known as classifier ensembles or multiple classifier systems, are based mainly on building a supervised classification model by integrating multiple classifier algorithms (Rokach 2009, 2010; Kuncheva 2014). For an unknown sample, the final prediction is obtained by combining individual votes of the base classifiers in the ensemble. Recently, there has been considerable growth in the literature about the use of decision tree–based ensemble learning algorithms in classification of remotely sensed images (Belgiu and Drăguţ 2016; Rokach 2016). The main working principle of decision tree–based ensemble algorithms, also known as decision forest algorithms, is to construct several decision trees by using different subsets of the original training data and to make a final prediction by combining their predictions. In the literature, prediction performances of the decision forest algorithms have been shown to be superior compared to use of single decision tree classifier (e.g., Briem et al. 2002; Miao et al. 2012; Kavzoglu and Colkesen 2013).

![Figure 1. Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Indian Pine hyperspectral data set. (a) Three-band false-color composite image; (b) available ground-truth data with 16 classes.](image-url)
As in any classifier ensemble model, diversity is the key issue for the overall performances of the decision forest classifier ensembles (Kuncheva and Whitaker 2003). According to Diez-Pastor et al. (2015), a suitable ensemble model should have accurate individual members, and at the same time their prediction errors should be in different samples. In other words, in order to construct an efficient ensemble model, not only is it necessary to construct base learners, but base learners must be diverse as well (i.e., the base learners return different predictions). Bagging and boosting, designed to create different subsets from the original training data set, are two popular methods for building a set of diverse classifiers. Within this concept, RF, a widely used decision forest method, provides diversity by applying two randomization procedures. First, bootstrap aggregating (i.e., bagging) is applied to create different subsets chosen randomly with a replacement for the training of each tree; second, randomly selected features are considered for a split in each tree node (Breiman 2001). In order to improve diversity between the classifiers, RoF, as introduced by Rodríguez et al. (2006), applies principal
component transformation to each randomly split subset, and resulting components are used for the training of each base classifier. More recently, CCF has been introduced as a classifier ensemble by Rainforth and Wood (2015). It is based on the construction of multiple oblique trees using CCA. In this study, the classification performances of RF, RoF, and CCF decision forest algorithms were assessed and compared using three benchmarked hyperspectral data sets.

**RF Algorithm**

The RF algorithm is one of the most widely used decision tree–based ensemble learning algorithms in remote sensing applications due to its ease of use and robustness (Belgiu and Drăguț 2016). The fundamental aim of the RF is to create a number of decision tree classifiers using random subsets of original input data and then to make a final prediction for a given test sample by combining individual tree predictions (Breiman 2001). Each individual tree within the ensemble model is trained using different subsets selected randomly with replacement using bootstrap aggregation procedure (i.e., bagging). Two-thirds of the randomly sampled data (i.e., in-bag samples) are used for training of decision tree, and the remaining one-third (i.e., out-of-bag samples) are used to estimate the prediction error rate (i.e., out-of-bag error). The second randomization process in the RF ensemble model is that rather than selecting the best split among all features, randomly selected features are considered at each node (Rokach 2016). As a result, a majority voting procedure is performed to make a final prediction of an unknown sample. Ensemble size or the number of trees in the ensemble and the number of features to be selected at each node are the main user-defined parameters for the application of RF algorithm.

**RoF Algorithm**

The RoF algorithm, introduced by Rodríguez et al. (2006), is a relatively new classifier ensemble method. It has been attracting considerable interest in the past few years due to its superior performance compared to bagging, boosting, and RF in many pattern recognition studies, such as classification of multispectral and hyperspectral imagery (Kavzoglu and Colkesen 2013; Xia et al. 2014; Du et al. 2015). RoF determines different training subsets for each base classifier (i.e., decision tree) by applying principal component analysis (PCA). For this purpose, the original input data set is randomly partitioned into K subsets, and then PCA is applied to each selected subset in order to construct a rotated feature space. In each iteration, the estimated components are used to build a new training data set, and each individual tree in the ensemble is trained with this transformed data set. In the prediction phase, confidence for each class is estimated by the average combination technique, and the final prediction is assigned to the target class with the largest confidence value. The number of individual trees (i.e., ensemble size) and the number of features in a subset are the main user-defined parameters of the RoF algorithm.

**CCF Algorithm**

The CCF algorithm was introduced by Rainforth and Wood (2015) as a novel decision tree–based ensemble method. The main idea behind the CCF algorithm is quite similar to the RoF algorithm. It is based application of canonical correlation analysis (CCA) before the training of each individual tree within the ensemble. Resulting correlation components are used to hyperplane splits in binary decision trees called canonical correlation trees. CCA is one of the multivariate techniques used to investigate the total correlation between two sets of variables (Richard et al. 1992). CCA is applied to find feature projections, providing a maximum correlation between the features and the target class labels for the construction of the CCF ensemble model. Although the ensemble size is a user-defined parameter of the algorithm, experimental results of recent studies have shown that parameter setting is not required during the CCF ensemble model construction (Rainforth and Wood 2015; Colkesen and Kavzoglu 2017, 2019; Xia et al. 2017).

**Performance Evaluation**

In order to compare classification performances of the RF, RoF, and CCF ensemble algorithms, confusion matrix–derived accuracy measures, namely OA, kappa coefficient (κ), and F-score were taken into account. The F-score, one of the most widely used class-level measures, is estimated as the harmonic mean of the user’s accuracy and the producer’s accuracy. For further evaluation of the classification results, the McNemar test, based on a comparison of two paired proportions, was also used to analyze statistical differences between the classifier performances (Feody 2004). If the estimated probability value (p-value) is less than the significance level of 0.05 (i.e., p < 0.05), the null hypothesis is rejected, and it can be concluded that there is a significant difference between two classification results. In addition, required computation time (CT) in seconds for training of each algorithm was also estimated using a desktop computer having a Core i7 quad-core (3.40-GHz) processor with 16 GB of RAM. Due to the accuracy of each base algorithm in the ensemble model and the diversity among algorithms being important for the performance of ensemble methods, three diversity measures—two averaged pairwise measures (i.e., Q statistic and correlation) and one nonpairwise measure (i.e., interrater agreement)—were estimated and analyzed for RF, RoF, and CCF, respectively. More detailed information and formulas related to the diversity measures can be found in Kuncheva and Whitaker (2003). In addition, kappa-error diagrams suggested by Kuncheva (2013), which are scatter plots with LLI = 1/2 points where L is the number of individual classifiers in the ensemble, were plotted and analyzed, thoroughly.

