

OPEN ACCESS

Neural network-based segmentation of satellite imagery for estimating house cluster of an urban settlement from Google Earth images

To cite this article: P D Wardaya and S Ridha 2014 *IOP Conf. Ser.: Earth Environ. Sci.* **18** 012019

View the [article online](#) for updates and enhancements.

You may also like

- [Research on Road Route Selection Design Based on Google Earth Intelligent System](#)
Xuetao Zhao
- [Spatial Temporal Land Use Change Detection Using Google Earth Data](#)
Adi Wibowo, Khairulmaini Osman Salleh, F.Th.R Sitanala Frans et al.
- [A review of Google Earth Engine application in mapping aquaculture ponds](#)
Arvinth Rajandran, Mou Leong Tan, Narimah Samat et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Neural network-based segmentation of satellite imagery for estimating house cluster of an urban settlement from Google Earth images

P D Wardaya¹ and S Ridha

Geosciences and Petroleum Engineering Faculty, Universiti Teknologi PETRONAS,
Tronoh, Perak Darul Ridzuan 31750, Malaysia

E-mail: ponk_2x@yahoo.co.id

Abstract. In this paper a backpropagation neural network is utilized to perform house cluster segmentation from Google Earth data. The algorithm is subjected to identify houses in the image based on the RGB pattern within each pixel. Training data is given through cropping selection for a target that is a house cluster and a non object. The algorithm assigns 1 to a pixel belong to a class of object and 0 to a class of non object. The resulting outcome, a binary image, is then utilized to perform quantification to estimate the number of house clusters. The number of the hidden layer is varying in order to find its effect to the neural network performance and total computational time.

1. Introduction

Object segmentation is one of the being actively investigated topics in remote sensing. Here, the segmentation term discussed is which related or similar to the object extraction or object detection (identification). The process is basically aiming at extracting the desired object from the un-desired surroundings. This is an important process as the object in the original data (e.g. satellite airborne imagery) has the continuous characteristics (the pixel values are continuously distributed) that make the direct quantification and characterization difficult to undertake. Performing the segmentation process allows one to isolate the object from its surrounding and therefore further action is an easier task.

In object segmentation, the algorithm used basically performs the two class classification process. It classifies pixel values of the images to be in the class of object or non object (surrounding). Usually, the final output is the binary images in which pixel belongs to object has value 1 meanwhile the surrounding has value 0. In display, this binary image is generally presented as the black and white, where white is object and black is for non object. The choice of which the algorithm decides to classify whether a certain pixel belongs to object or surrounding is the core operation of the process and is part of the machine learning subject.

Numerous methods exist for the object segmentation. Among them, artificial neural network (NN) is one of the most widely used techniques. It is likely applying the broader pattern recognition subject in the specific context of image processing analysis. In numerous pattern recognition-image processing literatures, neural network have been intensively used for solving problems related to image segmentation. These are for example can be found in ref [1-3]. Neural network variants including multilayer perceptron, radial basis function, hopfield network, and self organizing map have been widely used in remote sensing image classification task. As the remote sensing classifier, Tso and Mather [4], called the neural network variants to be more powerful and attractive than their neighbour, the statistical classifier (e.g. Maximum Likelihood or principal component analysis). The

¹ To whom any correspondence should be addressed.



attraction of the neural network lies on the fact that first, based on the existing literatures, it outperforms the statistical classifier in term of accuracy, and second it requires no assumption on the statistical data distribution. The latter is crucial as for instance, when the training data is limited, equal estimate of statistical parameter (although difficult to obtain) is needed as in the statistical classifier [4].

In this study, neural network is utilized to perform the house segmentation on the satellite imagery acquired from the Google EarthTM which the specific target is the urban settlement. The supervised method with the multilayer perceptron architecture by using backpropagation learning technique is employed to perform the two class classification, namely object and non-object (surrounding). Here object refers the house cluster and non-object refers to other image constituents. The objective of this study is to segment or extract the house cluster in the image to deliver the binary image containing only the classified house clusters. The resulting binary image is then quantified to estimate the number of the clusters. The performance of the neural network is also analyzed with the varying number of hidden layer. As for comparison to the previous researches, in the remote sensing literatures, the works on object segmentation (detection), especially for building-related object can be found in ref [5-7].

