

Paper ID: 8402058 Reviewers Comments

Paper title: GEOSPATIAL ANALYSIS OF FLOOD RISK AND VULNERABILITY ASSESSMENT ALONG RIVER BENUE BASIN OF KOGI STATE

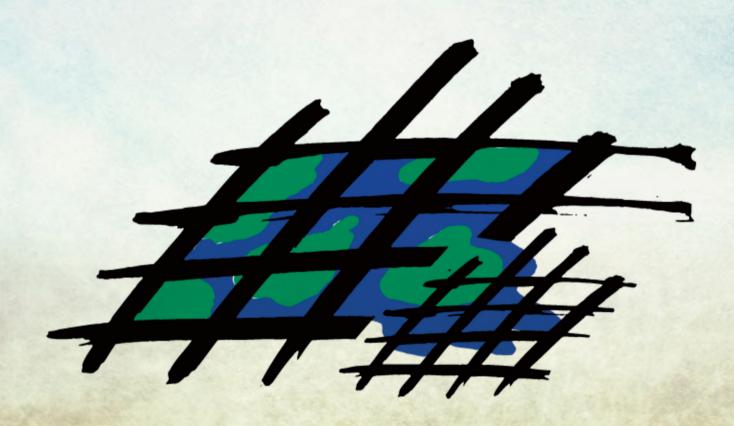
The study investigated flood risk and assesses vulnerability to flood along the river Benue basin of Kogi state using geospatial techniques and survey. The study delineated the areas at flood risk, classified the vulnerable areas and pointed out the socio-economic factors which increase the population vulnerability in the river Benue basin of Kogi state. Digital Elevation Model (DEM) and Landsat Thematic Mapper (TM) images for 2008, 2012 and 2015 were used. These images were processed, digitized, interpolated, classified, overlaid and buffered using the ArcMap 10.3.1 GIS software to generate classified land use/cover maps, topographic map and flood physical vulnerability maps. Household questionnaires which contained structured questions were distributed and obtained to analyse the social-economic vulnerability of the study area using descriptive analysis. The study revealed that areas within 6 kilometre buffer from the River Benue are most vulnerable to flood hazard with the vulnerability decreasing towards the southern part. This work is interesting and has the potential of publication. However, a revised version needs to consider the following issues:

- 1. One of the concerns is the images processed, which extend only to 2015, while we are currently in the last quarter of 2019. Perhaps this is a very old paper that is being resubmitted.
- 2. In the Introduction part, there is no attempt to link the issue discussed in the paper to already existing studies and highlight how it is similar or distinct from existing research. Authors should be explicit on the work of other researchers and their deficiencies, then establish the need for the present work.
- 3. The Maximum Likelihood method used in the research needs some explanations. It would give a better idea to readers if the author adds a flow chart to the methodology.
- 4. Some obvious grammatical errors should be fixed before publication e.g. "Table 6 and 7 shows that,..." should be "Table 6 and 7 show that,...".
- 5. To the conclusion, I suggest adding a concluding topic in which the author provides the limitations of the research and opportunities for future research.





Journal of Geographic Information System





Journal Editorial Board

ISSN: 2151-1950 (Print) ISSN: 2151-1969 (Online)

https://www.scirp.org/journal/jgis

Editorial Board

Dr. Alia I. Abdelmoty Cardiff University, UK

Dr. Tofael Ahamed University of Tsukuba, Japan

Prof. Dafer Algarni King Saud University, Saudi Arabia

Prof. Carlos Marcelo Di Bella Instituto de Clima y Agua CIRN-INTA Castelar, Argentina

Prof. Christophe Claramunt Naval Academy Research Institute, France

Prof. Weihong Cui International Europe and Asia Academy of Sciences, China

Prof. Fuchu Dai Chinese Academy of Sciences, China

Prof. Anrong DangTsinghua University, ChinaDr. Ahmed F. ElaksherCal Poly Pomona, USAProf. Hossein GhadiriGriffith University, Australia

Prof. Shuanggen JinShanghai Astronomical Observatory, ChinaProf. Saro LeeKorea Institute of Geoscience and Mineral Resources, South Korea

Dr. Xiang Li East China Normal University, China

Prof. Michel Mainguenaud Institut National des Sciences Appliquées, France

Prof. Jordi Martí-Henneberg University of Lleida, Spain

Prof. Hervé Martin Joseph Fourier University, France Dr. Kaveh Ostad-Ali-Askari Islamic Azad University, Iran

Prof. Christopher James Pettit The University of Melbourne, Australia

Dr. Saied PirastehUniversity of Waterloo, Canada

Prof. Biswajeet Pradhan Dresden University of Technology, Germany

Prof. Michael D. Proctor University of Central Florida, USA

Prof. Xiaosheng QinNanyang Technological University, Singapore

Dr. Himan ShahabiUniversity of Kurdistan, Iran**Prof. Shih-Lung Shaw**University of Tennessee, USA

Dr. P. S. Sriraj University of Illinois at Chicago, USA

Prof. Eddy Lynn Usery Center of Excellence for Geospatial Information Science, USA

Dr. Herna Viktor University of Ottawa, Canada

Prof. Jianhua XuEast China Normal University, ChinaProf. Xiaohuan YangChinese Academy of Sciences, China

Prof. Axing Zhu University of Wisconsin, USA

Dr. Liang-feng Zhu East China Normal University, China



ISSN Online: 2151-1969 ISSN Print: 2151-1950

Table of Contents

Volume 12	Number 1	February 2020
Geospatial Analy River Benue Basi	rsis of Flood Risk and Vulnerability Assess in of Kogi State	ment along
E. T. Umaru, A.	Adedokun	1
•	Applied to the Extraction of Road Geomet sifiers, Northeastern Brazil	ric Features Based on Optimum
M. Macedo, M.	Maia, E. Kohlman Rabbani, O. L. Neto	15

Journal of Geographic Information System (JGIS) Journal Information

SUBSCRIPTIONS

The *Journal of Geographic Information System* (Online at Scientific Research Publishing, https://www.scirp.org/) is published bimonthly by Scientific Research Publishing, Inc., USA.

Subscription rates:

Print: \$89 per issue.

To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements

Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)

Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.

E-mail: sub@scirp.org

COPYRIGHT

Copyright and reuse rights for the front matter of the journal:

Copyright © 2020 by Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

http://creativecommons.org/licenses/by/4.0/

Copyright for individual papers of the journal:

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.

Reuse rights for individual papers:

Note: At SCIRP authors can choose between CC BY and CC BY-NC. Please consult each paper for its reuse rights.

Disclaimer of liability

Statements and opinions expressed in the articles and communications are those of the individual contributors and not the statements and opinion of Scientific Research Publishing, Inc. We assume no responsibility or liability for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained herein. We expressly disclaim any implied warranties of merchantability or fitness for a particular purpose. If expert assistance is required, the services of a competent professional person should be sought.

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:

E-mail: jgis@scirp.org



ISSN Online: 2151-1969 ISSN Print: 2151-1950

Geospatial Analysis of Flood Risk and Vulnerability Assessment along River Benue Basin of Kogi State

Emmanuel T. Umaru¹, Adeniyi Adedokun²

¹Department of Urban and Regional Planning, Federal University of Technology, Minna, Niger State, Nigeria ²Centre for Disaster Risk Management and Development Studies CDRM & DS, Federal University of Technology, Minna, Niger State, Nigeria

Email: eumaru@futminna.edu.ng

How to cite this paper: Umaru, E.T. and Adedokun, A. (2020) Geospatial Analysis of Flood Risk and Vulnerability Assessment along River Benue Basin of Kogi State. *Journal of Geographic Information System*, 12, 1-14.

https://doi.org/10.4236/jgis.2020.121001

Received: October 4, 2019 Accepted: January 14, 2020 Published: January 17, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/





Abstract

River flooding has become a widely distributed and devastating natural disaster that has caused significant damages both economically and socially in Nigeria. Although, the Nigeria Meteorological Agency (NiMet) are always on time to provide flood warnings prior to flood disasters, spatial information pertaining to the areal extent vulnerable to the hazard are not usually made available. This study focuses on the application of geospatial techniques and quantitative research method for the assessment and analysis of flood risk and vulnerability assessment along the River Benue Basin of Kogi State. Digital Elevation Model (DEM) and Landsat Thematic Mapper (TM) images for 2010, 2014 and 2018 were used. These images were processed, digitized, interpolated, classified, overlaid and buffered using the ArcMap 10.3.1 GIS software to generate classified land use/cover maps, topographic map and flood physical vulnerability maps. Household questionnaires which contained structured questions were distributed and obtained to analyse the social-economic vulnerability of the study area using descriptive analysis. The study revealed that areas within 6-kilometre buffer from the River Benue are most vulnerable to flood hazard with the vulnerability decreasing towards the southern part. The two urban areas in the study area (Abejukolo and Oguma) are vulnerable to flood threat. Abejukolo is vulnerable due to high water footprint in the town and its surroundings. The factors that increase flood vulnerability of the inhabitants are low elevation, high illiteracy rate, high numbers of dependent population, poverty, low quality housing, poor land use and planning, inadequacy or absence of risk communication tools. The study recommends the development of guidelines for land use, continuous vulnerability mapping of the study area especially in the flood risk probability months (April to October), construction of dams to contain release water from the upstream, collaboration of NEMA with the local Government and traditional rulers for effective risk communication and relocation of communities lying on the river Benue basin in other to reduce the population vulnerability to flood impacts.

Keywords

Geospatial, Vulnerability, Flood, Disaster

1. Introduction

The National Oceanic and Atmospheric Administration [1] described flood as a discharge of water more than the carrying capacity of the available channel. It is a natural disaster which has considerable consequences on human life and the neighbouring environment. Flooding has claimed several lives, forced millions of people to flee their habitual homes, and caused serious ruins of assets and downgrading of many farmlands, wetlands, forest and other natural resources of the environment worldwide. According to [2], flood disaster is regarded as the most common disaster in the world and evenly spread worldwide. In Nigeria, flooding has led to considerable socioeconomic wreck than any other natural disaster. Several literatures have shown that these three types of flood are the common flood disasters that have taken place in the Nigeria in years past. Among the highlighted flood types, river flooding is the most huge and usually most damaging. River flooding is notably extensive with regard to its drainage area, unlike coastal and urban flooding. For example, the major rivers in Nigeria (Niger and Benue) emerge by many urban areas. According to [3], river flooding takes place in the floodplains of the great rivers. The effects of global warming and unregulated human activities contribute a great deal to this occurrence. Flood disaster is intensifying and quickly turning to an annual occurrence in Nigeria especially in the coastal regions while the dry northern regions of the country are not excluded too. According to a report by the Nigerian Hydrological Service Agency [4], devastating flood disasters have occurred in the last 5 years in more than 15 states of Nigeria. The state includes Lagos, Oyo, Rivers, Niger, Kwara, Benue, Edo, Anambra, Ondo, Borno, Kogi, Osun, Delta, Sokoto, Ogun, Oyo, and Ekiti among others. It was also recorded that these flooding occur mostly between the month of July and October when excessive rainfall is experienced in the country. However, many of these flooding were linked to dam release/failure or other disequilibrium effects on the river channels such as tidal waves. Flooding has led to severe loss of lives, and properties along the river Niger and Benue basin. Many communities along the basin are constantly relocating and each time they relocate, they usually relocate to another flood risk zone which often leads them to another loss. Nigeria as a country has paid less attention to this risk. The approach to flood management requires a change of paradigm. In other to effectively manage future flood disaster worries in flood prone areas of the federation, there is need for the National Emergency Management Agency (NEMA) in collaboration with other disaster management stakeholders to be able to forecast, predict and relocate people who at the risk of flood disaster to safe location before flooding. To achieve this, it is important to embrace a competent model with geo-visualization proficiency to investigate, model and create figures for accurate and effectual judgement. Such task can be achieved through the use of geospatial technology. Geospatial technology involves the synthesis of Geographical Information System (GIS) and remote sensing (RS) methods. It is designed in a way that allows wide varieties of data to be inserted, manoeuvred, examined, scrutinised, and presented as information attributed to the earth. Moreover, [5] [6] [7] all affirmed that the use of information gotten remotely in GIS system has proven to be the best approach for evaluating river flooding and risk determination.

The paper aims at examining the area at flood risk while also determining the factors that increase the vulnerability of communities at flood risk.

2. The Study Area

The study area covers the river Benue basin in Kogi State, as shown in **Figure 1**. The areas covered are Bassa and Omala Local Government Area. Bassa Local Government is located between latitude 7°54'N and 7.90°N and longitude 7°03'E and 7.05°E while Omala Local Government is located between latitude 7°49'N and 7.817°N and longitude 7°31'E and 7.517°E. The two local governments are located in the eastern part of Kogi State and traversed by the River Benue. According to the 2006 census figure by NPC, the population figures of Bassa and Omala LGA are 139,993 and 108,402 respectively with area coverage of 1925 km²

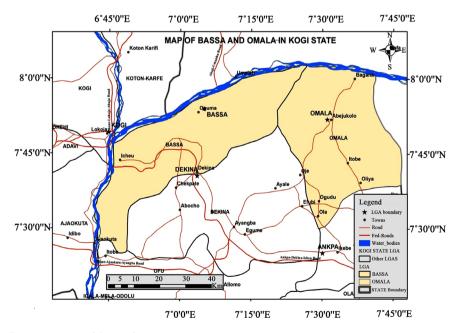


Figure 1. Map of the study area.

and 1667 km² respectively at the 2006 census. Annual rainfall in the study area is constantly high, with the yearly mean usually between 1100 mm and 1300 mm [8]. The temperature is seasonal, relatively high and ranges between 22°C and 32°C, having February through to March as the months with highest temperature [3]. Bassa and Omala Local Government Area are inhabited by different tribes. The Igallas and Nupes are the dominant tribe in the two local governments' area. The two local governments area has about 100 settlements which includes populated towns such as Oguma, Abejikolo, Ife, Ogodu, Akishamishi, Echeu, Akakana, Akabasisi, Odeyyi, Bagana, Gboloko and Ozonogulo. The villages and settlements are mostly lineal; people live according to clans thereby giving way to the circular style.

