Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Brasil, 21-26 abril 2007, INPE, p. 1205-1210.

Linear features detection in CCD/CBERS-2 image using neural network

Carlos Frederico de Sá Volotão [1]
Claudio Gelelete [2]

[1] 5ª Divisão de Levantamento – 5DL
Rua Maj Daemon, 81 – Rio de Janeiro - RJ, Brasil
volotao@hotmail.com

[2] Instituto Militar de Engenharia - IME
Pça Gen Tibúrcio, 80 – Rio de Janeiro - RJ, Brasil
cgelelete@yahoo.com.br

Abstract. This work presents a method to detect linear features in remote sensing images using feed-forward neural network and pixel contextual information. The net is trained using a 5x5 window. The input layer includes information extracted from the pixel and its neighborhood, including geometric details. A CCD/CBERS-2 band has been tested. The results indicate the feasibility of this method to detect linear features and contour detection together.

Keywords: backpropagation neural network, contextual information, linear detection, border detection, CCD/CBERS-2 image.

1. Introduction

Data processing involves the use of softcomputing techniques, including artificial neural networks, genetic algorithms and fuzzy technology. Although designed for real-world data processing tasks, these techniques are not yet fully penetrated into the world of remote sensing (Backer, 2002). The brain mechanism inspired the neural networks and can solve several sorts of problems such as classification of images. Menendez (2001) explains the object recognition process for images using neural networks and the neural networks fundamentals is presented by Haykin (2001).

Neural networks can achieve texture information (Greenspan and Goodman, 1993; Berberoglu et al., 2000). A single band can let to detect roads as one class, but Bosch (1999) results needed to add road information to infer membership.

Mena (2003) presents a bibliographic surveying of automatic road extraction from aerial and satellite imagery. Fischler et al. (1981) presented a technique for linear structure and road detection in low-resolution aerial imagery based only on the detected geometry, using a generic representation. Three-dimensional objects are not considered in this kind of approach and some works makes topology inferences by, for example, closing gaps (Mayer et al., 1998).

Road detection and extraction is used by mapping agencies to maintain the map heap up-to-date. “Arial and satellite images are promising data sources for map generation and update of available maps to support activities and missions of government agencies and consumers. Full exploitation of these data sources depends on automatic techniques for object extraction from satellite and aerial imageries”(Mokhtarzade and Zoej, 2006).

Satellite images are becoming available at lower cost and can be used for several purposes. The 30m resolution remote sensing imagery may be applied in cartography works if features are precisely detected. We are proposing a linear context input schema to detect image features for mapping purposes using artificial intelligence. This work presents a single CCD/CBERS 2 band implementation of feed-forward neural network to indicate whether this approach meets its goal. The detection of linear features with the procedures presented in this work will indicate that we can define a new input layer using multispectral data (as in Gelelete and Volotão, 2006) to achieve a better classification mechanism.

The use of other bands may be used to define imprecise spectral analysis and, for the method presented, we would like to detect the linear features, including rivers and roads. The limitations we have are mainly related to the output classes' definition and the training datasets.

Gelelete and Volotão (2006) presented an artificial neural network (ANN) and classified two kinds of land cover, land use, clouds, shadows and water classes in a multispectral 7-band Landsat image, comparing the results with a maximum likelihood classifier result, both classifiers using the same training set. The neural network gave the best results in generalization and accuracy aspects by visual investigation and it uses 3x3 and 7x7 pixel neighborhoods to achieve contextual information.

The linear subpixel-wide features like roads were not intended to be well detected because there was no information making the net able to understand this class, therefore the goal of the present work is to continue the work and make the ANN classify linear features, i.e., straight and curved segments, but it is limited by the use of one sole band. This was expected to improve the road classification deficiency presented by Gelelete and Volotão (2006) by enhancing the detection of the contextual information shape, i.e., the linear detection.

