Elsevier Editorial System(tm) for Landscape and Urban Planning
Manuscript Draft

Manuscript Number:

Title: Contextual Object-based Image Classification of Urban Areas

Article Type: Research Paper

Keywords: Classification; feature extraction; high spatial resolution imagery; LiDAR; mapping; urban areas.

Corresponding Author: Mr. Txomin Hermosilla,

Corresponding Author's Institution: Universidad Politécnica de Valencia

First Author: Txomin Hermosilla, MPhil

Order of Authors: Txomin Hermosilla, MPhil; Luis A Ruiz, Dr; Jorge A Recio, Dr; María Cambra-López, Dr

Abstract: This paper presents a technique for land-use object-based image classification of urban environments that combines high spatial resolution multi-spectral imagery and LiDAR data. Cadastral or land registry plots are used to divide the image and define objects. A set of descriptive features is presented, describing the objects at different urban aggregation levels. Objects are characterised by means of image-based (spectral and textural), three-dimensional, and geometrical features. In addition, contextual features describing two levels of the object are defined: internal and external. Internal contextual features describe the land cover object types (buildings and vegetation) inside the object. External contextual features describe each object while considering the common properties of neighbouring objects, which in urban areas usually coincide with the urban block. The proposed descriptive features emulate human cognition by numerically quantifying the properties of the image elements that enable their discrimination. The land-use classification accuracy values show that the proposed descriptive features enable an efficient characterisation of urban environments. The complementariness between the features derived from different aggregation levels is noticeable. Image-based features are highly discriminative, and the addition of internal and external contextual features significantly increases the classification accuracy of the urban classes considered in this study.

Contextual Object-based Image Classification of Urban Areas

Txomin HEMOSILLA[a], Luis Ángel RUIZ[a], Jorge Abel RECIO[a], Maria CAMBRA-LOPEZ[b]

[a] Geo-Environmental Cartography and Remote Sensing Group.
Universidad Politécnica de Valencia. Camino de Vera s/n, 46022 Valencia, Spain.

[b] ICTA. Universidad Politécnica de Valencia. Camino de Vera s/n, 46022 Valencia, Spain.

Txomin Hermosilla (**Corresponding author**)
txohergo@topo.upv.es
Geo-Environmental Cartography and Remote Sensing Group.
Universidad Politécnica de Valencia. Camino de Vera s/n, 46022 Valencia, Spain.
Tel: 00 34 963877000 ext. 75576
Fax: 00 34 96 387 7559

Luis Ángel Ruiz
laruiz@cgf.upv.es
Geo-Environmental Cartography and Remote Sensing Group.
Universidad Politécnica de Valencia. Camino de Vera s/n, 46022 Valencia, Spain.

Jorge Abel Recio
jrecio@cgf.upv.es
Geo-Environmental Cartography and Remote Sensing Group.
Universidad Politécnica de Valencia. Camino de Vera s/n, 46022 Valencia, Spain.

María Cambra-López
macamlo@upvnet.upv.es
ICTA. Universidad Politécnica de Valencia. Camino de Vera s/n, 46022 Valencia, Spain.

*Highlights

Research Highlights

- Characterisation of dynamic urban areas is complex but necessary.
- Object-based features from LiDAR and imagery are extracted for classification.
- New internal and external urban object context features are proposed.
- Contextual information clearly improves the classification of certain urban types.
- These techniques are suitable for geo-spatial database updating.

1. Introduction

Urban areas concentrate most of the socio-economical activities, jobs, educational and health services, and many cultural and leisure activities. These concentrations are important financial locations for business development and, consequently, for economic growth. These centres attract population because they offer greater opportunities for development. Approximately half of the world's population live in cities (United Nations, 2007) and this proportion is expected to increase progressively to 70% by 2050 (United Nations; 2010). The global increase in urban population has been produced by the rapid urbanisation processes experienced in developed countries in the middle of the twentieth century.

Fast growing cities produce urban sprawl with diverse consequences: mobility problems, atmospheric pollution, unplanned development, social exclusion, etc. At an environmental level, urban sprawl increases the dependence on cars, and the resulting reliance on fossil fuel causes a rise in pollution and greenhouse gas emission. Eventually, new transit infrastructures are required. Uncontrolled building and impervious surface construction leads to an increase in flood risk and a less effective absorption of rainfall into ground water aquifers, producing a decrease in land and water quality. As a consequence, it is necessary to develop technologies and methodologies that permit monitoring the effects of the various problems that are partially caused by urban sprawl. These technologies would help enable the rapid adoption of policies that minimise the negative effects of urban sprawl. Solutions require a precise knowledge of the current urban environment to enable the development of more efficient urban and territorial plans.

Urban areas are composed of different materials and objects (concrete, asphalt, plastic, glass, trees, grass, etc.) arranged in complex structures (transportation systems, recreational zones, residential, industrial, and commercial areas, etc.), (Welch, 1982). Analogously to both levels – material and structure – the terms *land cover* and *land use* are defined. Land cover is a biophysical indicator that describes the materials on the surface of a territory. Land use is an abstract concept that represents a socio-economic criterion referring to the dominant activity of a place, and may include category subdivisions with differing levels of detail. Urbanisation has been an important component of land use and land cover change, and its significance will undoubtedly continue to increase as the majority of the world's population move to cities (Breuste et al., 1998; Pickett et al., 2001; Whitford et al., 2001). The high dynamism of urban areas produces a continuous alteration of land cover and use, and consequently, cartographic information is quickly outdated. Therefore, the availability of detailed and up-to-date cartographic and geographic information is imperative for an adequate management and planning of urban areas. The amount of geographical data currently available is much higher than several years ago. New massive acquisition techniques generate high volumes of information with a constant increase in frequency. In addition to the spectral response of land covers, altimetric information, and information about the roughness of the surface are commonly acquired using laser scanners and radar sensors. However, this volume of data requires processing prior to being added to land use/land cover geospatial databases.

Usually the process of creating land-use/land-cover maps of urban areas involves field visits and classical photo-interpretation techniques using aerial imagery. These methodologies are expensive, time consuming, and also subjective as they require skilled operators with a knowledge of the area being studied. Digital image processing techniques help reduce the volume of information that needs to be manually interpreted. These techniques satisfy current demands for continuously precise data that accurately describes a territory. As a result, the international cartographic community aims to develop useful methodologies for the automatic processing and/or updating of spatial information in urban areas.

Early attempts to automatically derive land use information using digital image processing techniques failed in the precision and level of detail required for urban planning because of the low spatial resolution of the satellite imagery. The subsequent availability of high resolution spatial multi-spectral imagery could not fulfil expectations for increased classification accuracies. This problem, referred to as *'scene noise'* (Gastellu-Etchegorry, 1990), is related to the spatial heterogeneity in the spectral response of urban areas. Pixel-level analysis of high resolution imagery makes the extraction of robust descriptive features representing urban land use extremely difficult, because these cities are composed of different cover types that produce different spectral responses (Barnsley et al. 1991). This spatial variation of the spectral response is partially conditioned by size, shape, and spatial organisation of the buildings in intra-urban open spaces. However, spectral heterogeneity may constitute a useful feature for providing information about urban areas. According to Barnsley and Barr (2000) the main disadvantage for remote sensing is that while there is often a simple direct relationship between land-cover type and spectral reflectance, the same is rarely true of land use. Therefore, the image classification process to produce land-cover maps in urban areas can be considered straightforward when compared to the problematic process of deriving information on urban land use (Eyton, 1993).

Various methodological solutions dealing with high spatial resolution data suggest analysing the area at different levels, or scales, by using geo-referenced ancillary information (Sadler et al., 1991). After a preliminary classification of land cover and the recognition of key urban elements, urban land-use classification is achieved by applying object-based classification techniques over cartographic units. In an object-based approach, image analysis is performed by considering objects instead of pixels. An image object, or simply an object, is a group of pixels with common characteristics created by means of a determined segmentation criterion (Blaschke, 2010). The segmentation method employed is key in the descriptive features of objects because the resultant objects will differ depending on the algorithm and selected parameters. Plot-based image classification is a particular object-based classification case that uses cartographical limits to create objects. These limits better enable the definition of significant objects in the real world than automatic pixel aggregation. This is an especially suitable methodology for anthropogenic environments such as urban areas, where landscape units present unambiguous boundaries that are relatively stable over time.

The human recognition techniques employed for identifying elements in maps or images are performed by means of an intuitive analysis of individual characteristics and the spatial context of topological features within the overall environment (Hussain et al., 2007). The analysis and interpretation of spatial phenomena is a difficult task. According to Anders et al. (1999), the aim of retrieving structured information translated into more meaningful homogeneous regions can be achieved by identifying meaningful structures within the initial random collection of objects and by understanding their spatial arrangement. Urban areas can be decomposed in different aggregation levels, based on the categorisation, relationships, functions, and attributes of their various elements (Thomson and Béra; 2008): buildings, plots, and urban blocks. The urban cadastral plot, or simply a plot, represents a distinguishable administrative unit in terms of land ownership of an urban area. Buildings correspond to basic elements of urban areas and the analysis of their particular characteristics enables the establishment of morphological differences between urban zones at an internal plot level. The aggregation of contiguous plots produces higher level units: urban blocks. These blocks are groups of plots, surrounded by public roads, that combine open spaces and built-up areas whose geometrical shape and topological relationships significantly determine the appearance of urban environments, influencing spatial experience and defining local particularities related to a spatial identity (Laskari et al., 2008). The analysis of urban blocks enables the definition of urban morphology at a higher level than plots.

