Applied Geomatics (2021) 13:619–643
https://doi.org/10.1007/s12518-021-00379-y

ORIGINAL PAPER

Unsupervised mapping of a hybrid urban area in South Africa

Guy Blanchard Ikokou[1] · Julian Smit[2]

Received: 15 June 2020 / Accepted: 29 April 2021 / Published online: 7 June 2021

Abstract

Hybrid urban areas are dominated by important spectral mixtures from formal and informal housing units which make them difficult to map even for the most robust classifier. Proposals to introduce other descriptive features, such as size, shape, texture, and context into the classification process, come with another drawback which is how to ensure the selected feature thresholds are optimal. Image segmentation which is the backbone of object-based analysis depends on a range of parameters including scale parameter, shape, smoothness, colour, and compactness weighting factors. Current techniques to select optimal segmentation scales only give the remote sensing analyst control over one parameter out of five (20%). This study proposes a classification strategy that gives the analyst control of 60% of the parameters to ensure an acceptable segmentation outcome. The study also proposes a feature selection approach that eliminates feature overlaps within the feature space which may not be observable within the original data. An automatic optimal parameter selection function is also proposed in this study. Tested on a SPOT5 resolution merge image, the approach overpowered the accuracy metrics of (Kemper et al. in Int Arch Photogramm Remote Sens Spat Inform Sci 40(7): 1389, 2015) with overall, sensitivity, specificity, precision, true skill statistic accuracies of respectively 0.97, 0.96, 1, 0.942, 0.95 against 0.97, 0.804, 0.98, 0.477, and 0.781. Similar trends are observed with the smallest average error of omission for built-up and non-built structures at 0.042 and 0 against to 0.196 and 0.164. The errors of commission for built-up and non-built-up structures were 0.060 and 0.008 respectively compared to 0.523 and 0.585.

Keywords Urban mapping · Object-based image analysis · Informal settlement · Image segmentation parameters · Principal component · Chebyshev's rule

Introduction

Modern cities are generally composed of well-structured housing units and street patterns. The quest for better living conditions in main cities has resulted in the development of informal settlements within or at peripheries of towns, creating a town or city with a mixture of formal and informal housing structures (hybrid urban areas). Planning for service delivery in these informal areas requires detailed land cover information. Image classification has been widely used to extract land information from satellite and aerial images (Herold et al. 2003; Du Plessis 2015; Kohli et al. 2016; Debbage et al. 2017). The improvements in sensor resolution has resulted in high detailed geospatial information. However, one drawback from such improvement is the high spectral variability within individual classes that can lead to intra class confusion due to similarities in spectral signatures. The situation becomes even more exacerbated with complex environments such as urban areas, which have complex and unclear object spectral signatures (Ben-Dor et al. 2001; Herold et al. 2004; Kuffer and Barrosb 2011). For instance, Heiden et al. (2001) reported that tile materials such as polyethylene, bitumen, and concrete mainly dominate roofs of residential buildings while zinc materials dominate non-residential structures such as commercial buildings. The authors reported that low reflectance in long wavelengths was the main characteristic of roof made of zinc material while a strong reflectance in these wavelengths would characterise roof materials made of chains of hydrocarbons such as polyethylene or bitumen. In addition,

✉ Guy Blanchard Ikokou
IkokouGB@tut.ac.za

1 Geomatics Department, Tshwane University of Technology, Pretoria, South Africa

2 Geomatics Division, University of Cape Town, Cape Town, South Africa

rooftops made of concrete have a low reflectance in short wavelengths.

Herold et al. (2002) and Herold et al. (2004) reported that roads have increasing reflectance with long wavelengths and roads made of concrete and gravel can reach reflectance peaks in the infrared band. Concrete roads were described with high reflectance in the visible light spectrum while new asphalt roads show low reflectance in short wavelengths and high reflectance in long wavelengths including the visible and infrared. The authors reported that red tile and wood shingle roofs exhibit high reflectance in the infrared band. It is also argued that the presence of iron oxide in the red tile roofs increase the absorption in the visible light (Weng and Quattrochi 2006). Moreover, high reflection in the green band is a key attribute of water bodies while urban vegetation shows high reflectance in the red and infrared (Herold et al. 2003; Jilge et al. 2016).The development of very high spatial resolution sensors have also made objects' shapes and contextual attributes crucial in image classification. (Reigber et al. 2007; Meng et al. 2009; Chen et al. 2012). It was reported in Steiniger et al. (2008) that similar spatial advantages can also be made available from high-resolution aerial photographs.

Numerous image classification strategies of urban areas have been reported in the literature. Novack and Kux (2010) proposed an object-based classification strategy of an informal settlement in Sao Paulo in Brazil using a high-resolution Quickbird image. The approach used the Segmentation Parameter Tuner algorithm for the selection of optimal scale parameters. The principle underlying the selection process is that a parameter search is performed based on the fitness between a training sample drawn by the user and the segmentation produced by the algorithm (Costa et al. 2008). The authors extracted geometrical, spectral and textural segment features to train the object-based classification. The technique achieved a classification accuracy of about 70% with a kappa agreement of 65%. One drawback of the scale parameter selection process is the restriction in user defined search range as the proposed range may not include certain optimal scale parameters for the segmentation (Meyer and Niekerk 2016). An alternative image classification approach was presented in Odindi et al. (2012) who performed a land cover classification of Port Elizabeth using a Landsat image. The authors used the statistical K-means and ISODATA pixel-based classifiers to extract built up structures, green vegetation, water bodies, dune, and bare ground. Although the strategy has been widely used for land cover mapping (Abbas et al. 2016), the centroids estimated by the K-means are not always representative of their respective classes. The other limitation of the algorithm pointed out in Singh et al. (2011) is the presence of empty clusters during the classification. Furthermore, a comparison between the Iterative Self Organizing Data Analysis and eCognition has highlighted the superiority of the object-based algorithm as the latter relies on meaningful objects rather than individual pixels (Manakos et al. 2000). Similarly, Fundisi and Musakwa (2017) classified high-resolution Pleiades images of an urban area in Gauteng province (South Africa) using ISODATA algorithm in ENVI software. The mapped urban area was dominated by vegetation cover, which explains the high overall classification accuracy of 85.5% and kappa agreement of 77%. More recently Gxumisa and Breytenbach (2017) also pointed out the superiority of object-based classification approaches over their pixel-based counterparts when classifying a SPOT5 multispectral image covering the Soshanguve area in Gauteng province, South Africa. Object based classification strategies have also been reported as more suitable for heterogeneous areas such as urban areas (Kemper et al. 2015; Degerickx et al. 2017; Kuffer et al. 2017; Ouerghemmi et al. 2017; Van der Linden et al. 2019). One major contribution to the good performance of the proposed strategy in Gxumisa and Breytenbach (2017), with regard to building extraction, was the introduction of elevation data into the classification process to minimize the influence of pixel similarity among classes. However, one suggestion made by the authors to improve the accuracy was the use of techniques such as principal component analysis that identifies suitable spectral bands that would best describe each individual class (Gxumisa and Breytenbach 2017).

This study proposes a feature selection method combining a principal component analysis which identifies the possible optimal objects' descriptive features which are then projected onto a Chebyshev's matrix to eliminate possible features' overlaps which may not be observable in the original data. The final candidate features are qualified as unique after "optimization" in order to ensure an acceptable object identification during the classification process. The selection approach is tested on inter-class distance features to produce the more potent measures that separate each object from its neighbours. The proposed segmentation parameters' selection approach includes a local search through which possible optimal combinations of scale parameter/compactness/colour are identified to produce good quality segments at a single level segmentation. The obtained parameters are further processed through a global search tool to identify how many segmentation levels are needed to represent most of the land use/land cover classes in the imagery and reveal the associated parameter combinations. An additional scale/compactness search function is proposed to automatically identify best parameter matches to achieve acceptable segmentation outcomes.

Study area and methods

Study area and data

To test our image classification strategy, we chose the city of Stellenbosch due to its small size (Fig. 1), its challenging landscape with mountains and hills as well as its diversity in urban vegetation cover, building footprints and water bodies including dams, swimming pools, and reservoirs. Stellenbosch is a small town in the Western Cape Province located at 33.9321° S and 18.8602° E. Stellenbosch municipality covers an area of 831 km^2. To the west and southwest, it extends as far as the urban edge of the Cape Town metropolitan area while to the east and southeast it is bounded by mountain ranges. The western part of Stellenbosch municipality and the eastern part of Franschhoek valley are separated by mountains. With a population of 77,476 inhabitants, about 50% of residents live in suburbs including but not limited to Idasvallei, Coatesville, Die Boord, Brandwatch, Jamestown, Paradyskloof and the shantytown of Kayamandi at the North West periphery of the city. The university is located near the city centre while some schools are spread within the city and the shantytown of Kayamandi. Land use and land cover in Stellenbosch is similar to most South African cities including but not limited to roads, residential, urban vegetation, commercial, industrial, educational buildings, and water bodies.

The data used in this study includes a multispectral SPOT5 10 m spatial resolution and a 2.5-m spatial resolution panchromatic images. The historical imagery

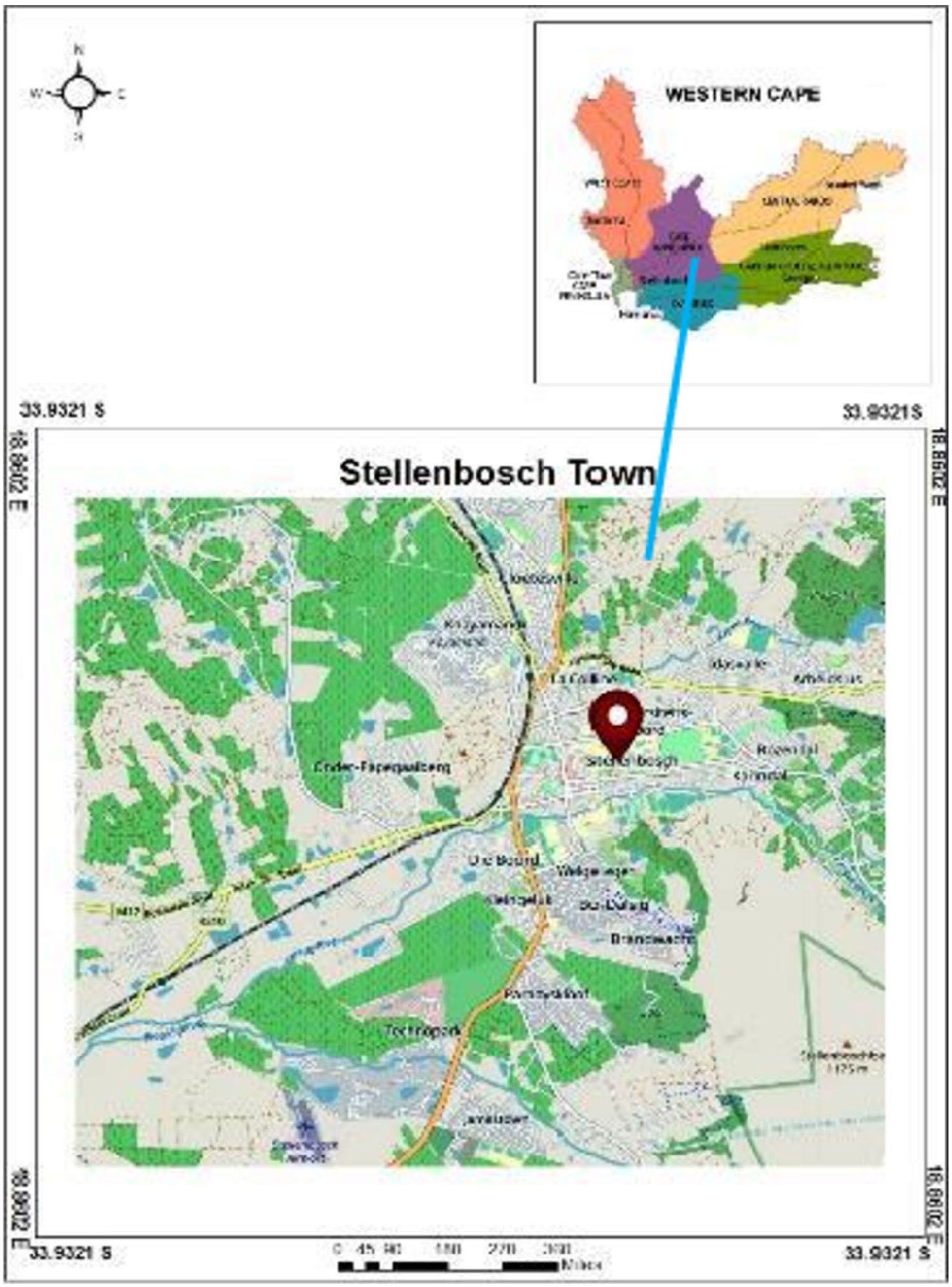

Fig. 1 A presentation of the study area of Stellenbosch

was captured on 20th November 2008. In addition to the satellite imagery, a 0.5-m high resolution aerial photograph covering the study area was acquired from the National Geospatial Information office in Cape Town. All satellite data was supplied with metadata files by the South African National Space Agency (SANSA). The date of image capture did not matter for the study since the focus was more on the land use/land cover classes which obviously may have expanded but this has no technical implications on the presented methodology.

