

## Urban road classification in geometrically integrated high-resolution RGB aerial and laser-derived images using the artificial neural network classification method

Tatiana Sussel Gonçalves Mendes & Aluir Porfírio Dal Poz

To cite this article: Tatiana Sussel Gonçalves Mendes & Aluir Porfírio Dal Poz (2019) Urban road classification in geometrically integrated high-resolution RGB aerial and laser-derived images using the artificial neural network classification method, International Journal of Image and Data Fusion, 10:1, 58-78, DOI: [10.1080/19479832.2018.1469547](https://doi.org/10.1080/19479832.2018.1469547)

To link to this article: <https://doi.org/10.1080/19479832.2018.1469547>



Published online: 02 May 2018.



Submit your article to this journal [↗](#)



Article views: 65



View Crossmark data [↗](#)



ARTICLE



# Urban road classification in geometrically integrated high-resolution RGB aerial and laser-derived images using the artificial neural network classification method

Tatiana Sussel Gonçalves Mendes <sup>a</sup> and Aluir Porfírio Dal Poz <sup>b</sup>

<sup>a</sup>Department of Environmental Engineering, São Paulo State University (Unesp), São José dos Campos, Brazil; <sup>b</sup>Department of Cartography, São Paulo State University (Unesp), Presidente Prudente, Brazil

## ABSTRACT

The problem of automated urban road network extraction is extremely complex because roads in urban scenes strongly interact with other objects. This problem can be simplified if road regions are first isolated using a classification procedure. The isolated road regions can be posteriorly used in tasks of refinement and reconstruction of the road network. This article addresses only the problem of road regions' detection using Artificial Neural Network as classification method. However, in urban areas, the use of spectral data alone commonly leads to the confusion of the road class with other classes in RGB images, such as building roofs and concrete, because these objects may present similar spectral characteristics. To overcome this problem, it is proposed the integration of a high-resolution RGB aerial image with laser-derived images. The classification results showed that the integration of the geometric (height) and radiometric (laser pulse intensity) laser data significantly improved the classification accuracy, also contributing for the better detection of road pixel. The laser intensity data help to overcome the effects of road obstructions caused by shadows and trees. On the other hand, the laser height data help to separate the aboveground objects from those on the ground level.

## ARTICLE HISTORY

Received 2 October 2017  
Accepted 23 April 2018

## KEYWORDS

Artificial neural network;  
airborne laser data; RGB  
aerial image

## 1. Introduction

The extraction of cartographic objects from urban aerial or satellite images has several limitations, mainly because urban scenes generally exhibit objects that are arranged in an agglomerated manner and are directly related to one another. For example, roads are disturbed by obstructions caused by vehicles, trees, buildings and shadows casted by objects along the margins of these roads. To simplify this problem, the regions that correspond to roads should be first detected. This strategy allows the extraction method to focus on road regions, thereby improving the reliability of the results and drastically reducing the search area and, consequently, the computational effort. In this context, classification methods can be used to effectively detect road regions in images taken from complex urban scenes.

**CONTACT** Tatiana Sussel Gonçalves Mendes [tatiana.mendes@unesp.br](mailto:tatiana.mendes@unesp.br) Department of Environmental Engineering, São Paulo State University, Rod. Presidente Dutra, Km 137, 8 Eugenio de Melo, São José do Campos, 12247-004, Brazil

Various methods have been proposed for classification of remote sensing images taken from complex urban regions (Haala and Walter 1999, Benediktsson *et al.* 2003, Shackelford and Davis 2003, Chanussot *et al.* 2006, Bellens *et al.* 2008, Makarau *et al.* 2011, Pálsson *et al.* 2012). However, as emphasised by Bellens *et al.* (2008), the use of spectral data alone to classify this type of data frequently results in the overlapping of information classes. A possible strategy to minimise this problem consists in integrating Airborne Laser Scanning (ALS) data and multispectral images. Both data sources are complementary in the sense that the multispectral image provides high spatial resolution and multispectral reflectance, while ALS data provide three-dimensional geometric information and also the laser pulse return intensity data (Huang *et al.* 2008, Guo *et al.* 2011), thereby improving the classification results. Zeng *et al.* (2002) improved the accuracy of the urban land-use detection by means of classification of a multispectral high-resolution Ikonos-based orthoimage, for which height information produced by ALS data was associated to each pixel. Alonso and Malpica (2010) integrated ALS elevation data with SPOT 5 medium-resolution (multispectral and panchromatic) images to detect buildings and other objects via Support Vector Machine (SVM) classifiers, resulting in an improvement of 16.78% in the overall accuracy of classification when compared to the results obtained using only image data. Salah *et al.* (2011) used a set of 26 uncorrelated feature attributes derived from original aerial images and ALS data (intensity and height data) in classification trees models for classifying building, trees, roads and ground. The results showed that using attributes generated from the multispectral aerial or intensity images resulted in lower classification accuracies. On the other hand, more accurate classification results were reached when they used attributes generated from or including height data. Huang *et al.* (2008) demonstrated the benefits of the integration of aerial imagery with ALS data developing a knowledge-based classification system (KCBS) that included discrimination rules between road, vegetation and building, which are based on their characteristics presented on data and expert knowledge. The normalised Digital Surface Model (nDSM) from ALS data was examined considering the low-, mid- and high-height levels. At each level, the A–V–N (Asphalt, Vegetation and Non-vegetation) classification was used to simplify the categories. The results showed that the incorporation of ALS data, especially nDSM, significantly improves the accuracy. Huang *et al.* (2011) proposed the fusion of the spectral signal from aerial image with elevation features (elevation difference, maximum and minimum values, variance and the grey-level co-occurrence matrix (GLCM) textures) from ALS-based DSM in three levels: the feature level (vector stacking), the multiclass output level (reclassification), and decision level (post-processing) using SVM as classifier for all fusion schemes. The overall accuracy increased about 13% compared with the classification result using only spectral information. Although the vector stacking and soft reclassification methods presented promising results, they are dependent on the elevation features used and the post-processing method is dependent on the prior knowledge about the test area (conditional probability of height distribution) to improve the distinction between spectrally similar classes.

Image classification methods have also been used to detect roads in complex urban areas that, in some cases, serve as an initial step for more complex tasks regarding the road reconstruction (Shackelford and Davis 2003, Matkan *et al.* 2014). For example, Gao and Wu (2004) performed a road reconstruction in dense urban areas by firstly applying an unsupervised classification to multispectral IKONOS images, followed by post-processing steps applied to the resulting road class, including noise removal, skeletonisation and linking of

road segments. The key point of the method is in the post-processing, since the classification method is not exploited, resulting in noises widespread in the result binary image. Zhang and Wang (2004) developed a method to detect the road network in an urban image generated by the fusion of panchromatic and multispectral QuickBird images. First, the k-means classification method is used to isolate the road class in image data and a filtering strategy based in edge-aided classification is used to disconnect small parking lots and buildings misclassified as road. Then, shape-based segmentation and a segments filtering algorithm are employed to remove these false positives. Despite the good results achieved after filtering, the method is applied in suburban area and involving many operations to eliminate false positives, mainly buildings. Das *et al.* (2011) exploited two road salient features (i.e. distinct spectral contrast and locally linear trajectory) to design a multistage framework to extract roads from urban high-resolution multispectral satellite images. Basically, road segments are detected by using a probabilistic SVM, from which road segments are extracted and linked. The method produced high degree of accuracy for suburban scenes, while in urban area, buildings and other objects similar to roads made the extraction process somewhat more difficult and only major roads have been detected. Mancini *et al.* (2009) used the machine-learning algorithm AdaBoost (i.e. Adaptive Boosting) to classify multispectral aerial images that were integrated with ALS data. The Hough transform was then used to reconstruct roundabouts and the line growing based segmentation to roads. The method allows reducing the effect of occlusions in the roads and extracted roads with different directions.

