Journal of Geomatics Vol 9 No. 2 October 2015

K-means versus fuzzy c-means as objective functions for Genetic Algorithms-based classification from aerial images and LIDAR data

M. Al-Nokrashy[1], A. Esmat[1], M.S. Gomaa[2] and A. Hamdy[1]
[1]Department of Civil Engineering, Faculty of Engineering, Al-Azhar University, Cairo, Egypt
[2]Department of Surveying Engineering, Shoubra Faculty of Engineering, Benha University, Cairo, Egypt
Email: engmod2000@yahoo.com

(Received: February 9, 2015; in final form: Aug 17, 2015)

Abstract: Combining data from different sensors has the potential to result in more accurate classification than a single sensor. The availability of high quality RGB and LIDAR data provides efficient image classification using the complementary properties of these data sources. This work mainly integrates Genetic Algorithms (GAs) with different fitness functions to extract buildings, trees, roads and grass from aerial images and LIDAR data. K-Means (KM) and Fuzzy C-means (FCM) algorithms were tested and compared, as fitness functions for GAs. Three groups of data were applied which include: RGB group; RGB/LIDAR data group and RGB/LIDAR/attributes group. Error matrix and K-HAT (kappa) statistics were adopted as well as visual inspection to evaluate the validity and robustness of the proposed techniques. FCM proved to be a preferable fitness function for GAs-based classification from aerial images and LIDAR data with accurate average classifications of 87.84%.

Keywords: Unsupervised Classification, Genetic algorithm, K-means, Fuzzy C-means, Digital imagery, LIDAR data

1. Introduction

LIDAR has the advantage of accurately and rabidly capturing surfaces. On the other hand, photogrammetry is a well-established mapping; surface reconstruction and feature classification technique that is characterized by high redundancy feature observation in multiple images (Ghanma, 2006). Also, features or attributes created from aerial images and LIDAR data are commonly used for land cover classification. Most authors reported that the integration of multiple types of data has led to improvement in classification performance.

The accuracy improvement of unsupervised classification remains a critical issue needing much more efforts due to the great demand for efficient automation of image classification. Genatic Algorithms (Gas), introduced by Holland (1975), is a convenient method for heuristic unsupervised classification (Coley, 1999; Pham and Karraboga, 2000). GAs offer several advantages over the conventional unsupervised classification methods which makes it as one of the most powerful unbiased optimization techniques for sampling a large solution space. One advantage is the capability to handle solutions of high degree of complexity that often involves large, non-linear and discrete attributes. Plenty of studies have shown that the GA technique is efficient to deal with large datasets and has a large chance to avoid a local optimal solution than other techniques (Huang et al., 2006; Zhou et al., 2010). Moreover, GAs are less complex and more straightforward as compared to conventional algorithms (Tabassum and Mathew, 2014).

Image classification has been widely and successfully applied by optimization algorithms specifically GAs (Rothlauf, 2006) that confirm the potential of GAs to produce a high level of quality results especially when there is no ground truth (Coley, 1999).

GAs have been widely investigated in machine learning and pattern recognition fields in which the datasets used typically consist of a few data points. On the other hand, there have been relatively few applications of GAs for feature extraction from remotely sensed data.

Liu et al. (2003) presented a new approach to road extraction from high resolution satellite imagery based on GAs with fitness calculation of clustering. The proposed approach applies GAs to learn the parameters and pick up good clusters automatically. The approach is demonstrated on pan sharpened QUICKBIRD imagery and preliminary results are encouraging.

Yang et al. (2006) have adopted a heuristic method based on GAs to automatically determine the number of cluster centroids during unsupervised classification. The GAs unsupervised classifier was tested on an IKONOS satellite image. Based on independent ground truth, an overall accuracy of 71.1% was reached as compared to 65.1% when using the (iterative self-organizing data analysis techniques) ISODATA algorithm.

Yang (2007) classified SPOT satellite images by GAs and ISODATA and compared the results with a supervised classification. The comparison showed that the GAs have performed better than the unsupervised ISODATA and as good as the supervised classifier. A modified GAs using maximum likelihood, as a clustering criterion, was also tested and proved to be as good as the supervised classifier.

Van Coillie et al. (2007) have applied GAs for feature selection in object-based classification of IKONOS imagery for forest mapping. The proposed method is a three-step object-oriented classification routine that involves the integration of: 1) image segmentation; 2) feature selection by GAs; and 3) joint Neural Network based object-classification. Results show that with GAs-feature selection, the mean classification accuracy is significantly higher than without feature selection.

In Mahi and Izabatene (2011), a Radial Basis Function Neural Network (RBFNN) is applied for the purpose of QUICKBIRD satellite image segmentation. During the unsupervised learning of the RBF network, an unsupervised GAs is employed to automatically determine the hidden layer parameters. Experimental results show that the RBF network combined to the GAs is an attractive approach for segmentation of multispectral remote sensing imagery.

Stavrakoudis et al. (2011) have proposed a multistage genetic fuzzy classifier for land cover classification from IKONOS satellite imagery. The proposed method is a three-stage process: 1) the first stage iteratively generates fuzzy rules, 2) a simplification stage follows aiming at further improving the interpretability of the initial rule base; 3) finally, a genetic tuning stage fine tunes the fuzzy sets database improving the classification performance of the obtained model. The results indicate the effectiveness of the proposed system in handling multidimensional feature spaces, producing easily understandable fuzzy models.

Chu (2012) integrated feature selection, GAs and Multi-Classifier System (MCS) with Dempster-Shafer theory of evidence for classifying different datasets. Classification results revealed that the proposed technique resulted in significantly higher levels of accuracy than any single classifier.

Almeida (2012) studied combination of GAs with decision trees for the object-based land-cover classification. The study found a satisfactory performance for the automatic assessment of the optimal segmentation parameters. Nevertheless, some problems such as the shape complexity of some targets, the internal spectral variability of certain classes, and the diverse conditions of ageing and maintenance of some roof classes found in the study area led to an over-segmentation of some targets.

Wikantika et al. (2014) have compared the performance of three models of Artificial Intelligence, Neural Networks, Fuzzy Logic and GAs and Maximum Likelihood statistical models in the process of classification of satellite imagery to estimate wetlands with Landsat image data. The comparison showed that the accuracy of the model by Maximum Likelihood is 84.64%, Neural Network Model is 98.78%, GAs model is 94.94% and Fuzzy Logic models is 75.43%. Tyagi and Verma (2015) applied GAs for the optimization of cluster centres of Fuzzy C-means (FCM) clustering algorithm.