**Results and Discussion**

In this study, the classification performances of the three decision forest algorithms (i.e., RF, RoF, and CCF) were evaluated using three widely used data sets: the AVIRIS Indian Pine, the ROSIS University of Pavia, and the AVIRIS KSC. The hyperspectral data sets are regarded as a challenging LULC classification scenario due to the availability of a high number of mixed pixels and imbalanced training data set (i.e., the number of reference samples for LULC classes varies greatly among the classes). On the other hand, the decision tree–based ensemble learning algorithms (i.e., RF, RoF, and CCF) have been recommended as robust classifiers in solving classification problems having a limited or imbalanced training data set (Rodríguez et al. 2006; Mellor et al. 2015; België and Drăguț 2016; Maxwell et al. 2018). In order to assess the classification performances of the ensemble learning algorithms for the training data sets that are limited and imbalanced, a limited number of training samples were used in the all classification scenarios considered in this study. For this purpose, 10% of the ground-truth samples for each class were randomly selected as training, and the remaining 90% were used for testing in all classification problems considered in this study. The RF, RoF, and CCF algorithms require the setting of the ensemble size (i.e., the number of trees) parameter from the user side. For all three hyperspectral data sets, the number-of-trees parameter was determined using the out-of-bag responses for RF and fivefold cross-validation strategy for RoF algorithms, while ensemble size was set to 200 for the CCF algorithm as suggested by Rainforth and Wood (2015). While the number of features in a subset was set to 3 for the RoF algorithm as suggested by Rodríguez et al. (2006), the number of features randomly selected for a split at each node was set to be the square root of the total number of input bands for the RF algorithm as...
suggested by Gislason et al. (2006). It should be noted that the entire classification process in the evaluation of the ensemble learning algorithms was performed with MATLAB software (version 2016a).

### Classification Results of the AVIRIS Indian Pine Hyperspectral Data set

For the classification of the Indian Pine data set using the RF, RoF, and CCF algorithms, training and testing samples were randomly selected for each class from the ground-truth image based on the 10:90 sampling ratio (Table 1). By applying a cross-validation procedure, ensemble size was estimated as 100 for RoF. On the other hand, 340 trees were found to be optimal for the RF algorithm with respect to the out-of-bag error response. The number of features was calculated as the square root of the 200 bands (i.e., 14) for the RF algorithm, while the number of features in a subset was kept to 3 for the RoF algorithm. The ensemble models constructed with optimal parameter settings applied to the test data set and estimated accuracy results are given in Table 1. It can be seen from the table that calculated OA values were 75.50% (κ-value of 0.717), 84.13% (κ-value of 0.817), and 86.44% (κ-value of 0.844) for RF, RoF, and CCF, respectively. The highest OA value was produced with the use of the CCF algorithm, whereas the lowest accuracy was estimated by RF, which was significantly lower than the OA values of RoF and CCF by about 10%. In addition, the difference between the classification accuracy of CCF and RoF was about 2%. In order to evaluate the statistical significance between the classification performances of the ensemble algorithms, the McNemar test was applied. The estimated probability values (i.e., p-value) were found to be lower than the statistical significance level of 0.05 (p < 0.5) in all pairwise comparisons. In other words, in parallel to the above findings, classification performances of the RF, RoF, and CCF algorithms were statistically significantly different for the Indian Pine data set in terms of the McNemar test results.

When the class-level accuracies (i.e., F-scores) were analyzed, it was observed that especially for the classes with only a few training samples (e.g., grass-pasture-mowed and oats classes), the CCF algorithm showed superior performance compared to RF and RoF. The improvement in class-level accuracy reached 0.69 for the oats class in terms of F-score. This could be result of the use of CCA during the training stage of the CCF ensemble model. As can be seen from the table, the class-level accuracies of RoF and CCF were significantly higher than that of RF for the alfalfa, soybean-clean and buildings-grass-trees-drives classes. For better visualization of the class-level prediction accuracy, the classification maps produced by the ensemble models are also given in Figure 4. It is shown in the figure that the CCF resulted in a more accurate and smoother thematic map compared to others. The CT required for the training stage was also computed for all three algorithms and is given in Table 1. It is clear that the RF algorithm was the fastest (16.11 seconds) for constructing an ensemble model. However, the RoF and CCF required much more time (i.e., 39.43 and 24.67 seconds, respectively) for the training stage. This is due mainly to the application of the feature extraction process in the modeling stage of RoF and CCF.

To analyze diversity among the members of the ensemble model constructed by the RF, RoF, and CCF algorithms, three diversity measures—Q-statistic, correlation coefficient, and interrater agreement—were performed, and the estimated statistic values are given in Table 2. It should be noted that a lower value shown in bold in the table indicates higher diversity in the ensemble model. As can be seen from the table, the
Table 2. Estimated diversity values of the random forest (RF), rotation forest (RoF), and canonical correlation forest (CCF) for the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Indian Pine data set.

<table>
<thead>
<tr>
<th>Diversity Measure</th>
<th>RF</th>
<th>RoF</th>
<th>CCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-statistic</td>
<td>0.579</td>
<td>0.530</td>
<td>0.514</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>0.315</td>
<td>0.275</td>
<td>0.262</td>
</tr>
<tr>
<td>Interrater agreement</td>
<td>0.315</td>
<td>0.275</td>
<td>0.262</td>
</tr>
</tbody>
</table>

Table 3. Classification accuracies and computational costs obtained by random forest (RF), rotation forest (RoF), and canonical correlation forest (CCF) for the Reflective Optics Systems Imaging Spectrometer (ROSIS) University of Pavia data set.

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Training Samples (10%)</th>
<th>Test Samples (20%)</th>
<th>F-Score</th>
<th>RoF</th>
<th>CCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asphalt</td>
<td>663</td>
<td>5968</td>
<td>0.91</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Meadow</td>
<td>1865</td>
<td>16784</td>
<td>0.94</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Gravel</td>
<td>200</td>
<td>1890</td>
<td>0.73</td>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td>Tree</td>
<td>306</td>
<td>2758</td>
<td>0.92</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>Metal sheet</td>
<td>134</td>
<td>1211</td>
<td>0.98</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>Bare soil</td>
<td>503</td>
<td>4526</td>
<td>0.75</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>Bitumen</td>
<td>134</td>
<td>1196</td>
<td>0.81</td>
<td>0.88</td>
<td>0.90</td>
</tr>
<tr>
<td>Brick</td>
<td>369</td>
<td>3313</td>
<td>0.84</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>Shadow</td>
<td>94</td>
<td>853</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

| Overall accuracy (%) | 89.33 | 94.13 | 94.07 |
| Computation time (seconds) | 36.53 | 53.82 | 56.74 |

The lowest diversity values shown in bold were calculated with the CCF algorithm as 0.514, 0.262, and 0.262 for the Q-statistic, correlation coefficient, and interrater agreement, respectively. Thus, it can be said that the decision trees in the CCF ensemble model are more diverse than in the RF and RoF ensemble models for the classification of the Indian Pine data set.

In order to perform further evaluation of the ensemble learning algorithms, kappa-error diagrams were plotted to visualize individual accuracy and diversity. For this purpose, ensemble size (L) was set to 100 for all ensemble models and 4.950 (i.e., L(L – 1)/2) points, each corresponding to a pair of classifiers for each ensemble model, as shown in Figure 5.