This study however, instead of being intended to reproduce or retake the similar investigation as in the existing literatures, brings the new insight for performing the relatively simple but accurate image analysis with the use of freely accessible Google EarthTM application. The author expectation is that by employing the proposed technique one may easily characterize or quantify the desired object (e.g. man-made structure or natural object) captured as the RGB image from this widely used application. The overall motivations of this study, including the choice of the neural network as the core analysis tool is for delivering the method that meets three criteria; free, fast, and fascinating. Free in term of the freely obtained data, fast in term of computational time, and fascinating in term of the outcome accuracy.

2. Algorithm Description

2.1. Backpropagation Neural Network : Introduction and Concept

Neural network is developed to mimic the use of human brain for solving various problems which is inefficient when being undertaken by human, especially when dealing with large amount of data. Such topic that can be addressed with the neural network is the pattern recognition. Basically, for pattern recognition task, neural network works as how the human eye-brain or ear-brain system works. It gathers the information with specific pattern or characteristic, learns it, stores the learning subject, and then uses what've been learned to analyse the other information. This is similar to the process of the brain when for example a child is classifying the banana among the other fruits. Brain gathers information of the banana and friends that can be colour or shape, then the brain learns this characteristics, stores them in the memory, and finally at the other day this child is able to pick banana lying among the other fruits.

The neural network is generally non-parametric; it requires no assumptions regarding to the statistical distribution of the data. How well it performs a task depends greatly on the training process given. The choice of the important parameters including architecture and learning rate is on the user hand, which in turn determine the overall performance of the neural network. Unfortunately, as noted by [4], there are no clear rules with these parameters choice but only a heuristic rule able to be used. Further, [8] reports a survey of the influence of these parameters on the overall network's ability. According to the report, it is noted that backpropagation neural network beside the input and output layer requires one hidden layer whose number of neuron on it greatly affect the network performance.

Backpropagation neural network is probably the most known and widely used neural network model among the others. The term Backpropagation actually refers to the learning method used in the multilayer neural network model. Thus, saying backpropagation neural network is meaning multilayer perceptron neural network as these terms can be interchanging each other. The typical backpropagation neural network architecture with only one hidden layer is illustrated in figure 1. In this figure, input layer, which performs no computation receives 4 inputs (training sample) and then passes it to the hidden layer where the core computations are performed. The neuron activity is updated sequentially through the input layer to the output layer by using a certain mapping function.

This process is known as the forward step. During this step the neuron activities are updated layer by layer, starting at the input layer to the output layer. This aims to generate output, that is the activation of the neurons in output layer.

When the forward step accomplished, the activities of the output neurons are compared with targeted (or expected) activities. The differences are called error that in turn to be distributed backwardly to the network. This backward step starts at the output layer where the weights begins to be updated based on the error. This weight update is necessary to reduce the identification error of the network.

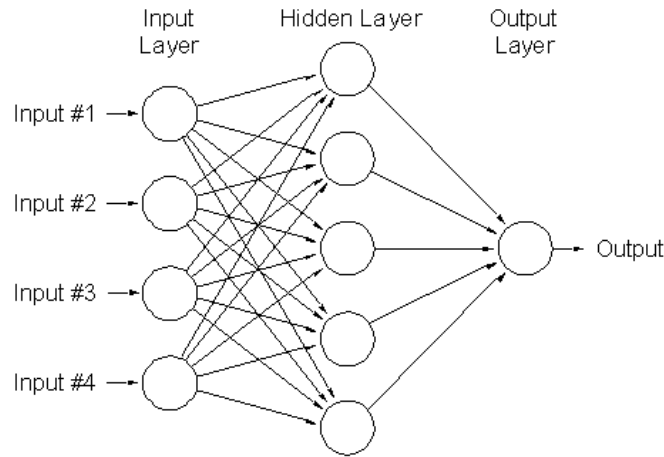


Figure 1. The typical multilayer perceptron with one hidden layer

Mathematically, the processes are described below. S_j denote the input received by neuron j , that reads

$$S_j = \sum_i w_{ij} a_i \quad (1)$$

Where a_j is the activities of the neuron I , and w_{ij} is the weight connection between neurons. This input value is then converted to the output value by using a certain mapping function, for example sigmoid function that reads

$$a_{j_m} = \frac{1}{(1 + e^{-S_{j_m}})} \quad (2)$$

The process of distributing forwardly the input is by updating the activities of the neurons. Once it is done, the error is calculated and expressed as

$$E_p = \frac{1}{2} \sum_n (t_{j_n} - a_{j_n})^2 \quad (3)$$

The equation (3) is needed to be minimized, therefore finding

$$\Delta w_{ij_x} = -\varepsilon \frac{\delta E}{\delta w_{ij}} \quad (4)$$

The gradient term in the right part can be expanded further to obtain

$$\frac{\delta E_p}{\delta w_{ij_x}} = \frac{\delta E_p}{\delta a_{j_z}} \frac{\delta a_{j_z}}{\delta w_{ij_x}} \quad (5)$$