So called indices or fitness functions are used to determine whether convergence has been reached. It is known that the choice of the index (classifier) has a major impact on the results of classification. In previous studies, a wide variety of indices have been proposed to be the fitness function of GAs. These include: K-means index (KMI); separation index (SI); partition and separation index (PASI); Davies-Bouldin index (DBI); Dunn's index (DI); partition coefficient index (PCI); partition entropy index (PEI); Fukuyama and Sugeno validity index (FSVI); Xie and Beni validity index (XBI); C-index (CI); and FCM index (FCMI).

Yang and Wu (2001) introduced the PASI and compared it with other five clustering indices, PCI, PEI, FSVI, XBI and DBI, for the clustering of SPOT-5 satellite image. The PASI proved to be superior to all other indices.

Bandyopadhyay and Maulik (2002a) integrated the DBI, DI, FCMI, and CI into GAs as fitness functions for clustering analysis of several artificial and real life datasets. The results showed that the proposed method is able to distinguish some characteristic land cover types in the image. Bandyopadhyay and Maulik (2002b) also integrated KMI into GAs for unsupervised clustering of the same datasets in order to improve the defect of KMI, needs the initial cluster numbers a prior. Bandyopadhyay et al. (2007) have modelled the problem of GAs-based clustering as simultaneous optimization of the XBI (Xie and Beni, 1991) and the FCMI (Bezdek, 1981). In this regard, IRS and SPOT satellite images have been classified and compared with the results of single objective GAs. The comparison indicated that better clustering performance is obtained if both of these indices are optimized simultaneously.

Yang et al. (2008) applied the GAs for unsupervised classification of multi-spectral IKONOS image. Three different indices, DBI, XBI and KMI, have been tested while varying a number of parameters of the GAs. The XBI proved to be the most accurate and robust index. However, it is much more sensitive to parameter tuning. Yang et al. (2014a) have evaluated three classification indexes, DBI, FCMI, and PASI, as fitness functions in GAs classification of SPOT-5 satellite image. The results showed that the FCMI has performed the best, followed by DBI and PASI. Yang et al. (2014b) integrated DBI and FCMI, so-called DBFCMI, in a GAs classifier to improve the accuracy and robustness of classification. For the purpose of comparison, well-known indices, DBI, FCMI and PASI, were employed. A SPOT-5 satellite image was applied for land use classification. As a result, the best overall accuracy of DBFCMI, DBI, FCMI, and PASI was 75.5%, 75.0%, 74.9%, and 74.2% separately.
Among all indices, FCMI has the advantage of robustness for ambiguity and maintains much more

information than other indices (Suganya and Shanthi, 2012). On the other hand, FCMI usually gives better result for overlapped data where data point is assigned membership to each cluster centre (Lu et al., 2013).

This work is probably the first attempt to use GAs for combination of multiple data sources, aerial images and LIDAR data, for improvement of land cover classification. To further improve the classification performance and overcome the shortcomings of the previous approaches, GAs have been proposed in this research.

The main objective of this research is to develop an accurate, time-effective and automatic method to classify buildings, roads, trees and grass by the fusion of aerial images, LIDAR data and attributes through GAs. This objective is achieved by comparing and combining the outputs from K-Means (KM) based GAs and FCM based ones in order to improve the quality of the classification results. In this regard, GAs were applied with KM and FCM as fitness functions. A combined method was applied to fuse the results from both techniques.

For the rest of the research, the KM based GAs will be referred to as KMGA while the FCM based one will be referred to as FCMGA. All the methods proposed in this research were implemented through a software package generated by the authors in a Matlab environment. After a detailed background in the following section, this paper is organized as follows: the classification methods used are described; then the results are presented and evaluated; and finally the results are summarized.

2. Background

GAs are adaptive methods which may be used to solve a variety of optimization problems by the principles of the evolution of a biological organism. Recently, GAs have been used in a wide variety of optimization problems, specifically in classifying digital data sets. GAs are very different from most of the traditional optimization methods. GAs generate a population of solutions at each iteration to approach an optimal solution. This means that GAs can process a number of designs at the same time. In addition it selects the next population by computations that involve random choices. The following sections describe how to establish a GAs classifier for automatically clustering a data set.

2.1 String representation

In GAs applications, the parameters of the searched space are encoded in the form of strings, so-called chromosomes. Each chromosome, representing an answer for the problem, is encoded by a binary, integer or real number. A variable string length is designed without assigning the number of classes a priori. A chromosome is encoded by positive real numbers and a negative integer '-1' which represents a non-existent cluster. The value of K (valid clusters) is randomly assumed in the range (K_{min}, K_{max}), where K_{min} is usually assigned to 2 unless special cases are considered. The length of a chromosome is taken to be K_{max} where each individual gene represents either a cluster centre or a non-existent cluster. The typical size of the population can range from 20 to 1000 (Coley, 1999). K_i are chosen randomly from the data set, and then are randomly allocated in the chromosome (Bandyopadhyay and Maulik, 2002a). As an example, assume an image including 3 bands, N pixels for each layer, K_{min}=2 and K_{max}=6, the number K_i equal to 5 for the chromosome i. Let the 3 cluster centres be: (120, 66, 228); (170, 66, 264); and (21, 114, 6). Randomly, the classification centres can be encoded into a chromosome as: -1 (120, 66, 228) (170, 66, 264) -1 (21, 114, 6) -1.

2.2 Selection and fitness function identification

The GAs should achieve two goals: maximizing the classification accuracy; and minimizing the number of selected features. These criteria used to create a single objective function as follows:

$$F = w * C(x) + (1-w) * \frac{1}{N(x)} \quad (1)$$

where x is the feature subset, $C(x)$ represents the classification accuracy, $N(x)$ is the size of selected feature subset and w is a parameter between 0 and 1 which adjusts the influence of each criterion. As the value of w is higher the weight of classification accuracy in fitness function is greater. On the other hand, reducing the value of w will give more penalties on the size of x (Tan and Fu, 2008). By adjusting w, a trade-off between the accuracy and the size of the feature subset obtained can be achieved. For this research, w was adjusted to 0.8 to avoid a large decrease in classification accuracy. Before a GA is operated, an objective function needs to be defined to measure the fitness of each chromosome. The fitness function assigns an adaptability degree to each chromosome in the population. Sections 3.3.1 and 3.3.2 summarize the used fitness indices KM and FCM.

2.3 Genetic operators

In general, GAs are composed of three main operations: reproduction; crossover; and mutation. Reproduction calculates a survival probability of each chromosome which is a criterion to reproduce better chromosomes for next generations. Crossover is a swapping process to create new chromosomes between the reproduced chromosomes. To avoid sticking to a local optimum, mutation is assigned to explore the possible optimums in all the space. The mutation probability is usually set smaller than the crossover, and controls the percentage to introduce new genes for trial. If the mutation probability is too low, some useful genes are never found out. On the other hand, if it is too high, there will be severely random perturbation (Gen and Cheng, 1997). These operations are repeated until the terminal criterion is satisfied, and the best string in the final generation is also obtained.

