

PERFORMANCE OF THE SUPPORT VECTOR MACHINE AND ARTIFICIAL NEURAL NETWORK CLASSIFIERS FOR ROADS IDENTIFICATION

A.C. de Andrade¹, M. J. Alixandrini Jr^{1*}, F. P. S. Carvalho¹, V. de O. Fernandes¹

¹ PPEC, Federal University of Bahia, Salvador/BA, Brazil

alexcurvandrade@gmail.com, Mauro.Alixandrini@ufba.br, Fernanda.Puga@ufba.br, vivian.fernandes@ufba.br

Commission III WG III/1

KEY WORDS: Classification, Roads, Artificial Neural Network, Support Vector Machine, DSM, Weka.

ABSTRACT:

The objective of this project was to compare two non-parametric classification methods (“Support Vector Machine” - SVM and “Artificial Neural Networks” - ANN) of road regions in high spatial resolution images and associated with data from Airborne Laser Scanning. The study aims to verify what kind of influence the layers of attributes have on the performance from respective classifiers: SVM and RNA. Our method based on tests of this classifiers on 4 bands of airborne images and normalization of the digital surface model (DSM) for showing only information on objects height in relation to ground and not of these in relation to the ground and relief, generating band 5. The samples were used to train chosen non-parametric classifiers (training sets for each different input image/landscape). All classifications had the same set of training samples and the same classification parameters. The optimal parameters for classifications were obtained through the existing library in the Weka mining package: LibSVM and LibMultiLayerPerceptron. Our results demonstrated the existence of a direct relationship between the elevation band of the targets in relation to the terrain (band 05) with the improvement of their performance and lower degree of between bands correlation can also be considered a factor that has a positive influence. As for Neural Networks, the experiment results demonstrate that the presence of the near infrared band (band 04) was decisive for the performance improving of certain combinations in relation to others.

1. INTRODUCTION

The objective of the project was to compare two non-parametric classification methods (“Support Vector Machine” - SVM and “Artificial Neural Networks” - ANN) of road regions in extremely high spatial resolution images. This work use also data from “Airborne Laser Scanning” – ALS associated with multispectral data. The study aims to verify what kind of influence the layers of attributes have on performance of the respective classifiers (SVM and ANN). The feature studied was the road network, which can have very peculiar characteristics with high diversity, mainly in urban environment. The intense dynamics of use and occupation of spaces in large urban areas in developing countries generate an accelerated process of transformations in land use of these cities. This scenario demands efforts to map new occupied or altered spaces and new roads created (regular or not) or, still, the necessary updating maps (MABOUDI,2017).

Associated with this, the reality found in cities in developing countries requires interventions by public authorities in order to reduce conflicts caused by irregular occupation and the lack of land use planning. In this context, the pathway is an important information for public planning. In recent decades, many methodologies have been proposed in relation to the challenge posed by the semi-automated extraction of road network using remote sensing techniques. This difficulty can be seen in the review of some works related to the theme and published in the last decade. The problem of semi-automated extraction from road regions (streets or highways) is widely analyzed by several areas of science and proposals based on genetic algorithms or specialist systems are increasingly recurrent. The non-

parametric methods appear as a strong trend. Therefore, the analyzes proposed in this project were carried out in a controlled testing and validation environment, where both classifiers received the same set of training samples, the same sets of attributes obtained through airborne images and with high spatial and radiometric resolution, as well as being validated from field truth image and through procedures consolidated in the literature, such as Kappa Coefficient. A simple and unrepeatable combination of available information plans allowed the analysis of influence of each layer on space of attributes and on performance of each of the classifiers.

1.1 SVM - Support Vector Machine

Some studies, in the last 20 years, demonstrate the great effectiveness of the image classifiers acquired by remote sensing based on the technique of Vector Support Machines or Support Vector Machines (HUANG et al., 2002; FOODY and MATHUR, 2004), mainly in high resolution images (DIXON and CANDADE, 2008) or other complex and noisy data.