The first term in the right part, by calculating the derivatives of equation (3) with respect to a_{j_n} is written as

$$\frac{\delta E_p}{\delta a_{j_z}} = (2) \left(\frac{1}{2} \right) (t_{j_z} - a_{j_z}) (-1) = -(t_{j_z} - a_{j_z}) \quad (6)$$

On the other hand, the second term reads,

$$\frac{\delta a_{j_x}}{\delta w_{ij_x}} = \frac{\delta}{\delta w_{ij_x}} \sum_n (w_{ij_n} a_{i_n}) = \frac{\delta}{\delta w_{ij_x}} (w_{ij_0} a_{i_0} + w_{ij_1} a_{i_1} \dots w_{ij_n} a_{i_n}) = a_{i_x} \quad (7)$$

The final expression of the modified weight function is obtaining by substituting equation (4), (5), and (6) to equation (4)

$$\Delta w_{ij_x} = -\varepsilon \frac{\delta E}{\delta w_{ij}} = \varepsilon \delta a_{i_x} \quad (8)$$

2.2. Data

The data for this study is acquired by using the free application Google EarthTM. Specifically, the studied location is the modern urban settlement, namely Bandar University Seri Iskandar. It is situated in Seri Sikandar Town, district capital of Perak Tengah in State Perak, Malaysia. The settlement was established in 2001 and according to Google Earth information the image acquisition was undertaken by satellite NOAA in 2007. The view of the data is presented in figure 2 below. In the figure, four possible classes are known namely, house, vegetation, road (asphalt), and ex-mining soil ground. The first is the desired object while the rest is regarded as the non-object.



Figure 2. The study area, a 1360x671 pixel RGB image.

3. Result and Discussion

3.1. Neural network performance and effect of the number of hidden layer

The performance of the neural network for a varying number of hidden layers is plotted and displayed in the figure 3 below. When only one hidden layer used the network reaches convergence relatively fast. The minimum MSE has been achieved in its 18th iteration out of 25. Increasing the number of hidden layer theoretically will increase the computational time but in contrary give the better outcome. By maintaining the training data constant in term of size, number, and selection, one hidden layer spends approximately 22 second, three hidden layers 26 second, and 5 hidden layers 56 second. This seems meets theoretical expectation that when the more the hidden layer involved the computational time will greatly increases. It is true but the increase is not linear. Thus we suggest that 5 hidden layers generate the so called *overfitting* which is associated with the failure in identifying the training data due to too many neurons involved [8].

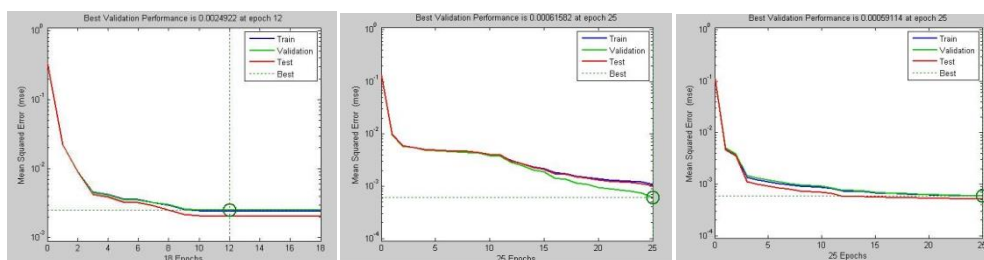


Figure 3. Convergence (MSE) performance plot of the network. Left: 1 hidden layer, middle: 3 hidden layers, right: 5 hidden layers

3.2. Segmentation Result

Figure 4 displays the segmentation result with the varying number of hidden layers. Overall, the network is able to extract the house cluster based only on the RGB value pattern in the image. The left and middle figure that is obtained by using one and three hidden layer show significant error that is classifying the non object as the house which in turn is also computed as the object. This contributes in the error in the house cluster calculation. One hidden layer relatively gives the optimal result in segmenting the left and middle part of the house cluster, indicated by the clear segmentation within house cluster in the region. However, it fails to do so for the rightmost part house cluster. It segments the originally 8 house clusters in this region into more than 12 smaller objects. In this case the network is called unable to recognize the regularities on the corresponding pixel.