3. Methodology

3.1 Study area and data sources

For this research a very high resolution digital aerial image was available. The image covers an approximate area of about 500x500m of the region surrounding the University of New South Wales (UNSW) campus, Sydney Australia. The area is a largely urban area that contains residential buildings, large campus buildings, and a network of main roads as well as minor roads, trees, open areas and green areas.

The color imagery was captured by film camera at a scale of 1:6000. The film was scanned in three color bands (red, green and blue) in TIFF format, with 15µm pixel size (GSD of 0.09m) and radiometric resolution of 16-bit as shown in Figure 1. The characteristics of image datasets are provided in Table 1.

On the other hand, the LIDAR data covers the study area including the first and last pulse as well as the LIDAR intensity data was used. Table 2 lists the characteristics of LIDAR datasets.

Table 1: Characteristics of image datasets

Size(km)	bands	pixel size (cm)	Camera	Look Angle (deg.)	
				along track	across track
0.5 x 0.5	RGB	9	LMK1000	±30	±30

Figure 1: An orthophoto of UNSW campus

Table 2: Characteristics of LIDAR datasets

Optech ALTM 1225	
Spacing across track (m)	1.15
Spacing along track (m)	1.15
Vertical accuracy (m)	0.10
Horizontal accuracy (m)	0.5
Density (Points/m^2)	1
Sampling intensity (mHz)	11
Wavelength (µm)	1.047
Average altitude (m)	1100
Laser swath width (m)	800

3.2 Data pre-processing

Data pre-preparation of the study area was implemented in several stages as follows:

3.2.1 Filtering of LIDAR data: LIDAR data filtering is the process of separating on-terrain points (DTM) from points falling onto natural objects. A grid method based on support vector machines has been applied for filtering of LIDAR Data (Salah and Trinder, 2010). After filtering of LIDAR points, it is converted into a digital terrain model (DTM) image. A digital surface model (DSM) was generated from the original LIDAR point clouds (first and last pulses).

Finally, a normalized digital surface model (nDSM) was generated by subtracting the DTM from the DSM.

3.2.2 Generation of attributes: Attributes are necessary to recover some common problems that emerge associated with high resolution image data, notably shadows caused by tall buildings or trees; and the spectral variability within the same land-cover class. These disadvantages may cause lower classification accuracy if the classification procedure cannot effectively handle them (Zhou et al., 2008; Lu and Weng, 2007). Before generating the attributes, the aerial photograph, already orthorectified by AAMHatch, was registered to the LIDAR intensity image using a projective transformation. The polymorphic texture strength based on the Förstner operator (Förstner and Gülch, 1987) has been generated and used as input for the classification process.

3.2.3 Generation of ground truth data: Accuracy assessment allows evaluating a classified image. In order to assess the performances of the proposed methods, two sets of class values for randomly selected points (900 points) in the classified image were applied. One set of class values was automatically assigned to these random points as they are selected, and the other set of class values (reference values) is input manually. These reference values are based on the digital aerial image of the test area. The accuracy assessment was performed by comparing the classification results with the ground truth data. The following sections illustrate the way the GAs can be applied and integrated with KM and FCM for unsupervised image classification.

3.3. Methods

3.3.1 KM -based GAs Classification (KMGA): Clustering algorithms can be broadly classified as hard, fuzzy, possibilistic and probabilistic. KM is one of the most popular hard clustering algorithms which partitions data objects into k clusters where the number of clusters, k, is decided in advance according to application purposes. KM is a simple and common clustering algorithm which can also be used within a GAs framework. It provides an iterative scheme that operates over a fixed number (K) of clusters, while attempting to simultaneously optimize centre locations and pixel assignments. KM represents the total variation disregarding the distance between different clusters. KM is computed as follows:

$$KM = 1/(\sum_{k=1}^{k} \sum_{i=1}^{N} \mu_{ik} ||x_i - v_k||^2) \quad (2)$$

K = total number of clusters,
N = total number of pixels,
μ_{ik}= membership function of each pixel x_i belonging to the K^{th}cluster,
x_i= pixel i with grey values x (one for each band),
v_k= average value of K^{th}cluster in the current iteration.

First, the k-means algorithm chooses the "centres" c_i of each cluster. In Bradley and Fayyad (1998), a refinement process is performed through a function probability associated with the set of pixels, in order to find a local minimum for faster convergence of the method. Secondly, a cluster is associated for each pixel of the image based on minimum distance between a given pixel x_i and the "centre", c_i of that cluster. Once the distance between the pixel and the centre of each cluster is calculated, the pixel is associated to the cluster that obtained the minimum distance. In this stage the Euclidian distance is used as metric. Consider a point x and a cluster centre c, where i indexes the spectral components of each cluster. Once associated a cluster for each pixel of the image, the clusters centres are recomputed. To compute the cluster centre, the mean of all pixels of that cluster are calculated.

The application of GAs in the area of classification takes advantage of extensive optimum search capabilities. General genetic procedure in the case of determining the best k centres for clusters consist of setting of parameters which include: number of clusters, population initialization, initial population fitness calculation and repeated selection, cross over and mutation operations until termination criteria are met (Hall et al., 1999).

For genetic KM selection of cluster number and other algorithm specific parameter values is required. Next, the population should be initialized with randomly created cluster centres. New populations are created by operations of selection, cross-over and mutation. For every solution in population, a fitness value is calculated according to the specific fitness function. Solutions with high fitness values come into the mating pool. The process is repeated until termination criteria are met. Below some implementation details are given:

Chromosomes: represent solutions consisting of centres of k clusters – each cluster centre is a vector of values in the range between 0 and 255 representing intensity of color component.

Population initialization and fitness computation: Cluster centres are initialized randomly to k points with values in the range 0 – 255. Next, a fitness value is calculated for each chromosome in the population.

Selection: The operation of selection tries to choose the best suited chromosomes from parent population that come into the mating pool. After cross-over and mutation operations, the child population is created. Most frequently GAs select into mating pool the best individual from a predefined number of randomly chosen population chromosomes. This process is repeated for each parental chromosome.

Crossover: The operation of crossover presents a probabilistic process exchanging information between two parent chromosomes during formation of two child chromosomes. Mostly, a two-point crossover operation is used.

Mutation: The operation of mutation is applied to each created child chromosome with a given probability. After the crossover operation children chromosomes that undergo the mutation operation flip the value of the chosen bit or change the value of the chosen byte to another in the range from 0 to 255.