Lorena and Carvalho (2007) describe a classifier based on machine learning techniques as classifiers based on investigative processes, as they are based on the principles of statistical inference. Vapnik (1999) says that machine learning algorithms aim to minimize errors, through construction of decision limits that allow for greater separation between classes. Originally, the Support Vector Machine (SVM) classifier was based on a model known as the “Best Margin Classifier” and considered only linearly separable data. Currently, these models start from the hypothesis of nonlinearly separable data entry, which gives rise to an N-dimensional space of attributes distinct

* Corresponding author

from original (MELONI, 2009). The non-linear kernel functions allow this adaptation for non-linear data sets, building an optimal space of higher dimensions, also called “characteristic space” or “attribute space”.

Even a binary classifier, based on the SVM method, can be adjusted to classify multiple classes. This is possible through association of several classifiers. For this, two ways can be used: one-against-one and one-against-all. In one-against-one, multiple classifiers can be arranged in $k(k-1) / 2$ SVM's algorithms, where k represents the number of classes. Thus, the input pairs of training sets need to be re-labeled and for each “classification problem”, where, each result will be considered a vote.

1.2 ANN – Artificial Neural Network

The high performance of the human brain in complex learning tasks inspired the elaboration of models of Artificial Neural Networks (ANN) and in an approximate way to functioning pattern of the human brain, these models are formed by a large amount of “cells”. That seek the best distribution of certain activation patterns (ZELL et al., 1995).

The RNA's architectures are organized in layers (BRONDINO, 1999). There are the simplest ones, which are formed by a group of neurons arranged in only one layer (perceptron) and those made up of several (feedforward). In this structure, the patterns are presented in the first layers, while intermediate layers do most of recognition process and output layers present the results obtained.

Therefore, the first layer of an ANN will be equivalent to the space of attributes or number of different layers of information available and the last will correspond to different classes of representation. Among these, lies the biggest challenge in using this method, which is the delimitation of the number of hidden layers and their respective nodes (GONZALES and WOODS, 2000).

And, in response to problems that include a high degree of non-linearity, the feednow RNA's stand out through algorithms known as “error backpropagation” (retro propagation), these networks correct error during the training itself, because they allow weight adjustments from one layer to another, as a way to minimize the sum of mean square error between expected and predicted output (BOCANEGRA, 2002).

The training of the “backpropagation” algorithm consists of two distinct phases and begins with arbitrary definition of weights at the network nodes. In initial phase, a training set is exposed to the network and propagated through different layers and for each processing element. In the last layers, both expected and predicted outputs are compared and errors are calculated. In the following phase, the reverse occurs, that is, the propagation returns to beginning and error is then passed on to each processing element for adjustments of initial weights and, considering a successful training, the average errors must decrease with increasing iterations until processing finds a convergence to a “stable” weight value.

2. DATASET

As previously described, the main objective of this project is to analyze the performance of some non-parametric methods in extraction of road regions in images of very high spatial

resolution and in view of increase in the attribute space. To achieve this goal, it was decided to use aerial survey images of very high spatial resolution, as well as the data obtained through laser profiling, from last mapping of the City of Salvador / BA. Final mapping results were used as orthophoto with 10cm geometric resolution and 16-bit radiometric and 4-band spectral: red (band 3), green (band 2), blue (band 1), infrared (band 4) corresponding to the sensor Ultracam-X and Lidar RIGEL VQ480 scanning with a density of 12 points per square meters and resampling for every 10 cm in its raster version. Two sections of this mapping were used in a situation of normal occupation (*Itaigara* neighborhood) and subnormal (*da Paz* neighborhood).



Figure 1. Itaigara neighborhood with road network in blue.

Both regions are characterized by a very dense occupation, having occurred in different ways, due to nature of this density:

- on one hand, we see a neighborhood with orderly occupation, very vertical and with high-income socioeconomic characteristics (*Itaigara*);
- on the other hand, we see a busy neighborhood in a disorganized manner, with little verticality and low socio-economic characteristics (*da Paz*).