After the pre-processing of the satellite imagery, we segmented the enhanced image using 8 randomly selected scale parameters to collect segments' brightness attributes. The series of segmentations were done keeping the shape/colour as well as the compactness/smoothness parameters at 0.5 for equal influence on the results and we gave each band a weight of 1 so that the multiresolution algorithm takes into account all the spectral information made available by each band in eCognition. The collected brightness measures were representative to all classes involved in this study according to the land use/land cover classes. The total number of samples collected within the study area represent a count of 1460 objects' spectral brightness measures selected across the image to avoid a bias representation of certain classes. It must be noted that the shape compactness we are referring to here is distinct from the segmentation parameter associated to the smoothness and will be computed as follows:

$$Shape\ \ compactness = \left[4\pi \times objectarea\right](object\ \ perimer)^{-2} \tag{1}$$

Because of spectral similarities across land use/land cover within urban areas, attributes such as inter-object separation distances, the perimeter measures, the shape compactness indices, object's lengths, width, length over width ratios will be considered in the classification process. However among the 7 mentioned attributes, the inter-object separation distance seems to be the most complex, due to the fact that in residential urban areas a large amount of buildings may share the same proximity distance. Moreover, classes such as urban trees which may be very close to residential buildings may exhibit similarity of proximity distance to residential buildings. This would give serious difficulties even to the most robust classifier to separate these classes under such circumstances. One solution we will explore and incorporate into the classification strategy is to propose an approach that could refine these distance topology relationships and offer optimised measures that could minimize misclassifications.

To minimize computer storage and the analysis time and reduce the image analyst effort, it is proposed to reduce the size of the images. The satellite images and the aerial photograph used in this study were cropped using Gimp 2.1.0.8 software compression free to not alter the pixel values of images statistics (Campbell 2006).

Study methodology

Geometric corrections

The imagery provided for the study was SPOT5 imagery of level 1A which consequently requires geometric corrections before any further analysis (Sowmya et al. 2017; MohanRajan et al.2020). Geometric correction of satellite imagery consists of modelling the relationship between coordinates on images and ground coordinates. The first-order polynomial model was disregarded for our area since it is more suitable for flat landscape and our study area is dominated by hills and mountains. Instead, the second degree polynomial model was selected for the geometric corrections (Mather and Tso 2016). All the satellite images were geometrically corrected using Lo 19 projection which is a local projection system, using the corrected 2008 colour aerial photograph as a reference. The panchromatic satellite image was first registered to the aerial photograph to preserve the high spatial resolution then the multispectral image was co-registered to the panchromatic satellite image using Erdas Imagine software. The nearest neighbour resampling method was used because according to Mather and Tso (2016), it does not alter the original pixel values and produces less distortions compared to cubic convolution and bilinear interpolation (Baboo and Devi 2010). We used a total of 10 ground control points which were identified in both satellite images and the aerial photograph. The total root mean square error of the correction of an image is estimated through as follows (Rocchini and Rita 2005):

$$Totalrms = \sqrt{\frac{1}{n\sum\left(u^2 + v^2\right)}} \tag{2}$$

Figure 2 presents the results of the image registration process. It can be observed that there is a continuation of linear features between the multispectral SPOT5 image on the right and the colour aerial photograph, revealing the success of the geometric correction strategy.

Reflectance normalisation

To collect remotely sensed data of lasting values, the data must be calibrated to physical units such as reflectance because the radiance recorded by the sensor for each pixel is an apparent radiance. This apparent radiance is the combination of the radiance of the object on the earth surface and atmospheric effects (Rani et al. 2017). The estimation

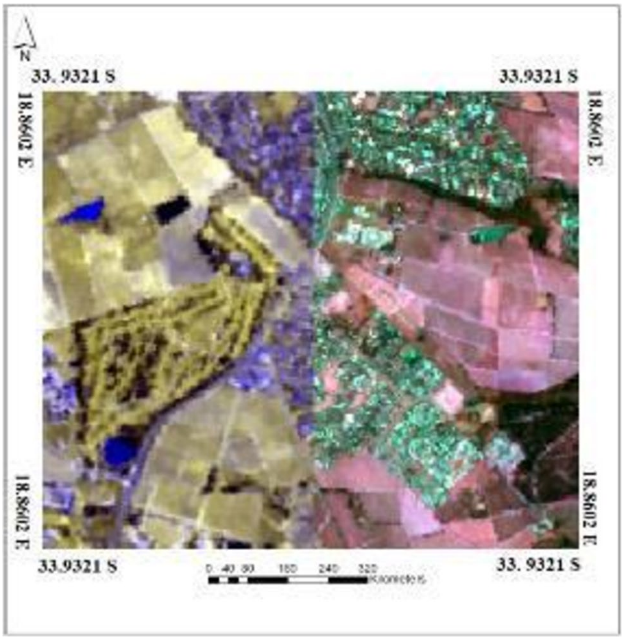

Fig. 2 Registration output showing a transparent overlay of the registered multispectral image over the colour aerial photograph

of ground reflectance requires the conversion of pixels' DN values to apparent radiance, then conversion of the apparent radiance to apparent reflectance and finally the apparent reflectance is converted to ground reflectance. For SPOT imagery, the conversion from pixel DN values to apparent radiance is done through the following (Mather and Tso 2016):

$$L_\lambda = \left[\frac{Gain}{Pixel\ \ DN\ \ value}\right] + Bias \tag{3}$$

With L_λ, the apparent radiance. The conversion of the apparent radiance to the apparent reflectance is done through the following equation (Mather and Tso 2016):

$$\rho_\lambda = \frac{\pi L_\lambda d^2}{ESUN_\lambda \cos\theta_s} \tag{4}$$

With ρ_λ the apparent reflectance, $ESUN_\lambda$ the exo-atmospheric solar irradiance in $Watts/m^2\mu m$, d the Earth-Sun distance in astronomical units and s the solar zenith angle. The Earth-Sun distance is estimated as follows (Mather 2004):

$$d = 1 - 0.01674\cos(JD - 4) \tag{5}$$

The target reflectance is estimated by multiplying the apparent reflectance by 400, rounded and encrypted back to 8bits radiometric resolution through the following piece of code implemented in Erdas function environment as follows:

$$\begin{array}{l} IF\ \ (ROUND(reflec\tan ce \times 400)))\rangle 255\ \ THEN \\ DN = 255 \\ ELSEIF(ROUND(refelc\tan ce \times 400))\langle 0 \\ DN = 0 \\ ELSE \\ DN = ROUND(reflec\tan ce \times 400) \\ ENDIF \end{array} \tag{6}$$

All the parameters used in the conversion of apparent radiance to the ground reflectance were provided in the metadata files of the satellite imagery. The result of the conversion from the pixel DN to ground reflectance is a range of pixel values from 0 to 255 grey levels in absence of atmospheric errors. A visual analysis of our imagery revealed the presence of water bodies within the study area, thus the lowest reflectance value expected should be zero or close to zero in the infrared band. However, the statistics of our multispectral image show that the image still has atmospheric distortions as illustrated by Fig. 3 as follows.

From the observation of Fig. 3, there is a need for further processing of the imagery to reduce the minimum pixel value attributed to water bodies to zero or a value very close to zero.

Atmospheric corrections

Several atmospheric correction models exist in the literature (López-Serrano et al. 2016; Sowmya et al. 2017**;** Boakye, et al. 2020.). Atmospheric correction methods can be related to the spectral resolution of the available multispectral satellite imagery and the availability of image capture data (Dutta and Das 2019; Lhissou et al. 2020; Miky 2019). Figure 4 details a workflow guiding the selection of the appropriate atmospheric correction method.

Observing both paths from the multispectral imagery the SPOT5 multispectral image satisfies the left path requirements. From the four correction models proposed it was noted that the empirical atmospheric correction model requires ground calibration data in the scene and the ancillary data provided with the imagery did not contain such information. The dark object subtraction model assumes that no atmospheric transmittance is lost and that there occurs no diffuse downward radiation at the surface (Song et al. 2001), but the hilly and mountainous landscape of Stellenbosch town does not satisfy these requirements. Moreover, a visual analysis of the land use/land cover classes on the satellite imagery and aerial photograph did not show the presence of dense vegetation cover, excluding the possibility of using the Dense Dark vegetation method. The radiative Transfer Model seems to be suitable for our study area, ATCOR2 and ATCOR3 available in the image pre-procession software PCI Geomatica, are such Radial Transfer Models. Since ATCOR2 is more suitable for flat landscape (Richter

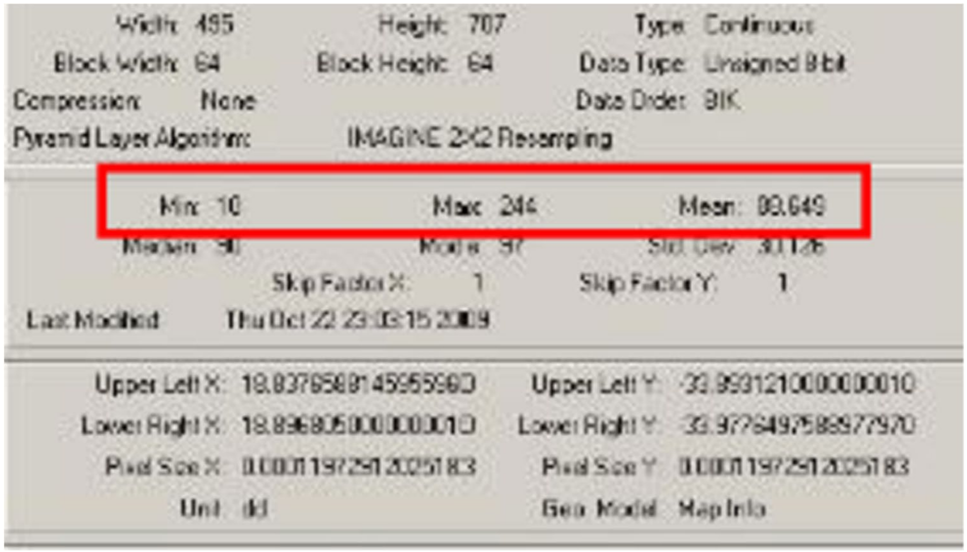

Fig. 3 The minimum and maximum reflectance statistics of the image respectively at 10 and 246

and Center 2004) ATCOR3 was selected for our study area and the tool has been used for atmospheric corrections of mountainous areas (Tan et al. 2012; Ateşoğlu and Tunay 2014). Since ATCOR3 requires the use of a digital elevation model, we used contour lines and GPS point's coordinates provided by the National Geospatial Information office to produce a digital elevation model using ArcGIS software. Figure 5 shows the outcome of the atmospheric correction process. After selection of a few water body segments, it can be observed that the lowest pixel value in the infrared band was reduced to 0.083018 while the brightest water body segment had a reflectance of 0.50404, which is expected in absence of atmospheric distortions on the objects' ground reflectance values.

Data fusion

Poor spatial and spectral resolutions have been a great challenge for urban mapping. Some authors have suggested a spatial resolution of a multispectral image finer than 5 m (Harold et al. 2003). High spatial resolution is required for a better description of metrics such as objects' shapes whereas different object and land surfaces are better identified if high spectral resolution is available (De Jong and Van DerMeer 2005). The SPOT5 panchromatic and multispectral resolutions are respectively 2.5 m and 10 m, offering respective pixel sizes of $6.25m^2$ and $100m^2$. Informal building units are reported to have sizes between 6 and $20m^2$ while formal residential building units are described to have sizes greater than $30m^2$ (Busgeeth et al. 2008). As a consequence, using an image that can provide rich spectral information about the objects on the earth's surface but provides coarse spatial resolution may not be the good combination to extract informal housing units. A solution to this dilemma is to bring together the high spectral resolution property of the multispectral image and the high spatial resolution property from the panchromatic image into one single image to benefit from both properties. Several resolutions merge approaches have been reported in literature with each technique offering its strengths and weaknesses (Simone et al. 2002; Ghassemian 2016; Pohl and Van Genderen 2016). Shamshad et al. (2004) investigated four resolution merge techniques including the Principal Component Analysis, the Multiplicative, the Brovery Transform and the Wavelet Transform resolution merge methods. The study reveals that all four techniques improved the image spatial resolution but only the Principal Component Analysis and the Wavelet Transform preserved the statistical parameters of the bands. For this study the Wavelet Transform method included in Erdas Imagine software was used. The choice on the Wavelet Transform over the principal component analysis was based on the fact that the method does not alter the image radiometric resolution (Shamshad et al. 2004; Mehra and Nishchal 2014). The outcome of this process is a multispectral image that possesses both the high spectral resolution and the high spatial resolution derived from the input images as shown in Fig. 6.

The visual interpretation of the original 10 m multispectral SPOT5 image reveals it would have been very difficult to extract high quality building outlines due to the poor spatial resolution. With the resolution merge performed, it can be observed that some building outlines are well represented.