Their major occupation includes agricultural practices, civil servant, trading and agrarian peasantry. The presence of sedimentary rocks and alluvium along the riverbeds, encourage fishing and farming. Each family produces different crops to meet their needs by farming at different altitudes of sierra. Notable among the farm produce in Kogi state include cassava, maize, coffee, cocoa, palm oil, cashews, groundnuts, yam, rice and melon. In their highlands, rearing of cattle is common among the people. Notable drainage features in the study area are Egeneja Warm Spring at Bassa Local Government Area and the Mimi River. Most towns and villages in the study area lack adequate water channel system as a result the low-lying nature and geology which results in water-logging in several areas.

3. Data and Research Methodology

The data used for this study was gotten from both primary and secondary sources. Primary data was collected through questionnaires, interviews and Geographic Positioning System (GPS) receiver in other to obtain coordinates of flood prone communities in the study area. Secondary data includes satellite imageries, topography maps and population data.

The questionnaire was aimed at assessing the socio economic vulnerability of some selected communities in the study area to flooding while The geographical information system (GIS) shape files gotten from digitized satellite imageries is aimed at determining areas at flood risk and assessing physical vulnerability of the areas to flood.

Images acquired from Landsat 8 on October 2010 (before 2012 flood event), October 2012 (during flood event) and 2015 (after flood event) were classified in order to assess the land cover change using ArcMap 10.3.1. Land use/cover change obtained from the classification of 2010, 2014 and 2018 images was processed and analysed within ArcMap 10.3.1 environment. A false colour composite of all the satellite imageries was created to distinct land use/cover which includes bare soil, vegetation, water bodies and built areas. With the use of Maximum Likelihood method, supervised image classification was conducted to obtain distinct land use and land cover of the terrain before and after flood. Digital Elevation

Model was processed and classified within ArcMap 10.3.1 environment to obtain the elevation map of the area under study. Spatial data, which include land use/cover, elevation, river networks and channels retrieved in raster format, was vectorised using ArcGIS 10.3.1 sofware. Analytical operations which include data layers coordinate systems projection was carried out to integrate these data from different sources.

3.1. Fieldwork and Primary Data Collection

The field work was carried out to validate the flood extent as captured by the satellite imageries and to identify other areas inundated by the flood. GPS receivers and printed copies of satellite imageries were used as the field tools to delineate inundated areas.

A survey of some selected communities in the study area was carried out during the field work and basic data on the socio-economic indices such as age, sex, literacy rate, employment status, etc. Were obtained using questionnaire and compiled.

3.2. Spatial Analysis Techniques

Spatial analyses of the study area carried out include buffer analysis and classification of terrain elevation. Vectorised river Benue basin including its tributaries was buffered. 20-kilometre buffer distance was marked on the Kogi state part of the Benue basin. Progressively, a 6-kilometre, 13-kilometre and 20-kilometre buffer distances were identified on the Kogi state part of the Benue basin. Terrain elevation was classified into two categories of 0 - 30 meters and 30 - 545 meters to represent low and high respectively.

3.3. Hazard Zoning

This different buffered layer was overlaid on the vectorised topographic strata using the ArcGis 10.3.1 software. The spatial intersect and erase overlay techniques was used to extract the potential flood risk areas from the overlaid buffered river basin and topographic map. The flood prone areas were classified with respect to magnitude based on elevation and distance from the river channels and water footprints in the study area. The study area was categorized into zones (high hazard, medium hazard, low hazard, and no hazard zones).

3.4. Assessment of Socio-Economic Vulnerability

Primary data gotten through household questionnaire deployed to some selected communities in the study area constitute the baseline data for this analysis. 364 households were sampled. Variables extracted from the baseline data were divided into demographic, socioeconomic and early warning system variables. These variables include age (Children and elderly population), gender (male and female), literacy rate, income rate, housing types, early warning systems forecast, gadgets and information dissemination.

4. Results and Analysis

4.1. Change in Land Use/Cover for 2010, 2014 and 2018

Figures 2(a)-(c) present the land-use/cover data for the year 2010, 2014 and 2018 respectively. The land use/cover classification comprises of four categories: soil or rock, vegetation, water bodies and built-up areas.

Table 1 showed the changes that have occurred on land use and land cover in the study area between 2010, 2014 and 2018. The rate of increase/decrease between 2010 and 2014 on soil/rock, vegetation, water bodies, and built-up areas were -5.9%, 1.1%, 6.1%, and 0.9% respectively. The highest positive change took place on water bodies, while the highest negative change took place on soil/rock in 2010/2014. This change could be attributed to the release of water from the Ladgo dam in Cameroon in 2012 which flooded the study area.

The rate of increase/decrease between 2014 and 2018 on soil/rock, vegetation, water bodies, and built-up areas were 8.3%, -8.3%, -8.0%, and 8.2% respectively while 2010 to 2018 increases were 2.4%, -9.4%, -1.9%, and 9.1% respectively. Thus, the highest positive change occurred on soil/rock and built-up areas while the highest negative change occurred on vegetation and water bodies in 2014/2018 as in 2010/2018. The positive changes can be attributed to overgrazing, increase in farming practices and rise in demand for shelter due to increasing human population. The negative changes could be attributed mainly to increasing deforestation, increase in drought event as a result of climate change effect, and intensified use of wetland for agriculture purposes.

Consequently, the increase in soil/rock, built-up areas, and decrease in vegetation due to population increase, economic activities, use of trees as fuel, logging coupled with low elevation can increase the vulnerability of households in the Benue basin to flood hazard.

4.2. Elevation of the Study Area

Figure 3 presents the elevation of the study area. Elevation order was obtained by classification of the DEMs and interpolation on base map of the study area. The figure shows that, the south eastern communities including Ojaji, Oliya, and Ebeje, are located within the high-elevated land. Naturally, low slope and low elevation are areas along the river basin watershed such as Oguma, Mozu, Ogba,

Table 1. Summary of land use and Land cover (in Hectares) for the study area in 2010, 2014 and 2018.

LULC Class	2010	2012	2018	% Change 2010-2014	% Change 2014-2018	% Change 2010-2018
Soil/rock	0.201	0.160	0.235	-5.9	8.3	2.4
Vegetation	0.595	0.623	0.548	1.1	-8.3	-9.4
Water body	0.048	0.106	0.033	6.1	-8.0	-1.9
Built-up	0.006	0.015	0.088	0.9	8.2	9.1

Source: Author's analysis, 2019.

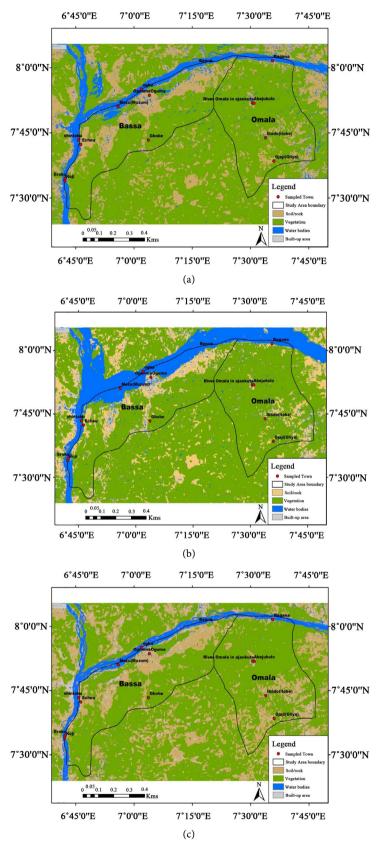


Figure 2. (a) Land-use/cover of 2010; (b) Land-use/cover of 2014; (c) Land-use/cover of 2018.

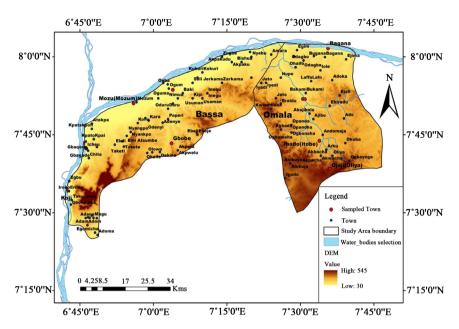


Figure 3. Elevation map of the study area.

Shintaku, Echeu, Eroko, Koji, Abejukolo, and Bagana and so are vulnerable to flooding. Therefore during preparation for flooding, people in the vulnerable communities can be evacuated and relocated to the communities in high elevated areas.

4.3. Distance from River Channel

Figure 4 presents the buffered map of the study area. This was performed to determine flood risk extent in accordance with proximity to river channel. It appears that areas, such as Oguma, Mozu, Bagana, Ogba, Icheu, Eroko, Koji and Akpaku, near the river network (6 km) are highly vulnerable to flood hazard. Areas such as Kofi, Kara, Odagba and Ohaife are at distance of 13 kilometres to the river, and are less vulnerable to flood hazard. The Southern region of the study area is not vulnerable to flood disaster by virtue of distance to the river Benue. They can be regarded as the safe zones where people can be advised to live or evacuation site during flooding in other to be safe from flood disaster. However, due to some active water channels such as river Omala in the southern eastern part, areas such as Abejukolo, Itobe, Ibado, Ajibu and it environment are vulnerable to flooding. Nevertheless, the effect of this parameter decreases in distances.

5. Analysis of Socioeconomic Vulnerability

5.1. Children and Elderly Population

Table 2 shows that children and elderly population constitutes 40.8% of the total population of the study area. The high percentage of the children and elderly population in the study area implies that the study area is highly vulnerable. This is as a result of increase in the burden of care following flooding and the ability

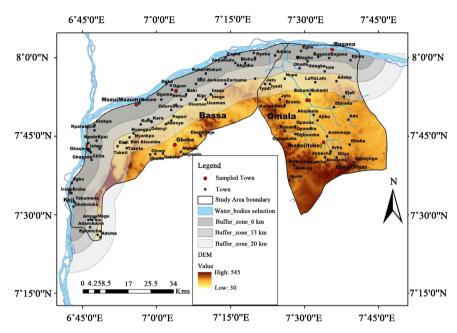


Figure 4. Buffer of river basin with base map.

Table 2. Age distribution of respondents.

Age Range (Years)	Respondents (%)
0 to 14	25.1
15 to 24	19.0
25 to 59	40.2
>60	15.7

to mobilise them out of harm's way during flooding. The elderly population is characterised by a decline in functional ability which could inhibit their state of alertness, risk perception, and mobility. This can considerably increase their vulnerability therefore making them potential victims of any natural disaster including flooding.

5.2. Gender

Table 3 shows that female population constitutes 59.4% of the total population of the study area shows the dominance over male population (40.6%). This implies that the study area's population is vulnerable to flood disaster. Factors that increase the female vulnerability include restricted movement, lack of formal education due to religious and cultural believe, working of women in low wage informal sectors and earning lower wage than men, high household responsibilities of women which ranges from being a mother and care giver for both their children and the elderly. Moreover, until recently, most females in the study area commit their lives to household activities from an early age and have lesser opportunities to be involved in educational and social activities than males. This makes the female population highly vulnerable to natural hazards.

Table 3. Gender distribution of respondents.

Gender	Respondents (%)		
Male	40.6		
Female	59.4		

5.3. Illiterate Population

Table 4 shows that 52% of the population in the study area (17.3% and 34.7% primary education and lack of formal education respectively) constitute the illiterate population. This implies that the study area is vulnerable to flood. Illiterate people lack understanding about hazard cause, trigger factors, severity and effects. They are also ignorant of how to prepare for emergency mostly depending on others to prepare and evacuate themselves during an imminent disaster. Moreover, the inability of an illiterate to read or write constitutes a significant difficulty for prompt action during emergency situation.

5.4. Housing Type and Location

Table 5 shows that 31% and 9.3% of houses in the study area were constructed with mud and Mud/wooden frame respectively while 63.5% of the houses were located in risk areas. The quality of the Cement and mud houses found in the study area are substandard, while wooden houses are inferior and therefore at risk of severe damage during flood. Generally, residents in the study area lack maintenance and renovation culture even as substandard buildings have an extremely high density in flood plain areas. The damage caused by natural hazards is likely to be enormous in the study area as a result of great vulnerability exhibit by buildings located in close proximity to water bodies and substandard houses built on flood plains. The results presented on Table 5 showed that housing quality were not only poor but mostly situated in high flood risk areas. The great numbers of human lives and properties that reside in such poorly constructed buildings and risk areas are at great risk of flood hazard. The infrastructure and poorly constructed houses in the low-lying places and flood plains of the study area will be easily flooded and damaged during flooding. These factors could turn the natural hazard into a disaster.