Termination criterion: A termination criterion determines when the algorithm completes execution and final results are presented to the user. The most frequent termination criterion is that algorithm terminates after a predefined number of iterations. Other possible conditions for termination of the KM algorithm depend on the degree of population diversity or situation when no further cluster reassignment takes place.

3.3.2 FCM-based GAs Classification (FCMGA): GAs searching capability can be used for the purpose of appropriately clustering a set of *n* unlabeled points in *N*-dimensions into *K* clusters. The following outline of the FCM-based GAs is based on Halder et al. (2011). Considering an image of size $m \times n$, the basic steps of the FCM-based GAs for clustering image data are as follows:

Encoding: For *N*-dimensional space, each cluster centre is mapped to *N* consecutive genes in the chromosome. Each chromosome represents a solution which is a sequence of *K* cluster centres. For image datasets each gene is an integer representing an intensity value.

Population initialization: In FCM-based Genetic method, the FCM is run P times for generating a population of P chromosomes; each chromosome is of size K. Each chromosome of the population is a potential solution by the FCM algorithm with number of clusters $c=K$. The FCM algorithm assigns pixels to each category by using fuzzy memberships. Let $X=(x_1, x_2,\ldots,x_N)$ denote an image with N pixels to be partitioned into c clusters, where x_i represents multispectral data. The algorithm is an iterative optimization that minimizes the cost function, a generalized least-squared errors function, defined as follows:

$$J=\sum_{j=1}^{N}\sum_{i=1}^{c} u_{ij}^{m}\left\|x_j - z_i\right\|^2 \quad (3)$$

where u_{ij} represents the membership of pixel x_j in the i^{th} cluster, z_i is the i^{th} cluster centre, $\|\ \|$ is a norm metric, and m is a constant. The parameter m controls the fuzziness of the resulting partition, m: any real number greater than 1, it was set to 2.00 by Bezdek (1981). The membership function represents the probability that a pixel belongs to a specific cluster. The cost function is minimized when pixels close to the centroid of their clusters are assigned high membership values, and low membership values are assigned to pixels with data far from the centroid. In the FCM algorithm, the probability is dependent solely on the distance between the pixel and each individual cluster centre in the feature domain. The membership functions and cluster centres are updated as follow:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\left\|x_j - z_i\right\|}{x_j - z_k}\right)^{\frac{2}{m-1}}} \quad (4)$$

where:

$$z_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} x_j}{\sum_{j=1}^{N} u_{ij}^{m}} \quad (5)$$

The objective function of FCMI is shown as (3), and the optimal cluster centres can be found by minimizing (3). The centre of the i^{th} cluster is determined by (5). Equation (4) is the membership function of x_i being assigned to the i^{th} cluster. Convergence of FCM can be detected by comparing the changes in the membership function or the cluster centre at two successive iteration steps.

Fitness computation: Two steps to accomplish the fitness computation. First, the pixel dataset is clustered according to the centres encoded in the chromosome under consideration, such that each intensity value x_i, i = 1, 2, ..., $m\times n$ is assigned to cluster with centre z_j, j = 1, 2,..., K.

$$if\left||xi - zi|\right| < \left||xi - zp|\right|, p = 1,2,\ldots\ldots k, and p \neq j \quad (6)$$

The next step involves adjusting the values of the cluster centres encoded in the chromosome, replacing them by the mean points of the respective clusters. The new centre z_i^* for the cluster c_i is given by:

$$z_i^* = \frac{1}{n_i}\sum_{x_j \in C_i} x_j, i = 1, 2, \ldots \mathrm{K} \quad (7)$$

Subsequently, the clustering metric M is computed as the sum of Euclidean distances of each point from their respective cluster centres and given by:

$$M = \sum_{i=1}^{k} M_i, i = 1, 2, \ldots\ldots K \quad (8)$$

where

$$M = \sum_{x_i \in C_1}\left\|x_j - z_i\right\| \quad (9)$$

The fitness function is defined as follow:

$$f = \frac{1}{M} \quad (10)$$

Hence the objective is to minimize the clustering metric M i.e. maximize f.

Selection: Fitness level is used to associate a probability of selection with each individual chromosome. Roulette Wheel selection was applied, if f_i is the fitness of individual c_i in the population, its probability of being selected is:

$$p_i = \frac{f_i}{\sum_{j=1}^{N} f_i} \quad (11)$$

N is the number of individuals in the population.

Crossover and mutation: In this research, a two-point crossover with a fixed crossover probability of μc is used. Also, each chromosome undergoes mutation with a fixed probability 0.05. A number δ in the range [0, 1] is generated with uniform distribution. If the value at a gene position is v, after mutation it becomes:

$$v \pm \delta * v, v \neq 0$$

$$v \pm \delta, \qquad v = 0$$

Termination criterion: for each iteration, the fittest chromosome is preserved elitism. Thus on termination, this chromosome gives the best solution encountered during the search. The algorithm is a two pass process. In the first pass the standard FCM algorithm is used to generate the population. In the second pass, the GAs algorithm is applied on the population generated by the FCM algorithm.

Validity measure for clustering: In order to determine the validity of the clustering on the given dataset, the cluster validity index for the fittest chromosome for a particular value of *K* is computed using equation 12. The cluster validity index used in this research is the one proposed by Turi (2001). It aims at minimizing the validity index given by the function as follow:

$$V = yx \frac{\text{int } ra}{\text{int } er} \quad (12)$$

The term *intra* is the average of all the distances between each pixel *x* and its cluster centroid z_i which is defined as:

$$\text{int } ra = \frac{1}{N} \sum_{i=1}^{k} \sum_{x \in C_i} \left\| x - z_i \right\|^2 \quad (13)$$

The *inter* term is the minimum distances between the cluster centroids which is defined as:

$$\text{int } er = \min(\left\| z_i - z_j \right\|^2), \quad (14)$$

where i=1, 2... K-1 and j=i+1, 2... K. This term is used to measure the separation of the clusters. Also, y is given as:

$$y = c * N(2, 1) + 1 \quad (15)$$

where c is a user specified parameter and N (2, 1) is a Gaussian distribution function with mean 2 and standard deviation 1, where the variable is the cluster number and is given as:

$$N(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(k-\mu)^2}{2\sigma^2}} \quad (16)$$

where k is the cluster number and μ=2 and σ=1. This validity measure serves the dual purpose of minimizing the intra-cluster spread and maximizing the inter-cluster distance.