The *da Paz* area was constituted by appropriation of urban space in an informal way, around 1980's. It is a heterogeneous landscape marked by a lack of infrastructure, common in peripheral neighborhoods, in addition to presenting a variety and complexity of targets that can be recognized as:

- In this scene, among classes of representation recognizable through visual interpretation of image, the following can be highlighted:
- Street regions with a very heterogeneous asphalt pavement (various dark, light tones and even partially or totally covered by “clay”);
- Various plant phytophysiognomies (creeping, shrub, tree, etc.);

- Different types of roofs / tiles (colonial, fiber cement, galvanized, etc.);
- Water tank of different colors;
- Some stretches of pond, etc.

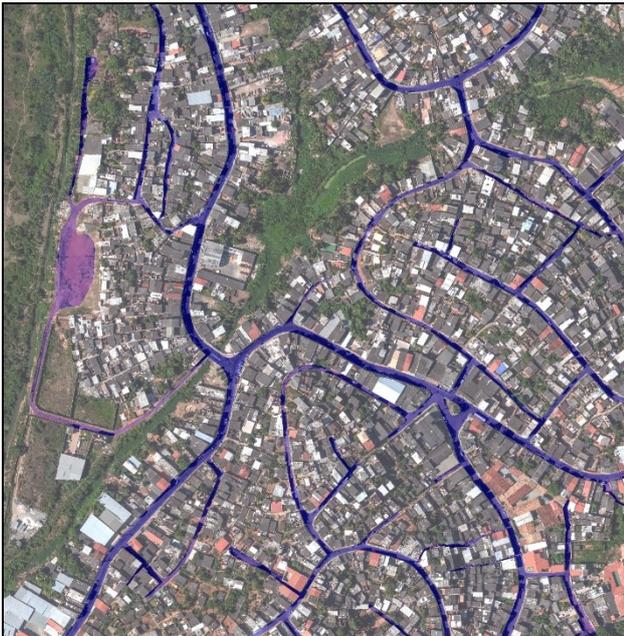


Figure 2. da Paz neighborhood with road network in blue.

3. METHODS

Initially, it was subtracted from Digital Surface Model (DSM) of the portion corresponding to the Digital Terrain Model (DTM). This operation sought to normalize DSM so that it presented only the information of objects height in relation to ground and not of these in relation to relief. Then, in the respective study areas, the training samples of representation classes necessary for isolation of “road” class, the target class of this research (positive class), were collected. In the same way, training samples were collected from classes of representation necessary for isolation of “non-road” class (negative class). The samples were used to train the chosen non-parametric classifiers (training sets for each different input image / landscape). All classifications will have the same set of training samples and the same classification parameters. However, with respect to samples, each landscape will have a certain set, while parameter settings will be the same for both landscapes. In this experiment, Kappa concordance coefficient was used, based on contingency matrices and the AUC (“Area Under Curve”), in order to verify classifications. The latter is the relationship between true positive rates versus rate of true negatives at different thresholds of classification. The general concept of this procedure was cascade refinement, analyzing behavior of the tested options and, from this analysis, proceed with a new round of tests, this time, with a smaller list of options than at the beginning of tests. These were divided according to existing library in Weka mining package - LibSVM and LibMultiLayerPerceptron, SVM and Neural Networks.

Parameters	valor
Kernel	Polynomial
Polynomial grade	2
Kernel function	1.000
Gamma	1/bands
Penalty	100.000
Pyramids	-
Limite de probabilidade	0.01

Table 1. SVM parameters obtained in Weka mining.

Parameters	value
Activation function	Logistic
Iterations	500
Hidden layers	-
Training threshold	0.9000
Training tax	0.2000
Training moment	0.9000
Output criteria	0.1000

Table 2. ANN parameters obtained in Weka mining.