The horizontal axis of the diagram represents the interrater agreement (nonpairwise kappa), and the vertical axis of the diagram represents the averaged individual error rate of the classifier pair. It should be noted that a measure of interrater agreement was derived from the contingency table of two classifiers using the equation $\kappa = (OA - AC) / (1 - AC)$, where $OA$ is the observed agreement (i.e., the probability that the two classifiers will both be correct or incorrect when classifying a randomly selected sample) and $AC$ is the probability that the two classifiers agree by chance on a randomly selected sample (Dietterich 2000; Kuncheva 2013). When the relative positions of the point clouds in the figure were analyzed, it was observed that the points representing the members of the CCF model were close to the lower-left corner of the diagram. Thus, it can be said that the highest diversity (i.e., lower kappa) and the highest accuracy (i.e., lower average pairwise error) was reached by the CCF algorithm compared to RF and RoF for the classification of the Indian Pine data set.

Classification Results of the ROSIS University of Pavia Hyperspectral Data Set

The classification performances of the RF, RoF, and CCF algorithms were also tested on the University of Pavia hyperspectral data set. For this purpose, training and testing samples were randomly selected for each class from the ground-truth image based on the 10:90 sampling ratio (Table 3). The optimal ensemble sizes were determined as 400 and 70 for RF and RoF, respectively. The number of features was calculated as the square root of the 103 bands (i.e., 10) for the RF algorithm, while the number of features was kept to 3 for the RoF algorithm. The test data set consisting of a total of 38,499 labeled pixels was classified using the optimized ensemble models of RF, RoF, and CCF, and calculated accuracy results are given in Table 3.

As can be seen from the table, the lowest OA was calculated as 89.33% ($\kappa$-value of 0.856) with the RF algorithm. However, the highest overall accuracies were estimated by the RoF and CCF algorithms as 94.13% ($\kappa$-value of 0.922) and 94.27% ($\kappa$-value of 0.924), respectively. These results clearly show that RoF and CCF exhibit similar performance on the University of Pavia data set and that both outperformed the RF.
algorithm in terms of OA results. The McNemar test was used to assess statistical significance between the performances of the ensemble learners. The statistical test results showed that the difference in classification performances between RoF and CCF classifiers was not statistically significant (i.e., $p > 0.5$), while differences in the performances of the remaining pairwise comparisons (i.e., RF-RoF and RF-CCF) were found to be statistically significant (i.e., $p < 0.5$). Therefore, statistical results confirmed the above findings that RoF and CCF showed statistically different classification performances compared to RF, while their performances were statistically similar for the classification of the University of Pavia data set.

From Table 3, it is observed that $F$-score values, indicating the class-level accuracies, estimated by RoF and CCF are above 0.80 for all land cover classes. However, the lowest $F$-score values were calculated with the RF algorithm as 0.73 and 0.75 for the gravel and bare soil classes, respectively. In addition, the RoF and CCF algorithms showed relatively better predictions on the bitumen, gravel, and tree classes having limited training samples except for the shadow and metal sheet classes. When the computational costs given in the table were analyzed, it was observed that the CT of the RF algorithm (36.53 seconds) was quite low compared to the estimated CT of the RoF and CCF algorithms. In order to visually display the classifications performances of the algorithms, the thematic maps were produced and are shown in Figure 6. As can be seen from the figure, relatively higher noisy estimations existed in thematic maps produced by the RF compared to others.

To further investigate the performance of decision forest ensemble algorithms (RF, RoF, and CCF), the Q-statistic, correlation coefficient, and interrater agreement measures were used, and derived diversity statistics are given in Table 4. It should be noted that a lower value shown in bold in the table indicates higher diversity in the ensemble model. It is clear from the table that diversity of the CCF was higher than that of RF and RoF in terms of the correlation coefficient and the interrater agreement measures, while the RoF algorithm was found to be more diverse than the others in terms of the Q-statistic measure. Therefore, it can be said that diversity of the individual members of the RoF and CCF ensemble models was higher than that of the RF ensemble model for the classification of the University of Pavia data set.

Table 4. Estimated diversity values of the random forest (RF), rotation forest (RoF), and canonical correlation forest (CCF) for the Reflective Optics System Imaging Spectrometer (ROSIS) University of Pavia data set.

<table>
<thead>
<tr>
<th>Diversity Measure</th>
<th>RF</th>
<th>RoF</th>
<th>CCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q$-statistic</td>
<td>0.711</td>
<td>0.700</td>
<td>0.669</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>0.335</td>
<td>0.288</td>
<td>0.270</td>
</tr>
<tr>
<td>Interrater agreement</td>
<td>0.335</td>
<td>0.288</td>
<td>0.270</td>
</tr>
</tbody>
</table>

Kappa-error diagrams of the RF, RoF, and CCF ensemble models are given in Figure 7. For each ensemble model, the ensemble size was fixed to 100 trees; therefore, there are 4,950 points in each plot. From the figure, it is observed that the point clusters of the RoF and CCF models located at nearly the same position in the figure and the points were much closer to the bottom-left corner of the diagram compared to those of RF. Therefore, it can be said that the accuracy and diversity of RoF and CCF are approximately equivalent to the University of Pavia data set. On the other hand, these algorithms reached better diversity and accuracy levels than RF with respect to the relative positions of the point clusters in the diagram.

Classification Results of the AVIRIS Kennedy Space Center Hyperspectral Data Set

The third hyperspectral data set Kennedy Space Center (KSC) was used to evaluate the classification performances of the decision forest algorithms. For the construction of each ensemble algorithm, training and testing data sets were formed by applying a stratified random selection strategy based on the 10:90 sampling ratio (Table 5). Optimum parameters of the RF algorithm, namely, the number of trees and the number of

---

**Figure 6.** Thematic maps of the Reflective Optics System Imaging Spectrometer (ROSIS) University of Pavia data set produced by (a) random forest (RF), (b) rotation forest (RoF), and (c) canonical correlation forest (CCF).
features, were determined as 400 and 13, respectively. Furthermore, the ensemble size was determined as 100, and the number of features in each subset was used as 3 for the RoF algorithm. The test data set, containing 13 land cover classes (totally 4,683 labeled pixels), was classified using optimized RF, RoF, and CCF ensemble models, and estimated OA, $\kappa$, $F$-score, and CT values are given in Table 5. When the calculated $\kappa$-values and $F$-values were compared, it was found that the highest classification accuracy of 93.26% ($\kappa$-value of 0.924) was obtained by the use of the CCF algorithm, whereas the lowest accuracy of 87.44% ($\kappa$-value of 0.860) was produced with the RF algorithm. Furthermore, with the use of CCF, the improvement in classification accuracy was about 2% compared to the RoF algorithm. This clearly showed that the RoF and CCF ensemble algorithms outperformed the RF algorithm for the classification of the KSC data set. In order to evaluate the statistical significance of difference among the RF, RoF, and CCF algorithms, pairwise comparisons were performed using the McNemar test. The results of the statistical test indicated that there were statistically significant differences between the classification result of the ensemble algorithms (i.e., $p < 0.05$). In other words, differences in classification accuracies between CCF and RF (about 6%), RoF and RF (about 4%), and CCF and RoF (about 2%) were found to be statistically significant. Thus, it can be said that the CCF algorithm showed statistically superior performance compared to others for the classification of the KSC data set. When the estimated CTs given in Table 5 were analyzed, it was observed that the RF was quite fast compared to CCF and RoF.