As noted in previous work, researches integrating ALS-based height data and optical images for object detection and delineation have been intensified in recent years. Most of the works use the height information from the ALS data as ancillary data such as DSM or nDSM. The height information favours the discrimination between aboveground objects of those in the ground level. On the other hand, the ALS-based intensity image has not been widely used in same context (Zhou 2013), mainly because this kind of image is too noisy. However, asphalt pavement is well contrasted in ALS-based intensity images. Thus, in the proposed method, this image is fundamental to improving the classification result for the road class, since it presents peculiar characteristics in relation to roads. We believe that combining optical images with both types of ALS data (geometric and radiometric) in a classification process can increase the accuracy and reliability of classification of the road class, especially in complex urban environments. Based on this motivation, we propose a study of different combinations between optical images and geometric and radiometric ALS data using the Artificial Neural Network (ANN) classification method.

ANN is a non-parametric method that have been widely used due to its advantages over statistical methods (Shao and Lunetta 2012, Song *et al.* 2012), such as no assumptions on the statistical distribution of the data, capability to work with large amounts of data (Pereira and Centeno 2017), and due to their robustness in handling noisy data (Sun *et al.* 2011, Weng 2012). Moreover, it allows the integration of different data sources, like optical images and ALS data, which are usually in different spectral ranges and do not present spectral nature. Mokhtarzade and Valadan Zoej (2007) and Benkouider *et al.* (2011) used ANN to separate the road class in high-resolution satellite images using spectral parameters in the neural network input, taking into to the spectral characteristics of roads in the image. Other different applications of ANN involving remote sensing data can be mentioned, which include land use/land cover classification (Pacifi *et al.* 2009, Pu *et al.* 2011, Ojaghi *et al.*

2015), change detection (Srivastava *et al.* 2012) and impervious surface estimation (Hu and Weng 2009, Du *et al.* 2015).

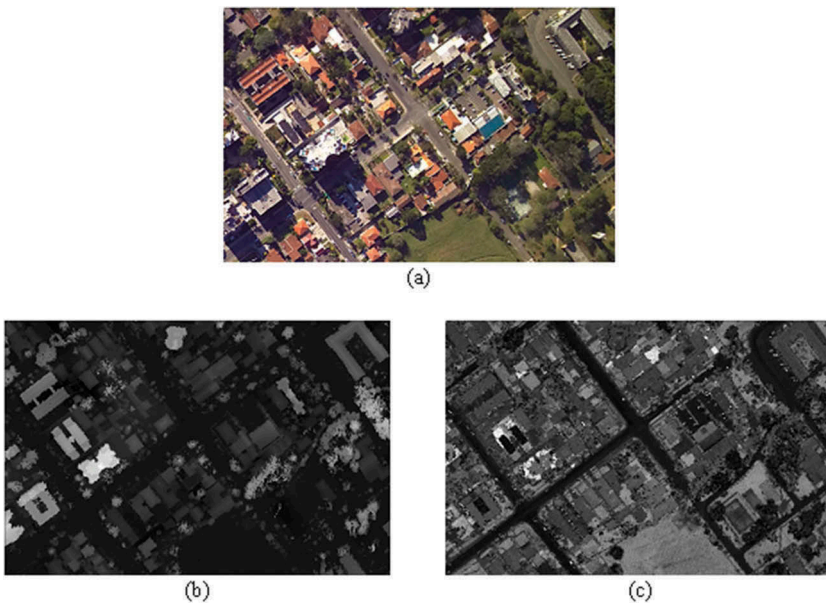
In this study, the proposed ANN classification method combines different image data, that is, the aerial high-resolution RGB image and ALS-derived images. Although several information classes (grass, trees, building roofs, asphalted roads, etc.) are defined in the classification problem, particularly when dealing with complex urban environments, our main focus is on the detection of the road pixels, as a first stage in the future extraction road network procedure. The present article extends the work of Dal Poz and Mendes (2013) and it is organized as follows. Section 2 describes the study area and dataset used. The proposed method is detailed in Section 3. Section 4 presents the experimental results and related discussions. Section 5 presents the major conclusions and future work.

## 2. Study area and datasets

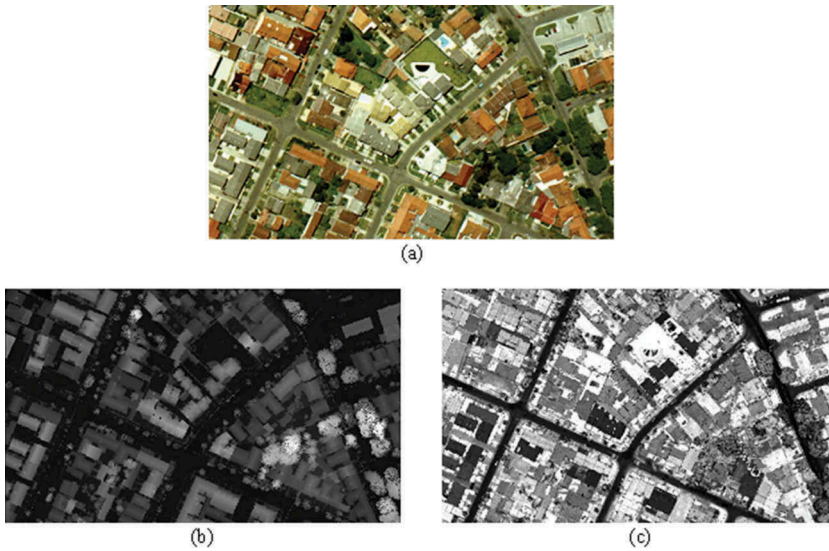
The study area comprised an urban region of the city of Curitiba, Southern Brazil. Two data sets were used to experimentally verify the proposed method. Each data set consists of an aerial high-resolution (GSD – Ground Sample Distance  $\sim 0.2$  m) RGB image and an ALS point cloud with a density of 4 points/m<sup>2</sup>.

The RGB image corresponding to test area 1 is showed in Figure 1(a). The nDSM and intensity images (showed in Figure 1(b, c), respectively) were generated from the ALS data through the method described in Section 3.1. It is worth mentioning that the RGB image shows road segments partially or totally obstructed by shadow and vegetation. On the other hand, these anomalies are minimised in ALS images, mainly in intensity image.

Figure 2(a) shows the RGB image corresponding to test area 2. This area presents road segments without obstructions by shadow, but the presence of building roofs with



**Figure 1.** Test area 1. (a) RGB image; (b) nDSM image and (c) intensity image.



**Figure 2.** Test area 2. (a) RGB image; (b) nDSM image and (c) intensity image.

similar characteristics to road may lead to confusion between these two classes. The corresponding nDSM and intensity images are showed in [Figure 2\(b, c\)](#), respectively. Trees obstructing roads can be observed near the right border of the three images. This effect is less severe in the intensity image.

### 3. Method

Our proposed study on the feasibility of using ALS data combined with optical images in a classification framework, using ANN, addresses three main steps: (3.1) geometric integration of ALS data and optical images and pre-processing; (3.2) ANN classification involving different combinations of geometrically integrated optical and ALS-derived images, with main focus on the separation capability of the road class; and (3.3) accuracy assessment for the classification results with the main focus on the road class.