3.4 Classifier combinations

Application of a multiple classifier system (MCS) in remote sensing has been discussed in detail in Benediktsson et al. (2007). The limitation of LIDAR data on hand in Egypt motivated the authors to combine the results obtained for RGB/KMGA and RGB/FCMGA in order to take advantage of each classifier and improve the overall accuracy. In this research, the RGB/KMGA and RGB/FCMGA were combined according to their reliability for each class. The class that receives the maximum Kappa Index of Agreement (KIA) is taken as the final classification. As a result, there are a number of possible overlapped pixels that could take place in the combined results. Most overlaps occur at the edges of classes. In order to compensate for these errors, the overlapped areas are compared and the class that receives the highest overall KIA is taken as the final classification.

3.5 Accuracy assessment

Accuracy assessments of the proposed system were undertaken using confusion matrices and Kappa statistics. The KIA is a statistical measure adapted for accuracy assessment in the remote sensing field by Congalton and Roy (1983). KIA is a means to test two images, if their differences are due to 'chance' or 'real disagreement'. It is often used to check for accuracy of classified images versus some 'real' ground-truth data.

$$k = \frac{N \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} (X_{i+} * X_{+i})}{N^2 - \sum_{i=1}^{r} (X_{i+} * X_{+i})} \quad (17)$$

r: number of row in cross classification table,
x_{ii} :number of combinations along the diagonal,
x_{i+} : total observations in row i,
x_{+i} : total observations in column i,
N : total number of cells.

For the per-category-KAPPA, the following algorithm was introduced to remote sensing by Rosenfield and Fitzpatrick-Lins (1986):

$$k_i = \frac{P_{ii} - P_{i+}P_{+i}}{P_{i+} - P_{i+}P_{+i}} \quad (18)$$

p_{ii} :proportion of units agreeing in row i / column i
p_{i+} :proportion of units for expected chance agreement in row i
p_{+i} :proportion of units for expected chance agreement in column i

3.6 Test description and workflow

In this research, a maximum chromosome length of K_{max}=8 was chosen, which is above the maximum number of clusters in the test image (Yang et al., 2006). Only two-point crossover operations were considered with fixed probability (Dunn, 1974; Chen and Lin, 2007). The other parameters were chosen to be 100 for population size, 0.8 for crossover probability (Yang et al., 2006) and 0.05 for mutation probability (Dunn, 1974; Pham and Karaboga, 2000; Kim and Kim, 2003). In this regard, three groups of data sources, as shown in Table 3, were tested and evaluated. First, the KMI as a fitness function was used and investigated. After that, the FCMI was tested and applied with GAs as a fitness function. Finally, an approach was proposed to combine the results from both KMGA and FCMGA. Figure 2 summarizes the workflow for the proposed techniques.

Table 3: Groups of dataset applied for the experiments

Group 1	Group 2	Group 3
- RGB	- RGB - DSM	- RGB - DSM - Intensity - polymorphic texture strength from nDSM

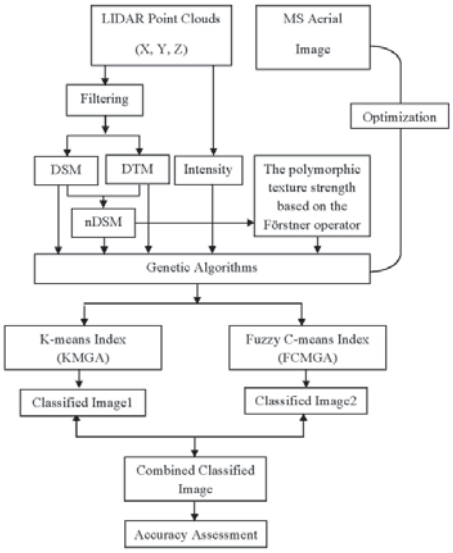


Figure 2: Workflow for image classification by image and LIDAR data fusion

4. Results and discussion

For the purpose of testing the efficiency of KMGA and FCMGA, the proposed algorithms have been applied to each of the three groups of data. Visual results from KMGA and FCMGA for the three groups of data are shown in Figure 3 and numerically expressed in figures 4 to 7.

Column 1 of Figure 3 illustrates the KMGA results while column 2 of the same figure shows the FCMGA results. Moreover, a close watch of Figures 3 reveals that both the approaches are able to yield the distinct clusters but the clusters determined by the FCMGA algorithm are found to be more compact compared to those provided by the KMGA algorithm. The performances of KMGA have been compared to those of FCMGA, in terms of the overall accuracy and their class-accuracy values which are discussed below.

4.1 Overall KIA

The results indicate a clear dependence on the range of input data included in the tests. Using only aerial images, many buildings were classified as roads because they have the same spectral reflectance and the class-accuracies were low. Using the aerial image and the LIDAR data increased the classification accuracy due to the suitability of LIDAR data for accurately detecting planar features, but still some errors occurred due to the poor horizontal accuracy of edge detection in the LIDAR data. The use of the aerial image, LIDAR data and extracted attributes increased the classification accuracy further since the attributes compensated for the weakness of LIDAR for edge detection.

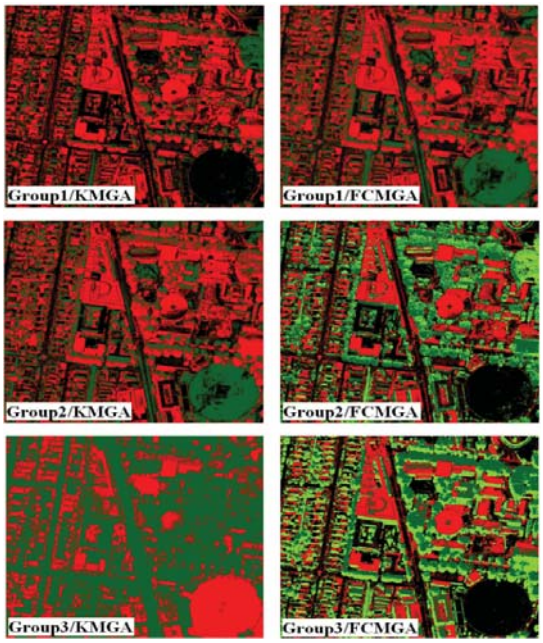


Figure 3: Classifications results. The colours indicate the different classes: Black stands for roads, dark Green for trees, light green for Grass and Red for buildings.

FCMGA was very stable in all situations achieving higher classification accuracy (fig. 4). The results also indicate the ability of the FCMGA to take advantage of the input data. The FCMGA outperformed the KMGA for the three groups of data. In the case of FCMGA, group three has yielded the best set of classes with 87.84% overall accuracy. One important note is that the results obtained for FCMGA improved when the elevation data was incorporated into the classification process. Once again, the results improved when the spectral attributes were applied.