5.5. Poverty

Table 6 and Table 7 show that, 57% constitute population of low income group, 41.6% constitute the population of the middle income group while only 1.3% of the population constitute the high income group. Unemployed members and the student population the study area constitute 34.5% and 28.2% respectively, while others (37.3%) are unemployed. Natural disasters impact more on the poor population because of their lack of resources to build or rebuild their homes and also meet other critical needs during a disaster cycle. This makes them unable to make adequate preparation for flood hazard thereby reducing their resistance and

Table 4. Distribution of literacy rate.

Literacy Rate	Respondents (%)
No Formal Education	34.7
Primary	17.3
Secondary	13.8
Post-Secondary	34.7

Table 5. Housing types distribution.

House Type	Respondents (%)	Location	Respondents (%)
Cement	37.9	Highland	36.5
Mud	31	Floodplains	21.4
Cement/Mud	20.1	Close to canal	18.1
Wood/Mud	9.3	Close River channel	23.9
Others	1.6		

Table 6. Income range distribution.

Income Class	Income Range (Naira)	Respondents (%)
Very low	<13,000	24.5
Low	13,000 to 45,000	32.5
Middle	45,000 to 100,000	41.6
High	>100,000	1.3

Table 7. Employment status distribution.

Employment Status	Respondents (%)
Employed	34.5
Unemployed	37.3
Students	28.2

coping capacity to flood hazard. This result shows that majority of the population in the study area are linked with poverty thereby making them more vulnerable to flood disaster.

5.6. Assessment of Flood Early Warning System (EWS)

Table 8 presents the assessment of flood forecasting, communication gadgets and alert dissemination of respondents in the study area. Analysis shows that majority (29.9%) of the households in the study area do not have access to any weather information, 36.5% access weather information through audio vision (radio and television), 17.9% access weather information through traditional methods while 25% rely on local authority for weather information.

Table 8. Flood forecasting, communication gadgets and alert dissemination.

Medium of EWS	Respondents (%)	Accessible Weather Elements	Respondents (%)	Forecast Frequency	Respondents (%)
Television	15.1	Temperature	32.1	Daily	11.8
Radio	21.4	Precipitation	30.8	Weekly	9.1
Newspaper	0	Relative Humidity	3.6	Monthly	2.7
Internet	0.5	Pressure	1.6	Occasionally	50.5
Local Authority	25	Cloud Cover	1.9	Non	25.8
Traditional Method	17.9	Non	29.9		
Telephone	0.5				
No Medium	19.5				

Analysis shows that majority (62.9%) of the households that have access to weather information can only access temperature and precipitation while majority (50.5%) of the respondents only access weather information occasionally. This shows the high rate of vulnerability as a result of inadequate early warning systems in the study area.

The residents of the study area have faith in the use of indigenous method for forecasting flood year. However this method will not be able to forecast flood caused by dam release or failure. This has been the cause of flooding in the study area over the years.

6. Flood Hazard Map

Figure 5 shows that, aside communities such as Mozu, Ogba, Bagana, Icheu, etc, with close proximity to the River Benue, Other communities such as Gbobe and Abejukolo are highly susceptible to floods. This is as a result of local River channels (River Omala) whose source is the river Benue flowing downstream toward Abejukolo and the surrounding areas. Other communities such as Ibado, Nupe, Taperi, Inyanpo, etc. has Medium to low flood risk. However, when there is huge discharge of water from upstream dams and the river exceeds it bank at full stage for days, the low and medium susceptible areas will become highly susceptible to flood.

The flood hazard map also provide existing risk information for land use planning and accurate development plans for new and existing urban areas such as Oguma and Abejukolo communities. Oguma and Abejukolo are major urban areas in the study area and are highly susceptible to flood. The hazard map will enable developers and engineers to perform satisfactory evaluation of costs of flood damages and flood reduction benefits.

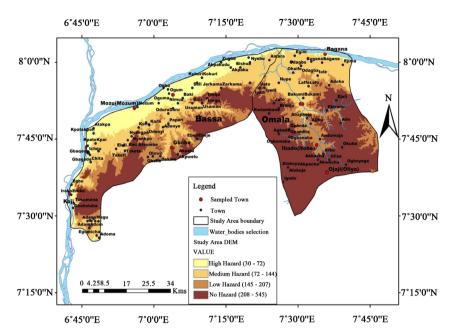


Figure 5. Flood hazard map of the study area.

7. Conclusion

The study has investigated flood risk and assesses vulnerability to flood along the River Benue Basin of Kogi state using geospatial techniques and survey. The study delineated the areas at flood risk, classified the vulnerable areas and pointed out the socio-economic factors which increase the population vulnerability in the river Benue basin of Kogi State. Many lives and property are liable to be damaged by flood as long as people continue to live in these areas prone to flood. The analysis shows that many communities settled too close to the river banks. In some cases, it seems the communities (Mozu and Bagana) are even on the river pathway. The influence of poverty on disaster cannot be overemphasised. To effectively manage risk and reduce vulnerability, economic safety nets and jobs opportunities should be created. This will eradicate poverty to a minimum level which in turn will help achieve the goal of disaster risk reduction. This study advocates the need for paradigm shift from the Nigeria conventional post disaster phases of flood disaster management which include response and resettlement to also include pre disaster phases of prevention and preparedness. A workable, proactive and people centred policy and plan together with results from this effort will greatly lower the vulnerability to flooding and provide an authentic control and dependable management.

Recommendations

Public awareness and education should be made in understandable languages by all relevant disaster management stakeholders in collaboration with the traditional institutions and faith based organisations in the communities on the dangers associated with constructing buildings on areas prone to flood and other practices that can alter the environment and originate risk to individuals.

There is need for constant flood stage monitoring and continual vulnerability mapping by the relevant agencies especially during the months of April to October when there is high probability of flooding.

Government must develop guidelines for land use in the river basins and ensure full observance of development protocols for the study area as regards the encroachment of flood plains particularly at Bagana, Eroko, Abejukolo, Oguma, Mozu and Icheu.

All the communities that are located at the banks and within the buffer distance of 6 km of River Benue and its major tributaries and which are prone to floods should be urgently relocated while enacting legislation that would ban new development in floodplains of the major towns such as Bagana, Mozu and Ogba.

NEMA should collaborate with the Local Government Council and the traditional rulers of the various communities on the Benue basin to provide flood warning systems and effective flood risk communication to the inhabitants.

Conflicts of Interest

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

References

- [1] National Oceanic and Atmospheric Administration (NOAA) Coastal Services Centre, Risk and Vulnerability Assessment Steps (1998) Environmental Analysis Risk and Vulnerability Assessment Tool (RVAT).
- [2] DMSG-Disaster Management Support Group (2001) The Use of Earth Observing Satellites for Hazard Support Group. Final Report, NOAA, Department of Commerce, USA.
- [3] Ologunorisa, T.E. (2004) An Assessment of Flood Vulnerability Zones in the Niger Delta, Nigeria. *International Journal of Environmental Studies*, **61**, 31-38. https://doi.org/10.1080/0020723032000130061
- [4] Nigeria Hydrological Services Agency (NHSA) (2016) National Flood Outlook for 2016. http://www.nihydro.gov.ng
- [5] Jayasselan, A.T. (2006) Drought and Floods Assessment and Monitoring Using Remote Sensing and GIS [Electronic Version]. Satellite Remote Sensing and GIS Application in Agricultural Meteorology, 2006, 291-313.
- [6] Ishaya, S., Ifatimehin, O.O. and Abaje, I.B. (2009) Mapping Flood Vulnerable Areas in a Developing Nigeria. *Journal of Sustainable Development in Africa*, **11**, 180-194.
- [7] Irimescu, A., Craciunescu, V., Stancalie, G. and Nertan, A.T. (2010) Remote Sensing and GIS Techniques for Flood Monitoring and Damage Assessment Study Case in Romania. 4th International Scientific Conference BALWOIS, Ohrid, 25, 29 May 2010, 1-10.
- [8] Clement, A.R. (2012) Causes of Seasonal Flooding in Flood Plains: A Case of Makurdi, Northern Nigeria. *International Journal of Environmental Studies*, 69, 904-912. https://doi.org/10.1080/00207233.2012.730668



ISSN Online: 2151-1969 ISSN Print: 2151-1950

Remote Sensing Applied to the Extraction of Road Geometric Features Based on Optimum Path Forest Classifiers, Northeastern Brazil

Márcia Macedo^{1,2*}, Maria Maia¹, Emilia Kohlman Rabbani², Oswaldo Lima Neto¹

¹Post-Graduate Program in Civil Engineering, UFPE, Pernambuco, Brazil ²Post-Graduate Program in Civil Engineering, UPE, Pernambuco, Brazil Email: *marcia.macedo@poli.br

How to cite this paper: Macedo, M., Maia, M., Kohlman Rabbani, E. and Neto, O.L. (2020) Remote Sensing Applied to the Extraction of Road Geometric Features Based on Optimum Path Forest Classifiers, Northeastern Brazil. *Journal of Geographic Information System*, 12, 15-44.

https://doi.org/10.4236/jgis.2020.121002

Received: January 22, 2020 Accepted: February 24, 2020 Published: February 27, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/





Abstract

One of the principal difficulties related to road safety management in Brazil is the lack of data on road projects, especially those on rural roads, which makes it difficult to use road safety studies and models from other countries as a reference. Updating road networks through the use of hyperspectral remote sensing images can be a good alternative. However, accurately recognizing and extracting hyperspectral images from roads has been recognized as a challenging task in the processing of hyperspectral data. In order to solve the aforementioned challenges, Hyperion hyperspectral images were combined with the Optimum Forest Path (OPF) algorithm for supervised classification of rural roads and the effectiveness of the OPF and SVM classifiers when applied to these areas was compared. Both classifiers produced reasonable results, however, the OPF algorithm outperformed SVM. The higher classification accuracy obtained by the OPF was mainly attributed to the ability to better distinguish between regions of exposed soil and unpaved roads.

Keywords

Roads, Multispectral Images, Hyperspectral Images, Optimum Path Forest Algorithm

1. Introduction

Traffic accidents in Brazil stand out in terms of magnitude, both in number of deaths and injuries as well as in their financial consequences for users and for society. According to the World Health Organization's world ranking [1], 37,345 traffic deaths were recorded in 2016 by the Ministry of Health's Mortality In-

formation System (SIM). The country remains far from the goal established by the United Nations (UN), which stipulates a 50% reduction in the number of victims over 10 years, beginning in 2011, and it ranks fifth among countries with the most traffic deaths, behind only India, China, the United States, and Russia. In 2016, 301,351 accidents were recorded, of which 169,163 occurred on federal highways inspected by the Federal Highway Police (PRF), representing approximately 56% of the total number of accidents. Of these accidents that occurred on federal highways, 4% had fatalities, 37% had injuries, and 59% were accidents where only property damage occurred. Approximately 67% of fatal accidents occurred in rural areas [2]. Due to the growth in traffic accidents, the sectors responsible are increasingly being questioned with regard to road quality and safety.

Although there are a variety of methods available to detect road safety problems, detailed analysis of road accidents remains one of the main indicators of network deficiencies. The decision-making process in road safety management depends on indicators that can objectively express the safety level of the components of a given transportation network [2]. In this sense, researchers have sought to relate historical series of traffic accidents to the geometric and operational attributes of the road using statistical regression models called Accident Prediction Models (MPA) or road safety performance functions. Although MPAs have been explored for more than two decades in countries such as Canada, the United States, England, and Sweden, in Brazil they are still in an incipient stage of development.

According to IPEA/ANTP [1], the particular characteristics of Brazilian traffic (pedestrians crossing outside of the crosswalk, only motorized vehicles stopping at traffic lights, lack of bicycle lanes, radars only reducing vehicle speed 100 meters before and after, among others) make it difficult to use studies and data from other countries as a reference source. It is necessary to search for data that express the Brazilian reality and that allow studies to be based on these data. It is therefore necessary to collect traffic accident data that allow these studies to be carried out, generate diagnoses, and indicate alternative solutions to the problems detected. One of the main difficulties, within this context, is related to the scarcity of computerized databases on accidents [3]. What does exist may also have flaws in collection and consequently little data reliability.

For many municipalities in Brazil, the only maps available showing their territories are those provided by the Brazilian Institute of Geography and Statistics (IBGE). These are often complex, outdated, and little known, causing the local reality and the cartography to be far apart.

When it comes to geometric road design, the problem is even greater. Around 70% of Brazilian highways were built in the 1960s and their projects are either on paper or digitalized as PDF files. According to Augusto Nardes, auditor of the Federal Audit Court (TCU), in an interview with the Globo newspaper [4], lack of planning is one of the TCU's biggest concerns and that this can be evidenced by the absence of necessary documentation (basic design, planning, and others)

in the execution of transportation projects. Another aggravating fact is that traditional methods for updating maps have not kept pace with the increase in the number of roads caused by the country's socioeconomic growth in recent decades, either due to the difficulty of accessing some locations, the difficulty in finding specialized technical personnel or the high cost.