#### 3.1. Data integration and pre-processing

The ALS data and aerial images can be integrated by projecting the ALS point cloud onto the aerial image followed by a regularisation of the projected points according to a grid with the same resolution of the aerial image. Briefly, the main steps are as follows: (1) transformation of the ALS point coordinates from the WGS84 (World Geodetic System) or E and N of the Universal Transverse Mercator projection with ellipsoidal height, where the ALS coordinates are typically made available, into the Local Cartesian Geodetic System (LCGS), the system in which the photogrammetric process of space resection is generally performed; (2) transformation of the ALS point coordinates from LCGS into the Photogrammetric Coordinate System (PCS) via collinearity equations; (3) transformation of the ALS point coordinates from PCS into the Image Coordinate System (ICS) in rows and columns, together with the introduction of image systematic errors (e.g. lens distortions)



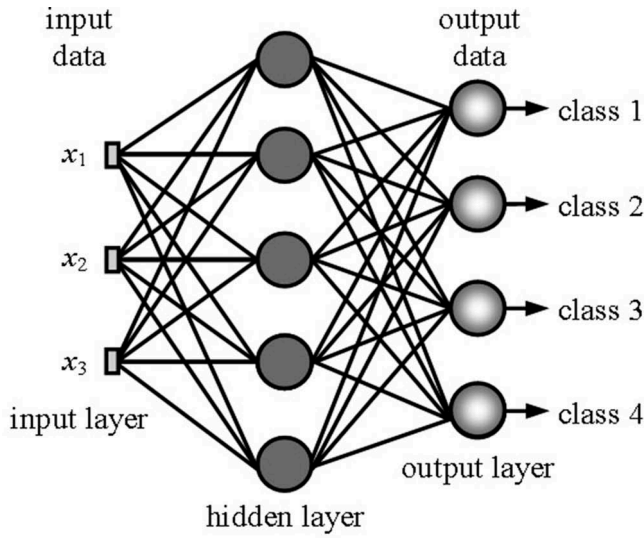
and 4) regularization of the projected point cloud according to grids with the same resolution of the aerial image. Details of the object-space coordinate transformations (step 1) and the photogrammetric processes (steps 2 and 3) can be found in Wolf and Dewitt (2000). In step 4, the following grids are firstly generated in the ICS: the DSM and the Digital Terrain Model (DTM). A third grid, the nDSM, is generated by subtracting the DTM from the DSM. The ground and aboveground objects can be separated in the nDSM by a thresholding process (nDSMThr) taking into consideration that aboveground objects (such as vegetation and building) have a minimum height. It is expected that the thresholded nDSM can supply a hard constraint to avoid classifying aboveground objects as roads and avoid height-related heterogeneity for aboveground objects.

As each ALS point has a radiometric response (i.e. the so-called laser pulse return intensity), together with the X, Y and Z Cartesian coordinates, the same procedure described above for the DSM and DTM generation in the ICS can be used to generate a laser pulse intensity image (or intensity image in short), also in the ICS and with same resolution of the input aerial image. Roads in intensity images possess homogeneous spectral responses and appear very dark due to the asphalt's low reflectivity of the laser pulse (approximately 17% according to Wehr and Lohr (1999)). Another significant characteristic of the intensity image is that it is not affected by shadows from other objects (Baltsavias 1999) as in optical images. Therefore, the intensity image can contribute considerably to refining information during the classification of images. Despite these advantages, the intensity images exhibit certain weaknesses. The first of these is the noise characteristic that directly affects the results of the classification (Yan *et al.* 2015), because the laser pulse scanning is a subsampling process. This issue can be minimised by applying a median filter or a morphological filter. An additional problem with this type of data is the presence of objects that show radiometric responses similar to those of roads, such as the building roofs and certain types of vegetation.

Data integration and pre-processing methods discussed above allow the generation of several geometrically integrated layers, as for example, R, G and B image layers; nDSM segmented using the thresholding method (the threshold used was 2.5 m), and the intensity image smoothed using the median filter. The DSM and DTM used to derive the nDSM were generated using different interpolation methods. The nearest neighbour interpolation method (El-Sheimy *et al.* 2005) was used for deriving the DSM because it preserves the sharpness of building breaklines. Points on the ground were automatically separated using the LAStools software and used to generate the DTM by using the kriging interpolation method (El-Sheimy *et al.* 2005). As already mentioned, the DSM and DTM were generated in the ICS at the same resolution of the RGB image.

### 3.2. Supervisioned classification through ANN

The data layers discussed in Section 3.1 are used to isolate the road class by employing the ANN image classification method. The ANN model used is the Multilayer Perceptron (Figure 3), whose structure is defined by an input layer, one or more hidden layers, and output layer (Tso and Mather 2009). The input layer contains the set of neurons that received the input data. The number of neurons in this layer corresponds to the dimensionality of the input data. The procedures about classifying are conducted within the hidden layer(s) and the results are passed to the output layer (Hu and Weng 2009).



**Figure 3.** The multilayer perceptron ANN model.

The number of neurons in the output layer is defined by number of classes. Usually, the number of hidden layers and neurons in each layer is experimentally determined.

The input to this network is the feature vector extracted from the data to be classified, and the output is typically a block code where one output is high indicating the class of the object and all other outputs are low (Ruck *et al.* 1990). Signals flow from the input layer through to the output layer by interconnections. These interconnections have associated numerical weights ( $w_{ji}$ ), which are adjusted during the training phase using some training rule with the set of feature vectors, the training set (Ruck *et al.* 1990, Tso and Mather 2009).

An efficient and commonly used algorithm for training the ANN is the Backpropagation algorithm (Haykin 1999). This algorithm is an iterative gradient descent training procedure and involves two phases, that is, forward and backward propagation (Tso and Mater 2009).

The first, a set of weights is randomly initialized and then, the input data (training pixels) are presented to the network and propagated forward to estimate the output value for each training pattern set (Kavzoglu and Mather 2003). During this process, the activities of neurons are sequentially updated through the input layer to the output layer, in order to generate output in the form of the activations of the output layer neurons (Tso and Mather 2009). Thus, each neuron receives the weighted input from the other neurons, sums these weighted inputs and then sends this output to the other neurons. The summation process can be expressed by Mather and Koch (2011):

$$x_j = \sum_i w_{ji} o_i \quad (1)$$

where  $x_j$  denotes the total input received by neuron  $j$  from all neurons to which it is connected in the preceding layer ( $i$ ),  $o_i$  is the activities of neuron  $i$ , and  $w_{ji}$  is the weight of connection from the  $i^{th}$  neuron to the  $j^{th}$  neuron. Value of  $x_j$  is converted to an output



value using a mapping function. This work used the logistic sigmoid function which can be expressed by Equation (2):

$$o_j = f(x_j) = \frac{1}{1 + e^{-x_j}} \quad (2)$$

In the second phase, the difference (error) between known and estimated outputs is fed backward through the network, and the weights associated with the neurons are updating so that to minimise the difference between the current and desired outputs (Kavzoglu and Mather 2003).

The whole process is repeated, with weights being recalculated at each iteration, until the error is minimal (Kavzoglu and Mather 2003). Generally, the least mean square error is applied (Haykin 1999), as given by Equation (3).

$$E(w) = \frac{1}{N} \sum_{j,k} (d_{j,k} - o_j)^2 \quad (3)$$

where  $w$  is a set of weights in a network,  $d_{j,k}$  is the desired output and  $o_{j,k}$  is the actual output of unit  $j$  produced by the network,  $N$  is the sample size of the set of training data made up of  $N$  input-output examples.

In this work, the input data layers were combined in different ways to determine the best combination for separating the road class and to verify the effective contribution of input layers from ALS data. All of the input data layers in the network were scaled in the range 0–255. Table 1 shows examples of combinations of data layers for evaluating the contribution of ALS data combined with a RGB image into the classification procedure by ANN.

The first combination (RGB) only utilises layers of the RGB image and it is used to analyse and compare the contributions of the subsequent additions of the ALS data during the classification process. The second combination (RGB + nDSMThr) adds the geometric ALS data to the RGB image layers by using the image representing the nDSMThr in the ICS. The major objective of performing this combination was to check the hypothesis by which the geometric ALS data is useful for separating classes corresponding to aboveground objects (e.g. roofs and trees) from those that are on the terrain surface. The intensity image in the ICS was smoothed using the median filter (resulting in the IS image) and added to the RGB image layers, giving rise to the third combination (RGB + IS) in Table 1. The rationality of smoothing the intensity image is that it is very noisy. The IS image is expected to be useful in the classification process because roads are very well contrasted in it. Moreover, due to the capability of the laser pulse in penetrating the tree structures and reaching the ground, road obstructed by trees can be totally or partially visible in the IS image. In order to verify the joint contribution of the geometric and radiometric ALS data, when combined to the RGB image, the last combination (RGB + nDSMThr + IS) mixes the RGB image layers, the nDSMThr image layer and IS image layer.