The validation of classifications results was done by comparing them with thematic images of field truth that have only two classes of representation, “road” and “non-road”, just like the classifications and were generated through visual interpretation of corresponding images. to each study area, based on manual vectorization. Therefore, each field truth image served as a reference for a cross table, which generated a confusion matrix for each of experiments. These enabled the calculation of the Kappa coefficient of agreement, as well as overall accuracy index, omission and commission errors, producer accuracy and user accuracy for each classification.

4. RESULTS

The algorithms chosen for the experiments were: Support Vector Machine (SVM) and Neural Net or Artificial Neural Network (ANN), both implemented. The classification adjustment parameters for both classifiers were defined after simulations supported by data mining techniques, in addition to the recommendations indicated in the specialized literature. The definition of the parameters observed criteria that allowed, both the best performance in terms of classification results, as well as its generalization for each image. However, since these are supervised classifications, that is, methods that require a set of reference standards that will serve as a model for learning the predictions of the categories of representation, each image / landscape required stages of interpretation, definition and collection samples of the representation subclasses. Therefore, each image had its own set of training samples. However, in the case of classification parameters, the same adjustment parameters of a given model (SVM or ANN) were used for both images / landscapes. Therefore, 104 tests were performed.

Through simulations and considering the two images and their respective training sets, it was possible to notice that low correlation between the layers of information seems to produce some positive effect for the SVM method, in terms of classification results.

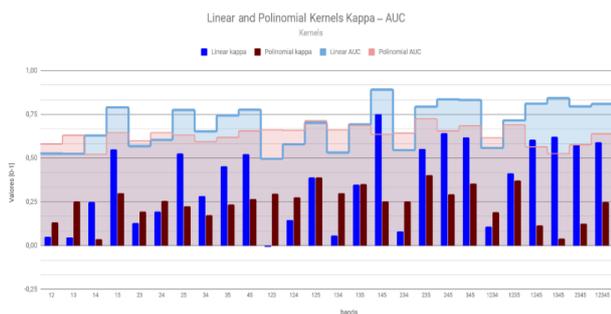


Figure 3. Performance of SVM kernels per band combination.

The “Linear” and “Polynomial” kernels obtained better results and demonstrated a behavior close to that of the previous landscape. The “Linear” kernel stood out in relation to the values achieved by the classification simulations (Kappa values), while the “Polynomial” kernel showed slightly lower values, however, a more homogeneous behavior. However, in this landscape, the values simulated with the “Linear” kernel behaved more irregularly, compared to the image of the Itaigara neighborhood. Through the simulations (Figure 3) and considering the two images and their respective training sets, it was possible to note that the low correlation between the layers of information seems to produce some positive effect for the SVM method, in terms of classification results. The differences between the tests of both landscapes did not show great variability, the “Polynomial” kernel proves to be the most appropriate for the real classifications to be performed, since, in the set 10, the tests suggest that this kernel has more predictable. The performance analysis of certain non-parametric algorithms based on change of the available attribute layers and the differences between the tests of both landscapes did not show great variability, the “Polynomial” kernel proves to be more appropriate for actual classifications to be performed, as, on the whole, the tests suggest that this kernel behaves more predictably, in addition to acceptable classification performance.

This step aimed to emulate the performance of the classification models, chosen for this work, using the algorithms available in the Weka library and related to the classifiers found. The association of the radiance information from the samples collected for training (in this case, represented by the DN – Digital Number) and definition of the prediction models of the positive (“Rua”) and negative (“non-Rua”) classes, considering the combinations described in Chart 8, made it possible to simulate the trends of the possible results that would be achieved using a certain set of adjustment values for the desired classification parameters. In this way, it was possible to “anticipate” which combination of parameters would improve the results of future classifications. In this experiment, the Kappa coefficient of agreement was used, based on contingency matrices and the AUC (“Area Under Curve”), with the objective of verifying the classifier. The latter is the relationship between true positive rates versus true negative rates at different thresholds. The concept of this procedure was the cascade refinement, based on more generic definitions (patterns found in the Weka mining package algorithms, for example), analyzing the behavior of the tested options and, based on this analysis, proceeding with a new round of tests, this time, a common list of options less than at the beginning of the tests. Therefore, this step fulfills the role of approaching a set of definitions more appropriate to the experiment and second a degenerative premise, because the focus of the project is the behavioral analysis, not the over-specialization of one or the other method.