As the precise characterisation of complex intra-urban patterns is a highly complex task it is common to use two stage approximation methods (Bauer and Steinnocher; 2001). Initially, the main land-cover types or significant elements in the image are detected and this information is then analysed in a spatial context to determine land use. Two methods have been principally employed to represent patterns and define contextual relationships: *fragmentation metric descriptors* (Alberti and Waddell, 2000; Zhang et al., 2004; Vanderhaegen and Canters, 2010), which are frequently used in ecological and landscape analysis (McGarigal et al., 2002); and *graph theory*, which extends the concept of relational graphs and enables the representation of both intrinsic features and extrinsic relationships. This approach has been used by Barnsley and Barr (1997), Barr and Barnsley (1998), Barnsley and Barr (2000), Zhan et al. (2002a), and Almeida et al. (2007).

Depending on the objective, urban characterisation has been focused on two units: buildings (particularly in cartographic generalisation issues) and urban blocks (especially in classification approaches using remotely sensed data). When working on cartographical generalisation issues, the absence of spectral and, frequently, three-dimensional information leads to the description of buildings using geometric features, i.e. size, main orientation, or shape complexity indices. Several contextual relationships are established, and these are based on adjacency (Hussain et al., 2007), spatial arrangement (Boffet and Rocca, 2001; Burghardt and Steiniger, 2005), ancillary thematic data (Boffet and Coquerel, 2000), zone building density (Boffet and Coquerel, 2000; Steiniger et al., 2009), or open areas (Boffet and Rocca, 2001). The neighbouring areas that provide context are defined using urban block limits, or by using distance buffers. However, buffer techniques produce misclassifications and identification errors in areas bordering different urban typologies (Burghardt and Steiniger, 2005).

Classification of urban blocks using remotely sensed imagery usually uses two-stage approximation methods. After classifying land-cover type or identifying significant urban elements – commonly buildings – a land use is assigned to each plot (Zhan et al., 2000) or urban block by examining their contextual relationships (Bauer and Steinnocher, 2001; Zhan et al., 2002b; Herold et al., 2003; Zhang et al., 2004b; Wijnant and Steenberghen, 2004; Herold et al., 2005; Laskari et al., 2008; Novack et al., 2010). Several descriptive features have been employed to characterise the land use of urban elements. The most frequently and successfully employed descriptor is the building-to-land ratio (BTL)(Van de Voorde et al., 2009). This feature is often complemented with height information and volumetric descriptors when three-dimensional data is available. Yoshida and Omae (2005) and Yu et al. (2010) define descriptor sets with a quantitative interpretation for the analysis of urban areas using LiDAR data. Vanderhaegen and Canters (2010) aim to classify urban land use by using metric descriptors in an indirect analysis based on deriving and studying the concentric and radial urban block profiles that characterise the volumetric distribution of buildings.

When urban environments are being analysed, due to the hierarchical structure of urban landscapes, it may be worthwhile considering the various aggregation levels of their elements. It has been shown that the consideration of the plot as an urban landscape analysis unit and its subsequent examination with lower and higher level aggregation units (represented by buildings and urban blocks) may provide information that is useful for a more accurate classification of land uses. Consequently, this paper aims to define and analyse context-based descriptive features for classifying land use in urban environments – using object-based image classification techniques and combining high spatial resolution imagery, LiDAR, and cartographic data. Context is described by analysing the plots at internal and external levels. At an internal level a comprehensive description of various land cover types contained inside the object is performed. The external level refers to the features of the upper units to which an object belongs. The meanings of defined feature groups, and their particular influence and contribution to classification accuracy, are studied in this paper.

2. Data and study area

The study area was defined in the city of Sagunto in the province of Valencia (Spain), as shown in Figure 1. Sagunto contains a variety of urban zones with urban industrial areas and several suburban areas. Large areas of citrus orchards and farmlands surround the city.

Imagery and LiDAR data were collected in the framework of the *Spanish Programme of Aerial Orthophotography* (PNOA), which provides periodic coverage (every two years) of very high resolution aerial orthophotography (10, 25, or 50 cm/pixel) of the entire national territory. Aerial images were acquired in June 2006 with a spatial resolution of 0.5 m/pixel and three spectral bands: infrared, red, green. The images were already orthorectified, geo-referenced, panchromatic and multi-spectral band fused, and radiometrically adjusted. LiDAR data was acquired in August 2009 with a nominal density of 0.5 points/m^2. The limits of the plots were provided by vectorial cadastral cartography at a scale of 1:1000, produced by the Spanish national land registry office (*Dirección General de Catastro)*.

3. Methodology

Urban land use classification was carried out following an object-based approach. The main steps of this approach were: class definition; sample selection; descriptive feature extraction; classification of the objects; and evaluation of the results. Objects were defined by means of cartographic boundaries derived from the cadastral geospatial database. These were exhaustively described through image derived features (i.e. spectral and texture features), three-dimensional features computed from LiDAR data, and geometrical features describing the shape of each object. In addition, a set of contextual features were defined at two levels: internal and external. Many of the features derived from both contextual levels are related to buildings, obtained using automatic building detection techniques.

3.1. Definition of classes and sample selection

The definition of urban land use classes was based on the specifications of the *Land Cover and Use Information System of Spain* (SIOSE) database, created using different criteria from different land-cover/land-use databases (urban, agricultural, forested, natural, and wetland areas). This data was generated by Spanish public administrations at a scale of 1:25,000. SIOSE divides territory in polygons that separate different environments or uses (Valcárcel et al., 2008).
The urban land use classes considered were: *historical*, *urban*, *open urban*, *detached housing*, *terraced housing* and *industrial* (Figure 2). The main characteristic of *historical* areas (Figure 2.a) is their irregularity, and that they feature long thin plots, very narrow roads, and few green zones. Buildings in this area are terraced, and grouped in compact urban blocks. *Urban* areas (Figure 2.b) represent zones designed to an urban plan, and usually developed around the

historical area. These are characterised by regular urban blocks, broad streets, and more extensive green areas than historical areas. Buildings are both commercial and residential, and attached together in compact and large urban blocks. *Open urban* zones (Figure 2.c) are planned areas composed of isolated buildings, commonly unrelated to the road network and surrounded by open and green areas. Suburban residential land uses are represented by *detached housing* (Figure 2.d) and *semi-detached/terraced housing* (Figure 2.e). The first group is composed of single family residential buildings; whereas the second group refers to semidetached or terraced houses. These constructions tend to appear in dispersed urban blocks that contain green zones. *Industrial* areas (Figure 2.f) are artificial zones populated with buildings and structures for manufacturing, transforming, repairing, storing, and distributing goods. Buildings are usually large and may be detached or attached. In addition to the urban classes, agricultural/vegetation related classes were defined into *orchards*, *bare/arable lands* and *croplands* in order to fully classify the study zone. These last two classes were finally merged in a single category. According to the internal variability of the defined classes, a total of 1309 samples were collected – distributed as shown in Table 1.

3.2. Data pre-processing

A normalised digital surface model (nDSM), i.e. the difference between the digital surface model and the digital terrain model (DTM), was generated from LiDAR data. An algorithm that eliminates points belonging to any above ground objects, such as vegetation or buildings, was used to generate the DTM, with minimum elevation points being selected in a series of progressively smaller windows. Firstly, an initial DTM was computed using the points selected. New minimum elevations were then chosen by using smaller windows that were compared with the initial DTM. The definition of a height threshold enabled the removal of ground points. This algorithm is fully described in Estornell et al. (in press).

A thresholding-based building detection approach was used. This method is founded on the establishment of two threshold values: one referring to the height, applied over the nDSM; and other referring to the presence of vegetation, defined using the normalised difference vegetation index (NDVI) image. The threshold value was determined in a semi-automatic manner by collecting samples of both classes to be differentiated. With the average and standard deviation values of both sample classes, Gaussian curves modelling their histogram were computed. The threshold value was defined as the point where both curves intersected. The binary images produced during the thresholding steps were softened using morphological opening and closing filters, and small objects were eliminated to remove noise. Finally, both binary images (vegetation and height) were intersected revealing the detected buildings. Buildings and vegetation masks were used to define several descriptive features. The building detection methodology is fully described and evaluated in Hermosilla and Ruiz (2009).

3.3. Definition of descriptive features

Visual techniques used by a photo-interpreter are based on the recognition of elements represented in images and the identification of their particular characteristics. These are related to shape, colour, texture, and also to the spatial context of the topological attributes of the internal components (spatial arrangement, land cover distribution) and the overall environment. The proposed descriptive features aim to emulate human cognition by numerically quantifying the properties of the image elements and so enable each to be distinguishable.

Descriptive features related to three different object aggregation levels were defined: object-based, internal context, and external context. Object-based features describe each object as a single entity based on several aspects that reflect the information typology used: multi-spectral, three-dimensional, geometry, etc. These features are computed using object-based image analysis FETEX 2.0 software, described in Ruiz et al. (2010). Object-based features are divided in two feature groups: image-based features (group I), and geometrical and three-dimensional features (group II). Internal context features (group III) describe an object with respect to the land cover types contained within the object (denoted as sub-objects), in this case were buildings and vegetation. External context features (group IV) characterise each object by considering the common properties of adjacent objects that when combined create an aggregation that is higher than plot level. These are termed super-objects and in urban areas these coincide with urban blocks.

