Journal of Basic and Applied Engineering Research
p-ISSN: 2350-0077; e-ISSN: 2350-0255; Volume 3, Issue 7; April-June, 2016, pp. 613-620

http://www.krishisanskriti.org/Publication.html

Fuzzy c- Means Classifier with Alpha-cuts in Application for Similarity and Dissimilarity Measures: A Literature Survey

Sayan Mukhopadhaya

Indian Institute of Remote Sensing, ISRO
E-mail: sayanmukhopadhaya@gmail.com

Abstract—*Fuzzy set theory has many applications. One of its important usage is in the classification of images using various classification methods like Fuzzy c- Means (FCM) classifier based on the fuzzy set theory. This paper analyses FCM using a literature review based on various articles ranging from 1987 to 2015 with the keywords to find out how FCM along with alpha-cuts and various similarity or dissimilarity measures have advanced in this period. On the basis of 75 articles, this work has classified the previous FCM classification works using the four categories such as: Land Cover Classification method; Fuzzy c- Classification; Measures of similarity and dissimilarity; and Fuzzy alpha- cuts in accordance with various research problems and domains. The advantages and outputs of the works are demonstrated in this work. Fortunately, this area of work gained its expected acknowledgment after being discovered about 30 years ago. Recent works have shown that many software and hardware are based on fuzzy logic. The main objective of this paper is to provide a comprehensive review of the techniques and resources used. Various usages of FCM along with alpha-cuts has been shown in this work. Classification structures for various applications of FCM have been defined. Selected bibliographies have been identified on the various topics.*

Keywords: *FCM, alpha- cuts, similarity and dissimilarity measures, land cover classification.*

1. INTRODUCTION

The works by Boyd et al.(2006); Foody et al. (2006) and Li et al. (2011) show that there is a need to have information of all the classes in the training set exhaustively, to determine a specific class by using supervised classification. But, this method may result in considerable error (Foody et al., 2006). Hence, supervised classification or hard classification is inappropriate for extracting a specific class (Foody et al., 2006). A problem like mixed pixel will also be faced by this conventional approach of classification (Upadhyay et al., 2013). Kumar et al.(2006 b) shows that mixed pixel problem are found on the boundary of two or more features in an image due to the pixel size compatibility with the class size. The mixed pixel problem can be solved by the fuzzy set theory, by using a membership function along with α-cuts/levels and quantifying the degree of belongingness of a pixel to a class (Dilo, 2006). The work by Foody (2000) show that Fuzzy c-Mean classifier can be used to solve the mixed pixel problem. This has been recognized in the past as well: "Fuzzy set theory provides a useful technique to allow a pixel to be a member of more than one category or class with graded membership" (Shankar et al., 2006)

Various work done by researchers show that distance norms have been used in image processing (Lee et al., 1996 ; Wang et al., 2005). The earlier work by Upadhyay et al.(2014) shows that distance norms like Euclidean, Mahalanobis and Diagonal Mahalanobis norms have been incorporated with FCM classifier. The work by Tyagi et al. (2015) shows that fuzzy classifier along with similarity and dissimilarity measures can be used to solve the mixed pixel problem. Earlier work done by Lee et al.,(2009) shows that if a similarity measure of a data-set has been found, it can also represent the dissimilarity, as a high level of similarity of data shows a low level of dissimilarity measure. The measure of similarity can be calculated on the distance between the data used and the overlapping area of two fuzzy membership functions also determines the similarity (Lee et al., 2009). There is a relationship between distance and similarity measures and the combination of similarity measure and distance measure shows the totality of information (Xuecheng, 1994).

2. LAND COVER CLASSIFICATION METHOD

The main purpose of image classification is that to classify every pixel into either on the basis of one to one classification (hard classification) or one to many classification (soft classification) (Mather and Tso, 2009). There are many classification methods to classify a remotely sensed image into different land cover types. According to Swain and Davis (1979) these methods can be categorized into: a) Methods based on whether a process of training is needed or not (i.e. supervised and unsupervised classification); b) Methods based on the usage and requirement of any parametric model (i.e. parametric and non-parametric). There are many algorithms developed for classifying images. Amid the prevailing algorithms, the most widespread are the maximum likelihood classifier (MLC), support vector machine (SVM, decision tress classifiers and neural network classifiers. Maximum Likelihood Classifier (MLC) algorithm is a supervised statistical approach for thematic mapping using pixel based information. MLC follows Gaussian rule approach and it becomes unreliable when the class size is small (Gopinath, 1998), but works fine for a large class size though there is high degree of computation. Though it has lot of limitations as it follows a normal distribution function for the signature of the classes (Swain and Davis, 1979), it is one of the most common and widely used classification algorithm (Wang, 1990 and Hansen et al.,1996).