simulations described above and some of the graphs generated for analysis. These are divided according to the library existing in the mining package Weka –LibSVM and LibMultiLayerPerceptron, SVM and Neural Networks respectively. For Neural Networks, it is important to emphasize that the focus of simulations analysis with LibMultiLayerPerceptron is concentrated, on the occasion of this study, in the adjustments with value “0” of hidden layers, as already commented. The increase in layers of information seems to contribute more to the performance of the Neural Networks method than the low correlation between them.

Thematic images of field truth supported the stage of validation of the results. These images have only two classes of representation, “Rua” and “non-Rua” and were generated through the visual interpretation of the images corresponding to each study area and from manual vectorization (QGIS software). Each field truth image served as a reference for a cross table, which generated a confusion matrix for each of the experiments. These made it possible to calculate the Kappa coefficient of agreement, as well as the overall accuracy index, omission and commission errors, producer accuracy and user accuracy for each test.

The global accuracy or global accuracy of a classification, in this case, concerns the relationship between pixels correctly predicted and the total pixels of an image. Therefore, the global accuracy values obtained by SVM classifications, for Itaigara and for da Paz neighborhoods, can be considered satisfactory and very little. However, the behaviors of classifications in relation to different metrics for evaluating the results obtained by SVM and Neural Network classifications, demonstrate that global accuracy is an index that overestimates accuracy of the classification, considering only ratio of pixels correctly classified in relation to total pixels.

Therefore, a more sensitive index should be used than global accuracy, as well as the variation of omission and commission errors of t “road” and “non-road” classes. The Kappa coefficient of agreement is a multivariate analysis and more sensitive to these variations, as figure 4 and figure 5 show graphs corresponding images of neighborhood landscapes of Itaigara and da Paz, respectively.

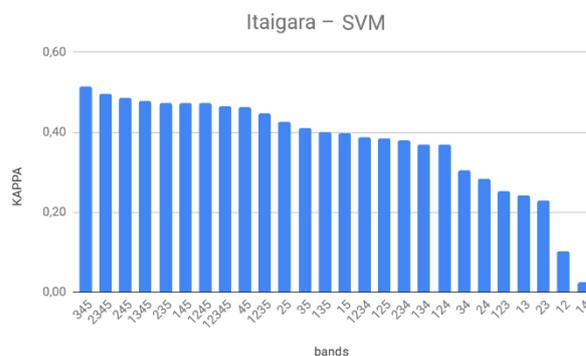


Figure 4. Variation of the cover index due to the combination of bands in the Itaigara.

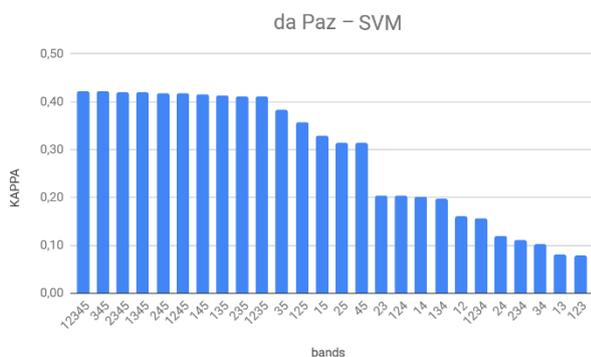


Figure 5. Variation of the cover index due to the combination of bands in *da Paz*.