Two different types of ***image-based features*** (group I) are used: spectral and textural. Spectral features provide information about the intensity values of objects in the different spectral bands. Mean, standard deviation, minimum and maximum descriptors have been computed for each object in the available bands and in the NDVI image. Textural features quantify the spatial distribution of the intensity values in the analysed objects. The following descriptive features are derived: kurtosis and skewness of the histogram; contrast, uniformity, entropy, covariance, inverse difference moment, and correlation, descriptors proposed by Haralick et al. (1973) and derived from the grey level co-occurrence matrix (GLCM), which are computed using a per-object approach (Balaguer et al., 2010); and the mean and standard deviation of the *edgeness factor* (Sutton and Hall, 1972), representing the density of the edges present in the neighbourhood of each pixel.

Group II is composed of ***geometrical and three-dimensional features***. Geometrical features describe the dimensions of the objects and their contour complexity. Area, perimeter, compactness (Bogaert et al., 2000) (see Equation (1)), shape index (see Equation (2)), and fractal dimension (Krummel et al., 1987; McGarigal and Marks, 1995) (see Equation (3)) descriptors are calculated.

$$Compactness = \frac{4 \cdot \pi \cdot Area}{Perimeter^2} \tag{1}$$

$$Shape\ Index = \frac{Perimeter}{4 \cdot \sqrt{Area}} \tag{2}$$

$$Fractal\ Dimension = 2 \cdot \frac{\log\left(\frac{Perimeter}{4}\right)}{\log(Area)} \tag{3}$$

Three-dimensional features are derived from the nDSM computed from LiDAR data. Each object is characterised by the mean, standard deviation, and maximum values of the heights. Table 2 summarises the object-based feature set computed.

Internal-context features (group III) describe an object by characterising the sub-objects contained within it. When applying the automatic building detection process explained in Section 3.2. and the vegetation mask produced in that step, two covers are considered: buildings and vegetation. Buildings correspond to basic elements of urban areas, and their characteristics shape our perception of the various urban morphological areas. Bi-dimensional and three-dimensional features describing the buildings inside each object were computed. Bi-dimensional features refer to built-up surface and built-up percentages in an object. This feature – usually referred to as building coverage ratio (BCR) or sealed surface – has been often used in literature (Yoshida and Omae, 2005; Van de Voorde et al., 2009; Yu et al., 2010), and is computed as described in Equation (4):

$$BCR = \frac{A_{Building}}{A_{Object}} \cdot 100 \qquad (4)$$

where $A_{Building}$ is the built-up area, and A_{Object} is the surface of the considered object. Building sub-objects were also characterised using a set of three-dimensional features describing their height using mean, standard deviation, and maximum values from nDSM.

The presence and density of vegetation is strongly related to the different urban areas. Analogously to Equation (4), the percentage of surface covered by vegetation within an object is defined. Additionally, statistical descriptors (mean and standard deviation) are computed to describe height and photosynthetic development of sub-objects identified as vegetation from nDSM and NDVI, respectively.

The ***external-context features*** (group IV) provide information about the properties of the super-object created by merging adjacent objects, and these produce new entities with a higher aggregation level (corresponding to urban blocks in urban areas). External context is described by considering the spatial relationships of adjacent objects by means of building-based, vegetation-based, geometrical and adjacency features.

Adjacency between objects was characterised using *graph theory*, based on the study of graphs, or mathematical structures used to model pairwise relations between objects from a collection. *Graph theory* (Laurini and Thompson, 1992; Almeida et al, 2007) has been described as an extremely valuable and efficient tool in storing and describing the spatial structure of geographical entities and their spatial arrangement. This theory was introduced for image classification purposes by Barnsley and Barr (1997), to describe the spatial relationship of adjacency – corresponding with edges in the graph – between geographical objects represented by vertices. To quantify the adjacency relationships between objects, several features were defined: the number of correspondences with surrounding objects; the mean distance of these adjacencies; and the standard deviation value of the distances between adjacent objects. These features are closely related to both object and super-object dimensions (Figure 3) and provide information about the spatial distribution of objects (plots) inside the super-object (urban block) by analysing the distances and variability of the edges.

According to Yoshida and Omae (2005), the shape, size, and number of buildings per block (often related to their socio-economic function) determine area and volume for an urban block. This implies the possibility that the land use of an urban block may be indicated by the quantitative observations related to the buildings present in it. These descriptors are often mentioned as urban morphology features. Super-objects are characterised with the built-up area and the BCR. The heights of the buildings contained in an urban block are described using the mean and standard deviation values. Features related with the volumetric information of buildings have also been computed. The volume of a building is given by Equation (5) (Yu et al., 2010):

$$V = \sum_{i=1}^{n} h_i r^2 \tag{5}$$

where r is the spatial resolution and h_i is the relative height obtained from nDSM for the pixel i in a surface detected as a building, composed of n pixels. Using the volume of each building, the mean volume is computed as the total volume of buildings divided by the number of buildings contained in an urban block as shown in Equation (6):

$$V_m = \frac{\sum_{i=1}^{n} V_i}{n} \tag{6}$$

where V_i is the volume of the building i and n the building total in the analysed super-object.

Equivalently to the internal context features, vegetation is characterised using the vegetation covered ratio, mean, and standard deviation values of nDSM and NDVI, from the vegetation detected within a super-object.

The geometrical properties of the polygons produced with the super-object are described using area, perimeter, compactness, shape index, and fractal dimension features. Table 3 summarises the internal and external feature set computed.

Figure 4 shows examples of the typical differences in building and vegetation coverage for the different urban classes considered. In general, buildings in the *historical* and *urban* classes include plots and urban blocks with small inner light wells. The *open urban* class usually has only a portion of built-up area in a plot or urban block; while a higher variability is found in the *industrial* class. The *detached housing* class tends to include several small buildings distributed in variable size plots and large urban blocks. The *semi-detached/terraced housing* class has larger built-up areas in small plots and urban blocks. Suburban residential areas show abundant vegetation. Little vegetation is found in industrial areas and in other urban classes.

At both internal and external levels, height (Figure 5) and volume are strongly related to the type of buildings. *Historical* class is mainly characterised by the irregularity of building heights and dimensions. *Urban* class contains taller buildings with more uniformity, larger dimensions, and higher volume values. *Open urban* class buildings have a diversity of dimensions and heights, but these are regular and lack internal variability. Individual *semi-detached/terraced housing* buildings normally have smaller dimensions, but taller buildings than the *detached housing* class. *Semi-detached/terraced housing* constructions are attached and so produce elongated building rows with high unitary volumes at the urban block level. *Industrial* class buildings are characterised by medium and constant heights and large dimensions that produce elevated unitary volume values. Building dimensions shape the geometrical aspect of urban blocks. *Historical* blocks are characterised by the extreme irregularity of their contours and by small and medium surface areas. In contrast, the *urban* class blocks show regular shapes with an abundance of perpendicular junctions that are similar to the *open urban* block. This class reveals especially variable dimensions. The *industrial* class blocks contain regular contours based on squared shapes and very large dimensions. Suburban single-family blocks also present a variety of sizes. *Detached housing* blocks are commonly square, while *semi-detached/terraced housing* reveals significantly elongated rectangular shapes.

3.4. Classification

To analyse the effect of using contextual features to classify urban land uses, four classification tests were applied. In the first test, a description of the objects was merely based on the image-based features (group I). In the second test, the geometrical and three-dimensional features (group II) were combined with the feature group I. In the third test, objects were described with features from group I and II, and combined with the defined internal context features (group III). In the final test, all the descriptive feature groups were combined by adding the external context features (group IV).

Objects were classified by applying the decision-trees obtained using the training samples. A decision-tree is a set of conditions organised in a hierarchical structure in such a way that the class assigned to an object can be determined following the conditions that are fulfilled from the tree roots (the initial dataset) to any of its leaves (the assigned class). The algorithm employed in this study was C5.0. The process of building a decision-tree begins by dividing the collection of training samples using mutually exclusive conditions. This algorithm searches partitions to obtain purer data subgroups, which are less mixed than the previous group from where they were derived. For each possible division of the initial data group, the degree of impurity of the new subgroups is computed; and the condition that gives the lowest degree of degree is chosen. This is iterated until the original data is divided into homogeneous subgroups by using the gain ratio as a splitting criterion until all the elements in a subgroup belong to the same class, or a stopping condition is fulfilled (Quinlan, 1993).

The boosting multi-classifier method was used. This methodology is based on the assignment of weights to the training samples. The greater the weight of a sample, then the greater its influence on the classifier. After each tree construction, the weight vector is adjusted to show the model performance. In this way, samples erroneously classified retain their weights, whereas the weights of correctly classified samples are decreased. Thus, the model obtained in the following iteration gives more relevance to the previously wrongly classified samples.

3.5. Methods for evaluation of feature influence and classification

The influence and usefulness of the proposed descriptive features for the particular classification problem was assessed using forward stepwise linear discriminant analysis (LDA). In this method, all variables are reviewed and evaluated at each step to determine which will contribute most to the discrimination between classes. That variable is included in the model and the process is iterated.