**Table 1.** Examples of combinations of input data layers for classification by ANN.

Combination	Input Layer Data
RGB	R, G and B
RGB + nDSMThr	R, G, B and nDSM with thresholding (nDSMThr)
RGB + IS	R, G, B and intensity image with median smoothing (IS)
RGB + nDSMThr + IS	R, G, B, nDSMThr and IS

### 3.3. Accuracy assessment

In order to numerically evaluate the classification results, the accuracy assessment was performed using the reference dataset (ground truth). Based on the error matrix, the accuracy statistics such as Overall Accuracy (OA) and Kappa coefficient (Congalton 1991) were used for evaluating the classification results. The OA is simply the percentage of correctly classified pixels, computed by dividing the total of correctly classified pixels by the total of reference pixels. Kappa is the percentage of agreement, that is, is based on the difference between actual agreement and the chance agreement in the error matrix.

The accuracy assessment for the road class separately was performed using the concepts of Producer's Accuracy (PA) and User's Accuracy (UA) measurement (Congalton 1991). The PA indicates the probability of a reference pixel being correctly classified, and the UA indicates the probability that a pixel classified in the map/image actually represents that category on the ground.

To determine if the classification results obtained for the combinations presented in Table 1 are significantly different, a test of significance is accomplished based on the test Z-statistic that uses the kappa values ( $K_1$  and  $K_2$  in Equation 4) and their associated variances (Im *et al.* 2012):

$$Z = \frac{K_1 - K_2}{\sqrt{\text{variance}(K_1) + \text{variance}(K_2)}} \quad (4)$$

where  $K_1$  and  $K_2$  are two kappa coefficients in comparison. The difference between them is considered to be significant at the 99% confidence level if the critical value of the Z-statistic exceeded 2.58.

## 4. Results and discussions

Classifications using ANN, which were performed with the IDRISI Andes software (Clark Labs, Worcester, MA, USA), tested different network architectures for each of the combinations of input data described in Table 1. In order to verify the capability of the ANN in separating the road class from other classes, using different data layer combinations, the following classes are defined: road, grass, trees, buildings with grey roofs, buildings with red roofs, concrete and bare soil.

For each of the combinations described in Table 1, the network was trained by testing different architectures with one or two hidden layers and varying the number of neurons for each layer. In network training process values for the learning rate of 0.01, a momentum factor of 0.5 and the number of iteration of 10,000 were adopted. The learning rate corresponds to the proportion of error that is passed for the neurons at each iteration and the momentum factor is used to optimise the network convergence (Tso and Mather 2009).

### 4.1. Test area 1

Training samples for each predefined class were manually collected by using polygons to delimitate homogeneous regions. The training data set obtained for test area 1 is the same for all combinations present for this area. Figure 4(a) presents the polygons representing the training and reference samples superposed on the RGB image for test area 1.



**Figure 4.** Teste area 1: (a) Training samples and (b) Reference samples.

After training the ANN, the classification was performed and the result was numerically evaluated. This requires new representative samples of each class, called reference data, which were collected using as reference the stratified random sampling strategy with 200 points. Thus, the polygons were manually collected in these points, discarding those that are located between two classes. The reference samples superposed on the RGB image for the test area 1 is shown in Figure 4(b).

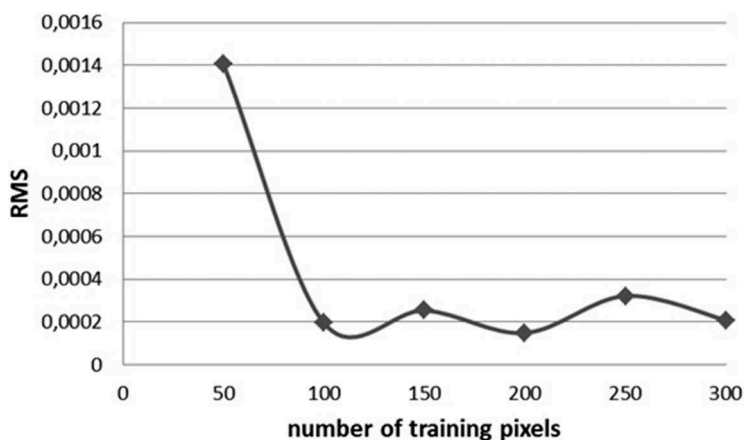
Table 2 lists the number of training and reference pixels sampled for each class.

Based on training data set presented in Figure 4 and Table 2, it defined the number of pixels per class that the software IDRISI selected for training and for testing. To choose this number, different values have been tested and at the end of each training process the Root Mean Squared Error (RMSE) of the network output was analysed. Figure 5 shows that the lowest value for the RMSE was obtained using 200 pixels; moreover, the increase of number of pixels does not present significant difference in the value for the RMSE and it increases the computational effort. Similar results have been found in Doma *et al.* (2015) for the sample size in training process. Thus, from all sample set (Table 2), 200 pixels were randomly selected for each class to be used as training pixels, and another 200 pixels were selected as testing pixels.

The number of hidden layer (one or two) and the number of the neurons for each hidden layer were defined experimentally (try and error method) based on the analysis the statistics of training provided by the software and on the results of classification by analysis of error matrix. Despite several strategies have been suggested to estimate the optimum number of hidden layer neurons, none of them has been universally accepted requirements (Kavzoglu and Mather 2003). Using a single hidden layer, a standard rule  $N_{\text{hidden}} = \text{int}(\sqrt{N_{\text{input}} * N_{\text{output}}})$  may be applied as used

**Table 2.** Training and reference data set for the test area 1.

Class	Test area 1	
	Training	Reference
Road	4657	5820
Grass	2311	3966
Tree	6165	6933
Grey roof	2512	3570
Red roof	2527	3434
Concrete	702	1540
Total	18,874	25,465



**Figure 5.** Selection of the number of training and testing pixels per class.

as default within IDRISI. This rule was applied in the architecture identified as A in Table 3, which present other architectures tested for the RGB + nDSMThr + IS combination for the test area 1 and their descriptive statistics, varying the number of hidden layers (one or two) and the number of neurons for each layer. For example, the architecture identified as D contains 5 neurons corresponding to the number of input layer data (R, G, B, nDSMThr and IS image), two hidden layers that contain 7 neurons in the first hidden layer and 11 neurons in the second hidden layer, and the output layer that contains 6 neurons, each one corresponding to a class (in short, we refer to this architecture as 5–7–11–6).

The accuracy rate and time training presented in Table 3 are based on training phase. Architectures with two hidden layers had effect on both classification accuracy rate and training time, with increase of the accuracy and the time of training (1–4 min). The architecture D in Table 3 presented the highest accuracy rate (97.08%) with processing time of 5'40". The architecture defined as default by software (with 5 neurons in a single hidden layer, in short, 5–5–6) presented the low accuracy rate (94.75%) in the training phase, and consequently, in the result of the classification, with 85.31% for OA and 0.82 for kappa.