Neural network classifier (NNC) algorithms avoid a few of the problems that are faced in MLC by choosing a non-parametric approach. They also do not follow a Gaussian rule approach. Neural networks have an advantage of high computation rate due to the presence of huge parallel networks, which resulted in the development of various other types of neural networks (Lippmann, 1987) such as : the most commonly used in the classification of remote sensing images is a collection of networks known as a Multi-Layer Perceptron (MLP) (Paola and Schowengerdt, 1995; Atkinson and Tatnall, 1997b). On the other hand, considering artificial neural network (ANN) which may be very complex as the learning rate can be very high for the data of higher dimensionality. Large training data is required for generalization as the data structure becomes complex on increasing the data dimensionality (Ablin and Sulochana, 2013). Decision tree classifier (DTC) uses a different approach for land cover classification. Safavian and Landgrebe (1991) in their work showed that a decision tree breaks a complex problem of classification into several stages of simple processes of decision making. There are univariate and multivariate decision tress which are determined based on the amount of variables used at each stage (Friedl and Brodley, 1997). At a global scale land cover classification is done using univariate decision tree (De Fries et al., 1998 ; Hansen et al., 2000). Multivariate decision trees are generally more compact than univariate decision trees and are also sometimes more accurate than univariate decision trees (Brodley and Utgoff, 1995). The hierarchical method provides an advantage that it is easily interpreted than ANN as the tree structure can be observed as white box. Another advantage is that it needs less complex training on comparison to ANN, but decision frames need to be framed for decision trees and they become complex when there is a large number of decision rules (Mather and Tso, 2009).

Support Vector Machine (SVM) classifier is a statistics based learning classification technique. It is used to allocate the labels as it was originally linear binary classifier (Mather and Tso, 2009). Construction of a separating hyperplane based on the properties of the training samples is the core operation of SVM. SVM has varieties of applications. Osuna and Freund (1997) has applied SVMs for human face detection along with digital image classification. Mukherjee et al. (1997) and Pal and Mather (2005) used SVM classifier for classifying remote sensing images. Huang et al. (2002) has showed that SVM gives higher accuracies than other classifiers like MLC, NNC and DTC. But, SVMs can be a bit time consuming as shown by Patra and Bruzzone (2011). Hard classifiers are poor in accounting information within mixed pixels and an analyzer has to adopt different methods like soft classifiers to get a proper way result. Soft classifiers results in different proportions of belongingness within a single pixel. These classifiers are generally based on fuzzy set theory, neural networks, etc. Fuzzy set theory classification takes heterogeneity and imprecise nature of the real world into account. It can also be used along with supervised classification algorithms. The next section provides a complete literature review on Fuzzy c- Means classification and the various distance measures that have been evaluated in the study.

3. FUZZY *C*- MEANS CLASSIFICATION

Fuzzy c- Means (FCM) is one of the popular fuzzy clustering method and this classification technique has been used for various applications for solving problems concerning remotely sensed data. This technique can be used with both supervised and unsupervised modes. Bezdek et al.(1984) showed that FCM can be incorporated with distance norms for clustering purpose with an unsupervised mode.

Various other works also show that FCM can be used to classify remotely sensed data. Bastin (1997) made a comparative study of soft classifiers like FCM, linear mixture model (LMM) and MLC, and it was concluded that FCM clustering algorithm can be applied widely as it does not make any statistical assumptions on the distribution of the training data. The work by Zhu (1997) shows how fuzzy logic can be used along with similarity algorithms to find out the uncertainty in a remotely sensed image. Thus, provides the areas where accuracy is high. Other works also show that fuzzy logic and fuzzy set theory can be used to classify remotely sensed images (Ji, 2003; Shalan et al. , 2003). The aforesaid works showed how mixed pixels are handled at the allocation stage, for class identification within a pixel. This is represented in the form of membership value of a class related to the class composition of the pixel. Wang (1990) used FCM approach with supervised mode to classify Landsat MSS and TM data consisting of seven land cover classes. FCM classification technique was able to distinguish the land cover classes in the areas with mixed pixels. The overall accuracy also improved by 5.11% while using FCM over conventional classification technique.