Because the updating of cartographic information is not just a topic of technical interest, but, above all, of economic and social interest, cartographic updating projects are being carried out with the use of geoprocessing and remote sensing techniques, beginning with Geographic Information Systems (GIS) and interpretation of satellite images, making data available in digital format. For Salbego *et al.* [5], the availability of information in digital format allows for the costs of the updating and replacement processes to be reduced, because products generated from GIS can be updated, edited, printed, and duplicated faster and more easily than those generated through traditional methods.

For more than 80 years, aerial photographs have been an indispensable tool for the development of a variety of engineering projects, such as highways, gas pipelines, transmission lines, and many others in Brazil. The spatial resolution of these images has greatly increased in recent decades, enabling their use in the implementation of highway projects. The automatic or semi-automatic extraction of roads can be the most convenient way to overcome the problem of the lack of project documentation for road safety.

Finding an efficient way to extract road networks automatically or semi-automatically is an important topic that has been discussed in many studies [6] [7] [8] [9] [10], in which different methods and algorithms have been used. Most studies agree that extracting roads from aerial images is a complicated task due to occlusion, shadows, and trees, as well as the different types of roads that appear in aerial images, and these conditions make it difficult to accurately extract roads [11] [12] [13] [14]. In addition, although many road extraction methods based on sensory images have been used in various studies, most of them were designed for urban or high-quality roads [15], while few are applicable to rural roads.

Despite the great potential for information extraction, image classification techniques face some challenges, such as the large amount of data to be processed, which can reduce classifier efficiency. Traditional pattern recognition methods for multispectral classification of remotely-sensed images are based on standard statistical techniques such as maximum likelihood [16]. These methods do not produce satisfactory results for hyperspectral imaging because they have a limited ability to resolve confusion between classes, especially classes with some spectral similarity.

The methods based on machine learning algorithms (Machine Learning Algorithms—MLA) have been applied to extract relevant information from hyperspectral data. The most popular approaches for hyperspectral image classification, such as SVM and ANN, have a considerable time limitation for large databases, especially in the training stage. To speed up SVMs to solve this problem,

some variations have been developed, such as LASVM and SVM without Kernel mapping. The former is limited to binary classification and the latter considerably reduces the classification accuracy in cases of class overlap [17].

A new classifier that has been highlighted in the literature is a technique called Optimum-Path Forest (OPF). The OPF proposes to classify patterns using graph theory concepts [18] [19]. This approach emerged as a generalization of the Image Foresting Transform (IFT) [17]. Due to its ease of use and efficiency during data training, it has been shown to be an interesting approach to classification problems [18]. As the technique is relatively recent in the literature and there are few studies on strategies to extract road segment information from hyperspectral images, this paper proposes to introduce the OPF algorithm to extract road geometric features from hyperspectral images and evaluate the effectiveness with a comparison between multi and hyperspectral images. The OPF classifier is not a substitute for deep networks, but can be used as a complement. The idea in this manuscript is not to act directly on feature extraction, as deep learning does, but on the classification stage, and the OPF can be used with the features learned by these networks.

In this method, a geometric base of rural roads will initially be built from road segments extracted from the image. Then, techniques for grouping and reconstructing missing segments from the road network will be used.

2. Background

2.1. Use of Remote Sensing Data to Identify Geometric Features of Roads

When classifying roads with remote sensing data, identification of small objects and linear features is important, while a high spatial resolution is required for more accurate classification [20]. Generally, satellite images with submetric spatial resolution are generated only available in a single panchromatic spectral range. Multispectral images generally have lower spatial resolution than panchromatic bands. This may not be sufficient to accurately distinguish the roads present.

Hyperspectral imaging systems are characterized by the division of the electromagnetic spectrum into a large number of bands (over 40) and are the result of technological advances in imaging. The narrow (typically 10 to 20 nm wide) and contiguous bands allow for the extraction of reflectance spectra at the pixel scale [21]. High spectral resolution allows for a more detailed analysis of land cover spectral signatures and human action than multispectral images [22]. The use of images to extract road features is not recent, however, separating land use classes such as roads and urban areas is not an easy task. The use of hyperspectral images for highway projects is still incipient and the perspective is that geometric characteristics of roads can be identified more simply with image spectroscopy.

When extracting highway information from data generated by remote sens-

ing, it is often more important to have high spatial resolution than high spectral resolution. However, for Taherzadeh *et al.* [23], hyperspectral remote sensing has great potential for application in the analysis of complex urban scenes. Low spatial resolution images make it difficult to identify roads. However, even for high spatial resolution sensors, if the materials present in the scene have a similar spectral response, they may be confused, depending on the number of bands used in the image acquisition process. In fact, the use of a hyperspectral sensor improves the chances of differentiation of these materials and their correct identification in the scene. The narrower the spectral bands, that is, the higher the spectral resolution of the sensor, the more detail about the spectral response of the targets can be extracted, making it possible to considerably reduce misclassification.

The characteristics of the road system's linear features in digital images are mainly linked to the spectral and geometric properties of the target, which directly influence the extraction processes. The pathways in a Remote Sensing image, from the geometric and spectral point of view, can be considered as narrow and continuous ranges of high brightness intensity, bordered by low intensity regions [24]. Xi and Weng [24] further state that the brightness intensity does not vary greatly over short distances along roads, due to the fact that their spectral properties are similar over short stretches. Geometrically, a pathway is usually composed of straight and curved segments, most commonly in the form of circular arcs. Eslami and Mohammadzadeh [25] describe rural or non-urban road sections of a digital image as having characteristics such as constant width, continuous curvature change, and homogeneous local distribution.

In low-resolution images, roads appear mainly as lines that form a more or less dense network, and are directly related to the degree of anthropogenic occupation of the region [26]. In this type of image, the pathways are expressed as lines about 1 to 3 pixels wide and then modeled as lines in the extraction process [27]. In high resolution images, geometric features such as structure and shape play a crucial role in recognizing the road network. Road network extraction over hyperspectral images has great advantages over multispectral images in that it increases the ability to discriminate the materials that make up the road surface from most other types of materials that make up the landscape.

The main advantage of hyperspectral remote sensing in road studies is the ability to measure spectral features or identify unique absorption bands present in materials in this environment. The large number of contiguous bands allows for the extraction of information about the chemical and physical properties of materials [28]. A target that is difficult to identify by traditional multispectral sensors can therefore be discriminated by hyperspectral sensors depending on the spectral features present in the pixels.

However, due to the spatial variability of spectral signatures, the extraction of hyperspectral image characteristics is widely recognized as one of the most challenging tasks in processing hyperspectral images [29] [30]. Many existing methods use manual classification [19] [27] [28] [31] [32] [33], which involves the

experience of specialists. Gabor filters [34], adaptive filters [35], and Markov chains [36] are often adopted. In recent years, pattern recognition methods have caused widespread interest in remote sensing [37] [38] [39]. They are considered to have great potential for classifying hyperspectral images. From training sets, the pattern recognition methods can effectively describe the characteristics of the data. Automated information extraction from images attempts to reproduce how the brain can interpret features. One way to make image interpretation simpler is to separate groups of pixels with similar spectral characteristics, that is, to recognize patterns. Pattern recognition involves techniques for assigning patterns to their respective classes automatically with minimal human interference.

The classification algorithm needs training to classify an image based on categories of interest. This training defines whether the process is supervised or unsupervised. When there is already prior knowledge of the area and it is possible to identify points as samples, the classification is called supervised. Unsupervised is when the definition of the most frequent groups in the image is done automatically by the computer, analyzing the distribution of the digital values. Supervised classifiers can be either parametric or nonparametric. Parametric supervision presupposes a knowledge of the statistics of the data, while nonparametric does not require an a priori statistical distribution.

When unsure of the Gaussian distribution of data, nonparametric algorithms are most often recommended for classification, such as SVM (Support Vector Machine) and Decision Tree classifiers, which are supervised rather than parametric. Although widely used to classify images from remote sensors, they require a lot of computational effort to achieve acceptable accuracy rates in a test suite. They often become unfeasible in situations that require constant data retraining and especially in applications with large data volumes.

Detailed studies of deep learning models for processing remote sensing data have been carried out. Chen *et al.* [29] propose a classification strategy based on deep belief networks (DBN). The multilayer DBN model is designed to learn the characteristics of the hyperspectral data, and the characteristics learned are then classified by logistical regression. Ding *et al.* [40] propose a method for classifying hyperspectral images based on convolutional neural networks (CNN), where convolutional nuclei can be automatically learned from the data through grouping. Wu *et al.* [41] propose a convolutional recurrent neural network (CRNN) for classifying hyperspectral data. Convolutional layers are used to extract locally invariant resources, which are then fed with some recurring layers to further extract contextual information between different spectral bands. Li *et al.* [42] propose a CNN-based pixel pair extraction structure for classifying hyperspectral images. A pixel pair model is designed to exploit the similarity between pixels and ensure a sufficient amount of data for CNN.

Some set learning methods based on the support vector machine [20] [31] [43] have achieved good classification performance with hyperspectral imaging systems. However, one of the biggest difficulties lies in estimating the parameters

of a statistical classifier from a generally limited number of training samples. In the classification process, the accuracy typically increases as additional information, i.e., new spectral bands, is added. The accuracy of the classifier reaches a maximum at a certain point, where the introduction of additional bands causes it to decrease. This is the well-known Hughes phenomenon, resulting from the uncertainty in the estimated values for the classifier parameters, an effect caused by a small number of training samples compared to the dimensionality of the data.

Gao et al. [44] used Random Multi-Graphs (RMG), which are a graph-based set method that uses systematically-constructed trees created from random-ly-selected subsets of resources. In other words, the trees are built in randomly chosen subspaces. Inspired by this randomness, the performance of the hyper-spectral image classification can be improved to mitigate the well-known Hughes phenomenon. However, graph-based set learning methods have rarely been considered for hyperspectral image classification.

Although there are many studies and methods for extracting roads based on remotely-sensed images [45]-[50], most were designed for urban roads or high-quality images, making few of them applicable to rural roads [15]. In their study, Jian *et al.* highlight three problems related to rural roads. First, the variability of materials used for pavement (asphalt, cement, gravel, stone, etc.) that have different spectral signatures can be a problem when the above methods are applied. Second, rural roads are generally narrow, and some road segments can be completely obscured by shadows from clouds, buildings, or other elements of the network. Third, rural roads have more curves than urban roads. It is difficult for existing methods to extract complete roads. Another problem with rural roads, specifically in Northeast Brazil, is the predominantly caatinga vegetation. In periods of drought, these plants can be easily confused with exposed soil and unpaved roads.

This study makes use of the graph-based machine learning model, which was rarely considered for classification of hyperspectral images. In particular, this method introduces OPF in hyperspectral classification for use in geometric design of roads.

2.2. OPF Classifiers

An optimal path forest classifier (OPF) was presented by Papa *et al.* [19], with two proposed variants: unsupervised and supervised, which are subdivided into full graph OPF and nearest k-neighbor graphs (k-NN), the most widely-used graphs. This model has produced good results when applied to other problems [17] [18] [51]. For this classifier, a forest is created where the nodes are descriptors. A complete adjacency relationship is considered for these nodes, and a distance function is used to define the relationships between the nodes. Within the forest, trees are classified in such a way that each tree is associated with a class. The same class can be represented by more than one tree.

This form of constructing the classifier allows non-linearly separable classes to be represented, with spatially dispersed samples [51]. Every supervised classifier is built from an initial training phase, in which a set of samples representative of each class to be characterized is provided. Subsequently, the classifier must be able to determine to which class a new sample presented belongs. In the case of the OPF classifier, the set of training samples is modeled using a labeled graph in which each sample is represented by a vertex. These vertices are organized into a forest, that is, a set of trees. Each tree in the forest contains only vertices of the same class, and for the same class there may be several trees. In this way, each class is characterized by a set of trees within the built forest.

To generate the set of representative trees for each class, this method applies a process of maximizing the connectivity map between samples of the same class. In this process, each vertex receives a value characterizing the cost of its connectivity to its group. This value is associated with the lowest cost path of that vertex with some prototype vertex of its tree. The prototype vertices are those within the tree that are closest to the vertices of another class.

To classify a new sample, the vertex having the lowest connectivity cost with the new sample is sought from within the entire forest. The class of that vertex is assigned to the new sample.

The OPF presented results similar to SVM and better than neural networks and Bayesian classifiers. The biggest difference is the execution time, which can be faster depending on the size of the database [52] [53].

Using CBERS-2B CCD satellite imagery covering a sedimentation area of the Pedras river in the city of Itatinga, Sao Paulo, Brazil, Souza, Lotufo and Rittner [31] compared the OPF classifier with the ANN-MLP, BC, and SVM classifiers. Although the results were similar to SVM, OPF was 65 times faster, especially relevant for large data volumes. Freitas et al. [32] estimated rainfall in agricultural areas using GOES satellite weather images, and compared OPF classifiers with SVM, ANN-MLP, and K-NN, with OPF demonstrating a superior runtime. OPF recognized collapsed areas from GeoEye-MS satellite imagery, yielding results similar to cutting-edge techniques [33]. Papa et al. [54], proposed a method combining the OPF classifier with three optimization algorithms (PSO, HS, and GSA) to mitigate the problem of reducing hyperspectral image data through band selection. The combination of OPF with HS and GSA has produced promising results. Macedo et al., [43] used the combination of Hyperion hyperspectral imagery with the Optimum Forest Path (OPF) algorithm for supervised classification of areas affected by desertification and compared the efficacy of the OPF and SVM classifiers when applied to these areas. Validation of the land cover thematic maps was performed based on confusion matrix analysis using the same set of validation points for consistency. Both classifiers produced reasonable results, however, the OPF algorithm outperformed SVM. The higher classification accuracy obtained by OPF was attributed principally to the ability to differentiate better between degraded areas (DA, DOC, and DP) and preserved areas (PGP and PDP).