The evaluation of the four classifications performed is based on the analysis of the confusion matrix (Congalton, 1991), by comparing the class assigned to each evaluation sample with the information contained in the reference database. The overall accuracies of the classifications were computed, as well as the producer and user accuracies for each class (which respectively reveal the errors of omission and commission). In addition, a specific confusion index was defined to quantify the confusion between a pair of classes, computed as the sum of their mutual errors divided by the total objects from that pair of classes. Confusion index value ranges between 0 (absence of per-class-pair errors) and 1 (all the objects of both considered classes are misclassified).

To improve the efficiency of the number of samples, the leave-one-out cross-validation technique was employed. This method uses a single observation from the original sample set as validation data, and the remaining observations as training data. This is iterated until each observation in the sample set is used once as validation data.

4. Results and discussion

4.1. Feature analysis

The predicted overall classification accuracy evolution for the 25 first variables included in the LDA model, considering descriptive features from all the groups defined, is shown in Figure 7. Several variables coming from the four different groups considered are selected among the most relevant features included in the model: image-based features (*IDM*, *Entropy*, *MeanG*, *MeanIR*, *StdevNDVI*, *MinR*, *StdevIB*, *MinG*); geometrical and three-dimensional features (*Perim_O*, *Fractal_O*); internal-context features (*VCR*, *MeanH_B*, *BCR*); and external-context features (*BCR_SO*, *Volume*). This illustrates their complementary nature, as well as the possibility of increasing the efficiency of the classification in terms of accuracy and reducing the number of variables by using only a selected and highly discriminant group of features. See Table 2 and Table 3 for feature code description.

The distinctive aspects of the different urban classes that enable their discrimination – analogously to the human interpretation process – are numerically expressed by means of the defined features. In Figure 6, four examples of the distribution of classes according to the ranges of values of different context-based descriptive features are shown. Thus, when analysing the per plot distribution of BCR and VCR feature values (Figure 6.a and Figure 6.b), the *historical* and *urban* classes reveal buildings covering almost the entire area of their plots with low vegetation coverage. In contrast, *semi-detached/terraced housing* and, particularly, *detached housing* had less built-up zones and more vegetation. The *industrial* class showed a high variability for BCR feature values and reduced values of VCR features. At the urban block level, significant differences between urban classes were also found. As seen in Figure 6.c, the *detached housing* class had the lowest values for mean volume of buildings, and *semi-detached/terraced housing* reached slightly higher values. The remaining classes generally showed high volumes. *Urban* and *historical* classes (Figure 6.d) were located in small urban blocks, whereas the *industrial* class usually appeared in the largest urban blocks. The suburban classes (*detached housing* and *semi-detached/terraced housing*) were distributed in urban blocks with highly variable sizes.

4.2. Urban land use classification

As shown in Table 4, the progressive addition of feature groups increases the classification accuracy, indicating the complementary nature of these feature groups. The lowest values were obtained when only image-based object features (group I) were considered. Three-dimensional data offered valuable information. Internal and external context features also produce noticeable increases in accuracy.

Per class user and producer accuracies for the various feature group combinations are shown in Figure 8. Analogously to the overall accuracy values, the least accurate performances were achieved when image-based object features were considered. The combination of different feature groups increases accuracy values. This increase was especially irrelevant in the case of the agricultural classes: *bare soil/arable and croplands* and *orchards*, which performed well when only considering feature group I. Among the urban classes, the highest accuracy result with the lowest number of descriptive features was obtained in the *industrial* class, attributable to the homogeneity of textures and the particular spectral response shown by this type of construction. Due to the high initial accuracy values, the subsequent inclusion of feature groups had little impact in this class, producing a slight land-use accuracy increase when adding external context features. Figure 10 a shows a classification result example in an industrial area. This figure shows that even though all the objects included in a super-object were characterised with identical features in group IV, their different classes were correctly assigned. In contrast, the lowest user and producer accuracies when considering feature group I were obtained in the *open urban* class, as it was confused with the *urban* class. The successive addition of the descriptive feature groups significantly enhanced the accuracy values for this class.

The pairs of classes *detached housing* and *semi-detached/terraced housing*, and *historical* and *urban* mutually revealed high levels of confusion due to their spectral similarities and the absence of a framework for contextualising differences. The per-class-pair confusion index (see Figure 9) noticeably decreased when three-dimensional and geometrical based features were considered, because plots contained in the *semi-detached/terraced housing* class are characterised by smaller dimensions and taller buildings than *detached housing* plots. The successive addition of contextual features –especially when these refer to the external context – reduces the confusion between both classes up to a value of 0.04. An example of the classification result of a suburban area with predominance of *detached housing* and *semi-detached/terraced housing* classes is shown in Figure 10.b.

Historical and *urban* classes also show an elevated initial per-class-pair confusion index – which was remarkably reduced as three-dimensional and contextual features were used in the classification. Objects belonging to both classes presented similar object level features, their main differences being found at super-object level. Super-objects of the *urban* class usually belong to a previously planned and ordered environment. Urban blocks of historical areas have irregular and complex shapes, as a consequence of a sporadic and unplanned growth over time. Figure 10.b graphically shows how *historical* and *urban* classes are in general efficiently discriminated, in spite of some minor errors produced in isolated objects, which may be decreased by applying a further analysis of objects that are isolated among different classes.

5. Conclusions

A set of context-based descriptive features for urban environment land-use classification is analysed in this paper. These features are computed from high spatial resolution imagery and airborne LiDAR data, and aim to imitate human cognition though the numerical quantification of the discrimant properties of image elements. The use of object-based image analysis facilitates the combination of information from different data sources and enables the multi-scale analysis of the images. By combining different data and aggregation levels, image objects are described in greater depth than in the pixel approach. This is true for diverse aspects of the objects (spectral response, geometry, altimetry, properties of internal elements, properties of the container object, etc). The results of the classification tests performed show that internal and external context features suitably complement the image-derived features, improving the classification accuracy values of urban classes – especially between classes that show similarities in their image-based and three-dimensional features. The proposed methodology, based on automated descriptive feature extraction from LiDAR data and images, is applicable for mapping cities, urban landscape characterisation and management, and updating geospatial databases, providing new tools to increase the frequency and efficiency of urban studies.

List of References

1. Alberti, M., Waddell, P., 2000. An integrated urban development and ecological simulation model. Integrated Assessment 1, 215–227.
2. Almeida J.P., Morley, J.G., Dowman, I.J., 2007.Graph theory in higher order topological analysis of urban scenes. Comput. Environ. Urban Syst 31 (4), 426-440.
3. Anders, K.-H., Sester, M., Fritsch, D. 1999. Analysis of settlement structures by graph-based clustering. Semantische Modellierung, SMATI 99, Munich, Germany, pp. 41–49.
4. Balaguer, A., Ruiz, L.A., Hermosilla, T., Recio, J.A., 2010. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification, Comput. Geosci 36 (2), 231-240.
5. Barnsley, M. J., Barr, S. L., 2000. Monitoring urban land use by Earth Observation. Surv. Geophys. 21, 269–289.
6. Barnsley, M., Barr, S., 1997. A graph-based structural pattern recognition system to infer land use from fine spatial resolution land cover data. Comput. Environ. Urban Syst 21, 209-225.
7. Barnsley, M.J., Barr, S.L., Sadler, G.J., 1991. Spatial re-classification of remote sensed images for urban land use monitoring. Proceedings of Spatial Data 2000, 17-20 September 1991, Oxford, U.K., 106-117.
8. Barr, S., Barnsley, M., 1998. A syntactic pattern recognition paradigm for the derivation of second-order thematic information from remotely-sensed image, in P. Atkinson and N. Tate (eds), Advances in Remote Sensing and GIS Analysis, John Wiley and Sons, Chichester.
9. Bauer, T., Steinnocher, K., 2001. Per-parcel land use classification in urban areas applying a rule-based technique. GeoBIT/GIS 6, 24-27
10. Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. 65(1), 2-16.
11. Boffet, A. and Rocca-Serra, S.; 2001: Identification of spatial structures within urban blocks for town characterization. Proceedings of 20th International Cartographic Conference, Beijing, China, pp. 1974–1983.
12. Boffet, A., Coquerel, C., 2000. Urban classification for generalization orchestration. The International Archives of Photogrammetry and Remote Sensing 38 (B4), 132-139.
13. Bogaert, J., Rousseau, R., Hecke, P. V., Impens, I., 2000. Alternative area-perimeter ratios for measurement of 2D shape compactness of habitats. App. Math. Comput. 111 (1), 71-85.
14. Breuste, J., Feldmann, H., Uhlmann, O., 1998. Urban Ecology. Springer, Berlin.
15. Burghardt, D., Steiniger, S., 2005. Usage of principal component analysis in the process of automated generalisation. Proceedings of XXII International Cartographic Conference, ICC2005, A Coruña, Spain, 11-16 July 2005, 12p.
16. Congalton, R., 1991. A review of assessing the accuracy of classications of remotely sensed data. Remote Sens. Environ. 37 (1), 35-46.
17. Estornell, J., Ruiz, L.A., Velázquez-Martí, B., Hermosilla, T., In press. Analysis of the factors affecting LiDAR DTM accuracy in a steep shrub area. International Journal of Digital Earth. http://dx.doi.org/10.1080/17538947.2010.533201