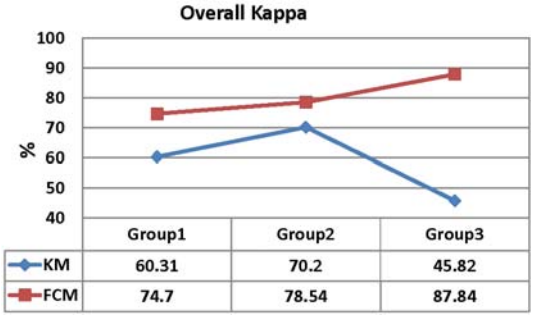


Figure 4: Overall Kappa results obtained for KMGA and FCMGA

On the other hand, the results obtained for KMGA improve when the elevation data are applied, and deteriorate suddenly to 45.82% when the spectral attributes are added. In this case, only two instead of four classes are clearly found. These results indicate that KMGA is inappropriate for real data sets in which

there are no definite boundaries between the clusters. Therefore, the authors believe that the poor performance of KMGA was not a result of any inadequate learning process of the algorithm but due to its own structure. This conforms to results obtained in Yang et al. (2008). They concluded that KMI seems to be unsuitable for the use in unsupervised classification of multi spectral image.

4.2 Per-class accuracy

An additional measure, per-class Kappa, was used to evaluate the performance of the two models. Unlike overall Kappa, per-class Kappa clearly shows how much the performance of the proposed method improves or deteriorates for each individual class. An assessment of the per-class Kappa confirms that the FCMGA performed the best in most cases as shown in Figures 5 to 7. Another advantage of the FCMGA over KMGA is that the achieved per-class Kappa errors are less variable.

In the case of RGB data, most of the per-class Kappa improved by the FCMGA. Whereas the application of FCMGA resulted in average per-class Kappa of 75.18%, the application of KMGA resulted in average of 60.29%. The roads class had the greatest increase from 42.6% using the KMGA to 76.36% with the FCMGA. Another advantage of the FCMGA over KMGA is that the achieved errors are less variable as shown in Figure 5. Whereas the application of FCMGA resulted in standard deviation of 6.37 for per-class Kappa, the application of KMGA resulted in a SD of 5.80.

In the case of RGB/LIDAR data, most of the per-class Kappa is improved by the FCMGA. Whereas the application of FCMGA resulted in average per-class Kappa of 81.68%, the application of KMGA resulted in average of 72.68%. The trees class had the greatest increase from 53.43% using the KMGA to 64.09% with the FCMGA. To the contrary, there was a decrease in per-class Kappa for the grass class from 98.93% using the KMGA to 97.86% with the FCMGA. However, those classes are still classified with relatively high per-class Kappa. Another advantage of the FCMGA over KMGA is that the achieved errors are less variable as shown in Figure 6. Whereas the application of FCMGA resulted in standard deviation of 17.59 for per-class Kappa, the application of KMGA resulted in a SD of 19.25.

In the case of RGB/LIDAR/Attributes data, most of the per-class Kappa is improved by the FCMGA. Whereas the application of FCMGA resulted in average per-class Kappa of 87.92%, the application of KMGA resulted in average of 60.96%. The grass class had the greatest increase from 28.58% using the KMGA to 89.06% with the FCMGA. Another advantage of the FCMGA over KMGA is that the achieved errors are less variable shown in Figure 7. Whereas the application of FCMGA resulted in standard deviation of 4.41 for per-class Kappa, the application of KMGA resulted in a SD of 13.56.

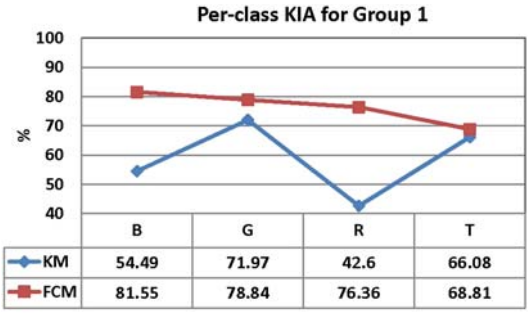


Figure 5: Per-class Kappa results obtained for KMGA and FCMGA in case of RGB dataset (B: Building; G: Grass; R: Roads; and T: trees)

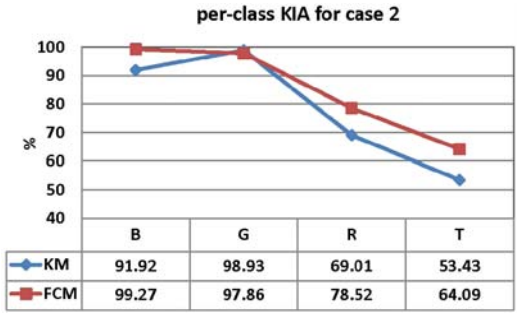


Figure 6: Per-class Kappa results obtained for KMGA and FCMGA in case of RGB/LIDAR dataset (B: Building; G: Grass; R: Roads; and T: trees)

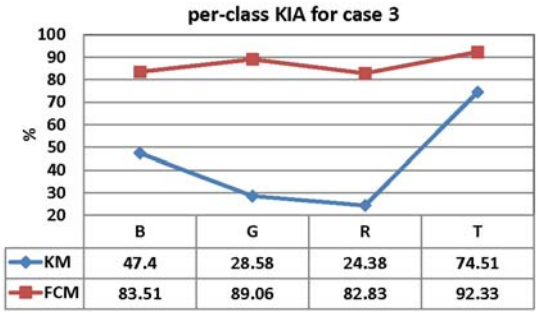


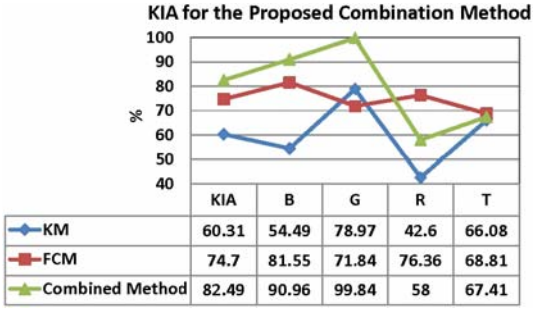
Figure 7: Per-class Kappa results obtained for KMGA and FCMGA in case of RGB/DSM/Intensity/nDSM/polymorphic texture strength dataset (B: Building; G: Grass; R: Roads; and T: trees)

4.3 Combined results

The improvement in overall Kappa achieved by the combination method compared with the KMGA and FCMGA was determined as shown in Figure 8. It is clear that the overall performances of the combined method are better than those of the KMGA and FCMGA. It can be seen that a considerable amount of the misclassified pixels have been recovered by the combined classification process. It can be concluded that the application of the combination process results

in the most significant improvement in classification accuracy. The strengths of each classifier have compensated for the weaknesses of the other. In general fusing KMGA and FCMGA improve classification accuracies. This demonstrates the benefit of combining different sensor sources at different classification levels.

