

# Fuzzy connectedness segmentation of remote sensing data in geological mapping by feature extraction

M. Abdelrazik<sup>1</sup>, H. Ayad<sup>2</sup>, M. Hammad<sup>3</sup> & T. Farag<sup>2</sup>

<sup>1</sup> Industrial Development Authority (IDA),  
Portsaid, Egypt

<sup>2</sup> Department of mathematics, faculty of science, Cairo University,  
Giza, Egypt

<sup>3</sup> Department of geology, faculty of science, Cairo University,  
Giza, Egypt

## Abstract

Remote sensing data is an important source of geological information. Several applications of different platforms were studied by many authors. One of the common applications was the geological mapping after applying various techniques of image processing. This paper aims to apply the fuzzy connectedness technique of image segmentation in solving the tasks of lithological characterization and trace various geologic contacts. This goal was achievable by using a new algorithm written by Matlab. The results after running the new code on the pilot area of G. Muwaylhah at Eastern Desert of Egypt were highly matches with those of the visual edge detection and mapping by running the classification algorithm of ENVI commercial package.

Keywords: *Fuzzy Connectedness, Lithological Characterization, Image Segmentation, Remote Sensing, Supervised Classification.*

## 1. Introduction

Lithological classification is the first basic step in geological mapping (Lisle *et al.* 2011). Defining the boundary between rock units has always been a tricky task. Most of rock masses have heterogeneous nature and change gradually into a different lithology. Thus, it's difficult to define a clear signature for each rock unit, and most classification techniques result in much noise than acceptable (Lu & Weng, 2007). Most of the used algorithms in the commercial packages are based on hard classifiers. Some of the down sides of such classifiers are the rigidity of the classes and the indifference to the spatial aspects of the image (Salati *et al.* 2011).

The concept of fuzziness was developed to deal with the problem of complex systems. It was firstly approached through the ideas of B. Russell in 1923. Russell noted that the models which are used to represent the physical world phenomenon are inherently different than the reality itself, for all the models are assume a crisp and deterministic nature of the universe, which is always not the case. The same issue is found in geological mapping because of the hard classifiers.

The adoption of fuzzy theory in digital image segmentation through several techniques, including characteristic feature thresholding and clustering, edge detection, and region growing (Bezdek, 1973; Cannon *et al.* 1986; Fu & Mui, 1981; Huntsberger *et al.* 1985; Zucker & Kant, 1981). Different algorithms were used individually and in collaboration with each other to produce better results such as (Cannon *et al.* 1986; Mitra *et al.* 2004; Pal *et al.* 2000) Most of the research was applied on abstract images. These techniques were in common usage on the images from the medical field (Udupa & Samarasekera, 1996) and did not extend to the images and multi-bands images from satellite optical data such as ETM+ , Aster, etc. (Salati *et al.* 2011)

Fuzzy theory operates on a membership function which inspects any objects and output a value between zero and one, in contrast to the yes-or-now model of probability theory. Thus for members of a fuzzy set, each object has a value of membership to the set that is strictly known but up to a whole one. This apparent nondeterministic of the function is originating from the sense of the vagueness in the data itself rather than imprecision in the system (Zimmermann, 1975).

Fuzzy Connectedness (FC) define an object through an inherited property of the images itself which is hanging togetherness. Hanging togetherness represents the tendency for a group of consecutive pixels to appear as a gestalt to a human viewer despite the heterogeneous nature of the image (Zimmermann, 2010). It could be translated into a mathematical formula through an affinity function which check possible link between any two pixels. For a group of pixels to be represented by a distinctive object, then all its members must have a value of affinity links over a certain threshold. Every affinity link value is estimated by the lowest value of affinity between a two adjacent pixels along the chain of pixels that represents this link. While the value of the affinity between any nonadjacent pixels is the value of the strongest link between the two (Udupa & Samarasekera, 1996).

The area under experiment consists of Metavolcanics, Granite pluton and Wadi deposits. Various samples were collected from the different lithologies and the algorithm was adjusted to separate the Metavolcanics into fuzzy class, and the rest of the image – comprising the granite pluton and the wadi deposits – were congregated into the other class.

## 2. Study area and the used data

The Central Eastern Desert (CED) of Egypt is vastly an arid area. It extends east from the river Nile valley to the Red Sea shore. Most of the exposed rock of the CED is igneous and metamorphosed Precambrian basement rocks, yet the some of the western side is covered by recent sediments (El Ramly, 1972; Greiling *et al.*, 1994; Shackleton, 1994) The study area is at the vicinity of the contact between sediments and basement. It lies at the midway between Idfu and Marsa Alam cites (Fig. 1).