18. Eyton, J.R., 1993. Urban land use classification and modeling using cover-type frequencies. Appl. Geogr. 13, 111-121.
19. Gastellu-Etchegorry, J., 1990. An assessment of SPOT XS and Landsat MSS data for digital classification of near-urban land cover'. ISPRS J. Photogramm. 11, 225–235.
20. Haralick, R.M., Shanmugan, K., Dinstein, I., 1973. Texture features for image classification. IEEE T. Syst. Man Cyb. 3, 610-621.
21. Hermosilla, T., Ruiz, L.A., 2009. Detección automática de edificios combinando imágenes de satélite y datos lidar (Automatic building detection combining satellite imagery and lidar data). Semana Geomática, 2-4 February, Barcelona, Spain. [in Spanish]
22. Herold, M., Couclelis, H., Clarke, K.C., 2005. The role of spatial metrics in the analysis and modelling of urban land use change. Comput. Environ. Urban Syst. 29 (4), 369-399.
23. Herold, M., Liu, X., Clarke, K.C., 2003. Spatial metrics and image texture for mapping urban land use. Photogramm. Eng. Rem. S.69 (9), 991-1001.
24. Hussain, M., Davies, C., Barr, R., 2007. Classifying buildings automatically: a methodology. Proceedings of the Geographical Information Science Research UK Conference, GISUK, 11-13 April, County Kildare, Ireland.
25. Krummel, J. R., Gardner, R. H., Sugihara, G., O'Neill, V., Coleman, P. R., 1987. Landscape patterns in a disturbed environment. OIKOS 48 (3), 321-324.
26. Laskari, S., Hanna, S., Derix, C., 2008. Urban identity through quantifiable spatial attributes: coherence and dispersion of local identity through the automated comparative analysis of building block plans. Proceedings of the Third International Conference on Design Computing and Cognition, Dordrecht, The Netherlands, 615-634.
27. Laurini, R., & Thompson, D. (1992). Fundamentals of spatial information (Vol. 5). London, UK: Academic Press.
28. McGarigal, K., Marks, B.J., 1995. FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351; Pacific Northwest Research Station, USDA-Forest Service, Portland.
29. Novack, T., Kux, H.J.H., Feitosa, R.Q., Costa, G. A., 2010. Per block urban land use interpretation using optical VHR data and the knowledge-based system interimage. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (4/C7), 6p.
30. Pickett, S.T.A., Cadenasso, M.L., Grove, J.M., Nilon, C.H., Pouyat, R.V., Zipperer, W.C., Costanza, R., 2001. Urban ecological systems: linking terrestrial ecological, physical, and socioeconomic components of metropolitan areas. The Annual Review of Ecology, Evolution, and Systematics 32 (1), 27-57.
31. Quinlan, J.R., 1993. C4.5. Programs for machine learning. San Mateo: Morgan Kaufmann.
32. Ruiz, L.A., Recio, J.A., Fernández-Sarría, A., Hermosilla, T., 2010. A tool for object descriptive feature extraction: application to image classification and map updating. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (4/C7), 6p.
33. Sadler, G.J., Barnsley, M.J., Barr, S.L., 1991. Information extraction from remotely sensed images for urban land analysis. Proceedings of the 2nd European conference on Geographical Information Systems EGIS'91, Brussels, Belgium, April, EGIS Foundation, Utrecht, pp. 955-964.

34. Steiniger, S., Taillandier, P., Weibel, R., 2010. Utilising urban context recognition and machine learning to improve the generalisation of buildings. Int. J. Geogr. Inf. Sci. 24 (2), 253-282.
35. Sutton, R.N., Hall, E.L., 1972. Texture measures for automatic classification of pulmonary disease. IEEE T. Comput. 21 (7), 667-676.
36. Thomson, M. K., Béra, R., 2008. A methodology for inferring higher level semantic information from spatial databases. Proceedings of the Geographical Information Science Research UK Conference, pp. 268–274.
37. U.N. Secretariat (2007) World population prospects, the 2006 revision. New York, The Department of Economic and Social Affairs, United Nations.
38. United Nations, Department of Economic and Social Affairs, Population Division: World Urbanization Prospects, the 2009 Revision: Press Release. New York, 2010
39. Valcarcel, N., Villa, G., Arozarena, A., García-Asensio, L., Caballero, M.E., Porcuna, A., Domenech, E., Peces, J.J., 2008. SIOSE, a successful test bench towards harmonization and integration of land cover/land use information as environmental referente data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (B8), 1159-1164.
40. Van de Voorde, T., Van der Kwast, J., Engelen, G., Binard, M., Cornet, Y., Canters, F., 2009. Quantifying intra-urban morphology of the Greater Dublin area with spatial metrics derived from medium resolution remote sensing data. Proceedings of the 7th International Urban Remote Sensing Conference, IEEE Geoscience and Remote Sensing Society, 20-22 May, Shanghai, China.
41. Vanderhaegen, S., Canters, F., 2010. Developing urban metrics to describe the morphology of urban areas at block level. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (4/C7), 6p.
42. Welch, R., 1982. Spatial resolution requirements for urban studies. ISPRS J. Photogramm. 3 (2), 139–146.
43. Whitford, V., Ennos, A.R., Handley, J.F., 2001. City form and natural process - indicators for the ecological performance of urban areas and their application to Merseyside, UK. Landscape Urban Plan. 57 (2), 91-103.
44. Wijnant, J., Steenberghen, T., 2004. Per-parcel classification of urban ikonos imagery. Proceedings of 7th AGILE conference on geographic information science, Heraklion, Greece, pp. 447–455.
45. Yoshida, H., Omae, M., 2005. An approach for analysis of urban morphology: methods to derive morphological properties of city blocks by using an urban landscape model and their interpretations. Comput. Environ. Urban Syst., 29 (2), 223-247.
46. Yu, B., Liu, H., Wu, J., Hu, Y., Zhang, L., 2010. Automated derivation of urban building density information using airborne LiDAR data and object-based method. Landscape Urban Plan 98 (3-4), 210-219.
47. Zhan, Q., Molenaar, M., Gorte, B., 2000. Urban land use classes with fuzzy membership and classification based on integration of remote sensing and GIS. International Archives of Photogrammetry and Remote Sensing 33 (B7), 1751-1759.
48. Zhan, Q., Molenaar, M., Tempfli, K., 2002a. Finding spatial units for land use classification based on hierarchical image objects. International Archives of Photogrammetry and Remote Sensing 34 (4), 6 p.

49. Zhan, Q., Molenaar, M., Tempfli, K., 2002b. Hierarchical image object based structural analysis toward urban land use classification using HR imagery and airborne lidar data. Proceedings of the 3rd Symposium on Remote Sensing of Urban Areas, 11–13 June, Istanbul, Turkey, pp. 251–258.
50. Zhang, L., Wu, J., Zhen, Y., Shu, J., 2004. A GIS-based gradient analysis of urban landscape pattern of Shanghai metropolitan area, China. Landscape Urban Plan. 69 (1), 1-16.

List of Tables

Table 1. Number of samples selected per class.

Class	Number of samples
Historical	170
Urban	244
Open urban	103
Detached housing	121
Semi-detached/terraced housing	161
Industrial	115
Orchards	157
Bare/arable and croplands	238
Total	1309

Table 2. Description and codification of image based and geometrical and three-dimensional object features.

Group I: image-based features	
	Spectral (for each band and NDVI image)
	Mean (*MeanIR*, *MeanR*, *MeanG*, *MeanNDVI*) Standard deviation (*StdevIR*, *StdevR*, *StdevG*, *StdevNDVI*) Minimum (*MinIR*, *MinR*, *MinG*, *MinNDVI*) Maximum (*MaxIR*, *MaxR*, *MaxG*, *MaxNDVI*)
	Texture
	Mean edgeness factor (*MeanEDG*) Standard deviation of edgeness factor (*StdevEDG*) Skewness Kurtosis Uniformity Entropy Contrast Inverse difference moment (*IDM*) Covariance Correlation
Group II: geometrical and three-dimensional features	
	Geometrical
	Compactness (*Compac_O*) Shape index (*Shape_O*) Fractal dimension (*Fractal_O*) Area (*Area_O*) Perimeter (*Perim_O*)
	Three-dimensional
	Height mean (*MeanH*) Height standard deviation (*StdevH*) Height maximum (*MaxH*)

Table 3. Internal and external context descriptive features compilation.

Group III: internal context features	
	Building related
	Height mean (*MeanH_B*) Height standard deviation (*StdevH_B*) Height maximum (*MaxH_B*) Building covered area (*BCA*) Building covered ratio (*BCR*)
	Vegetation related
	Height mean (*MeanH_V*) Height standard deviation (*StdevH_V*) NDVI mean (*meanNDVI_V*) NDVI standard deviation (*Stdev_NDVI_V*) Vegetation covered ratio (*VCR*)
Group IV: external context features	
	Connectivity
	Number of adjacencies (*NAdj*) Mean distance (*MeanDist*) Standard deviation of distance (*StdevDist*)
	Urban morphology
	Mean volume (*Volume_SO*) Building covered ratio (*BCR_SO*) Building covered area (*BCA_SO*) Mean height of buildings (*MeanH_SO*) Standard deviation of building height (*StdevH_SO*)
	Vegetation related
	Height mean (*MeanH_VSO*) Height standard deviation (*StdevH_VSO*) NDVI mean (*meanNDVI_VSO*) NDVI standard deviation (*StdevNDVI_VSO*) Vegetation covered ratio (*VCR_SO*)
	Geometric
	Compactness (*Compac_SO*) Shape index (*Shape_SO*) Fractal dimension (*Fractal_SO*) Area (*Area_SO*) Perimeter (*Perim_SO*)

Table 4. Overall classification accuracy values
when successively combining descriptive feature groups.