The limitation of the use of one instead of multiple bands should reduce significantly the computer processor time but the ability of detecting from distinct objects should become weak too. In the present work one can think there is a path crossing each image pixel and it needs to quantify how this path differs from its neighbors by a number of indexes. So we trained the net with the classes' indexes. Finally we compared all indexes for each image pixel with the trained net to allow the classification.

2. Input data and output classes

A CCD/CBERS-2 image was used in this work and we selected a 2000x2000 window because the computations proposed by this method and all the combination programmed for the training phase would be time consumer (Mokhtarzade and Zoej, 2006 used a 550x550 window for the same reasons). We did several adjusts using smaller images to analyze the detection of the linear class and using a 3 GHz Pentium 4 computer the final training, as described in the next section, it took about 16 hours.

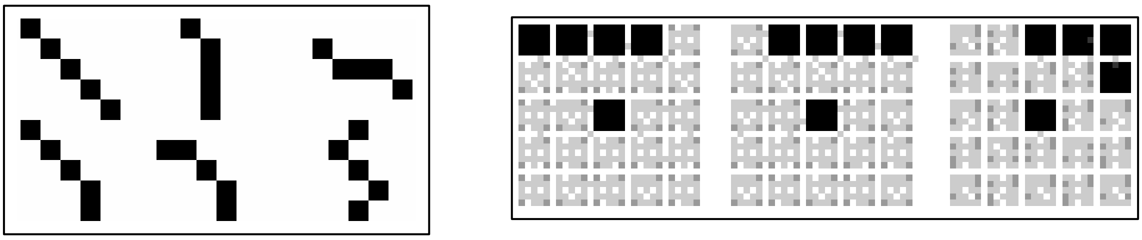
The 6 output classes were initially selected as being: (a) vegetation, (b) urban area, (c) water, (d) clouds, (e) shadows and (f) roads. It is important to consider that the collection of regions of interest was applied to one sole band, and therefore the digital number in several cases are not enough to detect the targets as well as the multispectral data do.

As getting the first results we noticed the shadows and the water was very close and we changed our sense in associating these both classes, for this one-band work, as homogeneous and dark areas. There are subtle characteristics among them but we preserved the number 6 for the output neurons, giving the classes a second corrected denomination: (a) heterogeneous area type I, (b) heterogeneous area type II, (c) homogeneous dark area type I, (d) homogeneous dark area type II, (e) homogeneous light area and (f) lines and contours, respectively.

We start designing the network with 6 classes but our focus was almost over the 6th class. The next problem was to define the inputs. About the input we made some statements: direction sensibility is undesirable; the pixel DN value should be used, but only in some cases; the neighborhood mean or median value may indicate if the pixel is darker or lighter;

heterogeneous and homogeneous area distinction is important and the relative context value represented by the neighborhood standard deviation may be used; the linear feature detection needs a highly sensible context analysis to indicate this kind of occurrence.

The input used was: the pixel DN, the 5x5 neighborhood mean value; the 5x5 neighborhood standard deviation; 4 entries corresponding to one of the 32 paths defined by the most similar values from border of the 5x5 neighborhood to the center and the path with biggest variance, using the path mean and deviation values for both paths; 4 additional entries corresponding to one of the 36 paths defined de same way but with linear curves and lines that crosses the central pixel, i.e., the pixel being analyzed; 4 entries corresponding to one of the 16 possible groups composed by 4 consecutive border pixels and defined by comparing its maximum and minimum variance values and computing the deviation and the mean values. Figures 1



(a) 6 examples from the 36 possibilities representing the possible paths crossing the reference (central) pixel.

(b) 3 in 16 possibilities of the 4 border sector pixels used to inform the net if there is any line crossing the 5x5 neighborhood.

Figure 1. Input geometry of pixels (the central pixel is always represented in black).

From the image we collected several pixel positions for each class corresponding to a minimum of 1000 points using the ENVI system. We produced one file for each class and used the above mentioned entries for training. All other phases involved were developed in MATLAB environment. We normalized all entries by using the mean and deviation values obtained from the training phase for each data entry.