The chosen area contains metavolcanic and granitic rock units. The former is comprised of meta-rhyolite, meta-andesite, and their tuffs. The pink granites form the conspicuous mass of Gabal Muwaylhah granitic pluton.

The Gabal Muwaylhah pluton has an oval shape with an aerial extent of about 7.5 km<sup>2</sup>. It is of moderate to high relief, showing some peaks separated by faults where the highest peak is 704 m above sea level.

The studied image represents the first band of the used Landsat-7 ETM+ scene (Fig. 2). The wavelength of this band ranges between 0.450±0.005 - 0.515±0.005µm with spatial resolution of 28.5m x 28.5m. The Image georeferenced according to the UTM system and represented by 598913.2500 to 606180.7500 E and 2756163.7500 to 2748896.2500 N coordinates of UTM Zone 36 North, Datum: WGS-84.



Figure 1 Location map of the study area

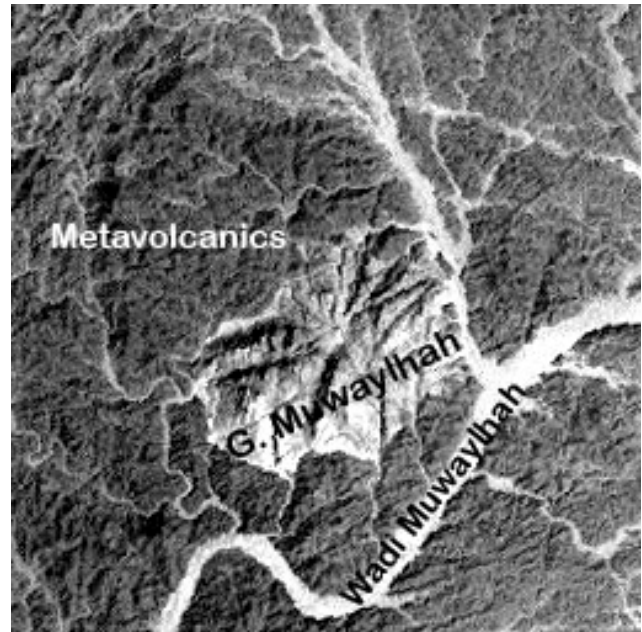


Figure 2 Input image

## 3. Methodology

The experiment starts by subsetting the ETM+ scene of the study area coordination. Then, the wadies were traced manually to produce the reference image. Simultaneously; three other images were produced: one by using the FC algorithm, and two by the commercial software – ENVI (Fig. 3). The last step was to compare the images with the referenced one (Fig.3c) and acquire the matching percentages. The images produced using ENVI are representing the two available possibilities of classification – unsupervised and supervised. For the supervised classification the Parallelepiped algorithm was used (Richards, 2013) whereas the IsoData was used for the unsupervised classification (Tou & Gonzalez, 1974). The image of FC was produced after processing of the ETM+ image of band 1 by the FC algorithm through Matlab software.

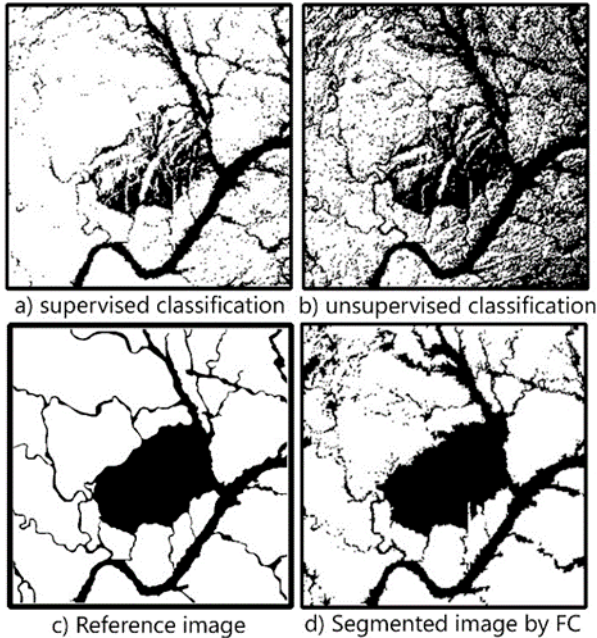


Figure 3 Images produced Using Envi (a-b) and Fuzzy connectedness (d) and the reference image (c)

The FC algorithm (Fig. 4) works in three main steps; every step has clear input and output;

1. Computation of the degree of connectedness (resemblance) between each candidate image element within image region and every image element in the input image. Output of the step: a set of images in which each one represents connectedness information between a region candidate element and all image elements.