3. Materials and Methods

3.1. Study Area

The scope of the analysis was the BR 232 Highway between km 141 and 356, latitudes 8°02'30"S and 8°39'27"S and longitudes 36°11'56"W and 37°48'57"W (Figure 1). This 255km stretch of rural highway passes through the municipalities of São Caetano, Pesqueira, Arcoverde, Cruzeiro do Nordeste, and Custódia, in northeastern Brazil. The principal economy of the region, commerce and services, is experiencing and economic crisis worse than the rest of Brazil. The Gross Domestic Product (GDP) of the northeast region, over the past year, grew at a rate only half that of Brazil's average. In 2018, the northeastern GDP grew only 0.6%, while Brazil's rose 1.1%. With a real average monthly worker income of US 385.00\$, Northeast Brazil still relies heavily on the federal public sector for investments in infrastructure and road safety [4].

The choice of areas was influenced by the availability of data for some road sections obtained from the DNIT road base and the Open Street Map (OSM) cartographic base. The diversity of the geometry (straight stretches and curves)

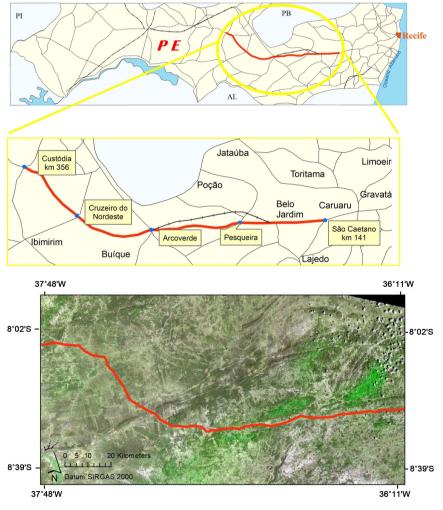


Figure 1. BR 232 Highway, between km 141 and 356.

of the features of interest (road network) made it possible to test the proposed methodology for different situations. The most recent images available that cover the areas of interest having good radiometric quality and/or being free of cloud cover were used.

3.2. Proposed Method

Figure 2 shows the architecture of the proposed approach. In the area to be processed, a Gaussian filter will be used to soften the original image and remove noise to improve the image quality. The edge detection algorithm will be used to detect high gradient regions within the image. The edge image will be generated and candidate road segments will be identified. Based on geometric knowledge, the adjacent candidate road segments will be linked together. For road segments missing because of image quality, occlusion, or other reasons, inference methods will be used and, finally, the complete road vectors will be extracted. The following sections describe the main modules of the architecture.

3.3. Data Pre-Processing Landsat 8

The digital image processing was carried out using satellite images from the TM/Landsat 8 sensor of orbits/points 215/66. The image used in the study was obtained free of charge through the database of the National Institute for Space Research (INPE), with a spatial resolution of 30 meters and six spectral bands in the reflected spectrum, from visible to shortwave infrared, dated from March 2015.

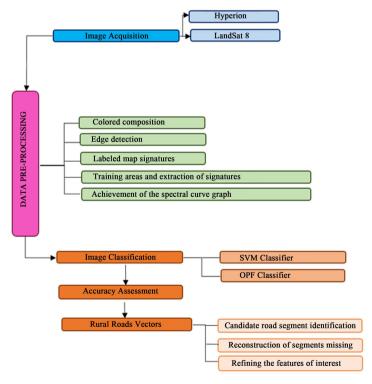


Figure 2. Proposed method.

At first, a study area was delimited using six bands (b1, b2, b3, b4, b5, b7, b8, and b9) from the TM/Landsat 8 sensor image, in which RGB (Red, Green, Blue) color compositions were created using different combinations of bands. These compositions made it possible to choose the best way to establish an interpretative analysis of the coloration of the study area, exposing different colors for different spectral behaviors of the targets within the image. The composition used to obtain the signatures was RB4 G (0.65 * [B3] + 0.35 * [B5]) B (0.85 * [B3] – 0.15 * [B5]), as it proved to be the most suitable for the study. This combination of bands is good for discerning variations in a landscape that does not contain an abundance of vegetation. The combination of geology bands has an important application for finding geological resources. This includes roads, degraded areas, and caatinga-type vegetation.

3.4. Data Pre-Processing Hyperion

The hyperspectral image used in the experiments was a section of the scene 216/65, supplied by the USGS, which covers a semiarid region of the state of Pernambuco between latitudes 8°02'30"S and 8°39'27"S and longitudes 36°11'56"W and 37°48'57"W, acquired by the Hyperion sensor. Hyperion has 242 10 nm-wide bands with 30m spatial resolution and covers a range from 400 to 2500 nm of the electromagnetic spectrum.

The bands sensitive to water absorption (bands 121 - 130, 166 - 180, and 233 - 242) and the uncalibrated bands (233 - 242) were removed by default, using the Erdas Imagine image processing software. After the analysis of the remaining 207 bands, the bands with many negative reflectance values were removed (1 - 12 and 58 - 76). The resulting image composed of 159 spectral bands was converted to ASCII format, resulting in a matrix of 159 columns and 75,452 instances.

3.5. Extraction of Road Geometric Features

The semiautomatic methodology developed for the extraction of the road network from images using the OPF supervised classifier is presented below.

3.5.1. Removal of Noise

To improve image quality, a Gaussian filter was adopted in order to identify a function to normalize the levels for each color. It is therefore necessary to determine the mean values and standard deviation for each color. The result of the processing of a given image with the Gaussian filter applied can be accompanied with the generation of a graph of the Gaussian distributions, in which each curve represents the RGB colors. According to [55], linear smoothing filters use a linear function capable of blurring and reducing noise in the image. This pre-processing is important to perform the extraction of larger objects. Noise can vary in brightness because of failures in the capture phase, failures in the transmission phase, or even because of steps performed in the pre-processing. The use of these linear smoothing filters calculates values for each pixel in the image by averaging the intensity levels of its neighbors, which are defined by a mask of size $m \times n$.

3.5.2. Roadway Detection

The detection of road elements requires, in addition to the satellite image itself (in RGB format), a mask where the streets, unpaved roads, and other places of interest are represented. Elements such as trees can interfere with the identification of objects of interest and, therefore, the first step of the method is to identify and subsequently remove them, so that the region of importance is reduced. For this purpose, the image is converted to the HSV (hue, saturation, and value) color spectrum. The purpose of this step is to improve accuracy when working with a color image.

In the filtering algorithm, where road identification is performed, the brightness value (V) (which varies between 0 and 1, with 0 meaning black and 1 meaning white) is assigned a certain threshold. Then, all pixels that have a value below the chosen threshold are assigned a value of 0. To improve and enhance the image, the saturation (S) was set to a value of 50%.

After the road identification process, other undesirable elements that can corrupt the data analysis are eliminated. Elements that represent buildings can be confused with traffic routes and for that reason they should also be removed. To do this, the image mask (mentioned at the beginning of the section) is required. The algorithm that performs this step consists of assigning the value 0 (black) to the spaces corresponding to the building elements in the analyzed image, based on the mask.

3.5.3. Edge Detection

After the initial pre-processing steps, the image is submitted to an edge detection algorithm. This step consists of looking for sudden variations (discontinuities) of color intensity, as this helps the objects and their characteristics (area, geometric shape) to be easily identified. In this study, it was decided to use the Sobel edge detection algorithm, as it has advantages over its competitors in the scope proposed by this study, giving more weight to the points close to the center pixel, which allows it to obtain more prominent edges.

3.6. Reconstruction of Segments Missing from the Road Network

Through the classification process based on defined rules, it was possible to delineate the most relevant sections of the road network present in the scenes under study [56]. However, due to occlusions from elements such as trees and buildings, there were still separate and non-continuous road sections after the initial stages. Some features of interest were not completely extracted and other features that were not of interest (noise) were detected. This is due to the limitations imposed by the images under study. Routines based on morphological operators were therefore introduced, which allowed non-extracted linear features of interest to be obtained, as well as preprocessing, filtering, and further refinement of solutions. All routines were developed in ArcGis software.

To recover the lost features of interest in the classification, a morphological

reconstruction operation was used, which consists of applying successive dilations to a marker image until it fits into a second mask image. As a marker image, the binary image produced from the results of object-based classification was used, where elements belonging to the roads are represented by the value one (1) and background elements by the value zero (0). This was initially used to enhance the roads, using a full mask structuring element of size 3×3 . This operation can be considered as a high-pass filter, which in addition to extracting high frequency information, also performs image smoothing [57]. After that, a thresholding operation was performed on the filtered image to separate the highlighted features of interest. The most appropriate threshold was determined by using the histogram of the filtered image.

3.7. Refining the Features of Interest and Obtaining the Road Network

The extracted features, according to the procedures described above, are linear segments in the form of polygons that represent the stretches of the road network. To obtain the paths in the form of simple lines, representative of the middle axes of the segments, it was at first necessary to apply operations to connect small disconnected sections and eliminate gaps as well as objects still present in the scene that are not part of the roads. Initially, the small discontinuous sections were connected using a dilation operation. Subsequently, the gaps were eliminated using the morphological closure operation. The area opening operation was the next step, applied to remove objects smaller than a certain number of pixels that are not part of the roadways. In all of these operations, a 3×3 full mask-shaped structuring element was used. To obtain the road network in the form of lines, the morphological thinning operation was applied until convergence, that is, each road segment is represented by a single line, with a width of one pixel.

The semi-automatic vectorization procedure was used for all of the highways present in the images. Layout was performed in the ArcGIS environment, setting the screen scale to 1:6000 in order to standardize target interpretation and obtain homogeneous and detailed reference line vectors. In this process, it was sought to delineate the center line of the roads.

3.8. Comparison between Classification Algorithms

To quantify the errors generated by the different classification techniques used in this study, error or confusion matrices were constructed, which allow for the data from the validation samples and the classification results to be compared. Among the most widely-used methods for assessing the reliability of a classification is the Kappa index, which is an accuracy measurement technique that can be used to determine whether one error matrix is significantly different from another. This index is based on the difference between the overall accuracy (indicated by the matrix diagonal) and the producer and consumer risk accuracy, which is indicated by the sum of the rows and columns of the confusion matrix.

Reliability is the main limitation in classifications from satellite images, generally with less than 90% accuracy [58]. Among the factors that interfere with accuracy are the mixing pixels, the overlap between reflectance data from different targets in space, the low representativeness of the training samples, and the classifier's own ability to deal with inconsistencies in the process [59].

4. Results

4.1. Experiments

The road classification process was performed using Erdas Imagine. As an initial step of this process, the objects were obtained through image segmentation. For image segmentation, a scale level of 40 was used, which implies discarding 40% of the lowest values of the gradient image, allowing only the most well-defined edges to be preserved. For grouping similar adjacent objects, three (3) spectral bands from the images were considered. The mask used was of size 3×3 . The values used in the scale and merge parameters mentioned above were defined after performing various tests on both images under study, choosing those that best made it possible to individualize the objects of interest in the two scenes.

The procedure for image classification used the following steps: color composition of the image (**Figure 3**), training areas and extraction of signatures (**Figure 4**), creation of the spectral curve graph, labeled map, segmentation, and final classification map.

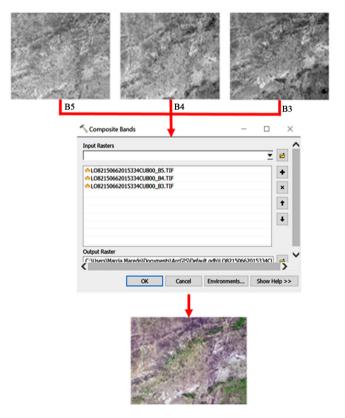


Figure 3. Composite bands RB4G (0.65 * [B3] + 0.35 * [B5]) B (0.85 * [B3] - 0.15 * [B5]).

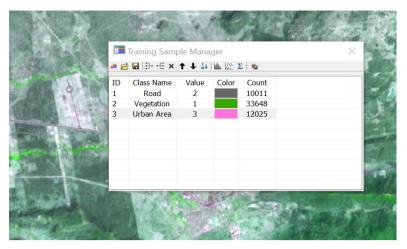


Figure 4. Training areas, extraction of signatures, and segmentation.

In the images Hyperion and LandSat 8 (**Figure 5** and **Figure 6**), 10 classes of land covering for better area discrimination were identified: 1) Water—W, 2) Exposed Soil—ES, 3) Dry Dense Caatinga—DDC, 4) Green Dense Caatinga—GDC, 5) Urban Area—UA, 6) Paved Road Asphalt—PRA, 7) Paved Road Concrete—PRC, 8) Unpaved Road—UR, 9) Clouds—C and 10) Shadows—S.

Samples from the 10 classes of soil covering, in proportion to the areas identified in the image, were used to generate three training sets. The first set was selected based on the location of points collected in the field with a GPS, identifying 8 classes. No field information was collected for the Clouds and Shadows classes. The second set was based on visual interpretation of the scene seeking similarity between the pixels. It was chosen based on multispectral images from the LANDSAT 8 OLI sensor, obtained at orbits/points 215/66, previously classified. The third set of samples was obtained through a spectral library without performing image classification.