Features identification

Spectral features

Spectral features play a vital role in describing objects on the Earth surface as discussed previously. To describe objects

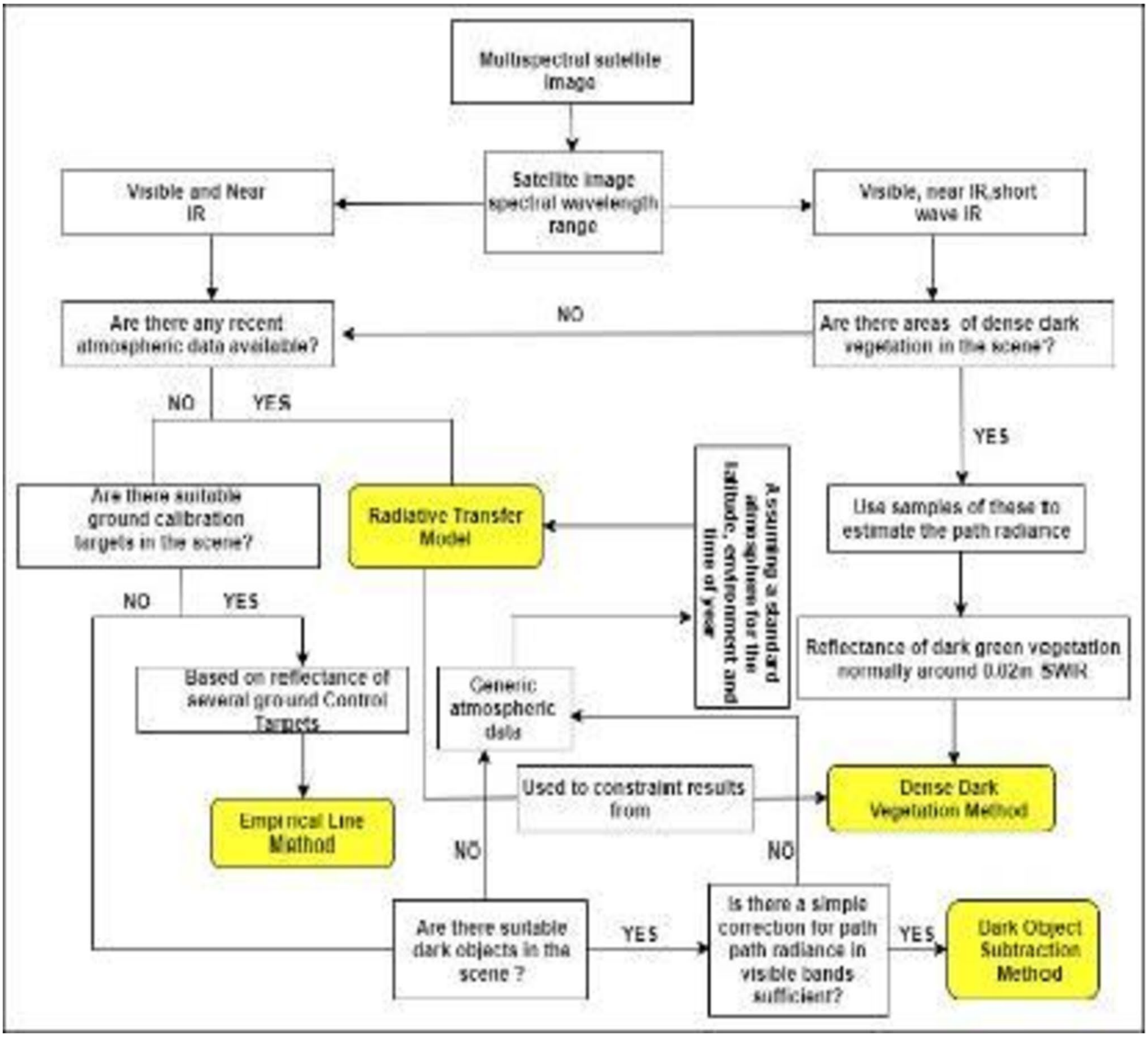

Fig. 4 Workflow showing various criteria to consider when choosing atmospheric correction methods (reproduced from Ncaveo 2005)

within our area few spectral metrics were estimated in order to separate land cover/land use classes from one another when running the classification algorithm. NDVI indexes have been proven efficient to separate vegetation from non-vegetation classes (Gandhi et al. 2015; Hashim et al. 2019). To separate vegetation from non-vegetation classes a threshold measure of the index was estimated to isolate vegetation class from other classes. In addition to the NDVI index, thresholds in the green band were also estimated to enhance the discrimination between the vegetation class and other classes. Spectral thresholds in the infrared band were estimated to isolate buildings from water bodies. Additionally, some spectral thresholds were estimated to separate residential from non-residential building classes.

Shape features

Shape can enable to separate tree patches from green building roofs. For that purpose shape compactness thresholds were estimated using the equation in (1). A certain number of shape compactness thresholds were estimated to separate residential from non-residential buildings and to separate informal housing from formal housing units.

Size features

Size characteristics including area sizes, perimeters, lengths, width, and length over width ratios were estimated. For instance, some length over width thresholds were estimated

Applied Geomatics (2021) 13:619–643

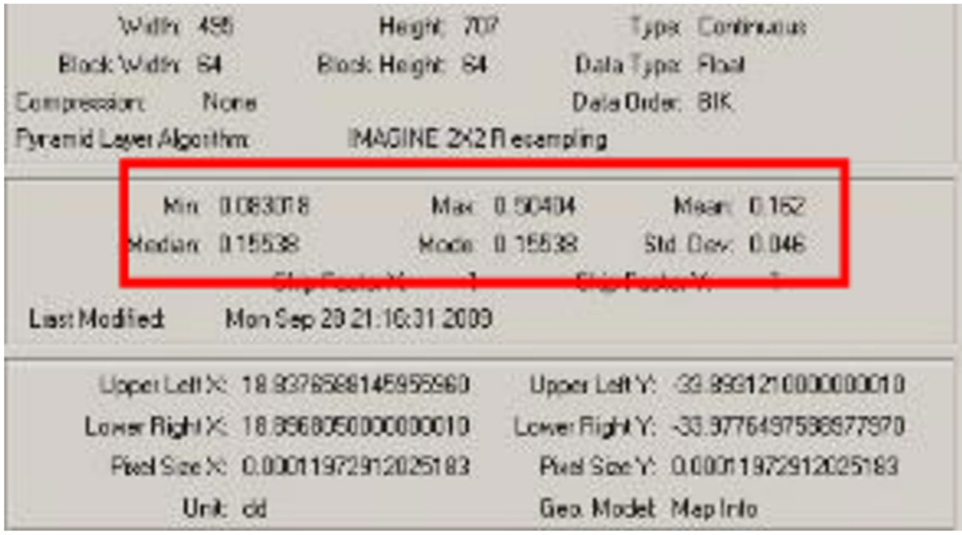

Fig. 5 The reflectance of water bodies was improved to reach a value near zero as minimum reflectance of the multispectral image in the infrared band

Fig. 6 Top: original SPOT5 10 m multispectral image and at the bottom the resolution merged image

to separate roads from non-roads classes. Length thresholds were also estimated to isolate roads from non-road classes. Objects' size measures were also estimated in order to identify the most meaningful measures that can separate land use/land cover classes from one another. Some area size thresholds were selected to separate buildings from water bodies as well as separating residential from non-residential buildings. Area size thresholds of parking were estimated to separate educational buildings from commercial buildings as well as informal housing units from formal housing units.

Distance features

High-dimensional data are very common in image classification when multiple features such as proximity distances between various components of land use/land cover classes are to be considered (Lever et al. 2017). Proximity between objects in urban areas is among the most diverse, making it very difficult to separate one class from others due to the large dimensionality of the data. One solution we opted in order to identify the most prominent distance features that enable the separation between classes and reduce the dimension of the data is to process the collected measures through a Principal Component Analysis (Ng 2017; Cushion et al. 2019). For the purpose, distance measures were manually collected after digitizing various classes samples in ArcGIS then recorded the different distances that separate classes using the measure tool in ArcGIS. Table 1 presents the averaged distance measures between the various land use/land cover classes.

In order to improve numerical stability in the subsequent PCA, a normalization was performed on the data (Gewers et al. 2018). Normalizing a dataset can optimize its discrimination power, which is central for image classification (Ng

Table 1 Averaged inter-class mutually separating distances

Reference class\variable distance to	Residential	Commercial	Industrial	Educational	Urban Vegetation	Roads
Residential	20 m	500 m	500 m	500 m	16 m	5 m
Commercial	500 m	1000 m	1000 m	500 m	5 m	150 m
Industrial	500 m	1000 m	20 m	1000 m	15 m	100 m
Educational	500 m	500 m	500 m	20 m	30 m	50 m
Urban Vegetation	6 m	5 m	15 m	30 m	3 m	5 m
Roads	5 m	150 m	100 m	50 m	5 m	50 m

2017). Each entry of the normalised matrix M_{norm} is estimated using the equation as follows:

$$x_{Norm} = \frac{x_i - \overline{x_i}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \overline{x_i}\right)^2}} \tag{7}$$

The produced matrix M_{norm} contains in its columns and rows the various cross-distances x_i between individual classes $_i$, where, i = {residential, commercial, Educational, urban vegetation, roads}, and $\overline{x}_i$ the mean cross-distance between a list of six classes with regards to one another.

The normalized matrix does not enable to identify and locate the greatest variance within the data; thus, a covariance matrix must first be estimated from the normalised data (Ng 2017; Hernandez et al. 2018). Equation (8) given as follows serve this purpose:

$$Cov(x, y) = \frac{\left(\rho - \overline{p}\right) \times \left(y - \overline{y}\right)}{N - 1} \tag{8}$$

With N, the number of cross-distances represented in each row or column of the matrix. The terms $\overline{x}$ and $\overline{y}$ represent the means estimates of x and y measures respectively. The identification and location of the greatest variances within the covariance matrix is done through the estimation of eigenvalues and their corresponding eigenvectors from the covariance matrix (Granato et al 2018; Cushion et al. 2019). Each eigenvalue was solved under the constraint that it satisfies the equation given in (9) as follows:

$$\det\left(A - \lambda_\kappa I\right) = 0 \tag{9}$$

Table 2 presents the estimated eigenvalues with their respective eigenvectors. The eigenvalues were rearranged in decreasing order with their associated eigenvectors.

The aim of applying principal component analysis on our cross-distance measures was to simplify and reduce the dimensionality of the original data with a minimum loss of the overall dispersion of the measures collected. Table 3 shows a summary of variance magnitude carried by each eigenvalue.

Adding the first four variance magnitudes carried by the respective first four eigenvalues reveal that they carry about 99.5% of the total cross-distance information.

The next step in our cross-distance data optimization is the selection of more potent eigenvalues as well as their associated eigenvectors. Several selection suggestions are available in the literature (Jolliffe and Cadima 2016). For this study, the cumulative variance magnitude approach was chosen for its simplicity. Each coefficient of the eigenvectors is a weighting factor for each cross-distance in the original data and the more weight is attached to a cross-distance, the more discriminative power it has. The products between the coefficients of eigenvectors with their associated cross-distances in the original data are called scores (Happ and Greven 2018). Each coefficient also describes the spatial relationship that exists between a reference class and the other classes within the same urban area. To estimate the data scores in this study, we will use the difference between cross-distances and their respective centre means in order to give equal influence to each cross-distance

Table 2 Computed eigenvalues with their respective eigenvectors (principal components' columns)

λ_i number	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
Classes\eigenvalues	3.413	1.172	0.964	0.418	0.033	-4.67E-7
Principal Components	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6
Residential	0.485	0.258	0.024	0.501	**0.656**	0.124
Commercial	0.531	-0.071	0.133	-0.194	-0.067	-0.808
Industrial	0.316	0.131	-0.731	-0.544	0.120	0.197
Educational	0.374	-0.313	0.579	-0.448	0.0786	**0.469**
Urban vegetation	0.142	0.857	0.254	-0.098	-0.398	0.117
Roads	0.472	-0.282	-0.220	0.452	-0.621	0.242

Table 3 Summary of variance magnitudes carried by each eigenvalue

Eigenvalues	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
Estimates	3.413	1.172	0.964	0.418	0.033	-4.67E-7
Proportion	0.569	0.195	0.161	0.070	0.006	-7.78E-08
Cumulated	0.569	0.764	0.925	0.995	1.00	1.00

measure involved in the analysis (Daszykowski et al. 2007). The process is labelled translation since it does not affect the interpretation of the data because the variances of our original cross-distance data are the same as those obtained after the translation. Since the aim of this analysis is to increase the discrimination power of each cross-distance measure, it is recommended to use in addition to the principal component analysis, a projection method to strengthen the PCA results (Gewers et al 2018) and for this study we selected the Chebyshev's projection method instead of the empirical rule because the estimated distance scores do not follow a normal distribution (Seresht and Ghassemian 2016).