Some of the per-class Kappa is improved by the combination method. Whereas the application of KMGA and FCMGA resulted in average per-class Kappa of 75.18% and 60.29% respectively, the application of the combined method resulted in average of 77.69%. The grass class had the greatest increase from 71.84% using the FCMGA to 99.84% with the combined method. To the contrary, there was a decrease in per-class Kappa for the roads and trees classes from 76.36% and 68.61% for roads and trees respectively, using the FCMGA, to 58% and 64.41% with the combined method. As a result, the achieved errors are more variable. Whereas the application of FCMGA resulted in standard deviation of 4.41 for per-class Kappa, the application of the combined method resulted in a SD of 13.56. However, no patterns can be derived that demonstrate that a certain classifier is better for a particular class. If a particular class is very important, FCMGA, KMGA and the combined results should be tested to select the best result for that class in a given study area. In general, the results indicate that the FCMGA is a more appropriate mapping technique for high resolution imagery. On the other hand, the KMGA method is still an unsatisfactory one to classify high resolution images.



	KIA	B	G	R	T
KM	60.31	54.49	78.97	42.6	66.08
FCM	74.7	81.55	71.84	76.36	68.81
Combined Method	82.49	90.96	99.84	58	67.41

Figure 8: Overall and per-class Kappa results obtained RGB dataset

The visual interpretation of the final classification result clearly shows a relatively high degree of noise in the KMGA and FCMGA results. In contrast to this, the classification that is based on the combined process appears more homogenous. Within the main land cover classes, some pixels in the aerial image are misclassified, whereas these pixels are correctly classified by the combination process. Figure 9, which is an enlarged portion from Figure 3, is a typical example showing the results from the KMGA classifier in Figure 9(*b*), FCMGA in Figure 9(*c*), and the combined results in Figure 9(*d*). The misclassified pixels are corrected and the noise is significantly reduced by the combined process. This clearly illustrates the different and complementary information provided by the two classifiers.

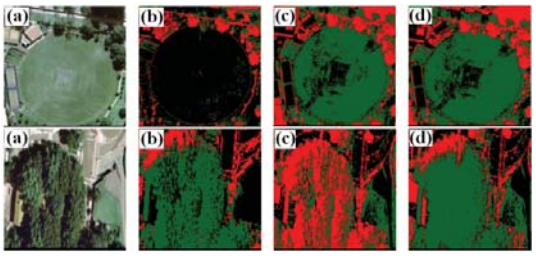


Figure 9: (a) Aerial Image, (b) Classification results of the KMGA; (c) classification results of the FCMGA; (d) Error correction after applying the fusion algorithm

5. Conclusions

This paper presented a GAs-based approach for unsupervised image classification. Experimental results were obtained by classifying a high resolution scene for a part of the region surrounding the University of New South Wales campus, Sydney Australia, depicting four different classes, namely buildings, roads, trees and grass. The study tried to optimize and validate GAs for unsupervised classification of high resolution digital imagery and LIDAR data. Two techniques have been applied: KMGA; and FCMGA. Also, the results obtained for both techniques were combined in order to improve the accuracy. Compared with the reference data, the results were evaluated based on a variety of criteria which includes visual inspection and the K-HAT statistics. Three groups of data were tested and evaluated. RGB dataset resulted in overall kappa of 60.31% in the case of KMGA, and 74.70% in the case of FCMGA. RGB/DSM introduced overall kappa of 70.20% in the case of KMGA, and 78.54% in the case of FCMGA. RGB/DSM/Intensity/nDSM/polymorphic texture strength achieved overall accuracy of 45.82% in the case KMGA, and 87.84% in the case of FCMGA. The combination of RGB/KMGA and RGB/FCMGA has improved the results and resulted in overall kappa of 82.49%. On the basis of the results drawn by this experiment it may be safely stated that the FCMGA outperformed the KMGA. In the future it is planned to process more and larger scenes and data sources in order to confirm the results found so far.

Acknowledgements

The authors would like to thank the AAMHatch (www.aamhatch.com.au) and the School of Surveying and Spatial Information Systems, The University of New South Wales, Sydney, Australia for providing the multispectral aerial imagery and LIDAR data.

References

Almeida, C.M. (2012) Genetic algorithms and data mining applied to optical orbital and LIDAR data for

object-based classification of urban land cover. Proceedings of the 4th GEOBIA, May 7-9, 2012 - Rio de Janeiro - Brazil. p. 649.

Bandyopadhyay, S. and U. Maulik (2002a). Genetic clustering for automatic evolution of clusters and application to image classification, Pattern Recognition, 35(6), pp. 1197–1208.

Bandyopadhyay, S. and U. Maulik (2002b) An evolutionary technique based on K-means algorithm for optimal clustering in R^N. Information Sciences, 146(1–4), pp. 221–237.

Bandyopadhyay, S., U. Maulik and A. Mukhopadhyay (2007). Multiobjective genetic clustering for pixel classification in remote sensing imagery. IEEE Transactions on Geoscience and Remote Sensing, 45 (5), pp. 1506 – 1511.

Benediktsson, J.A., J. Chanussot and M. Fauvel (2007). Multiple classifier systems in remote sensing: From basics to recent developments. MCS 2007, LNCS 4472, Springer Verlag, Berlin 2007, pp. 501-512.

Bezdek, J. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Plenum, 1981.

Bezdek, J.C., (1981). Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York, pp: 256.

Bradley, P.S. and U.M. Fayyad (1998). Refining initial points for K-means clustering. In Proceedings of the 15 International Conference on Machine Learning, pp. 91-99. Morgan Kaufmann.

Chen, C.C. and C.S. Lin (2007). A GA-based nearly optimal image authentication approach. International Journal of Innovative Computing, Information and Control, 3(3), pp. 631–640.

Chu, H.T. (2012). Combination of genetic algorithm and Dempster-Shafer theory of evidence for land cover classification using integration of SAR and optical satellite. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIX-B7, 2012 XXII ISPRS Congress, 25 August – 01 September 2012, Melbourne, Australia.

Coley, A.D. (1999). An introduction to genetic algorithms for scientists and engineers. World Scientific, Singapore, 188p.

Congalton, R.G. and R.A. Mead (1983). A quantitative method to test for fonsistency and correctness in photo interpretation. Photogrammetric Engineering and Remote Sensing. Vol. 49, Issue 1, PP. 69 – 74.

Dunn, J.C. (1974). A fuzzy relative of the ISODATA process and its in detecting compact well-separated clusters. Journal of Cybernetics, 3(3), pp. 32–57.

Förstner, W. and E. Gülch (1987). A fast operator for detection and precise location of distinct points, corners and centres of circular features. In Proceedings of the ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data, 2–4 June 1987, Interlaken, Switzerland, pp. 281–305.