A comparison between the accuracy from the classification result using the architectures tested shows that the higher values for OA were found for architectures D, G and J

**Table 3.** Examples of tested architectures for RGB + nDSMThr + IS combination for test area 1.

Architecture	Training phase		Accuracy assessment			
	Accuracy rate (%)	Time training	OA (%)	Kappa	PA (%) Road class	UA (%) Road class
A 5–5–6	94.75	4'00"	85.31	0.82	94.54	88.05
B 5–7–6	93.17	4'03"	89.65	0.87	95.21	93.02
C 5–11–6	94.83	4'09"	89.22	0.87	93.47	94.97
<b>D 5–7–11–6</b>	<b>97.08</b>	5'41"	<b>91.38</b>	<b>0.89</b>	<b>96.67</b>	<b>96.95</b>
E 5–9–16–6	96.83	6'32"	89.88	0.88	95.09	97.41
F 5–15–21–6	95.58	7'22"	89.52	0.87	90.36	96.20
<b>G 5–17–25–6</b>	96.75	7'55"	<b>90.73</b>	<b>0.88</b>	<b>95.15</b>	<b>97.09</b>
H 5–17–5–6	95.58	5'12"	85.12	0.82	96.98	91.62
I 5–25–9–6	96.25	5'40"	85.32	0.82	96.87	88.11
<b>J 5–25–18–6</b>	96.83	6'46"	<b>90.16</b>	<b>0.88</b>	<b>96.84</b>	<b>95.46</b>

(Table 3). Among them, the one with the higher values for PA (91.38%) was the architecture D, with kappa coefficient value of 0.89. However, it is verified in Table 3 that the values for OA, kappa coefficient, PA and UA for road class do not present significant variation in relation to the choice of the architecture, indicating that the choice of the architecture is not the central point of the method, but it is necessary to find a competent architecture to problem. This can be verified in Figure 6, which shows that the classification results using the architectures D, G and J are very similar.

Based on previous experiment, the architecture with two hidden layers containing 7 neurons (1st hidden layer) and 11 neurons (2nd hidden layer) was used for all combinations presented in Table 1. Figures 7, 8, 9 and 10 show the classification results for the test area 1. The accuracy assessment results are presented in Table 4.

Figure 7(a) presents the classification result for the RGB combination, which achieved an OA of 61.12%. The worst results were achieved for the tree, concrete and grey roof classes (corresponding PA were below 50%). On the other hand, the best result was found for the red roof class, for which the PA was 89%. The PA and UA achieved for the road class were 79.74% and 63.44%, respectively. Due to the above reason, the kappa coefficient was very low, that is, 0.53. Finally, in order to facilitate the visualisation of detected road regions, Figure 7(b) shows pixels classified as road in black and pixels from other classes in white. It is worth mentioning that many pixels inside the blocks were detected as belonging to the road class. Moreover, the road class shows several anomalies and suffers influence of the shadow present in the scene, with pixels classified as grey roof class.

Figure 8(a) shows the classification result for the RGB + nDSMThr combination. The ANN was trained similarly to the previous example, using the 4–7–11–6 architecture. As a result, the OA (85.19%) and the kappa coefficient (0.82) were much better than ones obtained using only the RGB image. The results for grey roof and concrete classes improved, corresponding

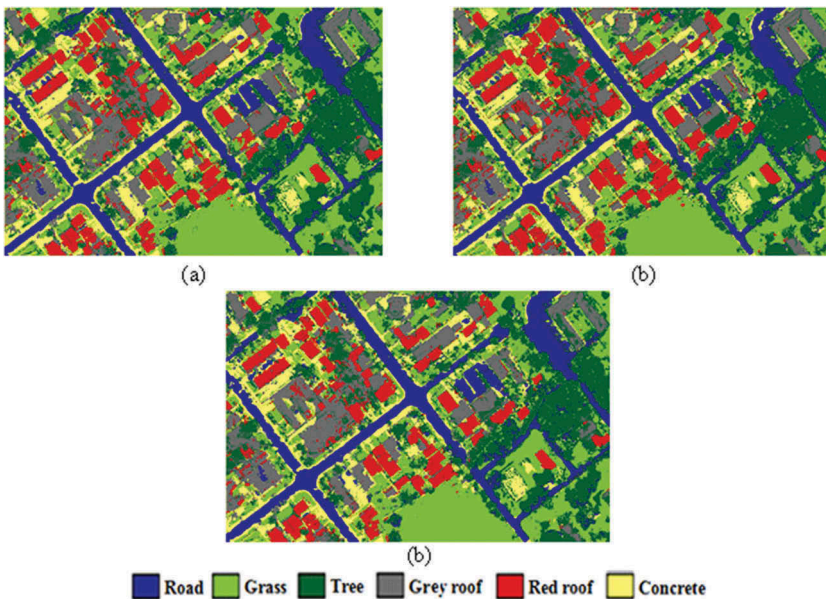
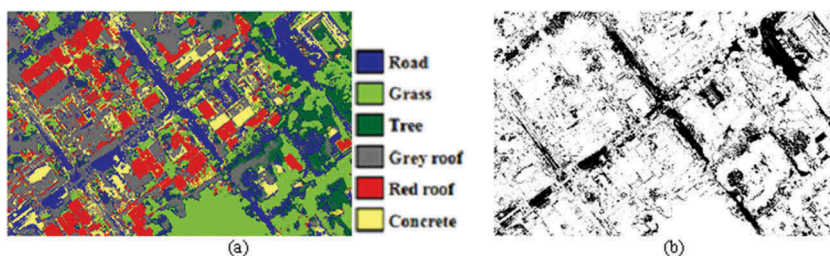
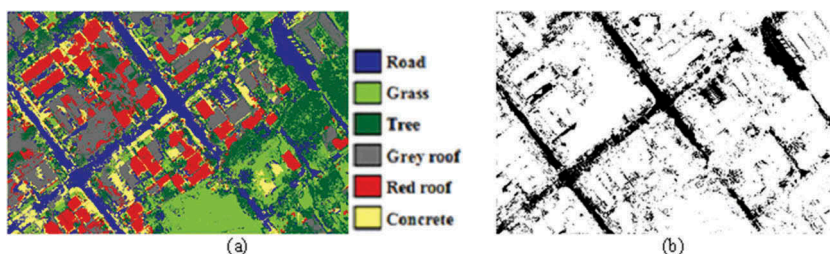


Figure 6. Classification results for test area 1. (a) Architecture D; (b) Architecture G and (c) Architecture J.

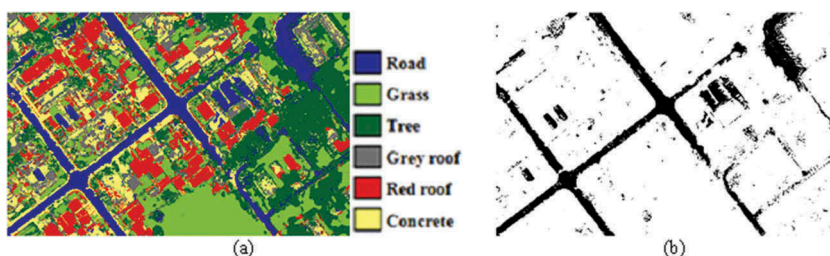




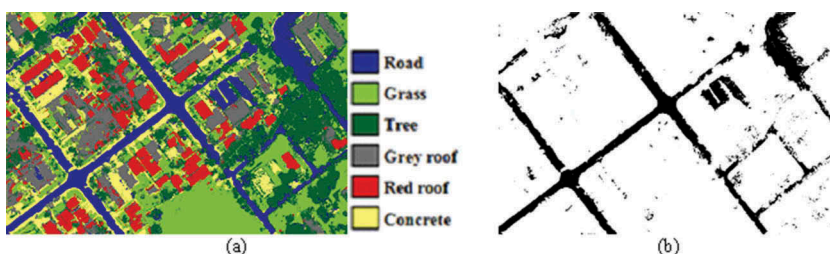
**Figure 7.** Test area 1 – Results obtained for the RGB combination. (a) Classification result and (b) Binary image showing road pixels in black.



**Figure 8.** Test area 1 – Results obtained for the RGB + nDSMThr combination. (a) Classification result and (b) Binary image showing road pixels in black.



**Figure 9.** Test area 1 – Results obtained for the RGB + IS combination. (a) Classification result and (b) Binary image showing road pixels in black.



**Figure 10.** Test area 1 – Results obtained for the RGB + nDSMThr + IS combination. (a) Classification result and (b) Binary image showing road pixels in black.