2. Thresholding of each output image from the first step with an appropriate threshold to yield a corresponding image that shows only those elements which hang together (highly connected) with the specified region candidate element. Output of the step: a set of images in which each one shows only a single region's image elements.

3. Combination of the output images from step 2 to yield a fully segmented image. Output of the step: an image that shows every region that is separated. The process flowchart is showed in (Fig. 5).

**Input:**  $\mathcal{C}, o, \kappa$

**Output:**  $K_o$ -scene  $\mathcal{C}_o = (C_o, f_o)$  of  $\mathcal{C}$ .

**Auxiliary Data Structures:** An  $nD$  array representing the  $K_o$ -scene  $\mathcal{C}_o = (C_o, f_o)$  of  $\mathcal{C}$  and a queue  $Q$  of spels. For the purpose of the algorithm, we refer to the array itself by  $\mathcal{C}_o$ .

```

begin
0. set all elements of  $\mathcal{C}_o$  to 0 except  $o$  which is set to 1;
1. push all spels  $c$  of  $C_o$  such that  $\mu_\kappa(o, c) > 0$  to  $Q$ ;
   while  $Q$  is not empty do
2.   remove a spel  $c$  from  $Q$ ;
3.   find  $f_{\max} = \max_{d \in C_o} [\min(f_o(d), \mu_\kappa(c, d))]$ ;
4.   if  $f_{\max} > f_o(c)$  then
5.     set  $f_o(c) = f_{\max}$ ;
6.   push all spels  $e$  such that  $\mu_\kappa(c, e) > 0$ 
7.     to  $Q$ ;
   endif;
   endwhile;
end

```

Figure 4 Fuzzy connectedness algorithm

## 4. Results and conclusions

The FC produced image was then compared to the reference photo against the two photos produced by ENVI using a simple matching algorithm. The FC image was found to have 91.67% degree of similarity to the manual reference, while the supervised classification images have 87.89% and the unsupervised classification image scored 70.04%. The algorithm was executed on a Dell® machine running on an intel core i3 processor of 2.6 GHz and 4GB of RAM. Although the FC algorithm may take longer time to run than the normal commercial software, it's free and could run faster if more computer power was involved in the experiment.

The experiment shows that FC has a better resemblance with the reference image than the images produced by the Parallelepiped and IsoData algorithms. The down sides of the FC are that it takes much more time to execute and it needs to be adjusted manually to suite the number and size of the samples collected for the classification. Yet, since the algorithm is still in the early stage of experimenting, these down sides could be avoided in future versions.