As can be seen from the estimated class-level accuracies given in Table 5, the highest $F$-score values were produced with the use of CCF in all cases. Furthermore, the CCF algorithm was more successful than the RF and RoF algorithms in distinguishing the hardwood swamp, slash pine, and oak/broadleaf hammock classes, which have very few training samples. On the other hand, $F$-score results indicated that class-level accuracy was high (above 0.90) on the scrub, cattail marsh, salt marsh, mudflats, and water classes. The thematic maps of the KSC produced by the ensemble learning algorithms are shown in Figure 8. It is obvious that the classification map produced by the RF algorithm has a noisier look (i.e., a salt-and-pepper look) compared to the maps produced by the RoF and CCF algorithms.

Table 5. Classification accuracies and computational costs obtained by random forest (RF), rotation forest (RoF), and canonical correlation forest (CCF) for the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Kennedy Space Center (KSC) data set.

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Training Samples (10%)</th>
<th>Test Samples (90%)</th>
<th>$F$-Score</th>
<th>RF</th>
<th>RoF</th>
<th>CCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scrub</td>
<td>77</td>
<td>684</td>
<td>0.93</td>
<td>0.93</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>Willow swamp</td>
<td>25</td>
<td>218</td>
<td>0.83</td>
<td>0.88</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Cabbage palm hammock</td>
<td>26</td>
<td>230</td>
<td>0.89</td>
<td>0.88</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Cabbage palm/oak hammock</td>
<td>26</td>
<td>226</td>
<td>0.63</td>
<td>0.70</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Slash pine</td>
<td>17</td>
<td>144</td>
<td>0.57</td>
<td>0.66</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Oak/broadleaf hammock</td>
<td>23</td>
<td>206</td>
<td>0.57</td>
<td>0.72</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Hardwood swamp</td>
<td>11</td>
<td>94</td>
<td>0.78</td>
<td>0.86</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Graminoid marsh</td>
<td>44</td>
<td>387</td>
<td>0.79</td>
<td>0.89</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Spartina marsh</td>
<td>52</td>
<td>468</td>
<td>0.87</td>
<td>0.92</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Cattail marsh</td>
<td>41</td>
<td>363</td>
<td>0.91</td>
<td>0.99</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Salt marsh</td>
<td>42</td>
<td>377</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Mudflats</td>
<td>51</td>
<td>452</td>
<td>0.91</td>
<td>0.96</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>93</td>
<td>834</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Overall accuracy (%) 87.44 91.53 93.26

In order to compare the classification performance of algorithms in terms of diversity, three diversity measures were used, and derived statistical values are presented in Table 6. It is clear from the table that the lowest statistic values shown in bold, indicating the highest diversity, were estimated for the CCF algorithm. These findings apparently showed that the CCF classifier ensemble exhibited superior performance compared to others in terms of accuracy and diversity results for the classification of the KSC hyperspectral data set.

The kappa-error diagram given in Figure 9 was prepared to analyze the relationship between the diversity and accuracy of base classifiers. From the figure, RF spans the largest diversity range, but much of the individual classifier pairs have large errors. CCF and RoF have similar diversity, but the former has better individual accuracy considering the relative positions of the closest point clouds to the left bottom-left corner of the diagram.
Conclusions

Recently, there has been considerable growth in the literature about the theme of the use of multiple classifier systems in the classification of satellite images. Within this context, decision forest algorithms, based on the construction of multiple decision tree classifiers in solving a given classification problem, have been widely used in many applications due to their robustness and simplicity. In this study, classification performances of advanced decision forest algorithms (i.e., RoF and CCF) were evaluated and their performances compared to that of the popular RF algorithm. Three benchmark hyperspectral data sets (i.e., Indian Pine, University of Pavia, and KSC) were used as data sets, and classification performances of the algorithms were analyzed using standard accuracy measures (i.e., OA, κ, and F-score), the nonparametric McNemar test, different diversity measures, and the kappa-error diagram.

Based on the comparative results of this research, some important conclusions can be drawn. First, for all three hyperspectral data sets, the RoF and CCF algorithms showed superior classification performances over the RF algorithm in terms of OA, κ, and F-score measures, and the differences in classification accuracies were statistically significant based on the McNemar test results. In addition, the largest differences in classification accuracies between RF and the others were estimated for the Indian Pine data set as 11% and 9% for RoF and CCF, respectively. Furthermore, although the highest classification accuracies were produced with the CCF ensemble algorithm for all benchmark hyperspectral data sets, there was no significant difference at the 95% significance level between the estimated accuracies of RF and RoF for classification of...
the University of Pavia data set. Second, results verifying the previous findings showed that the processing speed of the RF algorithm in the construction of the ensemble model was extremely fast compared to the processing speeds of RoF and CCF. Third, based on the Q-statistic, correlation coefficient, and interrater agreement statistics considered as diversity measures in this study, the diversity between the individual trees in the CCF ensemble model was found to be higher compared to that of the RF and RoF models for all hyperspectral data sets. Moreover, kappa-error diagrams also showed that generally CCF produces more accurate individual classifiers than the others, which were also more diverse than those in RF and RoF. Overall, the results of the study highlighted the potential usefulness of the advanced decision forest ensemble algorithms, namely, RoF and CCF, in hyperspectral image classification using limited training samples. This research confirms the previous findings and contributes the additional suggestion that diversity measures and the kappa-error diagram provide valuable information about the behavior of the base learner in ensembles and could be considered in assessing the accuracy of ensemble learning algorithms. Further studies are needed to validate the robustness of the advanced decision forest algorithms in the analysis of different data sets and classification problems.

References


Analyzing the Contribution of Training Algorithms on Deep Neural Networks for Hyperspectral Image Classification

Mehmet Akif Günen, Umit Haluk Atasever, and Erkan Beşdok

Abstract
Autoencoder (AE)-based deep neural networks learn complex problems by generating feature-space conjugates of input data. The learning success of an AE is too sensitive for a training algorithm. The problem of hyperspectral image (HSI) classification by using spectral features of pixels is a highly complex problem due to its multi-dimensional and excessive data nature. In this paper, the contribution of three gradient-based training algorithms (i.e., scaled conjugate gradient (SCG), gradient descent (GD), and resilient backpropagation algorithms (RProp)) on the solution of the HSI classification problem by using AE was analyzed. Also, it was investigated how neighborhood component analysis affects classification performance for training algorithms on HSIs. Two hyperspectral image classification benchmark data sets were used in the experimental analysis. Wilcoxon signed-rank test indicates that RProp is statistically better than SCG and GD in solving the related image classification problem.