**Table 4.** Accuracy assessment results for test area 1.

Combination	Accuracy assessment			
	OA (%)	Kappa	PA (%) road class	UA (%) road class
RGB	61.12	0.53	79.74	63.44
RGB + nDSMThr	85.19	0.82	88.83	85.58
RGB + IS	84.70	0.81	90.52	91.71
RGB + nDSMThr + IS	91.38	0.89	96.67	96.95

to PAs of about 88%. The value of the PA of 88.83% was obtained for the road class, increasing the PA in 9% when compared with the results of the previous combination (Figure 7), whereas AU increased in 22%, reaching a value of 85.58%. The visual inspection of Figure 8(a) (see also Figure 8(b)) shows that the nDSMThr image really helped to improve the detection of road pixels and the reduction of false detection of road pixels.

Next example (Figure 9) combines the RGB image with the ALS intensity image with a median smoothing (IS). The adopted architecture when applied to this combination results in 84.70% OA and 0.81 kappa coefficient. The classification result (Figure 9(a)) is slightly worse than one obtained by using the nDSMThr, meaning that the intensity ALS data is also useful for avoiding the confusion between the road class and other spectrally similar classes in the RGB image (like grey roofs). However, the grey roof class presented a very low PA (50%), indicating greater confusion with concrete class. The main observable difference in relation to the previous combination (RGB + nDSMThr) is the reducing of pixels misclassified as road class inside the block. Besides, the road class shows fewer failures, as depicted in Figure 9(b), in comparison with Figure 8(b). As a result, the road class PA was better (90.52%) for the test using the RGB + nDSMThr combination. The same occurs for AU, with a value of 91.71%.

Next combination (RGB + nDSMThr + IS) (Figure 10) mixes the three layers from the RGB image, one layer from the geometric ALS data (nDSMThr image) and one layer from the radiometric ALS data (IS image). The classification result obtained (Figure 10(a)) provided the best OA (91.38%). The kappa coefficient was 0.89, which was the better one. In general, the classification results were better for all classes (For instance, the producer accuracies were 81%, 92% and 96% for the grey roof, red roof and grass classes, respectively), but it is worthy to emphasise the very high PA (96.67%) and UA (96.95%) obtained for our interest class, that is, the road class. In fact, Figure 10(a) (please also see the result obtained for the RGB combination (Figure 7(a)) shows that roads were successfully classified in regions obstructed by trees and few pixels were wrongly classified as road inside the blocks. Finally, Figure 10(b) provides a visualisation of the result obtained for only the road class.

The results obtained for the test area 1 showed that ALS data layers were helpful in the detection of road pixels and in the reduction of false road pixels' detection inside the blocks. To verify if the results obtained for the test area 1 were statistically significant, the Z-statistic was calculated by using the Kappa coefficients and corresponding variances. Table 5 presents the Z-statistic for all results of classification accuracies and indicates a statistically significant difference, at the 99% confidence level, between the RGB and RGB + nDSMThr + IS combinations. When only ALS data with the RGB image (RGB + IS and RGB + nDSMThr combinations) are used, the results also present a statistically significant difference but the value obtained for Z lies within the limit of the critical value (2.58).

**Table 5.** Z-statistic test between classification accuracies to the test area 1.

Combination	RGB	RGB + nDSMThr	RGB + IS	RGB + nDSMThr + IS
RGB	-			
RGB + nDSMThr	63.48	-		
RGB + IS	61.36	2.58	-	
RGB + nDSMThr + IS	84.82	20.03	22.94	-

## 4.2. Test area 2

For the test area 2, training and reference samples for each class were collected and, for this area, it was included the bare soil class. Figure 11 exhibits the polygons representing the training (Figure 11(a)) and reference (Figure 11(b)) samples superposed on the RGB image for test area 2.

Details of the number of samples used as training and reference are presented in Table 6.

For the test area 2, the same architecture used in previous experiment was used (4–7–11–7). Table 7 shows the accuracy assessment results and Figures 12 and 13 only present the worst result (i.e. for the RGB image) and the best result (i.e. for the RGB + nDSMThr + IS combination), respectively.

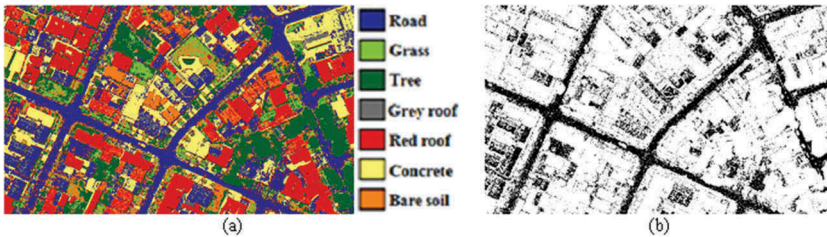
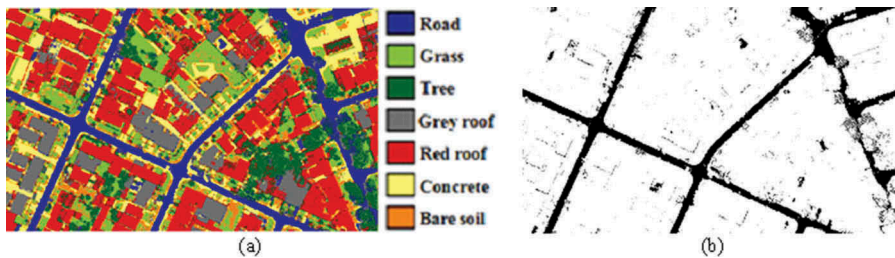
Figure 12 shows the classification results for the RGB image, for which an OA and kappa coefficient values of 74.02% and 0.67 were obtained, respectively. Figure 12(a) shows that most grey roofs were wrongly classified as roads. Figure 12(b) provides a better visualisation of road pixels wrongly detected inside the blocks. As a result, the PA for the grey roof class was very low, that is, 32%. Another important problem of classification of road class occurred with the road partially obstructed by trees near

**Figure 11.** Teste area 2: (a) Training samples and (b) Reference samples.**Table 6.** Training and reference data set for the test area 2.

Class	Test area 2	
	TRAINING	REFERENCE
Road	4771	7261
Grass	1284	2334
Tree	1344	1877
Grey roof	1268	2944
Red roof	1463	3740
Concrete	752	1459
Bare soil	536	1331
Total	11418	20946

**Table 7.** Accuracy assessment results for test area 2.

Combination	Accuracy assessment			
	OA (%)	Kappa	PA (%) road class	UA (%) road class
RGB	74.02	0.67	89.41	78.09
RGB + nDSMThr	84.70	0.81	83.60	94.08
RGB + IS	81.70	0.77	92.62	96.37
RGB + nDSMThr + IS	90.69	0.88	98.21	96.18

**Figure 12.** Test area 2 – Results obtained for the RGB combination. (a) Classification result and (b) Binary image showing road pixels in black.**Figure 13.** Test area 2 – Results obtained for the RGB + nDSMThr + IS combination. (a) Classification result and (b) Binary image showing road pixels in black.

the right border of the image. The PA and UA values of 89.41 and 78.09%, respectively, for the road class were achieved.

The classification result obtained for the combination between the RGB image and both geometric and radiometric ALS data (RGB + nDSMThr + IS) is presented in [Figure 13\(a\)](#). The OA (90.69%) and the kappa coefficient (0.88) are the best ones. It is worthy to note that the grey roofs were better detected, reducing the misclassification of road pixels inside the blocks (see [Figure 13\(b\)](#)). Also, the road near the right border of the image was much better detected than using only the RGB image. As expected, the better detection of road pixels is obtained, with PA and UA reaching values of 98.21% and 96.18%, respectively.

[Table 8](#) presents the Z-statistic for all classification accuracy results and indicates a statistically significant difference, at the 99% confidence level, between all combinations.