On the other hand, by using three hidden layer that being expected to give the more accurate outcomes the network fails to give the same result, especially for left and middle region of house cluster. However, on contrary, the rightmost part is segmented clearly. Further, miss-classification is less than the previous one hidden layer. Finally, the three hidden layer network in the right figure generates the better segmentation that the forerunners in the left and middle region. As compared the one hidden layer, the closely-touching clusters have more spacing that makes them better resolvable. However, for the rightmost cluster it generates the worst result than the previous two. Number of the quantified house cluster with the comparison of human classifier quantification is presented in table 1 below.

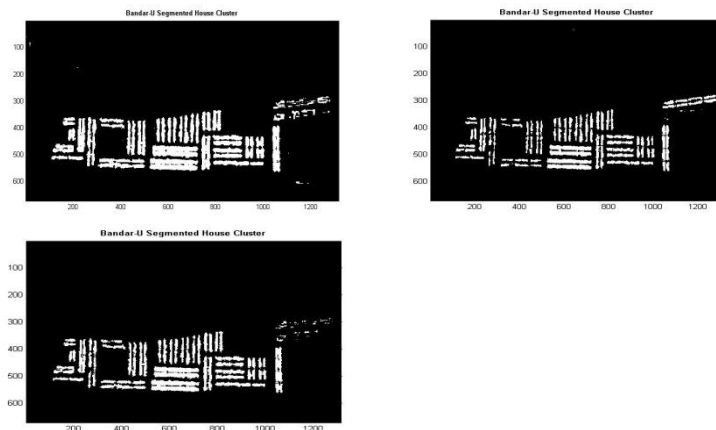


Figure 4. Segmented house cluster with varying number of hidden layers. Left: 1 hidden layer, middle: 3 hidden layers, right: 5 hidden layers

Table 1. Number of house cluster estimates and the corresponding computational time

Human	1 hidden layer	3 hidden layer	3 hidden layer
134	145	151	140
-	22	26	56

4. Conclusion

The neural network is able to perform house cluster detection based on the RGB pattern on each pixel of the image. Generally, the resulting error is that related to misclassifying the non-object to be house cluster and failure in segmenting the cluster properly. In term of computational time the network generates the acceptable outcomes within less than 1 minute that makes the method meets the expectation; fast. Overall 3 hidden layers network drives the closest result to the human classifier in term of number of observable house cluster.

5. Acknowledgement

Universiti Teknologi PETRONAS is gratefully acknowledged for providing facilitation and financial support to run this study.

References

- [1] A Ghosh, N R Pal and S K Pal 1991 Image segmentation using a neural network *Biol. Cybernet* **66** 151–8
- [2] A Ghosh, N R Pal and S K Pal 1992 Object background classification using Hopfield type neural network *Int. J. Pattern Recognition Artific. Intell.* **6** 989–1008
- [3] R H Silverman 1991 Segmentation of ultrasonic images with neural networks *Int. J. Pattern Recognition Artific. Intell.* **5** 619–28
- [4] B Tso and P M Mather 2009 *Classification Methods for Remotely Sensed Data* (Boca Raton: CRC Press)
- [5] N Shorter and T Kasparis 2009 Automatic Vegetation Identification and Building Detection from a Single Nadir Aerial Image *Remote Sensing* **1** 731-57
- [6] J R Jensen, F Qiu and K Patterson 2001 A Neural Network Image Interpretation System to Extract Rural and Urban Land Use and Land Cover Information from Remote Sensor Data *Geocarto Int.* **16** 19-28
- [7] J A Tullis and J R Jensen 2003 Expert System House Detection in High Spatial Resolution Imagery Using Size, Shape, and Context *Geocarto Int.* **18** 5-15
- [8] T Kavzoglu 2001 *An investigation of the design and the use of feed-forward artificial neural networks in the classification of remotely sensed images*. PhD thesis, Univ. Of Nottingham