Introduction
Advances in spectral imaging sensors have enabled the use of hyperspectral images (HSIs) in remote sensing research (Li et al. 2019; Liu et al. 2019). HSIs typically consist of image layers with a large number of different spectral features. HSIs are widely used in pattern recognition applications such as image segmentation and object identification because of the detailed information they provide about the spectral properties of objects. Image clustering and classification methods are the most commonly used image segmentation tools (Zhong et al. 2018).

The classification of HSIs can be performed using spatial, spectral, or spatial-spectral image features. K-means clustering (Filho et al. 2003), Fuzzy C-means clustering (Sigirci and Bilgin 2017), DBScan clustering (Datta Ghosh, and Ghosh 2012), expectation-maximization clustering (Marden and Manolakis 2003), and agglomerative hierarchical clustering methods (Medina Manian, and Chinea 2013) are commonly used unsupervised classification methods. In practice, spatial locations of training and test samples required for the classification process are determined using clustering methods (Guan, Yuen, and Coenen 2019; Nasiboglu, Tezel, and Besdok 2019). Hyperspectral images are data that contain two physical partitions called the encoder and the decoder. Encoder converts relevant data into coded-feature space (Lv, Peng, and Wang 2018; Makkie et al. 2019; Yu et al. 2018). Furthermore, the encoder changes the dimensional, homogeneity, smoothness, and continuity level of the respective data. The decoder is used to obtain the original data from the corresponding data converted to the coded-feature space. The AE is actually used for data reconstruction and data smoothing, unlike multi-layer perceptrons used in log-regression-based numerical prediction.

A DNN’s success in learning to solve the problem is sensitive to the complexity of the problem, the size of the data.
used, the level of homogeneity, the number of data, the model of data, and the architecture of the DNN used. The training algorithm used to train a DNN can radically affect the DNN’s problem learning success. The algorithm to be used to train the DNN should be able to optimize an excessive number of parameters arising from the nature of the DNNs. If the training algorithm used to solve this large scale optimization problem encountered in DNN has good memory management capability, this can reduce the computational platform cost and the computational complexity of the training process. It is necessary to analyze the extent to which the problem-solving success provided by an AE in HSI processing is influenced by the training algorithm used. The need for understanding the effect of AE training algorithms on classification motivates various researchers. This paper focuses on analyzing the contribution of training algorithms (i.e., scaled conjugate gradient (SCG), gradient descent (GD), and resilient backpropagation (RP)) on training of related AEs in order to solve HIS classification problems. The overall flowchart is shown in Figure 1.

Figure 1. Flowchart of the proposed method.

The contribution of this paper to the literature is listed below:

1. The contribution of SCG, GD, and RP training algorithms for AE was analyzed.
2. The advantages and disadvantages of SCG, GD, and RP in the training of AE were investigated.

This paper is organized as follows: the next section, “Deep Neural Networks”, further explores DNNs and the “Experiments” section explores the experiments used. The final section presents results and conclusions.

Deep Neural Networks

DNNs have been used to solve various high-level complex problems encountered in different application areas, such as hyperspectral image processing, video indexing, automatic pattern recognition, speech synthesis, natural language processing, analyzing of time series, healthcare applications, biomedical research, three-dimensional vision, etc. (Do et al. 2019; Grekousis 2019; Liu and Wu 2016; Shen, Wu, and Suk 2017). The success of DNNs in solving complex problems comes from their ability to significantly transform training data into limited-sized feature spaces (du Plessis and Broeckhoven 2019; Xu et al. 2019). DNNs learn complex problems at the state-of-art level because they use automatic, derived, limited-sized features instead of preorganized features of the related data, unlike conventional artificial neural networks. The DNNs’ superior learning ability requires over-sized example data and a highly complex computational load. The training of DNNs involves the computation of the optimal values of the hyper-sized parameters they have. As the layered structure of the DNN increases, the number of parameters it has increased becomes hypersized. Therefore, the DNN’s training processes require multiprocessor power. In practice, it is common to use the GPU to meet the calculation needs of DNNs. DNNs, unlike conventional artificial neural networks, consist of a large number of layers in which structurally relevant data is converted to different levels of features (Gunen, Atasever, and Besdok 2018; Shen, Wu, and Suk 2017). Classic neural networks contain a small number of layers, and the multi-layered feature extraction features are weak compared to DNNs.

AE, a type of DNN, is an unsupervised learning-based artificial neural network used to transform related input data into feature space. The corresponding input and output vectors are exactly the same in AE training. In other words, AEs also do not use a labeled output (Liou et al. 2014). Generally, the input vector converts an AE $y = (y_{1:K})$-sized input vector to a $d = (d_{1:N})$-sized feature vector. Multi-layered AEs create a stacked autoencoder (SAE). SAE improves the quality of related features by improving the structural parameters of AEs. The DNNs use the softmax function to assign the output from the SAE to the prelabeled classes. Using a supervised logistic regression process, the softmax layer completes the pattern recognition process by assigning the features of the corresponding SAE to the prelabeled class. Softmax classifier is a generalized form of logistic regression called multinomial or log-linear regression (Gunen, Atasever, and Besdok 2018). In contrast to the logistic regression that makes only binary classes, softmax is used for multiple classes without nonordered input.

The structure of the AE used in this article is shown in Figure 2. First, the AE input vector, illustrated in Figure 2, converts $y = (y_{1:K})$ to a $d = (d_{1:N})$ feature called vector code. Second, the AE illustrated in Figure 2 converts $d = (d_{1:N})$ to $d = (d_{1:N})$ vector code. The softmax layer in Figure 2 transfers the $d = (d_{1:N})$ features to the prelabeled output.

Autoencoder

AE is an unsupervised nonrecurrent, feedforward neural network consisting of an input, output, and hidden layer. AE consists of two structures: encoder and decoder. In encoder, the dimension of the input is changed to the same dimension as the hidden layer. In the decoder, maps data is changed to same dimension as the input data. Two structures use transfer function to represent new data (Han, Zhong, and Zhang 2017). $L$ is the number of hidden layers of the encoder in the first AE and $y = [y_{1}, y_{2}, \ldots, y_{L}]$ is the input data. Here $K$ is the dimension of input data and $(K,L)\in Z^{+}$. Equation 1 represents the maps’ input layer to hidden layer.

$$d = f(y; w, b), \tag{1}$$

where $f$ is an activation function and $b = [b_{1}, b_{2}, \ldots, b_{L}]^T$, $w = [w_{1}, w_{2}, \ldots, w_{L}]^T$ are biases and weights of the encoder, respectively.

The output vector from the encoder constitutes the input vector of the decoder. The number of hidden layers in the decoder is the same dimension as the input layer. As with the
encoder, it is transferred to the output with a transfer function. The transfer function is given as:

\[ \hat{y} = f(d; \hat{w}, \hat{b}) \],

where \( f \) is an activation function and \( \hat{b} = [\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_L]^T \), \( \hat{w} = [\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_K]^T \) are biases and weights of the encoder, respectively.