#### 4.3. Comparison to other methods

As discussed above, about 91% OA was achieved by the proposed classification method that integrates the RGB image with geometric and radiometric ALS data.

**Table 8.** Z-statistic test between classification accuracies to the test area 2.

Combination	RGB	RGB + nDSMThr	RGB + IS	RGB + nDSMThr + IS
RGB	-			
RGB + nDSMThr	29.21	-		
RGB + IS	20.33	9.04	-	
RGB + nDSMThr + IS	47.00	17.86	27.02	-

We compared this OA with OAs obtained by Huang *et al.* (2008). These authors also integrated the ALS data with aerial images using two classification methods. First, the maximum likelihood classification method was used for identifying urban object, achieving an OA of about 7% lower than the OA obtained by our method. Then, the authors applied to the same data set a KBCS that includes a three-level height, ‘asphalt road, vegetation, and non-vegetation’ (A–V–N) classification rule-based scheme. The KBCS achieved an OA of about 94%, which is slighter better than the 91% OA achieved by our method. However, one important drawback of the KBCS is the large number of thresholds needing determination. While the KBCS requires the setting of 18 thresholds, being only 2 determined automatically, our approach requires only one threshold (for the thresholding of the nDSM) and 5 training and architectural parameters of the ANN (number of hidden layers, number of neurons inside each hidden layer, learning rate, momentum factor, and number of pixels for training and testing). Moreover, our experiments show that these parameters can be kept almost unchanged.

Huang *et al.* (2011) used only DSM as ALS data and they combined it with spectral information from aerial image in an urban classification method composed of three levels. The results are dependent on the exhaustive pre-processing to obtain a set of elevation features (texture measures of GLCM, height difference, height variance and Maximum–Minimum values within a local area) and on the object-based segmentation process for generated the average height. Moreover, in the third level, prior knowledge about the test area (conditional probability of height distribution) is required. The method presents promising results, however, when comparing to the proposed method, it is verified that our method presents greater simplicity in the pre-processing, reached good results, mainly in regions that are confusion among similar spectral classes and regions that suffer shadow influence.

Salah *et al.* (2011) used feature attributes generated from RGB aerial image, intensity image, DSM and nDSM in the classification tree procedure to firstly obtain buildings, trees, roads and ground classes. The authors claim that using attributes generated from RGB aerial and intensity images resulted in lower classification accuracies, while using attributes generated from DSM and normalised resulted in significantly higher classification accuracies. This result disagrees with the results finding by our method, once that information of both ALS data (intensity and nDSM with thresholding) contributed to improve the classification accuracy.

It is worth mentioning that the proposed method, besides obtaining good results for the classification in general, improves considerably the accuracy for the road class, object of interest of the author, allowing to obtain road pixels in more complex areas, such as regions obstructed by shadows and trees.

## 5. Conclusions and future work

This article described an experimental study with the main focus on the classification of road pixels based on an ANN classification procedure of geometrically-integrated high-resolution RGB aerial and ALS-derived images.

The classification results showed that the RGB combination provided the worst performance due to well-known confusions among objects showing similar spectral responses, as for example, those related to roads and grey roofs. Of course, roads obstructed by trees in all input data were not detected. Combinations using the RGB and nDSMThr images, the RGB and IS images, or the RGB, nDSMThr and IS images clearly demonstrated that ALS data improved the classification as a whole and, particularly, the classification of our class of interest (i.e. roads). In fact, confusions between road class and other spectrally similar classes in the RGB image (i.e. grey roof and road) were reduced and roads obstructed by trees were partially or totally detected. However, the combination mixing the RGB image, the geometric ALS data (nDSMThr image) and the radiometric ALS data (IS image) provided the best classification result, achieving values of OA of about 90% and kappa values around 0.8. Especially roads were successfully classified in regions obstructed by trees and few pixels were wrongly classified as road inside the blocks, with values of 97% and 96%, respectively, for PA and UA.

Our future work will be focused on the development of post-processing techniques to refine the detected road regions and to reconstruct the road network using the refined road regions. These post-processing techniques would involve: regularisation of detected road regions using morphological operations; filtering of non-road regions; skeletonization of road regions and use of road knowledge to reconstruct the road network.

## Acknowledgements

We thank the LACTEC for providing the laser scanning data.

## Funding

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico [Grant MCT/CNPq n° 27/2007].

## ORCID

Tatiana Sussel Gonçalves Mendes  <http://orcid.org/0000-0002-0421-5311>

Aluir Porfírio Dal Poz  <http://orcid.org/0000-0002-2534-1229>

## References

- Alonso, M.C. and Malpica, J.A., 2010. Satellite imagery classification with LiDAR data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Kyoto, 38 (8), 730–735.
- Baltsavias, E.P., 1999. A comparison between photogrammetric and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 83–94. doi:[10.1016/S0924-2716\(99\)00014-3](https://doi.org/10.1016/S0924-2716(99)00014-3)

- Bellens, R., *et al.*, 2008. Improved classification of VHR images of urban areas using directional morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 46 (10), 2803–2813. doi:[10.1109/TGRS.2008.2000628](https://doi.org/10.1109/TGRS.2008.2000628)
- Benediktsson, J.A., Pesaresi, M., and Arnason, K., 2003. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (9), 1940–1949. doi:[10.1109/TGRS.2003.814625](https://doi.org/10.1109/TGRS.2003.814625)
- Benkouider, F., Hamami, L., and Abdellaoui, A., 2011. Use of the neural net for road extraction from satellite images, application in the city of Laghouat (Algeria). *Progress in Electromagnetics Research Symposium 2011-PIERS ONLINE*, 7 (2), 146–150.
- Chanussot, J., Benediktsson, J.A., and Fauvel, M., 2006. Classification of remote sensing images from urban areas using a fuzzy possibilistic model. *IEEE Geoscience and Remote Sensing. Letters*, 3 (1), 40–44. doi:[10.1109/LGRS.2005.856117](https://doi.org/10.1109/LGRS.2005.856117)
- Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing Environment*, 37, 35–46. doi:[10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B)
- Dal Poz, A.P. and Mendes, T.S.G., 2013. Road region detection in urban areas combining high-resolution RGB image and LASER scanning data in a classification framework. *Proceedings of International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40 (1), 53–56. doi:[10.5194/isprsarchives-XL-1-W1-53-2013](https://doi.org/10.5194/isprsarchives-XL-1-W1-53-2013)
- Das, S., Mirmaline, T.T., and Varghese, K., Oct 2011. Use of salient features for the design of a multistage framework to extract roads from high-resolution multispectral satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (10), 3906–3931. doi:[10.1109/TGRS.2011.2136381](https://doi.org/10.1109/TGRS.2011.2136381)
- Doma, M.L., Goma, M.S., and Amer, R.A., 2015. Sensitivity of pixel-based classifiers to training sample size in case of high resolution satellite imagery. *Journal of Geomatics*, 9, 53–58.
- Du, P., Xia, J., and Feng, L., 2015. Monitoring urban impervious surface area change using China-Brazil Earth Resources Satellites and HJ-1 remote sensing images. *Journal of Applied Remote Sensing*, 9, 0960941–09609415. doi:[10.1117/1.JRS.9.096094](https://doi.org/10.1117/1.JRS.9.096094)
- El-Sheimy, N., Valeo, C., and Habib, A., 2005. *Digital terrain modelling: acquisition, manipulation, and applications*. Boston: Artech House.
- Gao, J. and Wu, L., 2004. Automatic extraction of road networks in urban areas from ikonos imagery based on spatial reasoning. In: *Proc. International archives of photogrammetry and remote sensing*. Istanbul, 35, Comm. III/3, 331–337.
- Guo, L., *et al.*, 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 56–66. doi:[10.1016/j.isprsjprs.2010.08.007](https://doi.org/10.1016/j.isprsjprs.2010.08.007)
- Haala, N. and Walter, V., 1999. Automatic classification of urban environments for database revision using LiDAR and color aerial imagery. In: *Proc. International archives of photogrammetry and remote sensing*. Valladolid, 32, 76–82.
- Haykin, S., 1999. *Neural networks. A comprehensive foundation*. 2nd ed. NJ: Prentice Hall.
- Hu, X. and Weng, Q., 2009. Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks. *Remote Sensing Environment*, 113, 2089–2102. doi:[10.1016/j.rse.2009.05.014](https://doi.org/10.1016/j.rse.2009.05.014)
- Huang, M., *et al.*, 2008. A knowledge-based approach to urban feature classification using aerial imagery with lidar data. *Photogrammetric Engineering & Remote Sensing*, 74 (12), 1473–1485. doi:[10.14358/PERS.74.12.1473](https://doi.org/10.14358/PERS.74.12.1473)
- Huang, X., Zhang, L., and Gong, W., 2011. Information fusion of aerial images and LIDAR data in urban areas: vector-stacking, rec-classification and post-processing approaches. *International Journal of Remote Sensing*, 32 (1), 69–84. doi:[10.1080/01431160903439882](https://doi.org/10.1080/01431160903439882)
- Im, J., *et al.*, 2012. Impervious surface quantification using a synthesis of artificial immune networks and decision/regression trees from multi-sensor data. *Remote Sensing Environment*, 117, 102–113. doi:[10.1016/j.rse.2011.06.024](https://doi.org/10.1016/j.rse.2011.06.024)
- Kavzoglu, T. and Mather, P.M., 2003. The use of backpropagation artificial neural networks in land cover classification. *International Journal of Remote Sensing*, 10 (23), 4907–4938. doi:[10.1080/0143116031000114851](https://doi.org/10.1080/0143116031000114851)