Feature groups	Overall accuracy
Group I	72.9 %
Groups I+II	82.7 %
Groups I+II+III	87.1 %
Groups I+II+III+IV	91.8 %

List of Figures

Figure

Figure 1. Location of the study area (Sagunto).

Figure

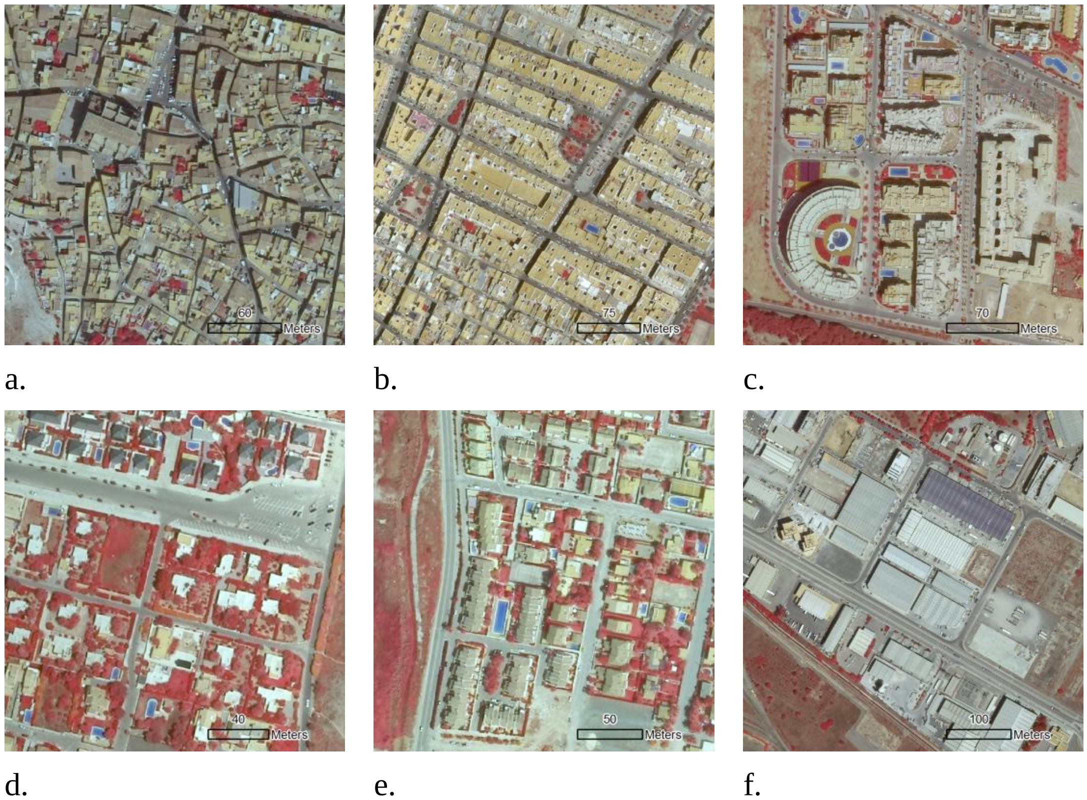
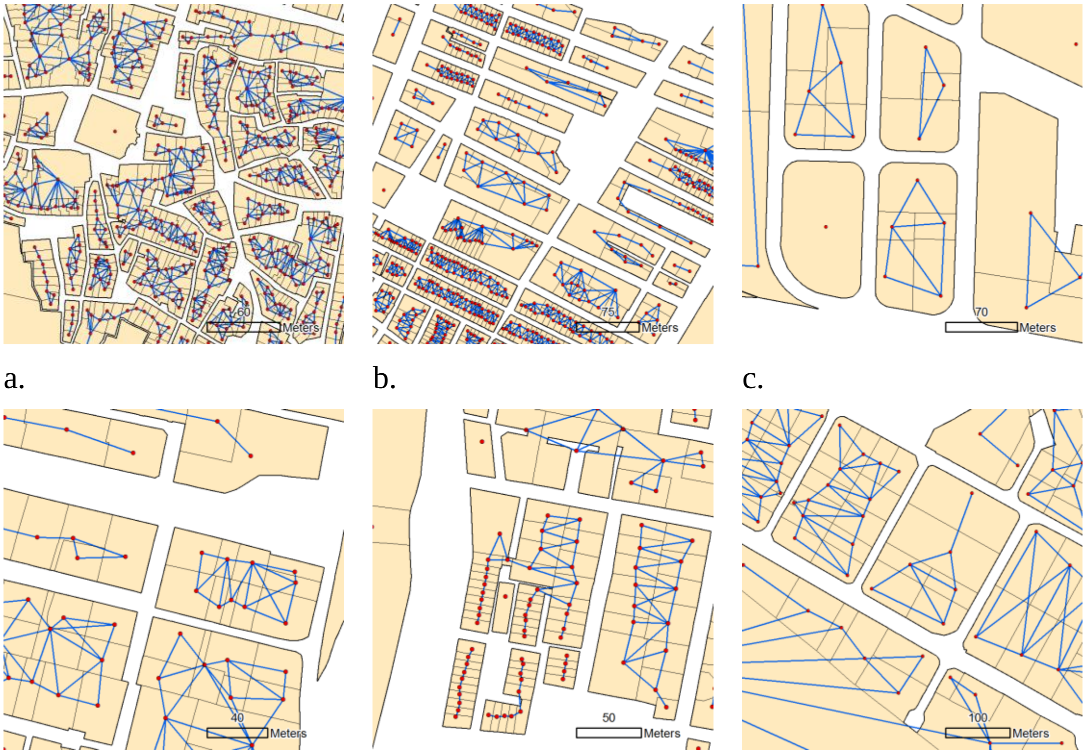


Figure 2. Examples of the urban classes defined in colour-infrared composition: a. *historical*; b. *urban*; c. *open urban*, d. *detached housing*; e. *semi-detached/terraced housing*; and. f. *industrial*.

Figure



a. b. c.

d. e. f.

Figure 3. Examples of adjacency relations derived using graph theory for the urban classes defined: a. *historical*; b. *urban*; c. *open urban*, d. *detached housing*; e. *semi-detached/terraced housing*; and f. *industrial*.

Figure

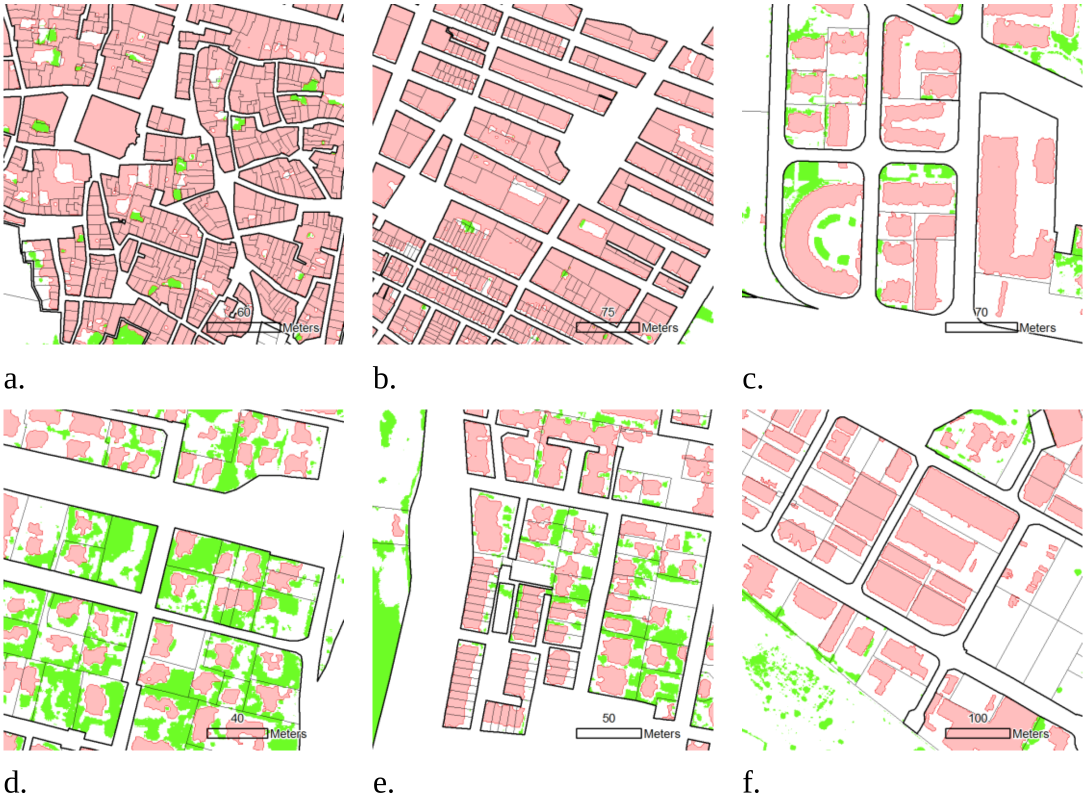


Figure 4. Examples of detected building (in pink) and vegetation (in green) for the defined urban classes: a. *historical*; b. *urban*; c. *open urban*, d. *Detached housing*; e. *semi-detached/terraced housing*; f. *industrial*.