The training sets were divided into two new disjoint sets for training and testing with 50% of the total samples in each one, 50% used to train the classifiers and 50% used to evaluate their precision (**Table 1**).

Three main experiments were performed. The first was to evaluate the effectiveness of OPF for classifying hyperspectral images with original hyperdimensional characteristics. In the second, a quantitative and qualitative comparison of OPF precision for data classification was made using multispectral images (Landsat 8) and hyperspectral images (Hyperion). In the third, a comparison is made between the OPF and SVM classification methods for hyperspectral images from the Hyperion sensor. For the OPF classifier, library C or LibOPF [54] was used, which is a design for optimal path forest classifiers. The same training set was used for the SVM classification.

To compare the results obtained between the classifications of both sensors and classifiers, the confusion matrix was generated in Erdas Imagine software and the kappa coefficient was calculated according to literature [58].

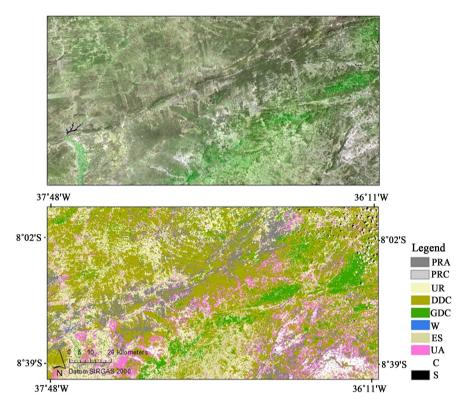


Figure 5. Hyperspectral Image classified using OPF algorithm (10 classes).

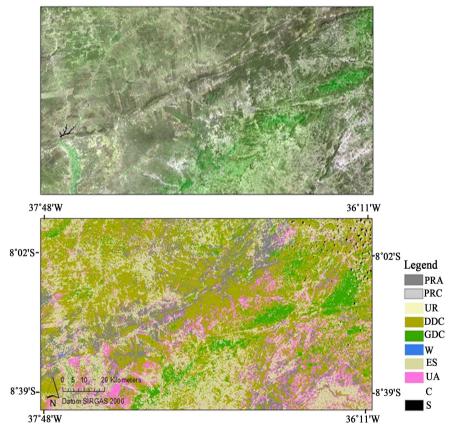


Figure 6. Hyperspectral Image classified using SVM algorithm (10 classes).

Table 1. The training and testing samples for hyperion imagery data in northeast Brazil.

Class Name	Samples			
Class Name	Dataset 1	Dataset 2	Dataset 3	
Water—W	42	49	74	
Exposed Soil—ES	54	196	356	
Green Dense Caatinga—GDC	574	984	2370	
Dry Dense Caatinga—DDC	138	786	1440	
Urban Area—UA	574	984	2370	
Paved Road Asphalt—PRA	76	1318	2278	
Paved Road Concrete—PRC	206	192	374	
Unpaved Road—UR	147	220	214	
Clouds—C	30	7	10	
Shadows—S	1020	2054	2752	

4.2. Supervised Classification

At first, the classes described above were grouped and five classes were considered for image classification and comparison between hyperspectral and multispectral images using the OPF classifier: 1) Water—W, 2) Exposed Soil—ES (Exposed Soil + Unpaved Highway), 3) Vegetation—V (Dry Dense Caatinga + Green Dense Caatinga), 4) Urban Area—UA and, 5) Road—R (Paved Road Asphalt). The classes of Paved Road Concrete, Clouds, and Shadows were not considered. This decision was based on the fact that the multispectral images do not present good separability between the constituent materials of the road (asphalt, concrete, gravel, etc.) [15]. Despite being a simpler experiment, it meets the needs of the Brazilian Northeast where most rural highways are asphalt.

In both classifications, misclassification between urban area and highways occurred. Although Hyperion's classification accuracy was statistically superior to Landsat and both commission errors and omission errors were lower for all classes studied, it can be said that the quality of both classifications was similar (Figure 7 and Figure 8). Although the OLI sensor collects approximately 75 times less data than the Hyperion sensor [60], the positioning and width of the spectral bands are sufficient to classify highways.

Subsequently, the image classification for the Hyperion sensor was performed, considering the 10 classes described above using the OPF classifier (**Figure 5**). All experiments were repeated using the SVM classifier (**Figure 6**) in order to compare the results.

Although low resolution satellite images are available for easy access, free of charge and on a large number of platforms, their use presents a real processing difficulty because they have a large amount of undesirable noise and imperfections that hinder feature extraction analysis. Pattern recognition is essential for remote sensing. However, in situations where the database is very large, the cost of training a classification algorithm may be unsatisfactory and time consuming.

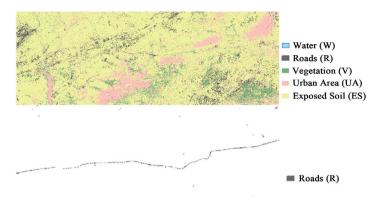


Figure 7. Hyperspectral Image classified using OPF algorithm (5 classes).

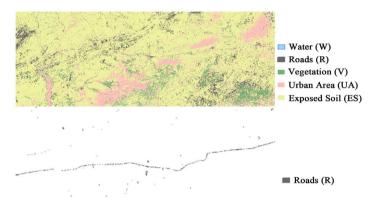


Figure 8. Multispectral Image classified using OPF algorithm (5 classes).

In this paper, multispectral images (Landsat 8) and hyperspectral images (Hyperion) were classified using the OPF classifier. Experimental results showed that the OPF obtained similar recognition rates. The study was a pioneer in the use of the OPF classifier to extract geometric road network features from hyperspectral images. These images store large volumes of data that ultimately reduce the efficiency of traditional classifiers such as SVM. OPF also had a longer execution time, but was 48 times faster than the classifiers tested in the training and testing stage. In the classification of the full image, OPF was about 180 times faster than other classifiers cited in the literature. Due to the greater amount of spectral information provided by the Hyperion sensor, the classification accuracy for this sensor's image was higher than that of the Landsat 8 sensor. The results were good with regard to both land use and land cover separation where spectral differences predominate when compared to classes in which the spectral signatures are very similar. The OPF for multispectral images also presented greater distinctions between regions of the studied areas (Urban Area and Roads). There was some small confusion between these areas, but this did not impact the final result of the classification.

The quality of land use and land cover classification obtained from the OLI sensor image was similar to that obtained from the Hyperion sensor. Having a higher level of spectral detail in hyperspectral images provides better capability to see the unseen and distinguish between urban areas and roads because of the high spectral resolution.

4.3. Separability of the Classes

Using the data from the six reflective bands of the OLI sensor, the accuracy in discriminating the five classes of interest in the studied scene was 96.5% with a Kappa value equal to 0.88.

Regarding the discrimination of classes of interest using the first eight bands of Hyperion, the classification accuracy was 97.9% and the Kappa value was 0.93. The separability of the classes obtained by OPF-Hyperspectral showed a level of accuracy better than OPF-multispectral. Specifically, for the classes related to Roads (R), the OPF-Multispectral obtained 92.78% accuracy, while the OPF-Hyperspectral reached 99.75% accuracy for these classes (**Figure 9**).

In the classification of multispectral data (Table 2), the smallest commission errors (pixels from other classes that were assigned to the reference class) and omission errors (pixels belonging to a reference class that were assigned to other classes) were observed for water (W), 1.34% and 2.88%, respectively, giving it the best performance of the classifier. Although 91.27% of the pixels belonging to the urban area class and 91.80% of the exposed soil class were correctly classified, the largest commission errors were observed for these classes. This indicates that 32.28% of the pixels that were classified as urban area and 31.42% of those that were classified as exposed soil belong, in fact, to other classes.

In the classification of hyperspectral data (**Table 3**), the best classification was also found for water (W), but with lower commission and omission errors (1.02% and 1.16%, respectively).

In the classification of hyperspectral data considering the 10 classes of interest, the classification accuracy was 97.9% and the Kappa value was 0.93. **Table 4** and **Table 5** show the confusion matrix obtained for the classification of variables using the OPF and SVM classifiers, respectively.

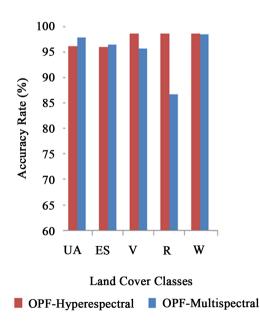


Figure 9. OPF-Hyperspectral × OPF-Multispectral separability.

Table 2. Confusion matrix for land use classes obtained from Landsat 8 images.

		Commission	Omission				
Class	R	W	V	UA	ES	(%)	(%)
R	86.35	1.37	4.68	1.39	6.32	10.33	13.65
W	1.04	97.16	1.74	0.35	0.21	1.34	2.88
V	12.73	0.41	86.21	0.00	0.00	17.27	9.79
UA	1.18	0.00	0.00	91.27	2.47	32.28	8.38
ES	6.50	1.54	0.27	4.51	91.80	31.42	6.19
Pixels	3283	1930	2326	432	797		

Table 3. Confusion matrix for land use classes obtained from Hyperion images.

		Commission	Omission				
Class	R	W	V	UA	ES	(%)	(%)
R	89.00	0.2	9.14	1.74	4.12	9.63	11.78
W	0.21	98.82	0.46	0.00	0.41	1.02	1.16
V	10.51	0.10	94.40	0.00	0.00	13.70	9.45
UA	1.18	0.00	0.00	93.01	0.41	23.77	6.99
ES	1.04	0.57	0.00	0.69	95.06	1.55	5.13
Pixels	3283	1930	2326	432	797		

 Table 4. Confusion matrix for land use 10 classes obtained from Hyperion images (OPF).

Class	W	ES	DDC	GDC	UA	PRA	PRC	UR	С	S
W	98.0	0	0	0	0	0	0	0	0	0.12
ES	0	65.22	2.38	0	0	0	0	19.88	2.32	0
DDC	0	7.11	80.0	1.66	0	0	0	6.31	0	0
GDC	0	0	0	90.0	0	0	0	0	0	0
UA	0	0.98	0	0	88.90	0.28	4.74	1.65	1.63	0
PRA	0	0	6.13	0	0	94.70	3.67	0	0	0.11
PRC	0	0	0	0	0.67	0.34	70.00	0	0	0
UR	0	14.34	0	0	1.12	0.11	1.22	58.35	0.36	0
С	0	0	5.97	0	0	0	0	0	90.43	0
S	0	0	0	0	0	0	0	0	0	94.0
Pixels	0.85	0.75	0.62	0.72	0.64	0.81	1.00	0.87	0.48	0.32
C^a	0	2.45	4.00	0.15	22.44	5.27	3.37	26.32	0.12	3.11
Op	0.12	15.29	5.88	4.72	20.13	1.13	15.21	0	0.41	1.12

a. Commission b. Omission.

Table 5. Confusion matrix for land use 10 classes obtained from Hyperion images (SVM).

Class	W	ES	DDC	GDC	UA	PRA	PRC	UR	С	S
W	100.00	0	0	0	0	0	0	0	0	1.27
ES	0	75.33	4.32	0	0	0	0	22.44	1.32	0
DDC	0	5.12	81.1	1.74	0	0	0	5.20	0	0
GDC	0	0	0	92.0	0	0	0	0	0	0
UA	0	0	1.13	0	86.01	0.12	7.12	0	1.54	0
PRA	0	0	0	0	0	97.01	0.73	0	0	0
PRC	0	0	0	0	0	0.37	73.04	0	0	0
UR	0	12.01	6.44	0	0	0	0	63.22	0	0
С	0	0	0	0	0	0	0	0	91.3	0
S	0	0	0	0	0	0	0	0	0	92.0
Pixels	0.85	0.75	0.62	0.72	0.64	0.81	1.00	0.87	0.48	0.32
C^{a}	0	0	5.00	0	20.44	4.76	2.34	23.48	0	2.17
Op	0	15	5.00	5.01	21.20	0	13.12	0	0	0

a. Commission b. Omission.

Classes W, PRA, C, and S were correctly classified, meaning that the classifiers were able to discriminate these classes. Among the road classes, separability was adequate for PRA (paved road asphalt), however for UR and ES, the spectral confusion was evident. The spectral similarity of the two classes most likely favored confusion in the classification. The same occurred for the UA and PRC classes.

It can be seen in **Table 4** and **Table 5** above that, of the 10 classes, three had a performance below 80% (ES 75%, UR 63%, and PRC 73%), three were between 80% and 90% (DDC 81%, GDC 92% and UA 86%), and four classes obtained an overall accuracy index greater than 90% (PRA 97%, C 91%, S 92% and W 100%).

4.4. Image Spectroscopy

These results show that hyperspectral images perform better than multispectral images on roads with very subtle spectral differences, such as lane discrimination and exposed soil. The highways class has a spectral signature similar to the urban spectral signature, but despite the similarity in the intensity of the absorption characteristics, they have different spectral curves with regard to the magnitude of reflectance (Figure 10). Imaging spectroscopy shows the ability of hyperspectral sensors to intercept electromagnetic energy in very narrow wavelength ranges and detect small absorption characteristics.