The relationship between input and output is

\[ \hat{y} = h(w, \hat{w}, \hat{b}, y) \].

The \( \hat{y} \) function indicates the relationship between the input and output of an AE. SAE is used to construct two or more AEs in succession, followed by a classifier (Ng 2011; Zabala et al. 2016). The SAE reconstruction determines the network parameters by optimizing the network. The reconstruction error of the SAE structure is given as:

\[ E_{ae} = \frac{1}{M} \sum_{i=1}^{M} e_i^2 + \frac{\beta}{2} \sum_{j=1}^{L} \left[ \frac{1}{M} \sum_{i=1}^{M} \| \hat{w}_j \|^2 \right] + \lambda \sum_{j=1}^{L} KL(\rho || \hat{\rho}_j), \]

where \( i = [1,2,\ldots,M] \) is the number of samples, \( e_i = \| y^{(i)} - \hat{y}^{(i)} \| \) is the error between the input and output layer, and \( \lambda \) is the weight regularization term used to prevent overfitting. \( \rho \) is the sparsity parameter and is constant. \( \hat{\rho}_j \) the mean activation value is given as:

\[ \hat{\rho}_j = \frac{1}{M} \sum_{i=1}^{M} f(y^{(i)}). \]

The Kullback-Leibler divergence, which is used to determine the similarity of the two distributions, restricts the sparsity of the output with the constant sparsity term \( \beta \). The Kullback-Leibler divergence is given as (Ng 2011; Zhuang et al. 2015):

\[ KL(\rho || \hat{\rho}_j) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_j}. \]

Training of Autoencoders

The success of the training algorithm used for training AE affects the success of AE’s ability to produce reasonable code features. The training algorithm affects both the accuracy of the model and its robustness. Training algorithms are designed to automatically tune the weights and biases of a network between input and output. Training algorithms aim to minimize the error, \( E_{ae} \), by updating the weight and bias values in each iteration after reading the input and output of AE training data. Various algorithms have been developed to adjust these parameters. It is still controversial to determine which algorithm will yield best results. Although there are many training algorithms in the literature, the SCG, GD, and RP algorithms have been preferred for training AE since they are relatively faster and use memory more effectively. In this study, the effectiveness of SCG, GD, and RP algorithms for training AE in hyperspectral image classification is examined in detail.

SCG is a supervised training algorithm used in large-scale constrained nonlinear problems and feed forward neural networks (FNN). The SCG algorithm was developed by Møller (Møller 1993), based on conjugate gradient algorithms. Second-order information from FNNs is used to determine the step size and the search direction. Unlike conjugate gradient algorithms, the SCG algorithm refrains line search by using the Levenberg-Marquardt method in each iteration. Line-search is time-consuming because it requires the network training inputs to be calculated several times in each iteration. The line-search requires fewer iterations to converge to the solution since the SCG requires more iterations by significantly reducing the number of computations from each iteration (Møller 1993; Zhou, Yang, and Sensing 2010).

GD, another important algorithm used in machine learning, was developed by Cauchy in 1847 (Barzilai and Borwein 1988). It is a first-order recursive iterative optimization algorithm to investigate the minimum value of a function. The backpropagation or gradient that propagates the error between the network input and output aims to reduce the network error as quickly as possible. In order to determine the local minimum, the gradient of the function at that point is proportional to the negative step. Choosing the appropriate step size is critical for the success of the algorithm. When small step size is selected, the computational cost increases because more iteration is required. However, if a relatively large step size is selected, it causes a skip from global minimum. Furthermore, it is effective to determine whether the network is stable or not and to achieve the desired result (Barzilai and Borwein 1988; Ruder 2016).

RP was developed by Riedmiller and Braun (Riedmiller and Braun 1993) in 1993 for the improvement of the GD algorithm. The GD algorithm does not use the magnitude of the gradient, which is an inherent disadvantage. RP updates the weight according to the local gradient sign. In the weight update, RP aims to not be too affected by the bad influence of the partial derivative. Faster convergence is made by changing the sign of the partial derivative of the corresponding weight each time. RP step size and other adaptive parameters are less than GD and SCG (Çiğiciolu and Beşdoğ 2006; Riedmiller and Braun 1993).

Experiments

In this paper, Salinas-A and Indian Pines HSI benchmark data sets were used. The Salinas-A and Indian Pines benchmark data sets also contain ground truth imagery. Both data sets were provided from the official website of the Computational Intelligence Group of the University of the Basque Country (UPV/EHU).

Indian Pines data set contains 12-bit, 145×145 pixels-sized 200 band images, obtained from the airborne visible/infrared imaging spectrometer (AVIRIS) sensor in the North West Indiana test area. The characteristics of the Indian Pines data set are given in Table 1.

<table>
<thead>
<tr>
<th>Class no.</th>
<th>Class</th>
<th>Number of samples (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>alfalfa</td>
<td>46</td>
</tr>
<tr>
<td>2</td>
<td>corn-notill</td>
<td>1428</td>
</tr>
<tr>
<td>3</td>
<td>corn-mintill</td>
<td>830</td>
</tr>
<tr>
<td>4</td>
<td>corn</td>
<td>237</td>
</tr>
<tr>
<td>5</td>
<td>grass-pasture</td>
<td>483</td>
</tr>
<tr>
<td>6</td>
<td>grass-trees</td>
<td>730</td>
</tr>
<tr>
<td>7</td>
<td>grass-pasture-mowed</td>
<td>28</td>
</tr>
<tr>
<td>8</td>
<td>hay-windrowed</td>
<td>478</td>
</tr>
<tr>
<td>9</td>
<td>oats</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>soybean-notill</td>
<td>972</td>
</tr>
<tr>
<td>11</td>
<td>soybean-mintill</td>
<td>2455</td>
</tr>
<tr>
<td>12</td>
<td>soybean-clean</td>
<td>593</td>
</tr>
<tr>
<td>13</td>
<td>wheat</td>
<td>205</td>
</tr>
<tr>
<td>14</td>
<td>woods</td>
<td>1265</td>
</tr>
<tr>
<td>15</td>
<td>buildings-grass-trees-drives</td>
<td>386</td>
</tr>
<tr>
<td>16</td>
<td>stone-steel-towers</td>
<td>93</td>
</tr>
</tbody>
</table>

The Salinas-A benchmark data set was obtained by AVIRIS displaying the Salinas Valley in California. Salinas-A is composed of 12-bit, 86×83 pixels-sized 224 band images. The characteristics of the Salinas-A data set are given in Table 2.

Due to the various limits of the optimization algorithms used in the training and the calculation equipment available at
hand, it is necessary to reduce the number of hyperparameters owned by DNNs. Decreasing the size of the DNN input vector is very effective in reducing the number of related parameters. In the experiments performed in this paper, NCA was used to reduce the related input-vector dimension of the DNN. In order to solve complex data mining and machine learning problems, there is still a need to develop a more effective method of selecting attributes. The classification of images containing hypersized data such as hyperspectral satellite images is very difficult due to excessive data size. Classifiers used to classify such data are also complex (Li et al. 2019; Liu et al. 2019). The removal of the image bands from the data set, which does not cause an increase in the classification success, increases the success of the classification. (Liu et al. 2019; Medjahed and Ouali 2018; Shukla and Nanda 2018; Wang et al. 2018).