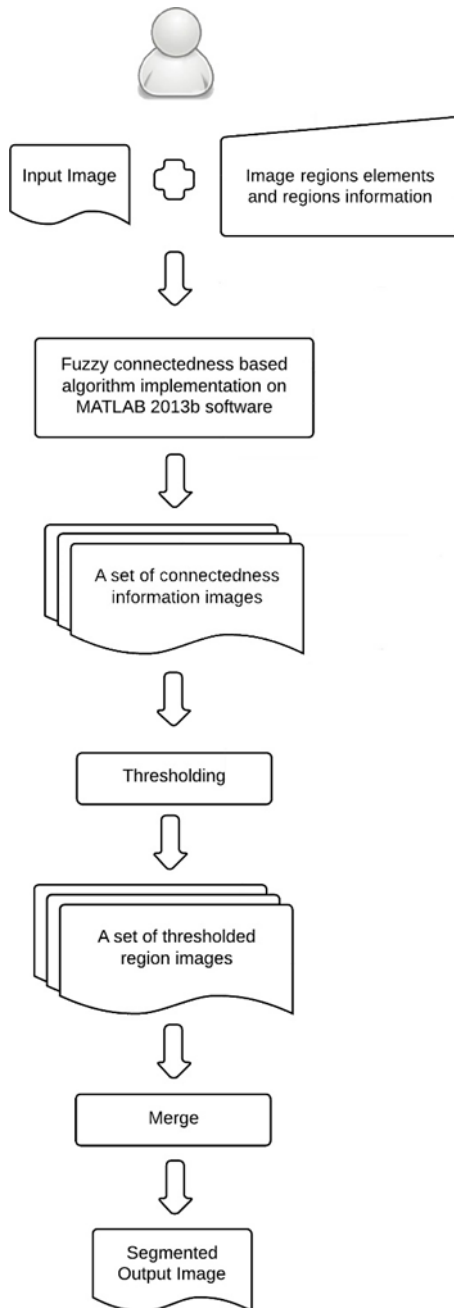


Figure 5 Fuzzy connectedness algorithm flowchart

## References

- Bezdek, J. C. (1973). *Fuzzy mathematics in pattern classification*. Cornell Univ., Ithaca, NY.
- Cannon, R. L., Dave, J. V., Bezdek, J. C., & Trivedi, M. M. (1986). Segmentation of a Thematic Mapper Image Using the Fuzzy c-Means Clustering Algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24(3), 400–408. <http://doi.org/10.1109/TGRS.1986.289598>
- El Ramly, M. F. (1972). Geological Map of the Basement Rocks in the Eastern Desert of Egypt. (1: 1000000).
- Fu, K. S., & Mui, J. K. (1981). A survey on image segmentation. *Pattern Recognition*, (13), 3–16.
- Greiling, R. O., Abdeen, M. M., Dardir, A. A., Akhal, H. E., El Ramly, M. F., Kamal El Din, G. M., ... Sadek, M. F. (1994). A structural synthesis of the Proterozoic Arabian-Nubian Shield in Egypt. *Geol Rundsch*, 484–501.
- Huntsherger, T. L., Jacobs, C. L., & Cannon, R. L. (1985). Iterative fuzzy image segmentation. *Pattern Recognition*, 18(2), 131–138. [http://doi.org/10.1016/0031-3203\(85\)90036-6](http://doi.org/10.1016/0031-3203(85)90036-6)
- Lisle, R. J., Brabham, P. J., & Barnes, J. W. (2011). *Basic Geological Mapping* (5th ed.). Wiley-Blackwell.
- Lu, D., & Weng, Q. (2007). A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *International Journal of Remote Sensing*, 28, 823.
- Mitra, P., Uma Shankar, B., & Pal, S. K. (2004). Segmentation of multispectral remote sensing images using active support vector machines. *Pattern Recognition Letters*, 25(9), 1067–1074. <http://doi.org/10.1016/j.patrec.2004.03.004>
- Pal, S. K., Ghosh, a., & Shankar, B. U. (2000). Segmentation of remotely sensed images with fuzzy thresholding, and quantitative evaluation. *International Journal of Remote Sensing*, 21(934204179), 2269–2300. <http://doi.org/10.1080/01431160050029567>
- Richards, J. A. (2013). Supervised Classification Techniques. In *Remote Sensing Digital Image Analysis: An Introduction* (pp. 247–315). <http://doi.org/10.1007/978-3-642-88087-2>
- Salati, S., Ruitenbeek, F. J. A. Van, Meer, F. D. Van Der, Tangestani, M. H., & Werff, H. M. A. Van Der. (2011). Lithological mapping and fuzzy set theory: Automated extraction of lithological boundary from ASTER imagery

by template matching and spatial accuracy assessment. *International Journal of Applied Earth Observations and Geoinformation*, 753–765.

Shackleton, R. M. (1994). Review of Late Proterozoic sutures, ophiolitic melanges and tectonics of eastern Egypt and north-east Sudan. *Geol Rundsch*, 537–546.

Tou, J. T., & Gonzalez, R. C. (1974). *Pattern recognition principles*. Image Rochester NY. Addison-Wesley Publishing Company, Reading, Massachusetts.

Udapa, J. K., & Samarasekera, S. (1996). Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation. *Graphical Models and Image Processing*, 58(3), 246–261. <http://doi.org/10.1006/gmip.1996.0021>

Zimmermann, H. J. (1975). Description and Optimization of Fuzzy Systems†. *International Journal of General Systems*, 2(1), 209–215. <http://doi.org/10.1080/03081077508960870>

Zimmermann, H. J. (2010). Fuzzy set theory. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), 317–332. <http://doi.org/10.1002/wics.82>

Zucker, S. W., & Kant, K. (1981). Multiple- level representations for texture discrimination. In *IEEE Computer Society Conf. on PRIP* (pp. 609–613). IEEE Computer Society Los Angeles, CA.