Image classification

Finding optimal local segmentation scale-compactness parameters combination

Post image enhancement, the image was segmented at respective scale parameters of 20, 40, 60, 70, 80, 100, 120, and 135 using eCognition. The collected segments' areas and brightness attributes were used to compute inter-segment heterogeneity ∂_{values}, using Eq. (10) as follows (Yang et al. 2019):

$$\partial = \frac{\sum \ell_i v_i}{\sum \ell_i} \tag{10}$$

With ℓ_i the area of the segment (i) and v_i the brightness attribute of the segments (i). The expression could hold with the numerator only but to minimize the instability caused by objects smaller than shacks' sizes recorded in our building size database, a division by the sum of segments 'sizes was applied to the numerator. Since the success of a segmentation process does not get credit from the scale parameter alone but from the correct balance with other parameters such as shape, colour and compactness weight(which is complementary to the smoothness weight), we performed a search of parameters that can optimise the performance of scale parameters. Each segmentation scale was then tested with various compactness thresholds ranging from 0.1 to 0.9(Smoothness weight decreasing from 0.9 to 0.1). To achieve internally homogeneous segments, the function in (10) must reach a peak (local maxima), characterizing segments with high heterogeneity attributes with reference to their neighbours at a given combination of scale parameter and compactness/smoothness weight. Although the proposed strategy differs in its formulation from those of Espindola et al. (2006), Kim et al. (2008) that incorporated the Moran's index in a Global function, it achieves the same objectives of identifying optimal scale parameters from the peak values of a curve representing objects' variances plotted against their respective segmentation scale parameters (Wang et al. 2019). Our proposed approach is a two-phase method which includes a local and a global search. The local search identifies the optimal compactness/smoothness thresholds that would produce a peak of the function which characterises high inter-segment heterogeneity at each one of the scale parameters considered. Table 4 shows compactness and normalized inter-segment local heterogeneity measures (Yang et al. 2019) tested with the scale parameter of 20. The test was repeated with all the remaining seven scale parameters.

To perform a local search for optimal parameter combinations, each normalized inter-segment heterogeneity (intra-segment homogeneity) is plotted against its associated compactness value to produce a search curve. A peak of the curve will mean there is a high inter-segment heterogeneity between segments. This means that each individual segment is internally homogeneous and can distinguish itself from its neighbours at the corresponding parameter combination thresholds that produced the peak(s).

Figure 7 shows an unsuccessful search for optimal compactness thresholds from the normalized inter-segment heterogeneity curve as there are no peaks along the curve.

Similarly, Table 5 shows data tested at scale parameter of 60 in order to identify local compactness thresholds that would produce homogeneous segments (distinct from their neighbours).

Figure 8 shows a successful search of optimal local compactness thresholds that would optimize the segmentation at

Table 4 Compactness thresholds and their corresponding normalized inter-segment heterogeneity values at scale of 20

Local compactness values	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Local normalized inter-segment Heterogeneity	1	0.987	0.905	0.821	0.654	0.477	0.390	0.336	0

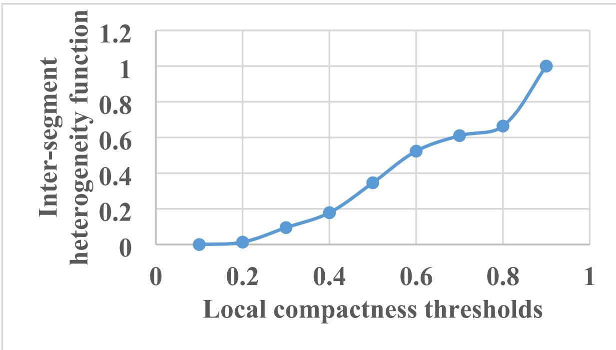

Fig. 7 Optimal local compactness search at scale parameter of 20

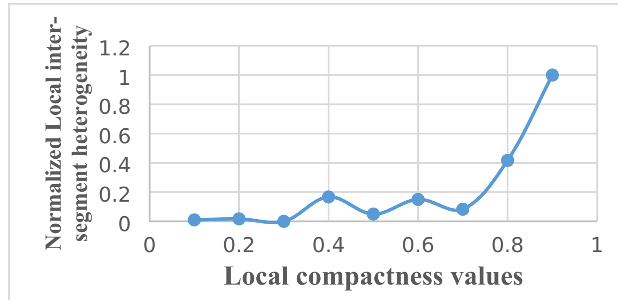

Fig. 8 Optimal local compactness search at scale parameter of 60

scale 60 and produce meaningful segments with shapes of segments as close as possible to their real world shapes.

The curve of normalized local inter-segment heterogeneity reveals two peaks at 0.4 and 0.6, meaning the combination of a scale parameter threshold of 60 would produce meaningful segments for at least two land use/land cover classes.

Finding the optimal number of segmentation levels needed for the study area

In order to determine the number of segmentation levels needed to obtain optimal objects outlines, we will look at the number of optimal compactness thresholds across the various scale parameters used in the search of the best fit combinations between scale parameter and compactness. Figure 9 presents the summary of scale and compactness parameters search results. In Fig. 9(a), (e), and (f), the respective scale parameters 20, 40, 120, and 135 did not find any best fit compactness thresholds, which would result in segments outlines far from their real world shapes while in Fig. 9(b), (c) and (d) the inter-segment heterogeneity curves produced some peaks, characteristics of homogeneous segments. Observing Figs. 8 and 9(b), (c) and (d), there is a redundancy of the compactness value of 0.6 at scale parameters 60, 70, 80, and 100; this could indicate a compactness threshold at which a certain type or groups of land use/land cover classes(s) is (are) "optimally" segmented. Figure 9(c) also shows an additional compactness threshold of 0.8 at scale parameter 80. With consideration of these observations we may hypothesize that there may exist only two segmentation levels needed to extract the different land use/land cover classes with acceptable segments' outlines from our study area. A global search of both parameters in the next subsection will attempt to verify this hypothesis.

Global search of optimal scale parameter-compactness combination

The global search process is similar to the local search approach but differs in the fact that the compactness thresholds are not assessed with reference to the inter-segment heterogeneity function but this time with reference to their respective scale parameters as illustrated in Table 6.

Due to the fact that the relationship between compactness thresholds and scale parameters is not linear, it was expected to obtain a curve which is not a sinusoid as shown in Fig. 10, but it can be seen that the shape of the curve shows two "summits" which seems to confirm the data requires only two segmentation levels for its classification.

Automatic selection of scale parameter and best fit compactness thresholds for optimal segmentation of the study area

Let α be a weighting compactness parameter, β represents the smallest segmentation scale parameter that would produce the smallest meaningful segment when α gets closer to zero. Let x_i be the optimal compactness threshold for the scale parameter y_i. Let n be the number of compactness thresholds that produced peaks in the local search process. Figure 10 shows that the relationship between scale parameters and compactness thresholds is not linear; thus, a linear approximation of the curve in Fig. 10 is needed in order to automatically estimate which scale parameter produces good quality segments if used with a certain

Table 5 Search for optimal compactness at scale parameter of 60

Local compactness values	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Normalized inter-segment heterogeneity values	0.010	0.017	0	0.017	0.050	0.150	0.083	0.417	1

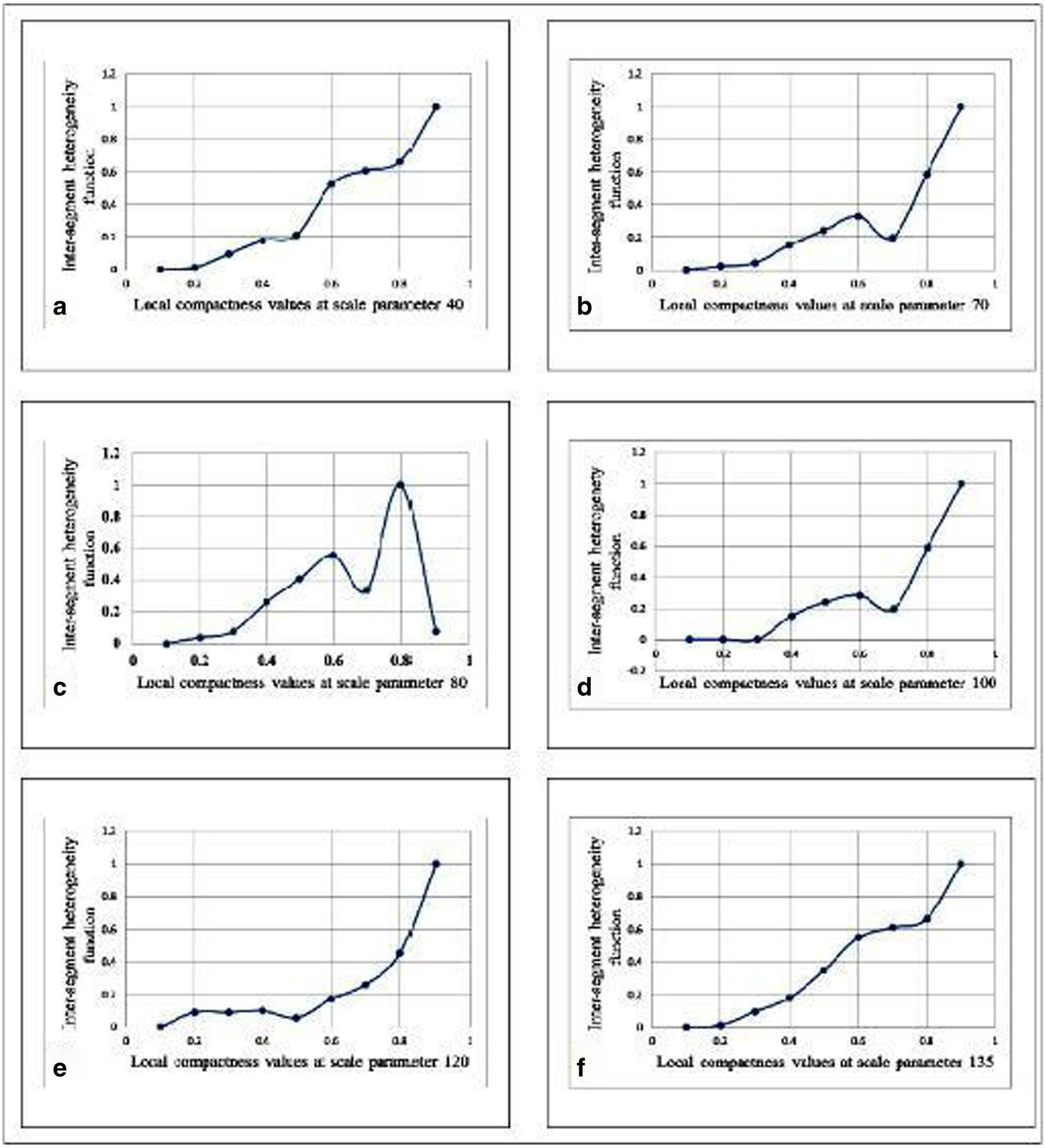

Fig. 9 Summary of local searches of best fit combinations between segmentation scale and compactness parameters. In (a), the output of a search at scale parameter 40, in (b) the research results at scale parameter 70 and in (c) the search result at scale parameter 80, (d) the search outcome at scale parameter 100, (e) the search results at scale parameter 120 and (f) the search results at scale parameter 135

Table 6 Scale parameters with their associated best fit compactness thresholds

Compactness thresholds	0	0	0.4	0.6	0.6	0.6	0.8	0.6
Segmentation scale parameters	20	40	60	60	70	80	80	100

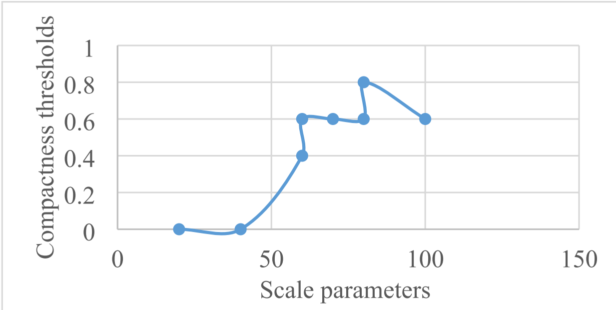

Fig. 10 Segmentation scale parameter and compactness thresholds in a global search process. The curve shows two dominant peaks at which scale parameter 60 with a compactness threshold of 0.6 and another peak at scale parameter 80 with a compactness threshold of 0.8. From the results in the figure, it can be argued that the various land use/land cover classes in our image can fairly be segmented at two segmentation levels at scale parameters 60 and 80 with respective compactness thresholds 0.6 and 0.8

compactness measure and vice versa. For the purpose, we will estimate the parameters α, β and n using linear Least squares method as follows:

$$\begin{cases} \alpha \sum x_i^2 + \beta \sum x_i = \sum x_i y_i \\ \alpha \sum x_i + \beta n = \sum y_i \end{cases} \tag{11}$$

The estimated function is expected to be a linear approximation of the relationship between scale parameters and their respective best fit compactness values and it will be of the form:

$$S = \alpha\psi + \beta \tag{12}$$

With S the scale parameter, ψ a weighting parameter equal to the compactness threshold. The first quantity after the equal sign matches the mathematical formulation of the scale parameter as define in the eCognition user guide, that describes the scale parameter as the product between the image variance and a weighting factor which may be the shape, the colour, the smoothness or the compactness(Definiens 2007). Our proposed formulation is more suitable for compactness parameters which the study focuses on, as well as segmentation scale parameters. The slight difference between our proposed approach and the mathematical formulation of the scale parameter in eCognition user guide is that we constrained the scale parameter to a minimum value of β which is the smallest scale parameter to produce meaningful segments outlines when the weighting factor is set to zero. This is to reduce the operator efforts and time in a trial and error parameter search since the function in (12) would enable to automatically find the best fit compactness threshold for any given scale parameter and vice versa. In the next section Fig. 11 presents the classification strategy workflow.