3. Training

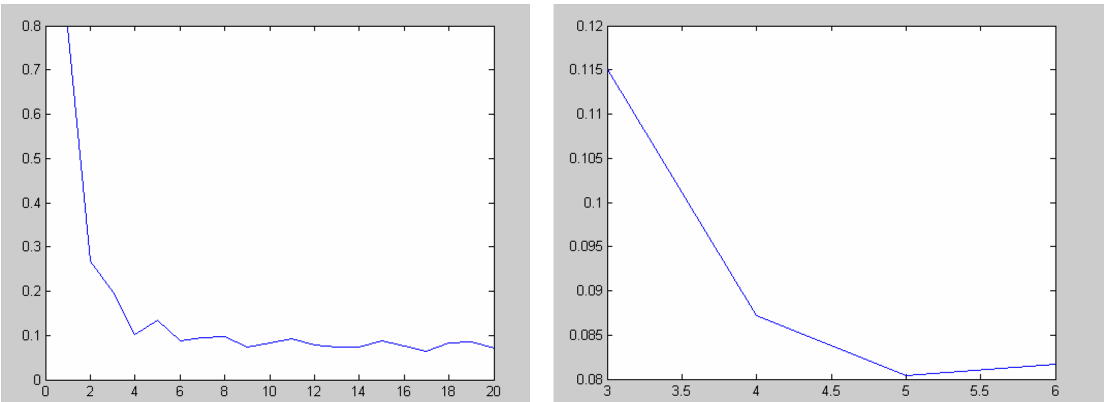
Initially we defined the linear features region of interest by all image pixels presenting radiometric flow caused by the linear features visually detected. The selected pixels were mixed with several surrounding objects and sometimes we selected about three pixels wide regions that were surrounding the linear feature desired. As both sides of the border were included in the same region of interest, the region became wider than the expected and as result this class became the dominant, including most of other targets pixels. This was not the intended result but indicated we should refine the regions of interest.

Starting the new region collection with a narrower definition of the linear feature pixels was hard because we had to take the decision about the best representing pixel, i.e., with big zoom we needed to find the best representing of an one-pixel wide curved or straight line that was best distinguishable from the background neighborhood. Sometimes it occurred that one side was darker and the other side was lighter than the background and we had to decide what best represented the line.

To understand this phenomenon the pixel must be viewed as a compound of two main factors: spatial geometry and electromagnetic radiation spreading. The spatial resolution can be easily calculated theoretically but the geometric calculus doesn’t compute the spreading phenomena as it occurs in real life. The spreading effect can be analyzed into parts and by the point spread function (PSF). The SPF is as a spectrally dependent function that introduces the blur from a variety of sources in the image formation process and it is a combination of atmosphere, platform, optics, detector and electronics effects. Ideally, each of these atmospheric turbulence, platform jitter, optics, detector and electronics effects would be in a flexible and functional form (Scigaj et al., 2004; Schott, 1996).

Mokhtarzade and Zoej (2006) presented a methodology for road detection using one single class and they made the tuning by computing several coefficients and the overall accuracy, but using a 3x3 neighborhood for a 3-band RGB high-resolution image. The number of neurons needed in the hidden layer was decreasing as increasing the number of neurons in the input layer. This happens because the detection ability of the network becoming enhanced by the input layer makes the network to become specialized and the number of hidden layer size was changed to prevent over-training problem.

The number of neurons to be used indicated in Fig.2a was about 3 to 6. After several training using distinct seeds and several repetitions the optimal number of neurons was 5 (Fig. 2b).



(a) (b)

Figure 2. Network hidden layer size definition.

4 . Experimental Results

A variety of networks with different neuron numbers in hidden layer were used. Each network defined was trained several times using different seed numbers. The original image, the final network classification and the precision assessment is shown in Fig. 3. We can see the roads are well defined. The ability of this net is proven efficient for detecting road pixels. The imprecision map shows us the less precise classes detected was clouds and urban area, because theses areas are shown in light gray color in the right image.