Figure

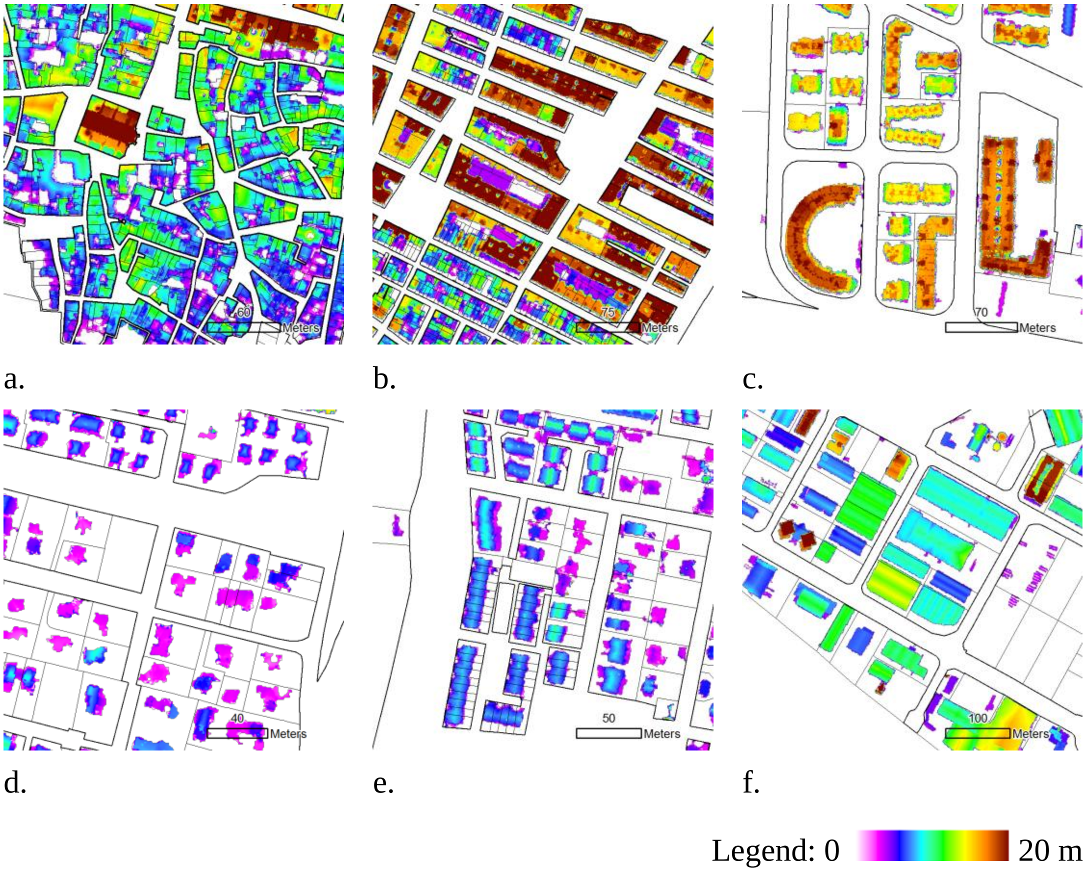


Figure 5. Examples of building height distribution for the urban classes defined: a. *historical*; b. *urban*; c. *open urban*, d. *detached housing*; e. *semi-detached/terraced housing*; and. f. *industrial*.

Figure

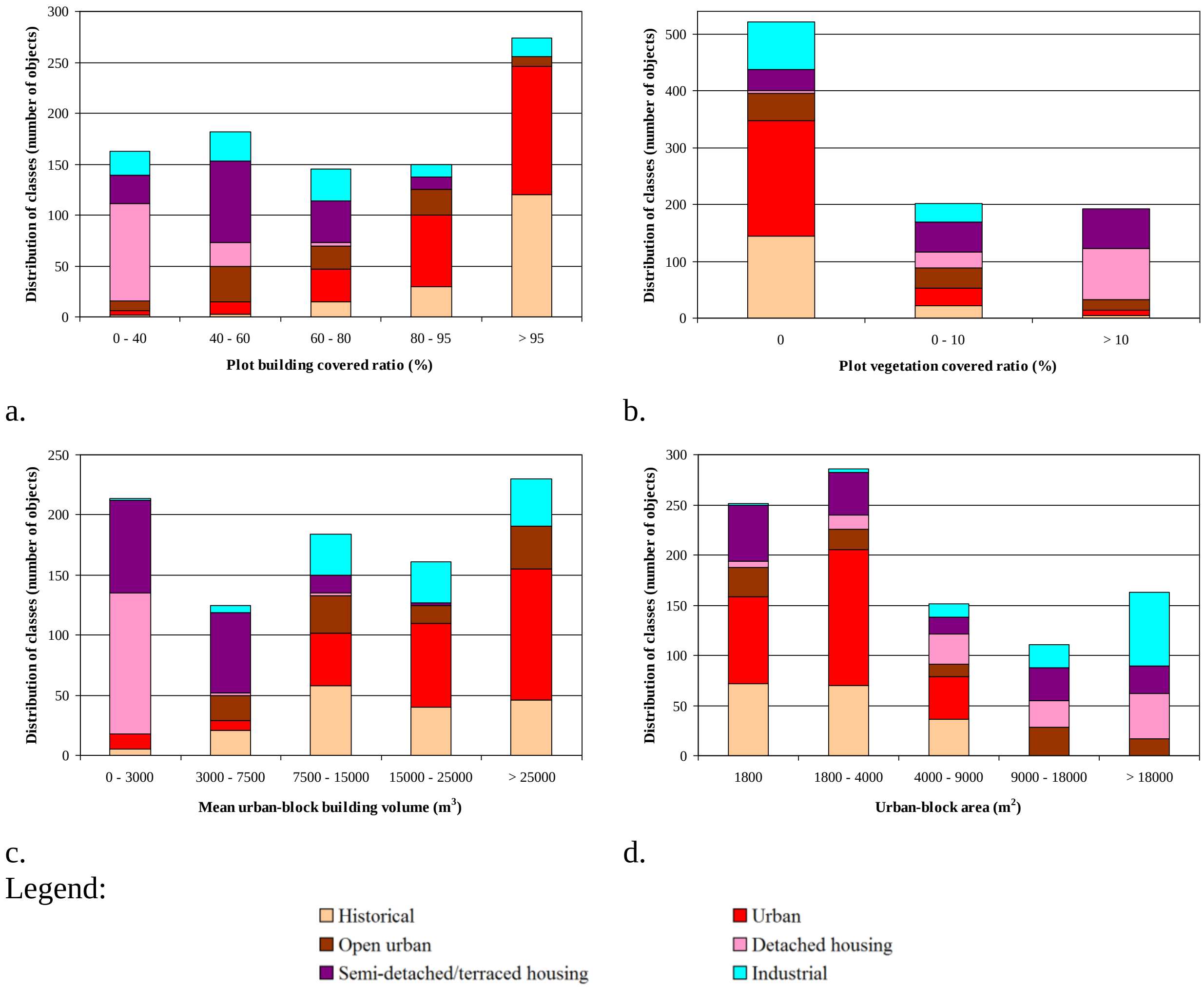


Figure 6. Distribution of classes according to the ranges of values of different descriptive features: (a) plot building covered ratio, (b) Plot vegetation covered ratio, (c) Mean urban-block building volume, and (d) urban-block area.

Figure

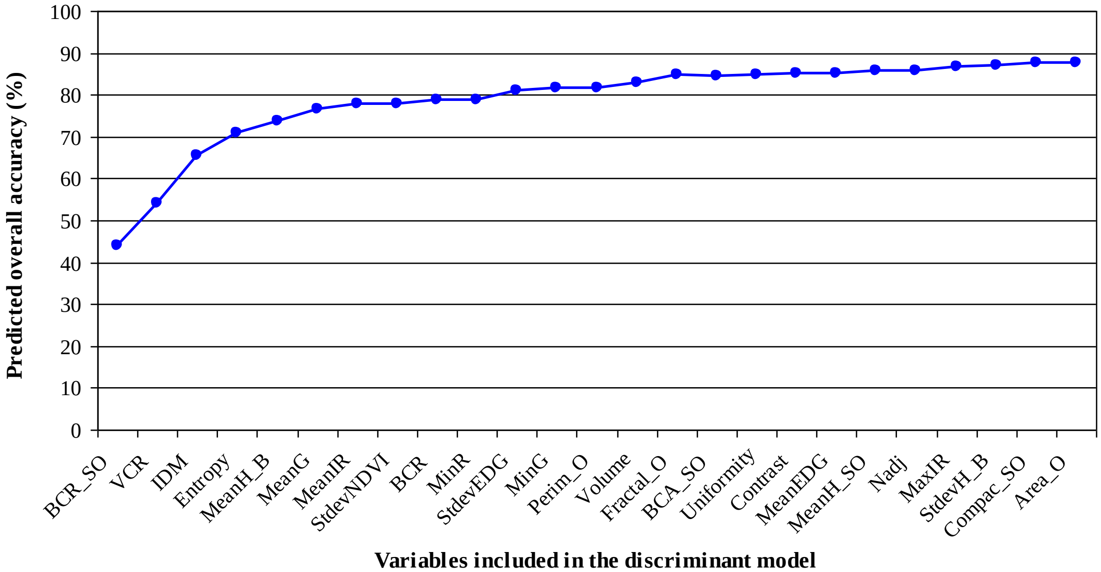


Figure 7. Predicted overall classification accuracy when the 25 first features are progressively included in the discriminant model. See Table 2 and Table 3 for feature code description.