Image classification

Figure 11 describes the various steps to be undertaken in order to perform our image classification using eCognition.

After identifying suitable scale parameters and associated shape compactness thresholds, two-level classifications were performed using eCognition object-based classification. The first-level classification aims to identify eight classes including formal, informal, educational, commercial, industrial buildings as well as water bodies, urban vegetation and road classes using spectral signatures, size, length, shape compactness measures as well as spatial context of each class. The second-level classification aims to separate informal housing units from all the other built up structures. The strategy employed was to group the formal, commercial, industrial and educational building classes into a single 'Formal' building super class. Water bodies, urban vegetation, roads and informal building classes were transferred to the second-level classification to build a final classification map of five classes. Table 7 presents the various objects'features to be used to separate individual segments from their neighbours.

Results

Spatial distance feature analysis

Table 8 presents the first four principal components and their respective loadings. An analysis of the first Principal Component shows that a large number of residential and commercial buildings are located near major roads. This principal component is strongly correlated to commercial buildings characterized by the largest loading value associated with this land use class and also seems to describe the town centre. The second principal component is mainly correlated to urban vegetation and reveals the absence of roads, commercial, educational buildings. This principal component seems to describe the periphery of the town and reveals that inter-class distance feature would easily enable to separate outer urban vegetation from residential, commercial, industrial, educational and road classes. The third principal component is strongly correlated to industrial and education buildings (large building infrastructures). The opposition in signs of these two land use classes reveals they are not correlated, meaning that with an increase of educational buildings there is a decrease

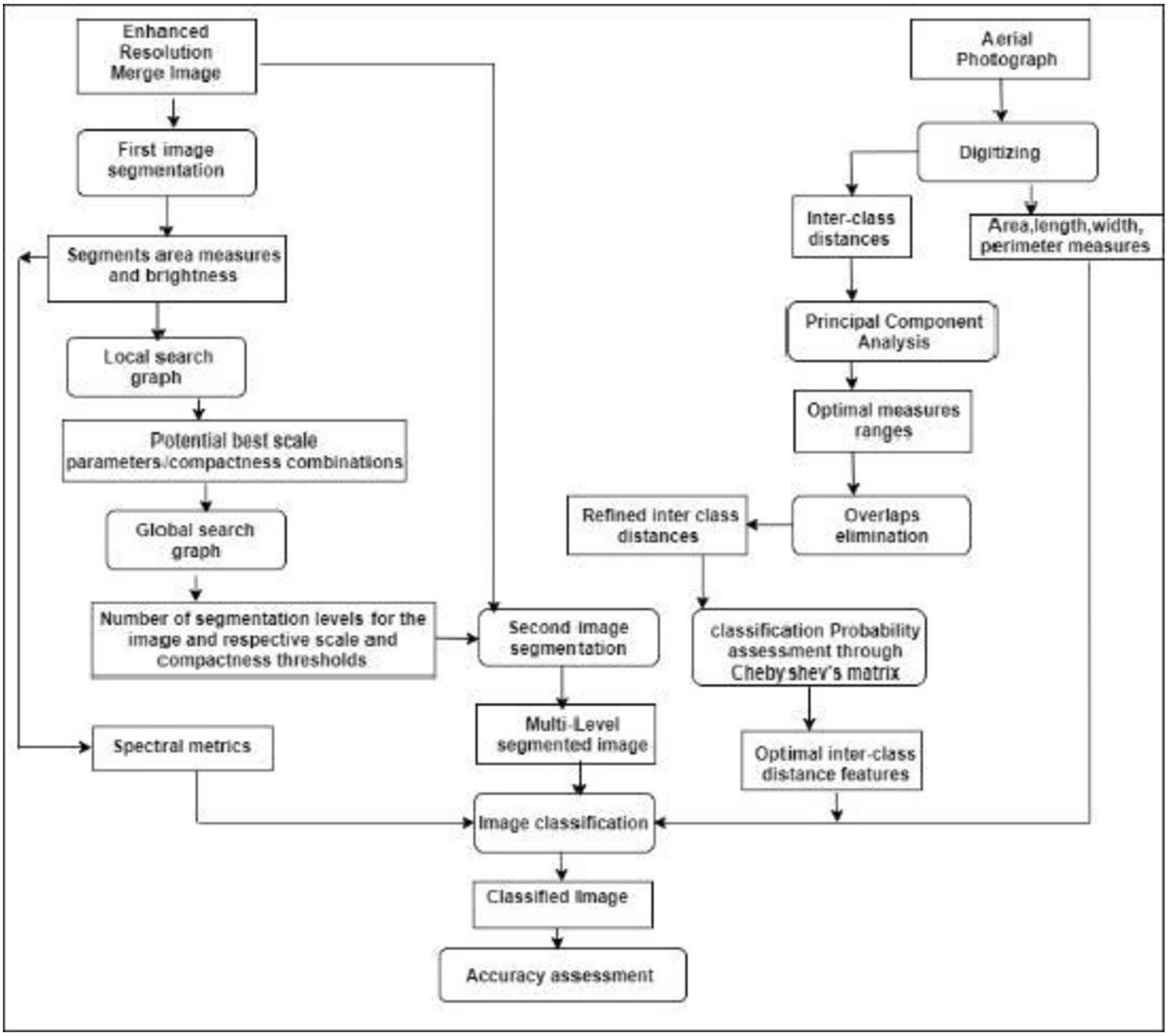

Fig. 11 The workflow has three main streams including a segmentation parameter search stream, an inter-class distance features optimization stream and the classification stream which gathers all the estimated attributes into the classification process

of industrial buildings in the neighbourhood. The fourth principal component is strongly correlated to residential building and road classes. The opposition in signs between residential buildings and commercial, industrial, educational building classes shows that with an increase of residential buildings there is as consequence a shortage in commercial, industrial, educational building classes. This seems to describe the informal settlement area.

The orthogonality property of the principal component matrix enables to keep any relationships that exist between land use/land cover classes in the original data (López–Bueno et al. 2018). The obtained matrix coefficients called scores represent the distances from the origin of the Principal Component coordinate system, along a principal component to the point where a variable is orthogonally projected onto the principal component as illustrated in Fig. 12.

Table 8 presents the projection of the origin distance feature from their original "coordinate system (X_1, X_2, and X_3) onto the principal component coordinate system. Each variable has a unique score along each principal component axis.

An observation at the second principal component scores reveals that commercial buildings and urban vegetation are located at almost the same distance from the origin of the principal component coordinate system; thus, it would be difficult to separate the two classes using the separation distance in the second Principal Component axis. The distance between educational buildings from the origin of principal components' coordinate system is the largest along

Table 7 The various objects features used for image classification

Land use/Lan cover classes	Descriptive attributes
Water class(Dams, swimming pools, water reservoirs)	Distances to residential buildings located between 10 and 25 m; Area sizes located between 200m2and 800m2;Perimeter measures located between 60 and 120 m; Reflectance in the red band smaller than 1; Reflectance in infrared band smaller than 1
Educational class(schools, university)	Distances from sport fields smaller than 677 m; Reflectance in the red band located between 90 and 190;Area sizes greater than 615m^2; Shape compactness = 0.704
Commercial buildings class	Distances from main roads smaller than 563 m; Reflectance in the red band located between 85 and 100; Area size smaller than 877m^2; Shape Compactness = 0.606
Industrial buildings	Distances from commercial buildings smaller than 899 m; Reflectance in the green band located between 100 and 192; Area sizes greater or equal to 700m^2; Length measures greater or equal to 55 m; Width measures greater or equal to 20 m; Shape Compactness = 0.391
Informal buildings class	Distance from commercial buildings greater or equal to 723 m;Reflectance in the green band located between 192 and 250; Area size measures greater or equal to 6m^2; Area size smaller or equal to 20m^2;Shape compactness = 0.785(smaller Units); Compactness = 0.775(Larger Units)
Formal buildings class(Blocks of flats, Town houses, RDP houses)	Distances from commercial buildings smaller than 577 m; Reflectance in the red band located between 100 and 250; Reflectance in the green band located between 98 and 192.;Area size smaller than 9576m^2 (Block of flats);Area sizes smaller than 117m^2 (Town houses); Shape compactness = 0.781(RDP units); Shape compactness = 0.7854(Town houses); Shape compactness = 0.786(large block of flats)
Urban vegetation	Distances from roads smaller than 133 m; Reflectance in the green band located between 23 and 94; NDVI index smaller than 0.3; Area size greater than 1500m^2(Sport fields); Area size smaller than 900m^2 (recreational parks); Shape compactness = 0.45(recreational parks); Shape compactness = 0.750) Sport fields)
Roads(Tare roads and gravel roads)	Distances to industrial buildings greater than 605 m;Reflectance in the red band located between 120 and 210; Length greater than 100 m; Width measures located between 9 and 11 m' Width measures located between 2 and 4 m;length over width ratios located between 1 and 17; Shape compactness = 0.2380

the second principal component. Along the third principal component, industrial buildings have the longest distance from the origin of the new coordinate system, while urban vegetation shows the longest distance from the coordinate origin along the fourth principal component. From the different scores of land use/land cover classes, we were able to estimate the inter-class distance features with reference to the new coordinate system origin. Table 9 shows the more potent distances separating the various land use/land cover classes.

The first, third, and fourth principal components seem to provide the most valuable inter-class distances. The most potent distance separating a majority of residential buildings from commercial buildings was estimated at about 723 m while some residential buildings were located near 601 m from some components of urban vegetation such as recreational parks, sport fields. Educational buildings were found located at a fair distance of about 647 m from some commercial establishments while some of the commercial establishments were located from the periphery of town at a distance approximating 1324 m. An approximate distance of 132 m separates some major roads from some recreational parks, urban trees, and sport fields. Commercial and industrial buildings are separated by an approximate distance of about 900 m as revealed along the third principal component. Some residential buildings were found located at an approximate distance of 475 m from educational buildings along the fourth principal

Table 8 The estimated class scores which are projection of the multi-dimensional inter-class distances onto 2D coordinate system

	PC_1 scores	PC_2 scores	PC_3 scores	PC_4 scores
Residential buildings class	28	-79	5	-313
Commercial buildings class	751	25	-317	-375
Industrial buildings	129	-174	583	86
Educational buildings class	104	194	-267	162
Urban vegetation class	-572	25	18	252
Roads class	-440	8	-22	188

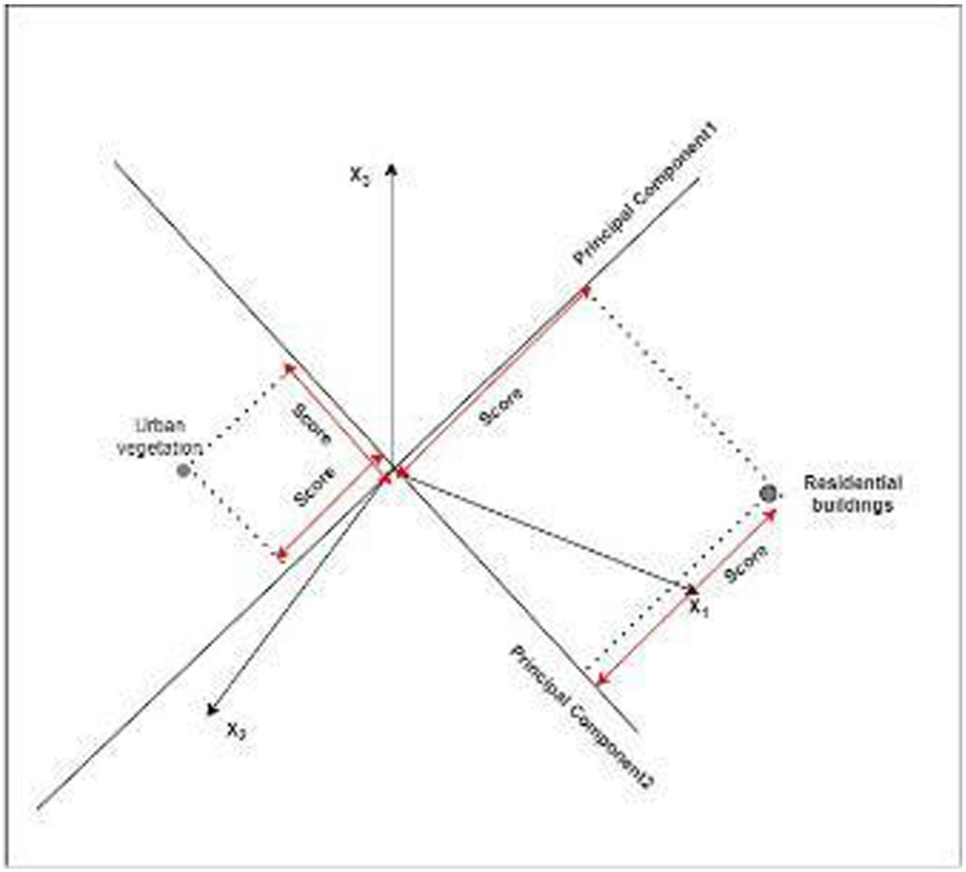

Fig. 12 Illustration of the score concept with examples for residential building and urban vegetation classes

component. In order to estimate the probability that a land use/land cover located 723 m from a residential building is a commercial establishment or a land use/land cover situated 475 m from an educational building is a residential building, the identified optimal distances are projected onto Chebyshev's matrix as presented in Table 10.