NCA is a method based on the nearest neighborhood technique, generating the weight value for each band by increasing the leave-one-out classification performance over the distance metric given by all data. (Goldberger et al. 2004; Yang, Wang, and Zuo 2012). NCA works faster than some popular feature selection methods and it is not sensitive to distribution of data (Goldberger et al. 2004).

Let \( \Psi = \{(input vector: x, output vector: y)\} \) \( \{(x_i, y_i), \ldots (x_n, y_n)\} \) be training sample set, where \([n, d] \leftarrow x_i\) and \([n, C] \leftarrow y_i\). Here, \(n, d,\) and \(C\) denotes the size of the samples, the dimension of the input vector, and the number of classes, respectively (Goldberger et al. 2004). The purpose of NCA feature selection is to calculate weighting vector, \(w\), to select the feature subset that optimizes the nearest neighbor classifier from all data. The weighted distance between two samples \(x_i\) and \(x_j\) is calculated as (Yang, Wang, and Zuo 2012):

\[
Dist_w(x_i, x_j) = \sum_{j=1}^{d} w_j |x_i - x_j|,
\]

where \(w_i\) is the weight belonging to the \(i^{th}\) band (feature). Then the probability of \(x_i\) selects \(x_j\) as its reference subset and is calculated as (Yang, Wang, and Zuo 2012):

\[
P_i = \frac{\zeta(Dist_w(x_i, x_j))}{\sum_{k=1}^{n} \zeta(Dist_w(x_i, x_k))} \quad \text{if} \quad i \neq j,
\]

\[
P_i = 0 \quad \text{if} \quad i = j
\]

where \(\zeta(z) = \exp(-z/\sigma)\), \(\sigma\) is an input parameter that affects the probability value of the selected reference point. The probability of the correctly classified query point \(x_i\) is found as (Yang, Wang, and Zuo 2012):

\[
P_i = \sum_j y_{iq} P_{iq}.
\]

After that, leave-one-out classification accuracy can be approximately calculated as:

\[
\gamma(w) = \frac{1}{n} \sum_i \sum_j y_{iq} P_{iq},
\]

if \(\sigma \to 0\), \(\gamma(w)\) is an accurate classification accuracy. Furthermore, both the feature selection procedure and the reduce overfitting expression given in Equation 10 can be used (Yang, Wang, and Zuo 2012):

\[
\gamma(w) = \sum_i \sum_j y_{iq} P_{iq} - \frac{1}{n} \sum_j w_j^2
\]
In Equation 10, \( \lambda \) represents the regulation parameter. Because \( \gamma(w) \) is differentiable, its derivative according to \( w \) can be computed (Yang, Wang, and Zuo 2012):

\[
\frac{\partial \gamma(w)}{\partial w_i} = -\left( \frac{1}{\sigma} \sum_i \left( \sum_{j \neq i} P_x |x_i - x_j| - \sum_j y_j P_x |x_i - x_j| - \lambda \right) \right) w_i \tag{11}
\]

Equation 11 is resolved iteratively and weights are calculated by using the related training algorithm. In order to examine the effectiveness of the NCA technique, different numbers of bands were selected for both data sets and the success of the training algorithms were examined in detail. In the experiments carried out in this section, the relevant standard AE structure given in Figure 2 is trained separately by using SCG, GD, RP, and related benchmark data sets. The random initial parameters were used in all experiments. The structure of used AE is summarized as (Epoch:1000; 1st AE hidden layer:50; 2nd AE hidden layer:500; 1st and 2nd AE transfer functions:Purelin; Weight Regularization:0.01; Sparsity Regularization:17; Sparsity Proportion:0.08). Each experiment was renewed using 30 different random seeds to be independent of the initial conditions of the relevant training process. The Mersenne Twister was used as a uniform random number generator in the experiments. The image bands were sorted from the largest to smallest according to the weights calculated with NCA and experiments were made according to this order. Band weights may be requested from the corresponding author. The test accuracy and the AE structure provided by a relevant AE were recorded in each 25 epoch to avoid the negativity of the over-learning effect during testing. The experimental results obtained for Indian Pines and Salinas-A are given in Table 3 and Table 4, respectively. Higher results than the accuracy value obtained by using all of the bands are shown in bold in Table 3 and Table 4.

When Table 3 is examined, it appears SCG, RP, and GD algorithms produce more successful results from the classification accuracy obtained by the use of all bands with only 150, 110, and 140 bands, respectively. When Table 4 is examined, it appears that SCG, RP, and GD algorithms produce more successful results from the classification accuracy obtained by the use of all bands with only 190, 60, and 40 bands, respectively. Figure 3 shows the best classification results for the Indian Pines benchmark data set.
have been obtained as $p = 1.86 \times 10^{-7}$, $R_+ = 30$, and $R_- = 0$. Therefore, the difference between the results of the GD and the results of SCG is statistically significant. Statistical test parameters obtained from a comparison of RP and SCG and have been calculated as $p = 0.06$, $R_+ = 9$, and $R_- = 21$. Hence, difference between the results of the GD and the results of RP is not statistically significant. Statistical test parameters obtained from a comparison of RP and SCG and have been calculated as $p = 1.85 \times 10^{-7}$, $R_+ = 30$, and $R_- = 0$. Thus, the difference between the results of the GD and the results of RP is statistically significant.

The results of the computational complexity analysis of the experiments performed using the RP, SCG, and GD for the Indian Pines benchmark data set are given in runtime seconds in Figure 5.

When Figure 6 is qualitatively examined, it is seen that the classification results obtained for the Salinas-A data set are consistent with ground truth imagery. Accuracy rate values obtained from related tests performed for the Salinas-A data set are shown in Figure 7. The overall accuracy values supplied by GD and SCG for the Salinas-A data set have been

Figure 5. Run time (in seconds) values provided by training algorithms for Indian Pines benchmark data set.

Figure 6. The best classification results of Salinas-A provided by different methods: (a) ground truth, (b) GD, (c) RP, (d) SCG, (e) NCA-GD, (f) NCA-RP, (g) NCA-SCG.

Figure 7. Accuracy rate values provided by training algorithms for Salinas-A benchmark data set.

Figure 8. Run time (in seconds) values provided by training algorithms for Salinas-A benchmark data set.
piecewise compared by using Wilcoxon signed rank test and the statistical test parameters have been obtained as $p = 8.67e^{-7}$, $R^+ = 28$, and $R^- = 2$. Therefore, the difference between the results of the GD and the results of SCG is statistically significant. Statistical test parameters obtained form a comparison of RP and SCG and have been calculated as $p = 1.8626e^{-9}$, $R^+ = 30$, and $R^- = 0$. Hence, difference between the results of the GD and the results of RP is statistically significant.