Foody (1996) had evaluated the execution of FCM and Fuzzy Neuron Network (FNN) techniques for classification of land cover using Airborne Thematic Mapper (ATM) data. A detailed study was carried out on the effects of fuzzy weight parameter ‘m’-value for the same dataset. It was observed that for m = 2.0, accurate fuzzy classification outputs were obtained for most of the cases. Thus, it was concluded that fuzzy classification technique gives more suitable results in land cover mapping than hard classification techniques. Atkinson et al.(1997a) had done a comparative study of Artificial Neural Networks (ANN), Mixture modelling and Fuzzy c-Means for mapping sub-pixel proportions of land cover classes for an area in the New Forest, U.K. It was observed that ANN was one of the accurate techniques; however, its successful implementation depends on accurate co-registration and availability of a training data set. Supervised Fuzzy c-Means classification gave slightly better results than mixture modelling.

Bastin (1997) made a comparative study between FCM, Linear Mixture Modelling and Maximum Likelihood Classifier for un-mixing pixels of low resolution present in combined Landsat TM data. The original TM data was operated as a referral map and the image was combined using mean and cubic filter having different kernel size, as the ground truth data was unavailable. Thereafter, the membership value for each classifier was calculated from the classified combined image. This result was then compared to continuous membership values of the proportional areas of sub-pixel available in the image with low-resolution. It was concluded for combined TM image at various scales Fuzzy c- means classifier provides a better approximation of sub-pixel land cover classes. Zhang and Foody (1998) applied FCM classification algorithm on SPOT HRV and Landsat TM data for classifying sub-urban area. It was inferred that the obtained outputs were advantageously accurate while applying fuzzy classification and evaluation methods over conventional hard classification or partially fuzzy methods. It was also observed that Kappa coefficient was more than twice while the fuzzy classification technique was applied as compared to the hard classification technique. Zhang and Foody (2001) presented two procedures for full fuzzy classification of remote sensing images; viz. a statistical method which is based on modified FCM clustering method along with a supervised approach and another is based on Artificial Neural Network (ANN). Both the procedures were used to deduce fully-fuzzy classifications of land cover, with fuzzy ground data, as it is very essential for both training and testing of the classifications. Results confirmed the superiority of fully -fuzzy over partially -fuzzy classification. Further, it was found that fully fuzzy class was more beneficial as it has more relaxed requirements for training pixels i.e. these need not be pure.

Lucas et al. (2002) used FCM and linear un-mixing techniques for sub-pixel habitat mapping of a coastal dune ecosystem from airborne imaging spectrometer image (CASI). It was observed that these techniques could be useful to find out land cover class proportions at sub-pixel level. Ibrahim et al. (2005) had done a comparative study of a three soft classification techniques – probabilistic maximum likelihood classifier and two fuzzy set theory based classifiers such as Fuzzy *c*- Means (FCM) and Possibilistic *c*- Means (PCM). Each classifier was evaluated on its uncertainty and accuracy measures by adopting fuzzy error matrix. Furthermore, the study concluded that to produce accurate and proper land cover classification there should be incorporated mixed pixels (which shows variability in the allocation of class) at all stages of the classifying process of remotely sensed images. Okeke and Karnieli (2006) showed methods to classify historical aerial photographs and also how fuzzy classification technique can be used to measure the accuracy of the outputs after classification. The results also showed the benefits of the usage of unsupervised or supervised mode of classification in combination with FCM algorithm and bootstrap method of resampling. These methods were concluded to be useful for classification and also for accuracy measurement for areas where training data and ground sample data were beyond the bound of possibility to obtain. Hore et al. (2007) proposed a simple FCM algorithm and measured the performance using image compression technique. It was found that this algorithm produced a better results compared to other clustering technique. Further, it produced excellent speed-ups in clustering and thus can be used even if the data cannot be fully loaded in to the computer memory.