- Makarau, A., Palubinskas, G., and Reinartz, P., 2011. Multi-sensor data fusion for urban area classification. In: *Urban remote sensing event 2011*. Munich, 21–24.
- Mancini, A., Frontoni, E., and Zingaretti, P., 2009. Automatic extraction of urban objects from multi-source aerial data. In: *Proc. CMRT09, IAPRS*. Paris, 38, Part 3/W4, 13–18.
- Mather, P.M. and Koch, M., 2011. *Computer processing of remotely-sensed images: an introduction*. Chichester: Wiley-Blackwell.
- Matkan, A.A., Hajeb, M., and Sadeghian, S., 2014. Road extraction from Lidar data using Support Vector Machine classification. *Photogrammetric Engineering & Remote Sensing*, 80, 409–422. doi:[10.14358/PERS.80.5.409](https://doi.org/10.14358/PERS.80.5.409)
- Mokhtarzade, M. and Valadan Zoej, M.J., 2007. Road detection from high resolution satellite images using artificial neural networks. *International Journal of Applied Earth Observation and Geoinformation*, 9 (1), 32–40. doi:[10.1016/j.jag.2006.05.001](https://doi.org/10.1016/j.jag.2006.05.001)
- Ojaghi, S., Ebadi, H., and Ahmadi, F.F., 2015. Using artificial neural network for classification of high resolution remotely sensed images and assessment of its performance compared with statistical methods. *American Journal of Engineering, Technology and Society*, 2, 1–8.
- Pacifici, F., Chini, M., and Emery, W.J., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing Environment*, 113, 1276–1292. doi:[10.1016/j.rse.2009.02.014](https://doi.org/10.1016/j.rse.2009.02.014)
- Pálsson, F., et al., 2012. Classification of pansharpened urban satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5 (1), 281–297. doi:[10.1109/JSTARS.2011.2176467](https://doi.org/10.1109/JSTARS.2011.2176467)
- Pereira, G.H.A. and Centeno, J.A.S., 2017. Assessment of training sample size for artificial neural network in supervised image classification using spectral and laser scanner data. *Bulletin of Geodetic Sciences*, 23 (2), 268–283.
- Pu, R., Landry, S., and Yu, Q., 2011. Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. *International Journal of Remote Sensing*, 20, 3285–3308. doi:[10.1080/01431161003745657](https://doi.org/10.1080/01431161003745657)
- Ruck, D.W., Rogers, S.K., and Kabrisky, M., 1990. Feature selection using a multilayer perceptron. *Journal of Neural Network Computing*, 2 (2), 40–48.
- Salah, M., Trinder, J.C., and Shaker, A., 2011. Performance evaluation of classification trees for building detection from aerial images and LiDAR data: a comparison of classification trees models. *International Journal of Remote Sensing*, 20, 5757–5783. doi:[10.1080/01431161.2010.507678](https://doi.org/10.1080/01431161.2010.507678)
- Shackelford, A.K. and Davis, C.H., 2003. Urban road network extraction from high-resolution multispectral data. In: *2nd GRSS/ISPRS joint workshop on remote sensing and data fusion over urban areas*. Berlin, 142–146.
- Shao, Y. and Lunetta, R.S., 2012. Comparison of support vector machine, neural network, and CART algorithms for the land cover classification using limited training data points. *ISPRS Journal of Photogrammetry and Remote Sensing*, 10, 78–87. doi:[10.1016/j.isprsjprs.2012.04.001](https://doi.org/10.1016/j.isprsjprs.2012.04.001)
- Song, X., Duan, Z., and Jiang, X., 2012. Comparison of artificial neural network and support vector machine classifiers for land cover classification in Northern China using a SPOT-5 HRG image. *International Journal of Remote Sensing*, 33 (10), 3301–3320. doi:[10.1080/01431161.2011.568531](https://doi.org/10.1080/01431161.2011.568531)
- Srivastava, P.K., et al. 2012. Selection of classification techniques for land use/land cover change investigation. *Advances in Space Research*, 50 (July), 1250–1265. doi:[10.1016/j.asr.2012.06.032](https://doi.org/10.1016/j.asr.2012.06.032)
- Sun, Z., et al., 2011. Estimating urban impervious surfaces from Landsat-5 TM imagery using multilayer perceptron neural network and support vector machine. *Journal of Applied Remote Sensing*, 5, doi:[10.1117/1.3539767](https://doi.org/10.1117/1.3539767)
- Tso, B. and Mather, P.M., 2009. *Classification methods for remotely sensed data*. 2nd ed. Boca Raton, FL: CRC Press.
- Wehr, A. and Lohr, U., 1999. Airborne LASER scanning – an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2–3), 68–82. doi:[10.1016/S0924-2716\(99\)00011-8](https://doi.org/10.1016/S0924-2716(99)00011-8)

- Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. *Remote Sensing of Environment*, 117, 34–39. doi:[10.1016/j.rse.2011.02.030](https://doi.org/10.1016/j.rse.2011.02.030)
- Wolf, P.R. and Dewitt, B.A., 2000. *Elements of photogrammetry: with applications in GIS*. Boston: McGraw-Hill.
- Yan, W.Y., Shaker, A., and El-Ashmawy, N., 2015. Urban land cover classification using airborne LiDAR data: a review. *Remote Sensing Environment*, 158, 295–310. doi:[10.1016/j.rse.2014.11.001](https://doi.org/10.1016/j.rse.2014.11.001)
- Zeng, Y. et al., 2002. Urban land-use classification using integrated airborne laser scanning data and high resolution multi-spectral satellite imagery. In: *Proc. Pecora 15/Land satellite information IV/ISPRS commission I/FIEOS*. Denver.
- Zhang, Y. and Wang, R., 2004. Multi-resolution and multi-spectral image fusion for urban object extraction. In: *Proc. XXth ISPRS congress geo-imagery bridging continents*, Istanbul, 35, Part B3, Commission 3, 960–966.
- Zhou, W., 2013. An object-based approach for urban land cover classification: integrating LiDAR height and intensity data. *IEEE Geoscience and Remote Sensing Letters*, 10, 928–931. doi:[10.1109/LGRS.2013.2251453](https://doi.org/10.1109/LGRS.2013.2251453)