4.5. Results for Extraction of Road Segments

Figure 11 and **Figure 12** show the result of the edge detection for the images from the Hyperion sensor and from the LandSat 8 sensor for the study area. The left side shows the original image, while the right side shows the result of the edge detection.

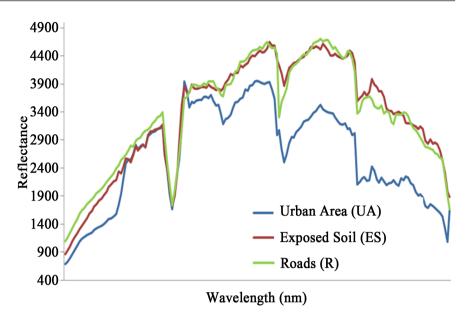


Figure 10. Spectral signature for the Urban Area (UA), Exposed Soil (ES), and Roads (R) classes.



Figure 11. Edge detection for the Hyperion sensor image for the study area.

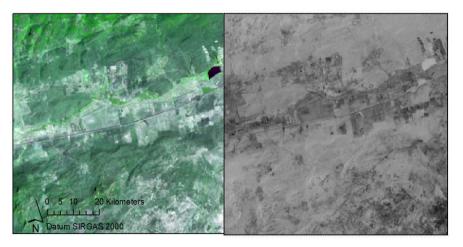


Figure 12. Edge detection for the Landsat 8 sensor image for the study area.

It can be seen that the algorithms accurately detected all of the obvious gradient changes in the image, however, the linear features are not so evident because of the low spatial resolution. Roads are identified more accurately in straight sections. Especially in Figure 12, where there are two types of roads, a concrete road and an asphalt road, the results of edge detection for both are good. However, on curved sections or near intersections, there are still a small number of incorrect features. It can be seen that the result is fragmented and few of the roads extracted are complete. This is probably because the spectral information limited to a single pixel cannot reflect the characteristics of rural roads. These results are best when the training samples are replaced by the spectral signatures of the targets. Another point that is highlighted is that the roads extracted incorrectly are mainly those with vegetation shadows or intersections, which lead to inaccurate inferences or interconnection of road segments. In general, OPF performed better in the extraction of roads with different materials and those having major changes in curvature.

Despite performance well in the identification of roads, the result from the classification alone is not sufficient to construct road bases that support the models of road safety, which need the geometric design for correct identification of straight and curved sections, as well as the radius of those curves. To complement this classification, the extraction process based on the inference of geometric characteristics is necessary [15]. A small test area (red rectangle) was selected to better show the results (**Figure 13**).

Figure 14 shows the result of the image segmentation (left) and the linear features extracted through OPF classification (right). **Figure 15** is the result of linking candidate segments and inference of missing road segments.



Figure 13. Selection of a test area (red rectangle).

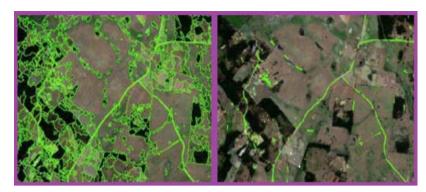


Figure 14. Image segmentation (left) and linear features extracted using OPF classification (right).



Figure 15. Linking candidate segments and inference of missing road segments.

4.6. Comparison of Results

The results were evaluated using real field information, provided by Embrapa Soils—UEP Recife. The experiments were performed 10 times for each set and achieved high accuracy, with Kappa index values ranging from 0.6712 to 0.7857, and Global Accuracy ranging from 90.07% to 91.52%. The results show that OPF in multispectral images was approximately 96.92% accurate, being superior to hyperspectral images. The OPF showed good performance when working with large data volumes. SVM for multispectral images was approximately 96.92% and lower for hyperspectral images. Table 6 shows the accuracy and execution time.

When the comparison performed was related to the performance of the data sets, it was demonstrated that, for all levels of detail, the ratings obtained using data from hyperspectral images showed significantly better results than the ratings obtained using data from multispectral images for both classifiers. However, for hyperspectral images, the OPF classifier was superior. The execution time for the OPF classifier was on average 11 times faster than for SVM.

This result shows that the accurate characterization of targets present in the urban environment requires a high spatial resolution, however, the combination of this important feature with the high spectral resolution can lead to more detailed and accurate results.

Table 6. Mean accuracy and mean total execution time for supervised classifiers

Method	Accuracy rate (%)	Time (s)		
OPF	96.92	± 0.47		
OPF	97.00	± 0.70		
SVM	94.92	± 7.70		
SVM	93.12	± 8.32		

5. Conclusions

Brazil ranks fifth in traffic accident deaths, according to the United Nations. To reduce this number, effective road safety management actions are necessary. This management depends on many factors associated with accidents and the places where they occur. For this, the geometric road base is indispensable. However, most Brazilian municipalities do not have maps. This reality is even worse in the Brazilian Northeast and in rural areas, due to the difficulty of access and the high cost of mapping by traditional methods or using high resolution satellite images. Although high resolution spatial images are commonly used worldwide for road extraction, they are not feasible due to the high cost of acquisition. Multi- and hyperspectral images show satisfactory results, mainly due to the asphalt composition of most of the principal roads. However, hyperspectral images, are better for distinguishing various constituent materials (concrete, asphalt, exposed soil, etc.) due to their high spectral resolution. This lack of information regarding road design can therefore be solved with the use of hyperspectral images. The goal of hyperspectral remote sensing in road studies is the ability to measure spectral features or identify unique absorption bands present in materials in the environment. The large number of contiguous bands allows for the extraction of information about the chemical and physical properties of materials. A target that is difficult to identify by traditional multispectral sensors can therefore be discriminated by hyperspectral sensors depending on the spectral features present in the pixels.

The results show that the OPF classifier obtained a better performance using hyperspectral images with an average hit rate of 97.90%. The use of satellite imagery can be an alternative in collecting information for accident databases.

In general, the classification of images alone does not meet the needs of road safety models, where it is important to reconstruct and accurately identify the elements of the road network (straight segments, curved segments, radius of curves, etc.) for later association with accident data (location, type of accident, climatic conditions, among others). Therefore, geoprocessing techniques are necessary to properly elaborate these bases. The methodology proposed in this study for identifying and extracting roads from low-resolution satellite images using digital processing and pattern recognition produced results with a good level of accuracy. The worst results were found on unpaved roads and exposed soil. Other geoprocessing techniques, such as reducing the excessive number of

vertices, reconstructing curved elements, and smoothing segments can be added to improve the geometric quality of the road base. Additional work could also be carried out to examine the possibility of applying the proposed method using neural networks to classify road networks.

Acknowledgements

The authors would like to thank the University of Pernambuco, its Polytechnic School of Engineering, and its Civil Engineering Master's Program for their financial and infrastructure support for the development, translation and publication of the article. The authors would also like to thank João Paulo Papa for his support in the utilization of OPF and acknowledge the meticulous and dedicated translation work by Simeon Kohlman Rabbani.

Conflicts of Interest

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

References

- [1] Organização Pan-Americana da Saúde (2016).
- [2] American Association of State Highway and Transportation Officials (2010) A Policy on Geometric Design of Highways and Streets. 5th Edition, AASHTO, Washington, D. C.
- [3] Boffo, G.H. (2011) Formatos e técnicas de modelos de previsão de acidentes de trânsito. Dissertação de Mestrado.
- [4] https://economia.uol.com.br/noticias/redacao/2019/04/06/pib-economia-nordeste.htm
- [5] Salbego, A.G., Giotto, E. and Madruga, P.R.A. (2006) Geoprocessamento aplicado ao diagnóstico e espacialização da infra-estrutura viária rural. *Ambiência*, **2**, 179-195.
- [6] Long, H. and Zhao, Z. (2005) Urban Road Extraction from High-Resolution Optical Satellite Images. *International Journal of Remote Sensing*, 26, 4907-4921. https://doi.org/10.1080/01431160500258966
- [7] Hormese, J. and Saravanan, C. (2016) Automated Road Extraction from High Resolution Satellite Images. *Procedia Technologies*, 24, 1460-1467.
 https://doi.org/10.1016/j.protcy.2016.05.180
- [8] Li, Y. and Briggs, R. (2009) Automatic Extraction of Roads from High Resolution Aerial and Satellite Images with Heavy Noise. *International Journal of Computer and Information Engineering*, **3**, 1571-1577.
- [9] Xia, W., Zhang, Y., Liu, J., Luo, L. and Yang, K. (2018) Road Extraction from High Resolution Image with Deep Convolution Network—A Case Study of GF-2 Image. Proceedings of the 2nd International Electronic Conference on Remote Sensing, 2, 325. https://doi.org/10.3390/ecrs-2-05138
- [10] Kahraman, I., Karaş, I.R. and Akay, A.E. (2018) Road Extraction Techniques from Remote Sensing Images: A Review. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018 International Conference on Geomatics and Geospatial Technology, Kuala Lumpur, Malaysia, 3-5 September 2018, 339-342.
 - https://doi.org/10.5194/isprs-archives-XLII-4-W9-339-2018

- [11] Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T. and Eklund, P. (2016) A Review of Road Extraction from Remote Sensing Images. *Journal of Traffic and Transportation Engineering*, **3**, 271-282. https://doi.org/10.1016/j.jtte.2016.05.005
- [12] Shackelford, A. and Davis, C. (2003) Fully Automated Road Network Extraction from High-Resolution Satellite Multispectral Imagery. *International Geoscience and Remote Sensing*, 1, 461-463. https://doi.org/10.1109/IGARSS.2003.1293809
- [13] Hu, X., Tao, C. and Hu, Y. (2004) Automatic Road Extraction from Dense Urban Area by Integrated Processing of High Imagery and LIDAR Data, Processing of High Resolution Imagery and LIDAR Data. *Remote Sensing and Spatial Information Sciences*, **35**, 288-292.
- [14] Saito, S., Yamashita, T. and Aoki, Y. (2016) Multiple Object Extraction from Aerial Imagery with Convolutional Neural Networks. *Journal of Imaging Science Technology*, 60, 10402-10402. https://doi.org/10.2352/J.ImagingSci.Technol.2016.60.1.010402
- [15] Liu, J., Qin, Q., Li, J. and Li, Y.P. (2017) Rural Road Extraction from High-Resolution Remote Sensing Images Based on Geometric Feature Inference. *ISPRS International Journal of Geo-Information*, **6**, 314. https://doi.org/10.3390/ijgi6100314
- [16] Petropoulos, G.P., Arvanitis, K.G., Sigrimis, N.A., Piromalis, D.D. and Boglou, A.K. (2012) Land Use Cartography from Hyperion Hypespectral Imagery Analysis: Results from a Mediterranean Site. 24th International Conference on Tools with Artificial Intelligence, Athens, Greece, 7-9 November 2012, 26-31. https://doi.org/10.1109/ICTAI.2012.184
- [17] Papa, J.P., Falcão, A.X. and Suzuki, C.T.N. (2009) Supervised Pattern Classification Based Optimum-Path Forest. *International Journal of Imaging Systems and Tech*nology, 19, 120-131. https://doi.org/10.1002/ima.20188
- [18] Papa, J.P., Nachif Fernandes, S.E. and Falcao, A.X. (2017) Optimum-Path Forest based on k-Connectivity: Theory and Applications. *Pattern Recognition Letters*, **87**, 117-126. https://doi.org/10.1016/j.patrec.2016.07.026
- [19] Amorim, W.A., Falcão, A.X. and Papa, J.P. (2018) Multi-Label Semi-Supervised Classification through Optimum-Path Forest. *Information Sciences*, **465**, 86-104. https://doi.org/10.1016/j.ins.2018.06.067
- [20] Melgani, F. and Bruzzone, L. (2004) Classification of Hyperspectral Remote Sensing Images with Support Vector Machine. *IEEE Transactions on Geoscience & Remote Sensing*, **42**, 1778-1790. https://doi.org/10.1109/TGRS.2004.831865
- [21] Kruse, F.A., Lefkoff, A.B., Boardam, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, J.P. and Goetz, A.F.H. (1993) The Spectral Image Processing System (SIPS)—Interactive Visualization and Analysis of Imaging Spectrometer Data. *Remote Sensing of Environment*, 44, 145-165. https://doi.org/10.1016/0034-4257(93)90013-N
- [22] Li, S., Wu, H., Wan, D. and Zhu, J. (2011) An Effective Feature Selection Method for Hyperspectral Image Classification Based on Genetic Algorithm and Support Vector Machine. *Knowledge-Based Systems*, 24, 140-148. https://doi.org/10.1016/j.knosys.2010.07.003
- [23] Taherzadeh, E. and Shafri, H.Z.M. (2011) Using Hyperspectral Remote Sensing Data in Urban Mapping over Kuala Lumpur. 2011 Joint Urban Remote Sensing Event, Munich, Germany, 11-13 April 2011, 405-408. https://doi.org/10.1109/JURSE.2011.5764805
- [24] Tong, X., Xie, H. and Weng, Q. (2014) Urban Land Cover Classification with Airborne Hyperspectral Data: What Features to Use? *IEEE Journal of Selected Topics* in Applied Earth Observations and Remote Sensing, 7, 3998-4009.