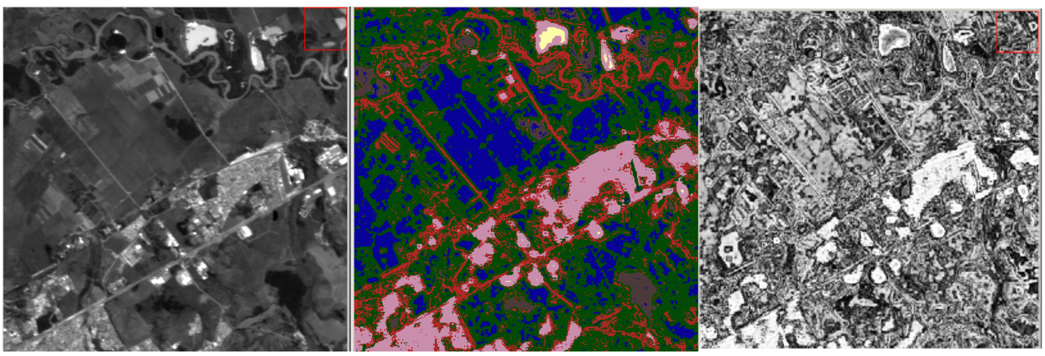


Figure 3. Region of original band, classified image and normalized imprecision map.

The borders are detailed in the zoom presented in Fig. 4, where the river was divided in two classes, one corresponding to the river dark and homogeneous central area and the other the river border. The transition region between classes was classified in the same class as the roads. The same class was attributed to the clouds' borders. This means the road and the river contour is related to a linear and border class, respectively road and shadow as it would be by the first denomination. The color table associated to the classes in the Fig. 3 was optimized to the visualization and is different from the one used below.

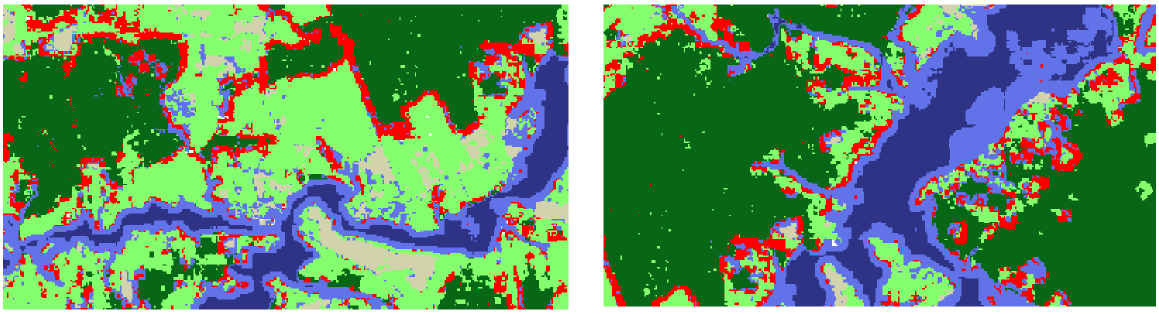


Figure 4. Neural network border classification details.

Fig. 5 shows the segmentation of the image by the linear class. The clouds' contours were classified as linear feature. Some pixels in the border of two different classes were classified as contour. The shadows were not detected with precision and the urban area is correctly represented as well as the clouds.