Although combination of bands with the best performance for *da Paz* landscape was B12345, in general, the lower correlation between the layers of information, the better performance of SVM method.

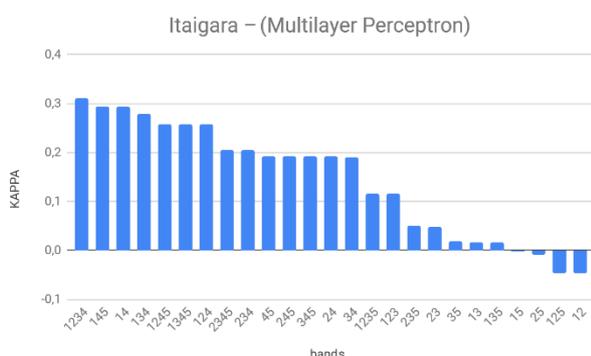


Figure 6. Variation of the cover index due to the combination of bands in *Itaigara*.

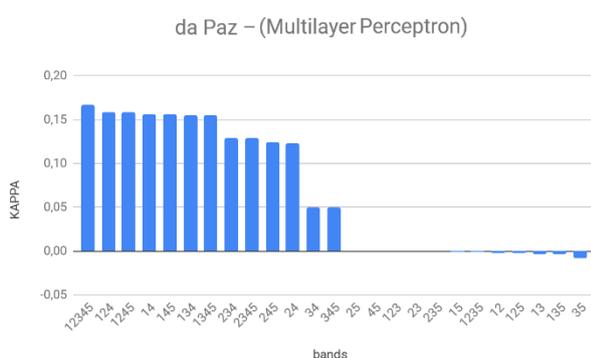


Figure 7. Variation of the cover index due to the combination of bands in *da Paz*.

Figure 6 and figure 7 show the results achieved by Neural Networks method, for neighborhood landscapes of Itaigara and *da Paz*, respectively. Band 4 (near infrared) presents itself as a minimum threshold of correlation between bands without prejudice to the performance of the method since, the apparent inverse relationship between bands and the classifier performance, for both landscapes, does not exert the same influence it had for SVM method. The interdependence between the bands proved to be necessary to Neural Networks method and can explain negative highlight for band 5 (altimetric

information), since this is the least correlated and interdependent among all.

5. CONCLUSION

The tables and graphs presented in the previous section show that the classifications show improvements according to the presence or absence of certain layers of information. The profile of land use and occupation also appears to exert influence, however, in a secondary way.

In the case of SVM method, in general, the results demonstrated that there may be a direct relationship between altimetric information band of targets in relation to terrain (band 05) and improvement of their performance and, therefore, the lower degree of correlation between the bands. It can also be considered a factor that has a positive influence. As for Neural Networks, the results of experiment demonstrate that the presence of the near infrared band (band 04) was decisive for improving performance of certain combinations in relation to others. The elevation band (band 05), in this case, in addition to not having a positive influence, is present in most of combinations with the worst performance. This fact may point to a “limit” to learning of that method and low correlation between the layers of information indicates that there is a balance between interdependence of bands and their correlation.

In order to create a controlled environment and where it was possible to compare two pattern recognition methods in different images, for two different landscapes, it was necessary to restrict both the size and the diversity of subgroups of the samples. However, the relative size of the samples, as well as the way they are taken, can affect both data mining and the results of the final classifications.

Although the algorithms used during the data mining stage to emulate the results of the classifications are based on the same source codes as the classifiers themselves, the results obtained in those simulations diverged, in part, from the real results achieved by the classifications. This divergence can be explored and, with this, propose better adjustments of the classification parameters in the future.

ACKNOWLEDGEMENTS

I thank the Research Group on Spatial Data Representation Analysis (GEPARD) and the Graduate Program in Civil Engineering (PPEC), for their collaboration and availability, as well as the Federal University of Bahia.