Gen, M. and R. Cheng (1997). Genetic algorithms and engineering design. John Wiley & Sons, New York. 379p.

Ghanma, M. (2006). Integration of photogrammetry and LIDAR. PhD thesis, Calgary University, Canada.
Halder, A., S. Pramanik and A. Kar (2011). Dynamic image segmentation using fuzzy C-means based genetic algorithm. International Journal of Computer Applications (0975 – 8887)Volume 28– No.6.

Hall, L., B. Ozyurt and J. Bezdek (1999). Clustering with a genetically optimized approach. IEEE Transactions on Evolutionary computation, Vol. 3, No. 2, pp 103-112.

Holland, J.H. (1975). Adaption in natural artificial systems. Ann Arbor, MI: The University of Michigan Press.

Huang, C.L. and C.J. Wang (2006). A GA-based feature selection and parameters optimization for support vector machines. Expert Systems with Applications, Vol. 31, pp.231-240.

Kim, J.B. and H.J. Kim (2003). GA-based image restoration by isophote constraint optimization. EURASIP Journal on Advances in Signal Processing, 2003(3), pp. 238–243.

Liu, H., J. Li and M.A. Chapman (2003). Automated road extraction from satellite imagery using hybrid genetic algorithms and cluster analysis. Journal of Environmental Informatics. 1: (2), pp. 40-47.

Lu, D. and Q. Weng (2007). A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing, Vol. 28, Issue 5, pp. 823 - 870.

Lu, Y., T. Ma, W. Tian and S. Zhong (2013). Implementation of the fuzzy C-means clustering algorithm in meteorological data. International Journal of Database Theory and Application, 6(6), pp.1-18.

Mahi, H. and H.F. Izabatene (2011). Segmentation of satellite imagery using RBF neural network and genetic algorithm. Asian Journal of Applied Sciences, 4: 186-194.

Pham, D.T. and D. Karaboga (2000). Intelligent optimization technique. Springer, London, Great Britain, 261p.

Rosenfield, G.H. and K. Fitzpatrick-Lins (1986). A coefficient of agreement as a measure of thematic classification accuracy. Photogrammetric Engineering and Remote Sensing, Vol. 52, Issue 2: PP. 223 – 227.

Rothlauf, F. (2006). Representations for genetic and evolutionary algorithms. Springer, Netherlands, 314p.

Salah, M. and J. Trinder (2010). Support vector machines based filtering of LIDAR data: A grid based method. FIG Congress 2010, Sydney, Australia, 11-16 April 2010.

Stavrakoudis, D.G, J.B. Theocharis and G.C. Zalidis (2011). A multistage genetic fuzzy classifier for land cover classification from satellite imagery. Soft Computing, Volume 15, Issue 12, pp 2355-2374.

Suganya, R. and R. Shanthi (2012). Fuzzy C- means algorithm- A review. International Journal of Scientific and Research Publications, 2(11), November 2012 Edition, ISSN 2250-3153.

Tabassum, M. and K. Mathew (2014). A genetic algorithm analysis towards optimization solutions. International Journal of Digital Information and Wireless Communications (IJDIWC) 4(1), pp.124-142. The Society of Digital Information and Wireless Communications, 2014 (ISSN: 2225-658X).

Tan, F. and X. Fu (2008). A genetic algorithm-based method for feature subset selection. Soft Computing, vol. 12, Issue 2, PP.111-120.

Turi, R.H. (2001). Clustering-based colour image segmentation. PhD Thesis, Monash University, Australia.

Tyagi, J. and N. Verma (2015). Optimization of fuzzy C means clustering using genetic algorithm for an image. International Journal of Computer Applications, 121(17), pp. 29-32.

Van Coillie , F., L. Verbeke and R. Wulf (2007). Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. Remote Sensing of Environment, Volume 110, Issue 4, 30 October 2007, Pages 476–487.

Wikantika, K., P.L. Chairuddin, A. Harto, D. Suwardhi and S. Darmawan (2014). Estimation of wetland in Indonesia with multitemporal satellite imagery extraction model-based neural network (NN), genetic algorithm (GA) and fuzzy logic. FIG Congress 2014, Engaging the Challenges - Enhancing the Relevance, Kuala Lumpur, Malaysia 16 – 21 June 2014.

Xie, X. and G. Beni (1991). A validity measure for fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13, pp. 841–847.

Yang, M.S. and K.L. Wu (2001). A new validity index for fuzzy clustering. In Proceedings of the 10th IEEE International Conference on Fuzzy Systems,3, pp. 89–92.

Yang, M., Y. Yang, T. Su and K. Huang (2014b). An efficient fitness function in genetic algorithm classifier for landuse recognition on satellite images. The Scientific World Journal. Volume 2014, p.12, Available at: http://dx.doi.org/10.1155/2014/264512, [Accessed on August 2015].

Yang, Y.F., P. Lohmann and C. Heipke (2006). Genetic algorithms for the unsupervised classification of satellite images. In Proceedings of the Symposium of ISPRS Commission III, Photogrammetric Computer Vision (PCV ’06), 36, pp. 179–184.

Yang, M. (2007). A genetic algorithm (GA) based automated classifier for remote sensing imagery. Canadian Journal of Remote Sensing. 33(3), pp. 203-213.

Yang, M., Y. Yang and Y. Chen (2014a). Evaluation of fitness functions of GA classification. Proceedings of the 2014 conference companion on Genetic and evolutionary computation companion, 12 -16 July 2014, Vancouver, Canada, pp. 1479-1480.

Yang, Y.F., P. Lohmann and C. Heipke (2006). Genetic algorithms for the unsupervised classification of satellite images. Symposium of ISPRS, Commission III Photogrammetric Computer Vision PCV ' 06. Bonn, Germany, 20 – 22 September 2006.

Yang, Y., P. Lohmann and C. Heipke (2008). Genetic algorithms for multi spectral image classification. In: Schiewe, J., Michel, U. (Eds.): Geoinformatics paves the Highway to Digital Earth, Festschrift zum 60. Geburtstag von Prof. M. Ehlers (2008), 8, pp. 153-161.

Zhou, W, A. Troy and J.M. Grove (2008). Object-based land-cover classification and change analysis in the Baltimore metropolitan area using multi-temporal high resolution remote sensing data. Sensors; Vol. 8: pp.1613–1636.

Zhou, M., J. Shu and Z. Chen (2010). Classification of hyperspetral remote sensing image based on genetic algorithm and SVM. Remote Sensing and Modelling of Ecosystems for Sustainability VII, Proc. of SPIE, Vol.7809, 78090A, doi:10.1117/12.860153.