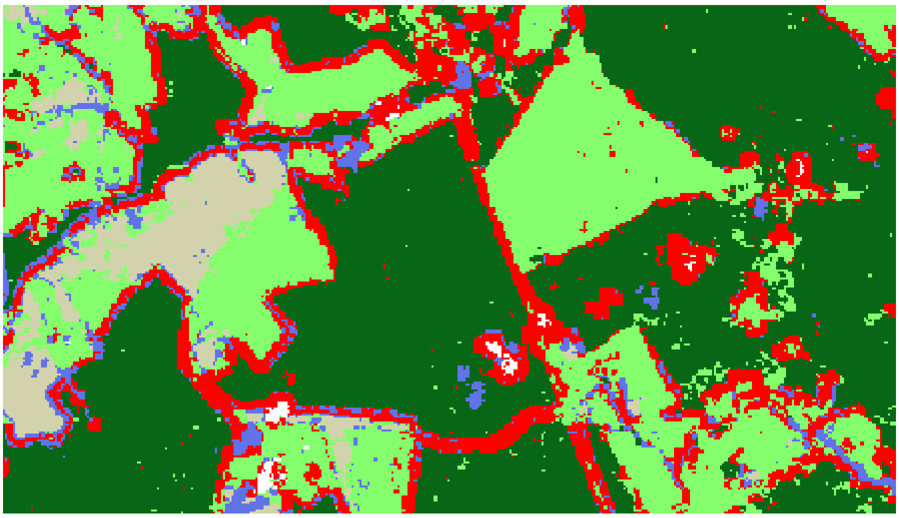


Figure 5. Neural network linear and border classification.

5. Conclusions

It has been presented a neural network contextual classification method that considers neighborhood geometry and statistics. It was used only one band of a CCD/CBERS-2 image to demonstrate the potential of the method.

The present work shows the feasibility of using line geometry from the pixel neighbors to give linear surrounded information to the network, but some optimization must be done to reduce the computer time involved with the present computer capacity.

Using other bands in input vector for the network training could be a natural enhancement for further study in order to improve network's ability in linear feature detection. The type and number of output classes may be studied in a way to result in the desired output. In resume, to improve the classification quality one could work in two main aspects: training data refinements and tuning the context algorithm to achieve specific results.

References

Backer, S. D. **Unsupervised Pattern Recognition: dimensionality reduction and classification**, Doctoral Thesis, University of Antwerp, 152p., Antwerp, 2002.

Menendez, A. **Image recognition using neural networks**, RTC Magazine, July 2001.

Haykin, S. **Redes Neurais: princípios e prática**. Trad. Paulo Martins Engel. 2.ed. Porto Alegre: Bookman, 2001.

Berberoglu, S.; Lloyd, C. D.; Atkinson, P. M.; Curran, P. J. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean, Elsevier Science Ltd, **Computers and Geosciences**, v. 26, pp. 385-396, 2000.

Bosch, E. H. **Texture information and supervised classification of hyperspectral imagery by means of neural networks**. In GeoComputation 99, 1999. Online http://www.geocomputation.org/1999/019/gc_019.htm

Fischler, M.A.; Tenenbaum, J.M.; Wolf, H.C. "Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. **Computer Graphics and Image Processing**, 15:201-223, 1981.

Greenspan, H. K.; Goodman, R. "Remote sensing image analysis via a texture classification neural network," in Advances in Neural Information Processing Systems 5, Hanson, S. J.; Cowan, J. D.; Giles, C. L. Eds. Morgan Kaufmann, pp. 425–432, San Mateo, 1993.

Mayer, H.; Laptev, I.; Baumgartner, A. "Multi-scale and snakes for automatic road extraction", in **Fifth European Conference on Computer Vision**, volume II, pages 720-733, 1998. http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MAYER/cvonline.html

Mokhtarzade, M; Valadan Zoej, M.J. "Road detection from high-resolution satellite images using artificial neural networks", in: **International Journal of Applied Earth Observationand Geoinformation,** 2006 (article in press, doi:10.1016/ j.jag.2006.05.001).

Schott, J. R. **Remote Sensing: The Image Chain Approach**, Oxford University Press, 424 pp. 1996.

Scigaj, C. ; Schott, J.; Brown, S.; Lee, P.; Faulring, J.; Sanders, N. **Sensor Modeling in DIRSIG**. 2004. Online < www.cis.rit.edu/~bmd1603/media/scigaj_lass04.ppt>

Gelelete, C.; Volotao, C. F. S. **Análise do potencial de classificação do uso e cobertura do solo através de redes neurais**. Submetido ao XIII SBSR. 2006.