From the analysis of the various distance ranges, it can be observed there are overlaps between certain distance ranges,

Table 9 A presentation of the various inter-class distances estimated from their respective scores with the optimal distances written in bold

Inter-class distance typology(m)	PC_1 scores	PC_2 scores	PC_3 scores	PC_4 scores	Mean	Standard Deviation
Residential-Commercial buildings	**723**	104	322	62	303	262
Residential-Industrial buildings	101	95	**577**	399	293	205
Residential-Educational buildings	76	273	273	**475**	274	141
Residential buildings-Urban vegetation	**601**	104	13	565	321	264
Residential-Roads	468	87	28	**501**	271	215
Commercial-Industrial buildings	622	198	**899**	461	545	254
Commercial-Educational buildings	**647**	170	49	536	351	248
Commercial buildings-Urban vegetation	**1324**	0.16	335	627	571	227
Commercial buildings-Roads	16	16	294	**563**	223	227
Industrial-Educational buildings	24	368	**850**	76	330	328
Industrial buildings-Urban vegetation	**701**	199	564	166	408	231
Industrial buildings-Roads	569	182	**605**	102	365	225
Educational buildings-Urban vegetation	**677**	170	286	90	306	225
Educational buildings-Roads	**544**	186	245	26	251	188
Urban vegetation-Roads	**132**	17	41	64	63	43

which constitutes a challenge for the classification algorithm. To address this challenge, an elimination of the overlapping distance windows is performed. For the purpose we kept the first inter-class range from 484 to 1089 m unchanged since it accommodates at least one of the identified optimal separation distances. Looking at the second distance range along the first Principal Component [473–1114 m], it can be observed that the end distance 1114 m is greater than the end distance of the first distance range 1089 m, resulting in the second distance range along PC1 becoming [1089–1114 m]. Similarly looking at the third distance range along the same Principal Component [374–916 m], it can be observed that the starting distance 374 m is smaller than 473 m, which is the starting distance in the second distance range [473–1114 m], this results in the third distance range of the first Principal Component becoming [374–472 m]. After eliminating all the overlapping distances and rearranging the Table 10 results in the following probability matrix in Table 11.

Relocating the potential inter-objects' distances pointed out in Table 10 into their respective refined ranges then applying Chebyshev's probability rules while relying on other attributes such as spectral reflectance, shape, size, it can be revealed that there is 94% of chances that a small or medium size building located at about 723 m and 577 m respectively from commercial establishments and industrial buildings, would be classified as a residential building. Similarly a land use/land cover located 132 m from a road has 75% of chances to be classified as urban vegetation while a linear feature located about 605 m from an industrial building has 94% chances to be classified as a major road. Ninety-four percent (94%) would be the probability of a building located about 677 m from a sports field to be classified as an educational building. There is 94% probability of a large building located about 563 m from a major road to be classified as a commercial building. A large building located about 899 m from a large building with brighter roofing material has 89% probability to be classified as an industrial building. Because of the small number of water bodies within the study area, they were not included in the principal component analysis; however, their locations with reference to residential buildings were between 10 and 25 m.

Spectral properties of features

An analysis of the collected spectral characteristics from the segmented objects show that some building roofs have digital numbers in the red band ranging between 100 and 250 while ranging between 98 and 192 in the green band. In contrast, urban vegetation class was well described by digital numbers ranging between 23 and 94 in the green band. Moreover, urban vegetation exhibits a Normalized Difference Vegetation Index greater than 0.3 to separate itself from non-vegetation classes. Reflectance lower or equal to one in the red and infrared bands characterized water bodies. In opposition, roads were found with high digital numbers in red and infrared bands.

Size and shape properties of features

Roads, which share similar spectral properties with some tiled roofs, were found with length measures greater than

Table 10 Projection results of the optimal inter-class distance features per Principal Component axis (in bold) onto Chebyshev's matrix. The best separation distances presented in Table 9 can be identified within their optimal distance ranges

Principal Components' axes	Chebyshev's inter-class distance(m) features' classification					
	$\|Mean-2\delta\|$	$\|Mean+2\delta\|$	$\|Mean-3\delta\|$	$\|Mean+3\delta\|$	$\|Mean-4\delta\|$	$\|Mean+4\delta\|$
First principal component	222	827	**484**	**1089**	746	1351
	208	849	**473**	**1114**	737	1378
	159	701	**374**	**916**	589	1131
	145	846	393	1094	**641**	**1342**
	118	1025	109	1252	**336**	**1479**
	54	869	284	1099	**515**	**1330**
	145	756	371	982	**371**	**982**
	125	626	312	813	**500**	**1001**
	23	150	66	193	**109**	**236**
Third principal component	117	703	323	908	**528**	**1113**
	36	1054	218	1308	**472**	**1563**
	326	985	**654**	**1313**	982	1641
	84	814	309	1038	**534**	**1263**
Fourth principal component	7	556	148	697	**289**	**838**
	231	676	**458**	**903**	685	1130

Table 11 The refined inter-class distance features' classification matrix after elimination of overlapping distances which could have resulted in mixed pixels during the image classification process

Principal Components' axes	Chebyshev's inter-class distance features' classification					
	$\lvert Mean - 2\delta \rvert$	$\lvert Mean + 2\delta \rvert$	$\lvert Mean - 3\delta \rvert$	$\lvert Mean + 3\delta \rvert$	$\lvert Mean - 4\delta \rvert$	$\lvert Mean + 4\delta \rvert$
First principal component	222	827	**484**	**1089**	746	1351
	208	849	**1089**	**1114**	737	1378
	159	701	**374**	**472**	589	1131
	145	846	393	1094	**1114**	**1342**
	118	1025	109	1252	**336**	**371**
	54	869	284	1099	**515**	**528**
	145	756	371	982	**371**	**374**
	125	626	312	813	**641**	**654**
	23	**150**	66	193	**528**	**554**
Third principal component	117	703	323	908	**472**	**484**
	36	1054	218	1308	**654**	**1001**
	326	985	**534**	**548**	982	1641
	84	814	309	1038	**289**	**336**
Fourth principal component	**548**	**556**	148	697	**556**	**641**
	231	676	458	903	685	1130

100 m and widths varying between 9 and 11 m near the city centre while varying between 2 and 4 m at the periphery especially within Kayamandi area. Roads are also characterized by a length over width ratios between 1 and 17. Some digitized sizes of sport grounds were found greater than 1500 square meters while some recreational parks were found with sizes approximating 900 square meters. Some educational buildings were found with sizes near 615 m^2 while some blocks of flats digitized sizes approximated 9576 m^2. Some town houses were found with area sizes approximating 117 m^2. Some industrial buildings had lengths close to 55 m while widths approximated 20 m, giving area size approximating 700 m^2. Some commercial establishments reached area sizes close to 877 m^2. The digitized area size of informal housing units (Fig. 13) varied between 6 and 20 m^2 and a field investigation revealed that some structures had roofing materials made of a mixture of wood, plastic and iron sheet.

The various objects' shape compactness thresholds were estimated from digitised objects. For instance smaller informal housing units were characterized by shape compactness values of 0.785 while larger informal housing units achieved 0.775 as shape compactness measure. Recreational parks which are accounted as part of urban vegetation class, were characterised by shape compactness close to 0.45 while Sport fields which presented more regular shapes were characterised by shape compactness of 0.75. Most residential Reconstruction and Development Program (RDP) houses were characterised by a shape compactness index of 0.781. Small residential townhouses were characterised with an average shape compactness index of 0.7854. Industrial buildings were characterised by shape compactness index of about 0.391 due to the non-regular roof shapes

Fig. 13 Informal housing units (shacks), the structures generally have heterogeneous roofing materials

for the majority of buildings while commercial buildings were characterised by an average shape compactness index of 0.606. Most of roads were characterised by shape compactness indices approximating 0.2380. Large blocks of flats were characterised by an average shape compactness index of approximately 0.786 while educational buildings were characterised by a shape compactness index of 0.704.

Automatic selection of best fit scale parameters and compactness thresholds

After solving Eq. (12), a slop value of 50 was obtained while the intersect measure was found at 40 to produce the best fitting linear scale model as follows:

$$S = 50\psi + 40 \tag{13}$$

With the constraint that the scale parameter S and the inter-segment heterogeneity value ψ must always be positive integers. The model enables to predict the optimal combination of scale parameter/compactness thresholds that would produce meaningful segments. The model was tested with the scale parameter 40 which revealed no meaningful segments would be produced since the resulting estimate of inter-segment heterogeneity values is zero which is an attribute of scale parameters that cannot successfully separate individual segments from their closest neighbours. The test with the scale parameter 20 did not satisfy the condition of positive inter-segment heterogeneity estimate and cannot produce meaningful segments as previously found in Fig. 7. However, the compactness threshold of 0.4 predicted a best match with the scale parameter 60 while the compactness 0.6 predicted a best match with the scale parameter 70 as previously found during the local search experiment.

Image segmentation

As per the experiment in Fig. 10, two segmentation levels were performed with scale parameters 60 and 80 with respective compactness thresholds of 0.6 and 0.8 in eCognition. Figure 14 illustrates a segmentation at scale of 60 with compactness of 0.3 and 0.6.

Table 12 Segmentation parameter settings**.** The level 1 segmentation was performed at scale parameter 60 with a compactness weight of 0.6 as experimentally estimated, the shape weight was set at its maximum so that the algorithm accounts for it more than colour during the segmentation

Shape settings					
Smoothness	Compactness	Colour	Shape	Scale Parameter	Segmentation Levels
0.4	0.6	0.1	0.9	60	Level 1
0.2	0.8	0.1	0.9	80	Level 2

For each segmentation, we put a high value on the shape weighting factor so that it is more accounted for during the segmentation than the colour weighting. The compactness weights were selected from the experimental results with their respective scale parameters as illustrated in Table 12.

It can be observed that the smaller the scale parameter the larger is the smoothness weight. The scale parameter of 60 requires a compactness weight of 0.6 and smoothness weight of 0.4 while the scale parameter 80 requires a compactness weight of 0.8 and smoothness weight of 0.2.

Image classification

The first classification level was performed with scale parameter 60 and compactness threshold of 0.6 to extract five land use/land cover classes including roads, formal, informal residential buildings, water bodies and urban vegetation in eCognition software. The second level classification was performed using the segmentation results at scale parameter 80 and compactness thresholds of 0.8 to extract three land use/land cover classes including commercial, industrial and educational buildings.

Merging both levels produced a final classification map of five classes in which industrial, commercial and education building classes were merged to produce one ‘Formal’ building super class since the goal was to separate informal structures from formal ones. The simple random sampling method was used for the accuracy assessment with a minimum of 50

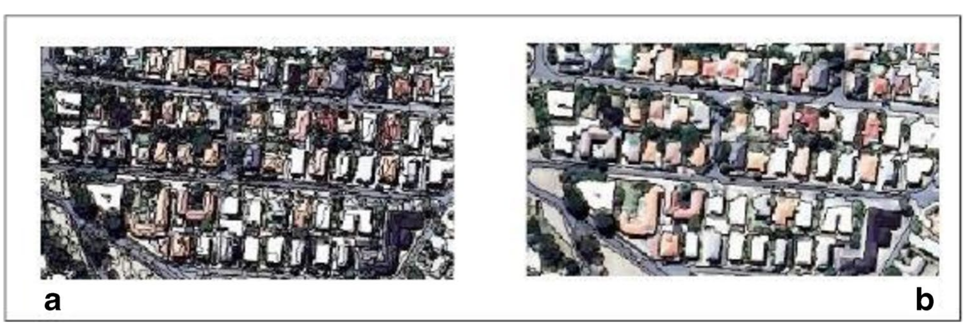

Fig. 14 Illustration of two segmentation results at scale parameter 60 with compactness thresholds 0.3 (a) and 0.6 (b)

Table 13 Level 1 classification confusion matrix

	Reference data									
		industrial	informal	formal	Commercial	educational	water	roads	Grassland	Total
Classification datadata	industrial	64	0	0	1	1	0	1	0	67
	informal	1	65	11	0	0	0	0	0	77
	formal	1	4	134	0	0	0	0	0	139
	commercial	2	0	0	56	1	0	1	0	60
	educational	1	0	0	2	54	0	0	0	57
	water	0	0	0	0	0	72	0	0	72
	roads	1	0	0	0	0	0	110	0	111
	Grassland	0	0	1	0	0	0	0	64	65
	Total	70	69	146	59	56	72	112	64	619
	%PA	91.429	94.203	91.781	94.915	96.429	100	98.214	100	
	%EO	8.571	5.797	8.2191	5.085	3.571	0	1.786	0	
	%UA	95.522	84.416	96.403	93.333	94.737	100	99.099	98.462	
	%EC	4.478	15.584	3.597	6.667	5.2632	0	0.901	1.538	

samples per class as suggested in (Barow et al. 2019). The analysis of the confusion matrix of the first level classification shows that 'water bodies' class produced the highest producer's accuracy of 1 due to its unique low reflectance in the infrared band while roads produced the second highest producer's accuracy of 0.982. The formal housing class was found with the third highest producer's accuracy while informal housing class reached a producer accuracy of 94.20. The highest error of omission was found with industrial building class at a value of 8.57%. The greatest error of commission was found with informal building class that achieved a value of 15.58%. The greatest user's accuracy was achieved with the class water while the road class takes the second place with 99.099%. Formal housing units achieved a greater user's accuracy than informal housing class with respectively 96.40% and 84.42% as summarised in Table 13. The overall classification of the first level segmentation was at 95.52% with kappa at 94.78%.