Figure

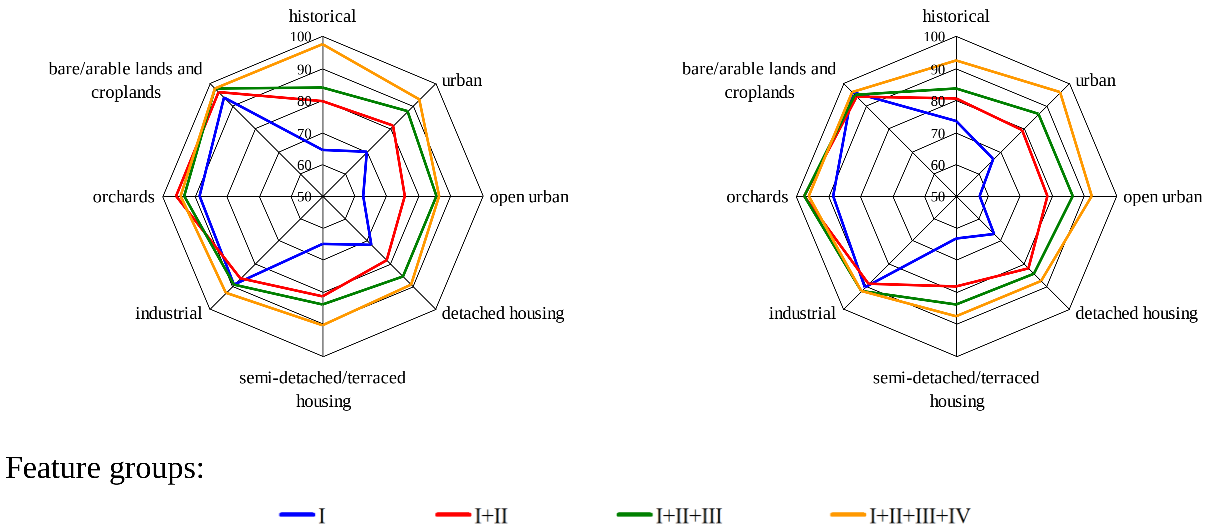


Figure 8. Per-class user (left) and producer (right) accuracies when different feature groups are combined.

Figure

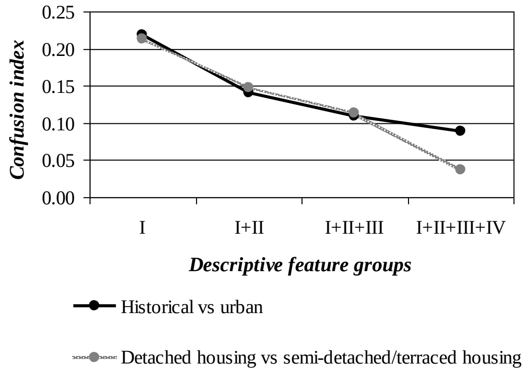
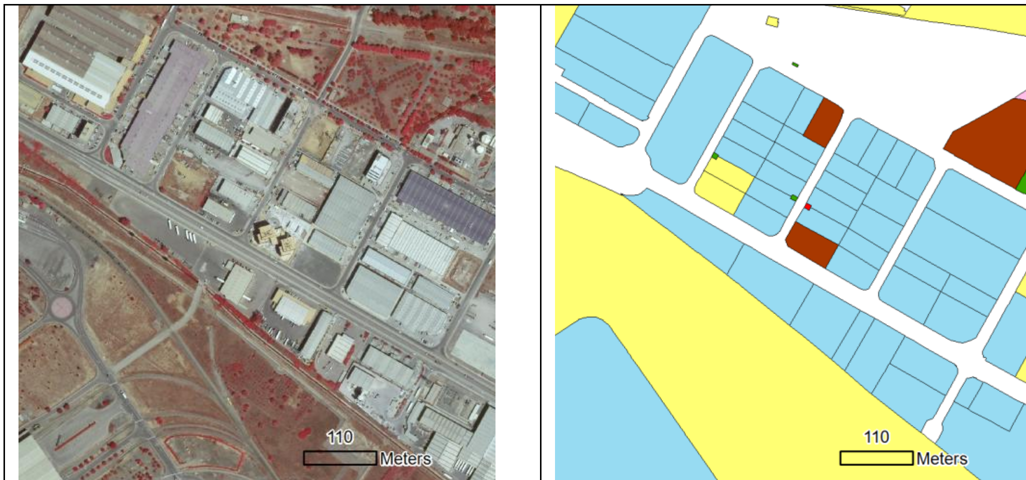
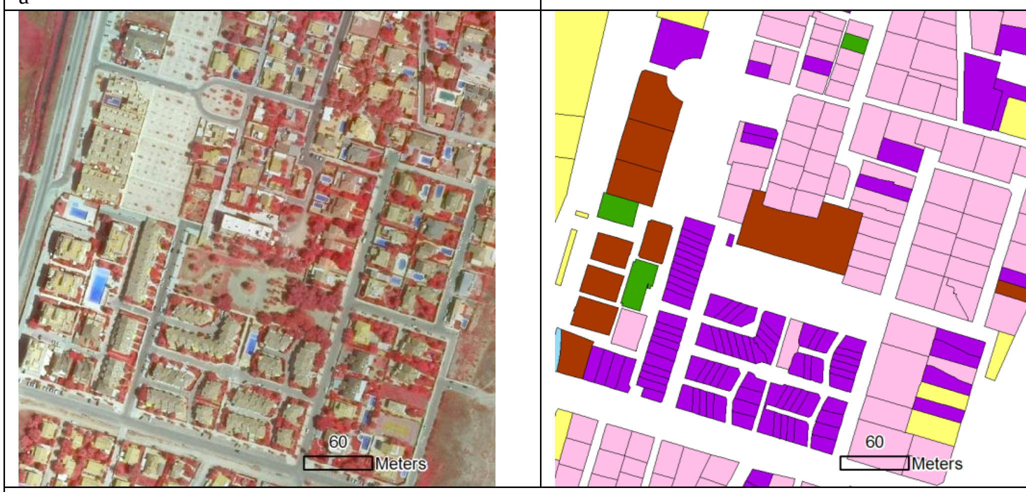


Figure 9. Per-class-pair confusion index as successive descriptive feature groups are combined in classification comparing historical vs. urban, and detached housing vs. semi-deteached/terraced housing classes.

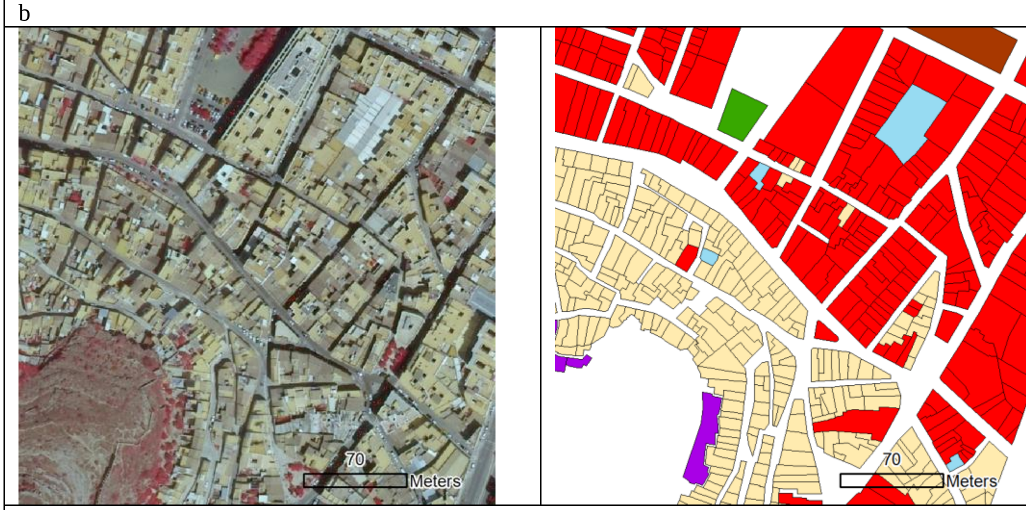
Figure



a



b



c

Legend:

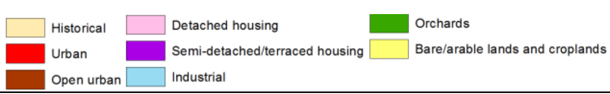


Figure 10. Three details of colour infrared images (left) and a land-use thematic map (right) derived from the classification using the most efficient set of features.