Statistical test parameters obtained form a comparison of GD and RP and have been calculated as $p = 8.65e^{-7}$, $R^+ = 28$, and $R^- = 2$. Thus, the difference between the results of the GD and the results of RP is statistically significant.

The results of the computational complexity analysis of the experiments performed by using RP, SCG, and GD for the Salinas-A data set are given in runtime seconds in Figure 8.

**Results and Conclusions**

In this paper, the effect of AE training algorithms on the success of the problem of classifying hyperspectral images using AE was investigated. Two hyperspectral image data sets and one standard AE structure were used in the experiments. Hyperdimensional input data is reduced by using NCA, RP, SCG, and GD were used for the training of related AE structures. Experiments were performed 30 times by using different initial conditions. The best results from each experiment (i.e., accuracy rate and runtime values) were compared by using the Wilcoxon signed rank test (significance level; $p$-value = 0.05). The NCA has improved classification accuracy for all training algorithms in both data sets. In addition, this increase in success in the GD algorithm is highest as percentage (7% for Indian Pines and 4% for Salinas-A data set). For Indian Pines data set, the standard deviation of the test results provided by SCG is larger. That is, the RP provided more robust test results than SCG. GD has less test success than RP and SCG. This shows that GD is not as successful as RP and SCG in hyperparameter optimization for the related data and problem. Also, it can be said that SCG and GD are working faster than RP. For Salinas-A data set, RP is more successful and more robust than SCG and GD for the related problem and AE. Furthermore, it can be said that SCG runs faster than GD and RP. Also, the velocities of GD and RP are identical. RP works slower than other methods due to its complex structure. According to the statistical results for the related problem and AE structure, the following results have been found:

1. The results provided by the AEs are sensitive to the training algorithm.
2. RP provides more robust results compared to SCG and GD for two data sets.
3. GD was not as successful as RP and SCG in the experiments conducted in this paper.
4. The performance of SCG and RP are quite similar for Indian Pines data set.
5. Computational complexity of RP is higher than SCG and GD, in general.
6. NCA improves classification performance and reduces computational requirements of related AE structures.

It may be interesting to examine the effectiveness of different feature selection techniques on different DNN structures and training algorithms. Feature selection techniques, especially developed using artificial intelligence optimization algorithms, are quite remarkable for HSI classification. Furthermore, training of DNN structures with evolutionary computational algorithms (partial swarm optimization, cuckoo search, weighted differential search algorithms, etc.) realization is also among the topics of interest.

---

**References**


The 3rd edition of the DEM Users Manual includes 15 chapters and three appendices. References in the eBook version are hyperlinked. Chapter and appendix titles include:

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Authors/Editors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction to DEMs</td>
<td>David F. Maune, Hans Karl Heidemann, Stephen M. Kopp, and Clayton A. Crawford</td>
</tr>
<tr>
<td>2</td>
<td>Vertical Datums</td>
<td>Dru Smith</td>
</tr>
<tr>
<td>3</td>
<td>Standards, Guidelines &amp; Specifications</td>
<td>David F. Maune</td>
</tr>
<tr>
<td>4</td>
<td>The National Elevation Dataset (NED)</td>
<td>Dean B. Gesch, Gayla A. Evans, Michael J. Oimoen, and Samantha T. Arundel</td>
</tr>
<tr>
<td>5</td>
<td>The 3D Elevation Program (3DEP)</td>
<td>Jason M. Stoker, Vicki Lukas, Allyson L. Jason, Diane F. Eldridge, and Larry J. Sugarbaker</td>
</tr>
<tr>
<td>6</td>
<td>Photogrammetry</td>
<td>J. Chris McGlone and Scott Arko</td>
</tr>
<tr>
<td>7</td>
<td>IfSAR</td>
<td>Scott Hensley and Lorraine Tighe</td>
</tr>
<tr>
<td>8</td>
<td>Airborne Topographic Lidar</td>
<td>Amar Nayegandhi and Joshua Nimetz</td>
</tr>
<tr>
<td>9</td>
<td>Lidar Data Processing</td>
<td>Joshua M. Novac</td>
</tr>
<tr>
<td>10</td>
<td>Airborne Lidar Bathymetry</td>
<td>Jennifer Wozencraft and Amar Nayegandhi</td>
</tr>
<tr>
<td>11</td>
<td>Sonar</td>
<td>Guy T. Noll and Douglas Lockhart</td>
</tr>
<tr>
<td>12</td>
<td>Enabling Technologies</td>
<td>Bruno M. Scherzinger, Joseph J. Hutton, and Mohamed M.R. Mostafa</td>
</tr>
<tr>
<td>13</td>
<td>DEM User Applications</td>
<td>David F. Maune</td>
</tr>
<tr>
<td>14</td>
<td>DEM User Requirements &amp; Benefits</td>
<td>David F. Maune</td>
</tr>
<tr>
<td>15</td>
<td>Quality Assessment of Elevation Data</td>
<td>Jennifer Novac</td>
</tr>
</tbody>
</table>

This book is your guide to 3D elevation technologies, products and applications. It will guide you through the inception and implementation of the U.S. Geological Survey’s (USGS) 3D Elevation Program (3DEP) to provide not just bare earth DEMs, but a full suite of 3D elevation products using Quality Levels (QLs) that are standardized and consistent across the U.S. and territories. The 3DEP is based on the National Enhanced Elevation Assessment (NEEA) which evaluated 602 different mission-critical requirements for and benefits from enhanced elevation data of various QLs for 34 Federal agencies, all 50 states (with local and Tribal input), and 13 non-governmental organizations.

The NEEA documented the highest Return on Investment from QL2 lidar for the conterminous states, Hawaii and U.S. territories, and QL5 IfSAR for Alaska.

Chapters 3, 5, 8, 9, 13, 14, and 15 are “must-read” chapters for users and providers of topographic lidar data. Chapter 8 addresses linear mode, single photon and Geiger mode lidar technologies, and Chapter 10 addresses the latest in topobathymetric lidar. The remaining chapters are either relevant to all DEM technologies or address alternative technologies including photogrammetry, IfSAR, and sonar.

As demonstrated by the figures selected for the front cover of this manual, readers will recognize the editors’ vision for the future – a 3D Nation that seamlessly merges topographic and bathymetric data from the tops of the mountains, beneath rivers and lakes, to the depths of the sea.

Co-Editors
David F. Maune, PhD, CP and
Amar Nayegandhi, CP, CMS

**PRICING**

- **Student (must submit copy of Student ID)**: $50 +S&H
- **ASPRS Member**: $80 +S&H
- **Non-member**: $100 +S&H
- **E-Book (only available in the Amazon Kindle store)**: $85
ASPRS Offers
» Cutting-edge conference programs
» Professional development workshops
» Accredited professional certifications
» Scholarships and awards
» Career advancing mentoring programs
» PE&RS, the scientific journal of ASPRS

asprs.org