Table 14 shows the confusion matrix of the second level classification. The formal structures super class is described as the merge of industrial class, commercial, educational class and formal residential class. The class achieved a producer's accuracy value of 0.951. Road class achieved the second highest producer's accuracy at 0.982 after water bodies and urban vegetation classes at the value of 1 each. The overall classification accuracy was obtained at 96.73% with kappa index (Brian 2016; Hoque and Lepcha 2020) at 95.45%. The classification of informal housing structures achieved a producer accuracy of 94.20%.

Table 15 represents accuracies comparison of our proposed methods and that of Kemper et al. (2015). Among the compared metrics are the overall accuracy of the classification, the sensitivity accuracy, the Specificity accuracy, the Precision accuracy, the True Skill Statistic, the Omission and Commission errors. The sensitivity is the probability of the approach to classify correctly a building pixel as part of the Formal structure class. It shows how good is the classification in correctly identifying a building based on the attribute measures defined by the reference dataset (Zhu et al. 2010). The measure can also be derived from the omission error as follows:

Table 14 Level 2 classification confusion matrix

	Informal	Formal	Water	Roads	Grassland	Total
Informal	65	12	0	0	0	77
Formal	4	251	0	2	0	257
Water	0	0	72	0	0	72
Roads	0	0	0	110	0	110
Grassland	0	1	0	0	64	65
Total	69	264	72	112	64	562
PA	0.942	0.951	1	0.982	1	
EO	0.058	0.049	0	0.018	0	
UA	0.844	0.977	1	1	0.985	
EC	0.156	0.023	0	0	0.015	

Table 15 Accuracy metrics comparison between our proposed strategy and that of Kemper et al. (2015)

Accuracy Measures	Kemper et al. (2015) method	The proposed approach
Accuracy	**0.97**	**0.97**
Sensitivity	0.804	0.958
Specificity	0.98	1
Precision	0.477	0.940
True Skill Statistic(TSS)	0.781	0.95
Average built up class Omission error	0.196	0.042
Average non-built up Omission error	0.164	0
Average built-up class Commission error	0.523	0.060
Average non-built-up Commission error	0.585	0.008

$$\text{Sensitivity} = 1 - \text{ error of omission} \quad (14)$$

Since the measure presented in (Kemper et al. (2015) defines the sensitivity of the built-up class, we estimated our measure as the average for the classified build up classes including formal building structure, informal building structures and road, which was done as follows:

$$\text{Sensitivity} = 1 - (0.058 + 0.049 + 0.018/3) = 0.958 \quad (15)$$

The obtained value positively deviates by 0.15461 from the result achieved by the first method in Table 15. Our proposed classification strategy has successfully identified non-built-up pixels which include water bodies and urban vegetation and this led to specificity accuracy measure of 1 which is slightly superior to the Kemper et al. (2015) metric. The correctness metric (Precision accuracy) achieved by our approach is close to double the measure achieved by the Kemper's approach. This measure reveals the ability of the classification strategy to correctly classify non-built up classes. For our results, we had 72 pixels of water bodies out of 72 reference pixels that were correctly classified as water bodies. Similarly 64 pixels of the urban vegetation class out of 64 in the reference data were successfully identified as urban vegetation. The measures can be recovered from the error of commission as follows:

$$\text{Precision} = 1 - \text{ error of commission} \quad (16)$$

In Eq. (16), the error of commission was considered as an average value for the built-up classes that includes Formal structure class, the informal structure class and the road class from our final classification. The metric was estimated as follows:

$$\text{Precision} = 1 - ((0.156 + 0.023 + 0)/3) = 0.940 \quad (17)$$

The True skill Statistic accuracy describes the agreement between the number of building pixels correctly classified with regards to reference pixels. It can be calculated by adding the sensitivity accuracy plus the specificity accuracy minus one (Kemper et al. (2015). An alternative way of estimating this metric is to add the number of correctly classified Formal building pixels plus the number of correctly classified informal building pixels divided by the sum of their respective reference pixel totals. Both methods produce two very close estimates of the metric. The value of the metric for the proposed classification strategy reached about 0.95 compared to 0.781 achieved by the first classification strategy.

Due to the fact that our non-built up pixels were successfully identified, the classification strategy achieved a zero error of omission in comparison to 0.164 achieved by the first classification method. The average built-up and non-built-up error of commission were found very low as outcome of our proposed classification strategy. Figure 15 presents the final classification map displaying the Formal structures class in red and the informal class in dark colour.

In the map, water class is represented by blue polygons while urban vegetation superclass is represented in green. The majority of water reservoirs were identified in the Northern part of the city which suggests some important agricultural activities. A certain number of water bodies were also found near the city centre while few others were successfully identified south of the city centre. The entire road network has been successfully classified and is represented in grey colour. The majority of informal housing units are found in Kayamandi shantytown in the North West of the city represented in black colour. Some dark spots were identified within the city centre and the field investigation revealed the objects were large cars and trucks with sizes approximating those of some informal housing units. This occurrence could be explained by the fact that contextual information relating cars to parking or roads were not accounted for through the principal component distance analysis and it is considered as a 'blind spot' during our image classification preparation. Some formal housing units were successfully classified within the shantytown of Kayamandi and these were identified as Reconstruction and Development Program houses. The majority of urban vegetation, mainly sport fields, were located at the periphery of some residential areas as predicted in the fourth Principal Component with both classes having opposite signs.

Discussion and conclusion

The study dealt with some of the challenges faced when mapping urban areas in general and hybrid urban areas in particular where formal and informal housing units are located within the same cadastral boundaries of the city. Image classification relying solely on spectral properties

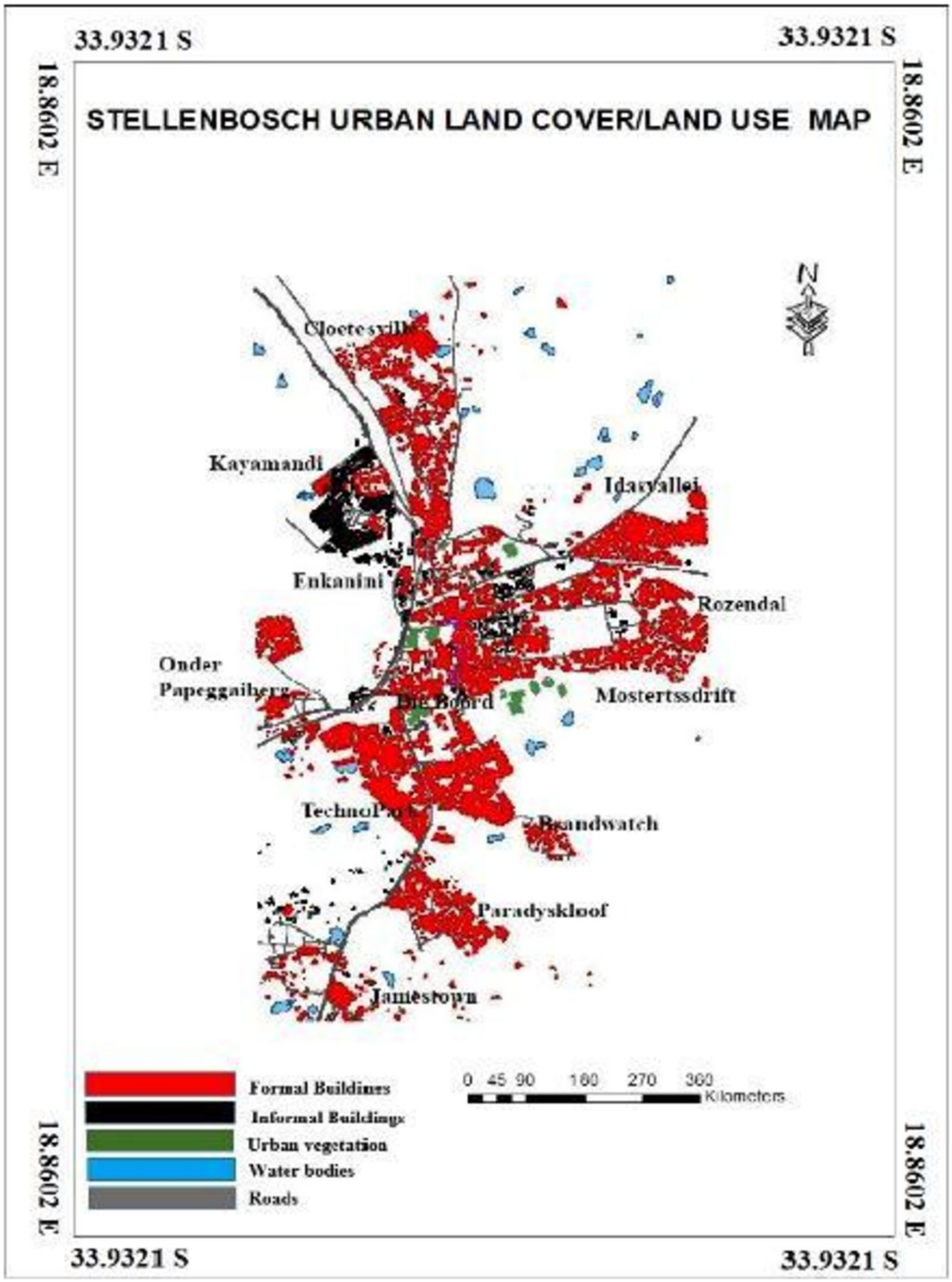

Fig. 15 Image classification of urban area. Most of the land use/land cover classes are clearly defined

has been proven insufficient to separate distinct classes with similar spectral properties. The introduction of contextual information such as spatial inter-class separation distance features, objects' sizes, into the classification process also comes with some drawbacks in the sense that selecting the optimal features within a larger dimensionality of data is not an easy task and is sometimes performed using a trial and error approach. Combining poor quality descriptive features can result in poor image classification results in object-based image analysis. The other challenge is how to balance the five parameters involved in the image segmentation process that include scale parameter, colour, compactness, shape and smoothness weights, in order to achieve an acceptable outcome which will ensure a good image classification outcome. Several studies in the literature have suggested methods to find the optimal segmentation scale parameter, however the success of a segmentation does not solely rely on the scale parameter but rather on a well-balanced adjustment of all the parameters. An approach to select optimal descriptive features was proposed in this study and tested on inter-distance measures in order to isolate the best measures and predict their impact on the image classification outcome. This was made possible through a combination of principal component analysis and Chebyshev's probabilities approaches. The selected distance features, combined with additional attributes such as but not limited to spectral reflectance, shape compactness, size properties, have ensured an overall classification accuracy of 97% and a kappa index of 95.45%. The proposed strategy outperformed some of the most robust classification strategies in certain accuracy assessment metrics. The proposed approach also gives image analysts the control over 60% (3 out of 5) of parameters involved in the image segmentation process including the scale parameter, the compactness and the smoothness weights as presented in Table 12, through local and global parameter searches which also determine the number of segmentation level(s) needed for a given image. Moreover, the approach proposed a scale parameter search function that can predict the optimal scale parameter/compactness weight combination that would guarantee a good intra segments' homogeneity. To the best of our knowledge, such a strategy has not yet been proposed in image analysis although one of the strategy components, namely the principal component analysis has widely been used to identify the best spectral bands that could well describe certain features.

Currently, there is no standard within the remote sensing community to define a best accuracy assessment threshold. However, for homogeneous areas such as vegetation areas, an accuracy of 90% would be acceptable while in complex areas such as urban areas, an accuracy of 85% would be adopted as a perfect agreement (Ye et al. 2018). Thus, since achieved overall accuracies of 96% and 97% for the first-level and second-level classifications respectively fall within this range, they can be considered as successes considering the complexity of the study area as illustrated in the Fig. 16.

Fig. 16 A subset of Kayamandi shantytown area with formal buildings surrounded by informal structures

The use of size, shape metrics, spectral metrics, and spatial context offers more advantages than a combination of texture, morphological and radiometric features in classification of urban areas as proposed in Table 7. The study also added new knowledge to the existing body of literature with regards to spectral and spatial metric attributes that describe urban areas. Further studies would look at incorporating elevation data into the classification procedure and test the approach with different image resolutions and on different study areas since the proposed technique can be repeated on a different area.

Author contribution The authors have contributed 100% in the content of this paper.

Data availability The data used in this study has been legally acquired free of charges for research purposes only.