REFERENCES

Alonso, M. C.; Malpica, J. A. **Satellite imagery classification with LiDAR data**. *Trees (A)*, v. 2106, n. 19, p. 11, 2010.

Bradley, Andrew P. **The use of the area under the ROC curve in the evaluation of machine learning algorithms**. *Pattern recognition*, v. 30, n. 7, p. 1145-1159, 1997.

Bocanegra, C. W. R. **Procedimentos para tornar mais efetivo o uso das redes neurais artificiais em planejamento de transportes**. São Carlos, v. 97, 2002.

Brondino, N. C. M. **Estudo da influência da acessibilidade no valor de lotes urbanos através do uso de redes neurais**. São Carlos-SP, 1999.

- Congalton, Russell G. **A review of assessing the accuracy of classifications of remotely sensed data.** Remote sensing of environment, v. 37, n. 1, p. 35-46, 1991.
- Dixon, B. and Candade, N. **Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?.** International Journal of Remote Sensing, v. 29, n. 4, p. 1185–1206, 2008.
- Foody, G. M.; Mathur, A. **Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification.** Remote Sensing of Environment, v. 93, n. 1-2, p. 107-117, 2004.
- Gonzalez, R. C. and Woods, R. C. **Processamento digital de imagens.** Pearson Educación, 2009.
- Huang, J.; Blanz, V.; Heisele, B. **Face recognition using component-based SVM classification and morphable models.** In: Pattern Recognition with Support Vector Machines. Springer, Berlin, Heidelberg, 2002. p. 334-341.
- Hudson, William D. **Correct formulation of the kappa coefficient of agreement.** Photogrammetric engineering and remote sensing, v. 53, n. 4, p. 421-422, 1987.
- Lorena, A. C.; Carvalho, A. C. P. L. F. **Uma Introdução às Support Vectors Machines - Tutorial.** Revista de Informática Teórica e Aplicada, v. 14, n. 2, p. 43-67, 2007.
- Liu, Li; Lim, Samsung. **Color Component-Based Road Feature Extraction from Airborne Lidar and Imaging Data Sets.** Journal of Surveying Engineering, v. 143, n. 1, p. 04016021, 2016.
- Matkan, Ali Akbar; Hajeb, Mohammad; Sadeghian, Saeed. **Road extraction from lidar data using support vector machine classification.** Photogrammetric Engineering & Remote Sensing, v. 80, n. 5, p. 409-422, 2014.
- Maboudi, Mehdi et al. **Object-based road extraction from satellite images using ant colony optimization.** International Journal of Remote Sensing, v. 38, n. 1, p. 179-198, 2017.
- Meloni, R. B.. **Classificação de Imagens de Sensoriamento Remoto usando SVM.** 2009. Tese de Doutorado. PUC-Rio.
- Mena, Juan B. **State of the art on automatic road extraction for GIS update: a novel classification.** Pattern Recognition Letters, v. 24, n. 16, p. 3037-3058, 2003.
- Song, Wenbo et al. **Automated geospatial conflation of vector road maps to high resolution imagery.** IEEE Transactions on image processing, v. 18, n. 2, p. 388-400, 2009.
- Vapnik, V. (1999). **An overview of statistical learning theory,** IEEE Transactions on Neural Networks 10:988-999.
- Zell, A., and Coauthors, 1995: SNNS: Stuttgart. **Neural Network Simulator.** User's manual, version 4.1, University of StuttgartRep. 6/95, 350 pp. [Available online at <http://www-ra.informatik.uni-tuebingen.de/SNNS/>].
- Zeng, Y.; Zhang, J.; Wang, G.; Lin, Z. **Urban land-use classification using integrated airborne laser scanning data and high resolution multi-spectral satellite imagery.** In: Proceedings of the ISPRS Commission I Symposium 2002. Denver, CO: ISPRS, 2002.
- Ziems, Marcel; Rottensteiner, Franz; Heipke, Christian. **Verification of road databases using multiple road models.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 130, p. 44-62, 2017.