https://doi.org/10.1109/JSTARS.2013.2272212

- [25] Eslami, M. and Mohammadzadeh, A. (2016) Developing a Spectral-Based Strategy for Urban Object Detection from Airborne Hyperspectral TIR and Visible Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sens*ing, 9, 1808-1816. https://doi.org/10.1109/JSTARS.2015.2489838
- [26] Weismiller, R.A., Van Scoyoc, G.E., Pazar, S.E., Latz, K. and Baumgardner, M.F. (1985) Use of Soil Spectral Properties for Monitoring Soil Erosion. Soil Erosion and Conservation. Soil and Water Conservation Society of America.
- [27] Wang, Y.Y. and Li, J. (2008) Feature-Selection Ability of the Decision-Tree Algorithm and the Impact of Feature-Selection/Extraction on Decision-Tree Results Based on Hyperspectral Data. *International Journal of Remote Sensing*, 29, 2993-3010. https://doi.org/10.1080/01431160701442070
- [28] Herold, M., Roberts, D.A., Gardner, E.M. and Dennison, P.E. (2004) Spectrometry for Urban Area Remote Sensing—Development and Analysis of a Spectral Library from 350 to 2400 nm. *Remote Sensing of Environment*, 91, 304-319. https://doi.org/10.1016/j.rse.2004.02.013
- [29] Chen, Y., Jiang, H., Li, C., Jia, X. and Ghamisi, P. (2016) Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Transactions on Geosciences and Remote Sensing*, 54, 6232-6251. https://doi.org/10.1109/TGRS.2016.2584107
- [30] Yuan, Y., Lin, J. and Wang, Q. (2016) Hyperspectral Image Classification via Multi-Task Joint Sparse Representation and Stepwise MRF Optimization. *IEEE Transactions on Cybernetics*, **46**, 2966-2977. https://doi.org/10.1109/TCYB.2015.2484324
- [31] Souza, R., Lotufo, R. and Rittner. L. (2012) A Comparison between Optimum-Path Forest and k-Nearest Neighbors Classifiers. 25th SIBGRAPI Conference on Graphics, Patterns and Images, Ouro Preto, Brazil, 22-25 August 2012, 260-267. https://doi.org/10.1109/SIBGRAPI.2012.43
- [32] Freitas, G.M., Ávila, A.M.H., Papa, J.P. and Falcão, A.X. (2009) Optimum-Path Forest Based Rainfall Estimation. 16th International Conference on Systems, Signals and Image Processing, Chalkida, Greece, 18-20 June 2009, 1-4. https://doi.org/10.1109/IWSSIP.2009.5367753
- [33] Pisani, R., Riedel, P., Costa, K., Nakamura, R., Pereira, C., Rosa, G. and Papa, J.P. (2012) Automatic Landslide Recognition through Optimum-Path Forest. 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22-27 July 2012, 6228-6231. https://doi.org/10.1109/IGARSS.2012.6352681
- [34] Jia, S., Hu, J., Xie, Y., Shen, L., Jia, X. and Li, Q. (2016) Gabor Cube Selection Based Multitask Joint Sparse Representation for Hyperspectral Image Classification. *IEEE Transactions on Geosciences and Remote Sensing*, 54, 3174-3187. https://doi.org/10.1109/TGRS.2015.2513082
- [35] Benediktsson, J., Palmason, J. and Sveinsson, J. (2005) Classification of Hyperspectral Data from Urban Areas Based on Extended Morphological Profiles. *IEEE Transactions on Geosciences and Remote Sensing*, 43, 480-490. https://doi.org/10.1109/TGRS.2004.842478
- [36] Tarbalka, Y., Fauvel, M., Chanussot, J. and Benediktsson, J.A. (2010) SVM and MRF-Based Method for Accurate Classification of Hyperspectral Images. *IEEE Geosciences and Remote Sensing Letters*, 7, 736-740. https://doi.org/10.1109/LGRS.2010.2047711
- [37] Santos, A., Araujo, A. and Menotti, D. (2013) Combining Multiple Classification Methods for Hyperspectral Data Interpretation. *IEEE Journal of Selected Topics in*

- Applied Earth Obsavations and Remote Sensing, **6**, 1450-1459. https://doi.org/10.1109/JSTARS.2013.2251969
- [38] Qi, C., Zhou, Z., Sun, Y., Song, H., Hu, L. and Wang, Q. (2017) Feature Selection and Multiple Kernel Boosting Framework Based on PSO with Mutation Mechanism for Hyperspectral Classification. *Neurocomputing*, 220, 181-190. https://doi.org/10.1016/j.neucom.2016.05.103
- [39] Zg, E., Zhang, X., Jiao, L., Li, L. and Hou, B. (2016) Spectral-Spatial Hyperspectral Image Ensemble Classification via Joint Sparse Representation. *Pattern Recognition*, **59**, 42-54. https://doi.org/10.1016/j.patcog.2016.01.033
- [40] Ding, C., Li, Y., Xia, Y., Wei, W., Zhang, L. and Zhang, Y. (2017) Convolutional Neural Networks Based Hyperspectral Image Classification Method with Adaptive Kernels. *Remote Sensing*, **9**, 618. https://doi.org/10.3390/rs9060618
- [41] Wu, H. and Prasad, S. (2017) Convolutional Recurrent Neural Networks for Hyperspectral Data Classification. *Remote Sensing*, 9, 298. https://doi.org/10.3390/rs9030298
- [42] Li, W., Wu, G., Zhang, F. and Du, Q. (2017) Hyperspectral Image Classification Using Deep Pixel-Pair Features. *IEEE Transactions on Geosciences and Remote Sensing*, **55**, 844-853. https://doi.org/10.1109/TGRS.2016.2616355
- [43] Macedo, M.R.O.B.C., Times, V.C., Cavalcanti, G.D.C. and Kohlman Rabbani, E.R. (2015) An Architecture to Classify Desertification Areas Using Hyperspectral Images and the Optimum Path Forest Algorithm. *Electronic Journal of Geotechnical Engineering*, 20, 1881-1895.
- [44] Gao, F., Wang, Q., Dong, J. and Xu, Q. (2018) Spectral and Spatial Classification of Hyperspectral Images Based on Random Multi-Graphs. *Remote Sensing*, 10, 1271. https://doi.org/10.3390/rs10081271
- [45] Sujatha, C. and Selvathi, D. (2015) Connected Component-Based Technique for Automatic Extraction of Road Centerline in High Resolution Satellite Images. *Image and Video Processing*, 2015, Article No. 8. https://doi.org/10.1186/s13640-015-0062-9
- [46] Shi, W., Miao, Z., Wang, Q. and Zhang, H. (2014) Spectral-Spatial Classification and Shape Features for Urban Road Centerline Extraction. *IEEE Geoscience and Remote Sensing Letters*, 11, 788-792. https://doi.org/10.1109/LGRS.2013.2279034
- [47] Li, M., Stein, A., Bijker, W. and Zhan, Q. (2016) Region-Based Urban Road Extraction from VHR Satellite Images Using Binary Partition Tree. *International Journal of Applied Earth Observation and Geoinformation*, 44, 217-225. https://doi.org/10.1016/j.jag.2015.09.005
- [48] Kamangir, H., Momeni, M. and Satari, M. (2017) Automatic Centerline Extraction of Covered Roads by Surrounding Objects from High Resolution Satellite Images. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W4, 111-116. https://doi.org/10.5194/isprs-archives-XLII-4-W4-111-2017
- [49] Abdollahi, A., Bakhtiari, H.R.R. and Nejad, M.P. (2017) Investigation of SVM and Level Set Interactive Methods for Road Extraction from Google Earth Images. *Journal of the Indian Society of Remote Sensing*, **46**, 423-430.
- [50] Abdollahi A., Biswajeet, P. and Nagesh, S. (2019) Extraction of Road Features from UAV Images Using a Novel Level Set Segmentation Approach. International Journal of Urban Sciences, 23, 391-405. https://doi.org/10.1080/12265934.2019.1596040
- [51] Alpaydin, E. (2010) Introduction to Machine, Massachusetts-USA. The Massa-

- chusetts Institute of Technology Press, Cambridge, MA.
- [52] Fernandes, S.E.N. and Papa, J.P. (2019) Improving Optimum-Path Forest Learning using Bag-of-Classifiers and Confidence Measures. *Pattern Analysis and Applications*, **22**, 703-716. https://doi.org/10.1007/s10044-017-0677-9
- [53] Fernandes, S.E.N., Pereira, C.R., Ramos, C.C.O., Souza, A.N., Gastaldello, D.S. and Papa, J.P. (2019) A Probabilistic Optimum-Path Forest Classifier for Non-Technical Losses Detection. *IEEE Transactions on Smart Grid*, 10, 3226-3235. https://doi.org/10.1109/TSG.2018.2821765
- [54] Papa, J.P., Falcão, A.X., de Albuquerque, V.H.C. and Tavares, J.M.R. (2012) Efficient Supervised Optimum-Path Forest Classification for Large Datasets. *Pattern Recognition*, 45, 512-520. https://doi.org/10.1016/j.patcog.2011.07.013
- [55] Barros, W.K., Kohlman Rabbani, E.R., Santana, C.V., Oliveira, D.M.G. and Macedo, M.R.O.B.C. (2015) Segurança Viária: Estudo Piloto na BR-101 no Estado de Pernambuco. XI CNEG—Congresso Nacional de Excelência em Gestão, 1.
- [56] da Costa, T.C.C., de Oliveira, M.A.J., Accioly, L.J. de O. and da Silva, F.H.B.B. (2009) Análise da degradação da caatinga no núcleo de desertificação do Seridó (RN/PB). Revista Brasileira de Engenharia Agrícola e Ambiental, 13, 961-974. https://doi.org/10.1590/S1415-43662009000700020
- [57] Rocha, L.M., Cappabianco, F.A.M. and Falcão, A.X. (2009) Data Clustering as an Optimum-Path Forest Problem with Applications in Image Analysis. *International Journal of Imaging Systems and Technology*, 19, 50-68. https://doi.org/10.1002/ima.20191
- [58] Hudson, W.D. and Ramm, C.W. (1987) Correct Formulation of the Kappa Coefficient of Agreement. *Photogrammetric Engineering & Remote Sensing*, **53**, 421-422.
- [59] Figueiredo, S.M. de M., Cavalcante, L.M., Valentim, J.F. and Figueiredo, E.O. (2009) Avaliação da acurácia de classificação digital de imagens no mapeamento de áreas de pastagens degradadas em Rio Branco, Acre. *Anais XIV Simpósio Brasileiro de Sensoriamento Remoto*, 5789-5796.
- [60] Thenkabail, P., Gamage, M. and Smakhtin, V. (2004) The Use of Remote Sensing Data for Drought Assessment and Monitoring in Southwest Asia. Research Report.



Call for Papers

Journal of Geographic **Information System (JGIS)**

ISSN 2151-1950 (Print) ISSN 2151-1969 (Online)

https://www.scirp.org/journal/jgis/

JGIS is an international refereed journal dedicated to the latest advancement of Geographic Information System. The goal of this journal is to keep a record of the state-of-the-art research and promote the research work in these fast moving areas. The journal publishes the highest quality, original papers included but not limited to the fields:

Subject Coverage

- Cartography and Geodesy
- Computational Geometry
- Computer Vision Applications in GIS
- Distributed, Parallel, and GPU Algorithms for GIS
 Spatial Data Structures and Algorithms
- Earth Observation
- Environmental Geomatics—GIS, RS and Other **Spatial Information Technologies**
- Geographic Information Retrieval
- Geographical Analysis for Urban and Regional Development
- GIS and Cloud Computing
- GIS and High Performance Computing
- Human Computer Interaction and Visualization
- Image and Video Understanding
- Location Privacy, Data Sharing and Security
- Location-Based Services
- Performance Evaluation
- Photogrammetry
- Similarity Searching
- Social Networks and Volunteer Geographic
- Spatial Analysis and Integration

- Spatial and Spatio-Temporal Information Acquisition
- Spatial Data Mining and Knowledge Discovery
- Spatial Data Quality and Uncertainty
- Spatial Data Warehousing, OLAP, and Decision Support
- Spatial Information and Society
- Spatial Modeling and Reasoning
- Spatial Query Processing and Optimization
- Spatial Semantic Web
- Spatio-Temporal Data Handling
- Spatio-Temporal Sensor Networks
- Spatio-Temporal Stream Processing
- Spatio-Textual Searching
- Standardization and Interoperability for GIS
- Storage and Indexing
- Systems, Architectures and Middleware for GIS
- Traffic Telematics
- Transportation
- Visual Languages and Querying
- Wireless, Web, and Real-Time Applications

We are also interested in short papers (letters) that clearly address a specific problem, and short survey or position papers that sketch the results or problems on a specific topic. Authors of selected short papers would be invited to write a regular paper on the same topic for future issues of the JGIS.

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-mail

https://www.scirp.org/journal/jgis E-mail: jgis@scirp.org

What is SCIRP?

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

What is Open Access?

All original research papers published by SCIRP are made freely and permanently accessible online immediately upon publication. To be able to provide open access journals, SCIRP defrays operation costs from authors and subscription charges only for its printed version. Open access publishing allows an immediate, worldwide, barrier-free, open access to the full text of research papers, which is in the best interests of the scientific community.

- High visibility for maximum global exposure with open access publishing model
- Rigorous peer review of research papers
- Prompt faster publication with less cost
- Guaranteed targeted, multidisciplinary audience





Website: https://www.scirp.org Subscription: sub@scirp.org Advertisement: service@scirp.org