Code availability The provided piece of code for radiometric correction can be used by anyone who has interest in it.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abbas AW, Minallh N, Ahmad N, Abid SAR, Khan MAA (2016) K-Means and ISODATA clustering algorithms for landcover classification using remote sensing. Sindh Univ Res J-SURJ (Science Series) 48(2)

Ateşoğlu A, Tunay M (2014) Evalauting digital elevation models in the atmospheric correction procedure for different satellite image data (Bartin case study, Turkey)

Baboo SS, Devi MR (2010) An analysis of different resampling methods in Coimbatore, District. Glob J Comput Sci Technol

Barow I, Megenta M, Megento T (2019) Spatiotemporal analysis of urban expansion using GIS and remote sensing in Jigjiga town of Ethiopia. Appl Geomatics 11(2):121–127

Ben-Dor E, Levin N, Saaroni H (2001) A spectral based recognition of the urban environment using the visible and near-infrared spectral region (0.4-1.1 μm). A case study over Tel-Aviv, Israel. Int J Remote Sens 22(11):2193–2218

Boakye E, Anyemedu FOK, Quaye-Ballard JA, Donkor EA (2020) Spatio-temporal analysis of land use/cover changes in the Pra River Basin Ghana. Appl Geomatics 12(1):83–93

Brian B (2016) Temporal and spatial patterns of urban sprawl and their implications on environmental planning in Mbarara Municipality. Appl Geomatics 8(3–4):201–216

Busgeeth K, Brits A, Whisken JB (2008) Potential application of remote sensing in monitoring informal settlements in developing countries where complimentary data does not exist

Campbell J (2006) Introduction to remote sensing, 7th edn. Taylor & Francis, London

Chen L, Zhao S, Han W, Li Y (2012) Building detection in an urban area using lidar data and QuickBird imagery. Int J Remote Sens 33(16):5135–5148

Costa GAOP, Feitosa RQ, Cazes TB, Feijó B (2008) Genetic adaptation of segmentation parameters. In Object-based image analysis. Springer, Berlin, pp. 679–695

Cushion EJ, Warmenhoven J, North J, Cleather DJ (2019) Principal component Analysis reveals the approximate to distal pattern in vertical jumping is governed by two functional degrees of freedom. Front Bioeng Biotechnol 7:193

Daszykowski M, Kaczmarek K, Vander Heyden Y, Walczak B (2007) Robust statistics in data analysis—A review: Basic concepts. Chemometr Intell Lab Syst 85(2):203–219

Debbage N, Bereitschaft B, Shepherd JM (2017) Quantifying the spatiotemporal trends of urban sprawl among large US metropolitan areas via spatial metrics. Appl Spat Anal Policy 10(3):317–345

Definiens AG (2007) eCognition user guide. http://www.DefiniensCom/index.php. Accessed 23 May 2020

Degerickx J, Okujeni A, Iordache MD, Hermy M, van der Linden S, Somers B (2017) A novel spectral library pruning technique for spectral unmixing of urban land cover. Remote Sens 9(6):565

De Jong, Van Der Meer FD (2005) Remote Sensing Image Analysis: Including the spatial domain, vol 5. Kluwer Academic Publishers, New York

Du Plessis DJ (2015) Land-use mix in South African cities and the influence of spatial planning: Innovation or following the trend? S Afr Geogr J 97(3):217–242

Dutta I, Das A (2019) Exploring the dynamics of urban sprawl using geo-spatial indices: A study of English Bazar Urban Agglomeration West Bengal. Appl Geomatics 11(3):259–276

Espindola GM, Camara G, Reis IA, Bins LS, Montero AM (2006) Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. Int J Remote Sens 27(14):3035–3040

Fundisi E, Musakwa W (2017) Built-up area and land cover extraction using high resolution Pleiades Satellite Imagery for Midrand, in Gauteng Province, South Africa. Int Arch Photogramm Remote Sens Spatial Inform Sci 42

Gandhi GM, Parthiban S, Thummalu N, Christy A (2015) Ndvi: Vegetation change detection using remote sensing and gis–A case study of Vellore District. Procedia Comput Sci 57:1199–1210

Gewers FL, Ferreira GR, de Arruda HF, Silva FN, Comin CH, Amancio DR, Costa LDF (2018) Principal component analysis: a natural approach to data exploration. *arXiv preprint arXiv:1804.02502*

Ghassemian H (2016) A review of remote sensing image fusion methods. Inform Fusion 32:75–89

Granato D, Santos JS, Escher GB, Ferreira BL, Maggio RM (2018) Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. Trends Food Sci Technol 72:83–90

Gxumisa A, Breytenbach A (2017) Evaluating pixel vs. segmentation based classifiers with height differentiation on SPOT 6 imagery for urban land cover mapping. S Afr J Geomatics 6(3):436–448

Happ C, Greven S (2018) Multivariate functional principal component analysis for data observed on different (dimensional) domains. J Am Stat Assoc 113(522):649–659

Hashim H, Abd Latif Z, Adnan NA (2019) Urban vegetation classification with NDVI threshold value method with very high resolution (VHR) PLEIADES Imagery. Int Arch Photogramm Remote Sens Spat Inf Sci, 237–240

Heiden U, Roessner S, Segl K, Kaufmann H (2001) Analysis of spectral signatures of urban surfaces for their identification using hyperspectral HyMap data. In IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (Cat. No. 01EX482) (pp. 173–177). IEEE

Hernandez W, Mendez A, Göksel T (2018) Application of principal component analysis to image compression. In STATISTICS (pp. 107–137). InTechOpe

Herold M, Clarke KC, Scepan J (2002) Remote sensing and landscape metrics to describe structures and changes in urban land use. Environ Plan A 34(8):1443–1458

Herold M, Gardner ME, Roberts DA (2003) Spectral resolution requirements for mapping urban areas. IEEE Trans Geosci Remote Sens 41(9):1907–1919

Herold M, Roberts DA, Gardner ME, Dennison PE (2004) Spectrometry for urban area remote sensing—Development and analysis of a spectral library from 350 to 2400 nm. Remote Sens Environ 91(3):304–319

Hoque I, Lepcha SK (2020) A geospatial analysis of land use dynamics and its impact on land surface temperature in Siliguri Jalpaiguri development region West Bengal. Appl Geomatics 12(2):163–178

Jilge M, Heiden U, Habermeyer M, Mende A, Juergens C (2016) Identifying pure urban image spectra using a learning urban image spectral archive (LUISA). In Remote Sensing Technologies and Applications in Urban Environments (Vol. 10008, p. 100080J). International Society for Optics and Photonics

Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. Phil Trans R Soc A Math Phys Eng Sci 374(2065):20150202

Kemper T, Mudau N, Mangara P, Pesaresi M (2015) Towards an automated monitoring of human settlements in South Africa using high-resolution SPOT satellite imagery. Int Arch Photogramm Remote Sens Spat Inform Sci 40(7):1389

Kim M, Madden M, Warner T (2008) Estimation of optimal image object size for the segmentation of forest stands with Multispectral IKONOS imagery. In: Blaschke T, Lang S, Hay GJ (eds) Object-based image analysis concepts for knowledge driven remote sensing applications. Berlin, 2008, pp 291–307, Springer, Berlin, 2008

Kohli D, Sliuzas R, Stein A (2016) Urban slum detection using texture and spatial metrics derived from satellite imagery. J Spat Sci 61(2):405–426

Kuffer M, Barrosb J (2011) Urban morphology of unplanned settlements: The use of spatial metrics in VHR remotely sensed images. Procedia Environ Sci 7:152–157

Kuffer M, Pfeffer K, Sliuzas R, Baud I, Maarseveen M (2017) Capturing the diversity of deprived areas with image-based features: The case of Mumbai. Remote Sens 9(4):384

Lever J, Krzywinski M, Altman N (2017) Points of significance: Principal component analysis

Lhissou R, El Harti A, Maimouni S, Adiri Z (2020) Assessment of the image-based atmospheric correction of multispectral satellite images for geological mapping in arid and semi-arid regions. Remote Sens Appl Soc Environ 20:100420

López–Bueno D, Pham QA, Montoro G, Gilabert PL (2018) Independent digital predistortion parameters estimation using adaptive principal component analysis. IEEE Trans Microwave Theory Tech 66(12):5771–5779

López-Serrano PM, Corral-Rivas JJ, Díaz-Varela RA, Álvarez-González JG, López-Sánchez CA (2016) Evaluation of radiometric and atmospheric correction algorithms for aboveground forest biomass estimation using Landsat 5 TM data. Remote Sens 8(5):369

Manakos I, Schneider T, Ammer U (2000) A comparison between the ISODATA and the eCognition classification methods on basis of field data. APRS 33(Suppl. B7):133–139

Mather PM (2004) Computer Processing of remotely-sensed Images. Wiley, West Sussex

Mather P, Tso B (2016) Classification methods for remotely sensed data. CRC Press

Mehra I, Nishchal NK (2014) Image fusion using wavelet transform and its application to asymmetric cryptosystem and hiding. Opt Express 22(5):5474–5482

Meng X, Wang L, Currit N (2009) Morphology-based building detection from airborne LIDAR data. Photogramm Eng Remote Sens 75(4):437–442

Miky YH (2019) Remote sensing analysis for surface urban heat island detection over Jeddah Saudi Arabia. Appl Geomatics 11(3):243–258

Meyer HP, Niekerk A (2016) Assessing edge and area metrics for image segmentation parameter tuning and evaluation

MohanRajan SN, Loganathan A, Manoharan P (2020) Survey on Land Use/Land Cover (LU/LC) change analysis in remote sensing and GIS environment: Techniques and Challenges. Environmental Science and Pollution Research International

NCAVEO (2005) Network for Calibration and Validation of Earth Observation data School of Geography, University of Southampton SO17 1BJ, UK

Ng SC (2017) Principal Component Analysis to reduce dimension on digital image. Procedia Comput Sci 111:113–119

Novack T, Kux HJH (2010) Urban land cover and land use classification of an informal settlement area using the open-source knowledge-based system InterIMAGE. Health Risk Soc 55(1):23–41

Odindi J, Mhangara P, Kakembo V (2012) Remote sensing land-cover change in Port Elizabeth during South Africa's democratic transition. S Afr J Sci 108(5–6):60–66

Ouerghemmi W, Gadal S, Mozgeris G, Jonikavičius D, Weber C (2017) Urban objects classification by spectral library: Feasibility and applications. In: 2017 Joint Urban Remote Sensing Event (JURSE). IEEE, pp 1–4

Pohl C, Van Genderen J (2016) Remote sensing image fusion: A practical guide. Crc Press

Rani N, Mandla VR, Singh T (2017) Evaluation of atmospheric corrections on hyperspectral data with special reference to mineral mapping. Geosci Front 8(4):797–808

Reigber A, Jager M, He W, Ferro-Famil L, Hellwich O (2007) Detection and classification of urban structures based on high-resolution SAR imagery. In: 2007 Urban Remote Sensing Joint Event (pp. 1–6). IEEE

Richter R, Center RSD (2004) ATCOR: Atmospheric and topographic correction. German Aerospace Center, Mars: Oberpfaffenhofen, Germany

Rocchini D, Di Rita A (2005) Relief effects on aerial photos geometric correction. Appl Geogr 25(2):159–168

Shamshad A, Hussin WMAW, Sanusi M (2004) *International Symposium on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences 2004*

Simone G, Farina A, Morabito FC, Serpico SB, Bruzzone L (2002) Image fusion techniques for remote sensing applications. Inform Fusion 3(1):3–15

Singh K, Malik D, Sharma N (2011) Evolving limitations in K-means algorithm in data mining and their removal. Int J Comput Eng Manag 12:105–109

Song C, Woodcock CE, Seto KC, Lenny MP, Macomber SA (2001) Classification and Change detection using Landsat TM Data: When and How to Correct Atmospheric Effects? Remote Sens Environ 75:230–244

Sowmya DR, Shenoy PD, Venugopal KR (2017) Remote sensing satellite image processing techniques for image classification: a comprehensive survey. Int J Comput Appl 161(11):24–37

Steiniger S, Lange T, Burghardt D, Weibel R (2008) An approach for the classification of urban building structures based on discriminant analysis techniques. Trans GIS 12(1):31–59

Tan KC, San Lim H, MatJafri MZ, Abdullah K (2012) A comparison of radiometric correction techniques in the evaluation of the relationship between LST and NDVI in Landsat imagery. Environ Monit Assess 184(6):3813–3829

Van der Linden S, Okujeni A, Canters F, Degerickx J, Heiden U, Hostert P, Priem F, Somers B, Thiel F (2019) Imaging spectroscopy of urban environments. Surv Geophys 40(3):471–488

Wang Y, Qi Q, Liu Y, Jiang L, Wang J (2019) Unsupervised segmentation parameter selection using local spatial statistics for remote sensing image segmentation. Int Appl Earth Observ Geoformation 81:98–109

Weng Q, Quattrochi DA eds (2006) Urban remote sensing. CRC press

Yang L, Mansaray LR, Huang J, Wang L (2019) Optimal Segmentation Scale Parameter, Feature Subset and Classification Algorithm for Geographic Object-Based Crop Recognition Using Multisource Satellite Imagery. Remote Sens 11(5):514

Ye S, Pontius RG Jr, Rakshit R (2018) A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. ISPRS J Photogramm Remote Sens 141:137–147

Zhu W, Zeng N, Wang N (2010) Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. NESUG proceedings: health care and life sciences, Baltimore, Maryland, 19, p. 67