Zhang et al. (2008) proposed a Similarity based Fuzzy and Possibilistic c-Means algorithm called SFPCM. UCI (Asuncion et al., 2007) repository data sets were used. Similarity was computed as the reciprocal of the Euclidean distance. The results showed that SFPCM has better classification accuracy than FPCM (Fuzzy Possibilistic c-Means) algorithm and also converges much quicker than FPCM on the UCI repository data-sets. Dwivedi et al. (2012) carried out a comparison of FCM (Fuzzy c-Means) and PCM (Possibilistic c-Means) using AWIFS, LISS-III and LISS-IV data sets from ResourceSat-1 (IRS-P6) satellite. Accuracy Assessment was done by using FERM, SCM and Fuzzy Kappa Coefficient; norms were considered namely Euclidean, Mahalanobis and Diagonal Mahalanobis only. The FCM-overall accuracy of was measured to be 97% with the MIN-LEAST operator with the optimized weighting exponent "m" value of 4.0. Kannan et al. (2013) showed that FCM clustering algorithm had not only been used in remote sensing images but also in medical image processing. A synthetic MRI image was used for the work. The used FCM methodology dealt with the uncertainty present in the data as the segmentation of the MRI images took place. The results showed that accuracy of the mentioned approach was higher than the other standard FCM methods and the results were obtained in lesser number of iterations.

4. MEASURES OF SIMILARITY AND DISSIMILARITY

Zwick et al. (1987) studied and compared the 19 measures of similarity or dissimilarity among the fuzzy sets. These measures were both geometric and set-theoretic, and they were compared on their behavioral performances. It was concluded that distance measures could be evaluated on one's interest and the best distance measure should be chosen on the basis of high correlations for the particular situation. Deer et al. (1996) studied the fuzzy classifier based on the conditioned reciprocal of the Mahalanobis distance. It demonstrated a general technique of image classification for a remotely sensed image of a rural region of South Australia containing six classes. The result of this experiment showed that the weighting exponent "m" at 1.25 showed best results. Liu et al.(2009) proposed improved FCM algorithm which is based on Standard Mahalanobis distance (FCM-SM). A comparative study of this algorithm with the other algorithms like Gath-Geva (GG) clustering algorithm, Gustafson-Kessel (GK) clustering algorithm, FCM with adaptive Mahalanobis distance (FCM-M), FCM with common Mahalanobis distance (FCM-CM) and FCM with Euclidean distance was done on three real data-sets. The results showed that FCM-SM showed better performance than the others.

Takahashi et al.(2011) studied on the process of handling uncertain data or missing data using clustering algorithm fuzzy c- means (FCM) by using cosine correlation (FCM-C) and tolerance vector (FCMT-C). The results showed that there were differences between FCMT-C and FCM-C, but none algorithms could cluster the data properly. Ye (2011) investigated a cosine similarity measure for the Fuzzy Multicriteria Decision Making (FMCDM) problem between two trapezoidal fuzzy numbers. This work had also been extended for alternatives of ranking and also practical example of the work was used to choose the alternatives of the investment. The results showed that proposed method had been simple and effective. Das (2013) analyzed how pattern recognition technique can be used with Fuzzy c-Means (FCM) classifier. In this work, the data analyzed was in the form of numerical vectors with a predefined cluster. Besides, Euclidean other distances like Canberra and Hamming were also used in FCM classifier to get the variation in the outputs of membership values of the objects in the different clusters. The results showed that Euclidean produced the fastest and the most expected outputs whereas the outputs with Canberra were slowest and the least expected. Charulatha et al. (2013) had done a comparative study on FCM classifier with various distance metrics like Euclidean, Manhattan, Canberra, TChebychev and Cosine. The results showed that the different distance metrics work differently with the variation of weighting exponent "m" and the author concluded that there is a need of exhaustive exploration of the distance metrics for different kind of data-sets on various clustering algorithms.

Jafar et al., (2013) studied comparatively the K-means clustering algorithm and FCM by incorporating distance measures such as Chebyshev, Chi-square and σ- distance measure on four datasets. The results showed that FCM based on Chi-square distance measure had better performance than Chebyshev distance measure and also the Chebyshev distance measure showed maximum partition coefficient and the partition entropy was the least for it. Kouser et al. (2013) had applied K-means clustering algorithm with distances measures like Euclidean, Manhattan and Chebyshev on flower data set. The experiment results showed that the overall accuracy for Chebyshev distance and Euclidean distance is comparable, whereas Chebyshev distance had the highest number of iterations. Dik et al. (2014) showed how fuzzy clustering results improve when a weighting factor is introduced in the inter-object distances. The distances considered were Euclidean, Manhattan, Spearman and Chebyshev incorporated with FCM and were tested on three datasets. The results showed that there was a significant improvement in the accuracy when weighted distances were considered over unweighted distances. Sinwar et al. (2014) studied two distance metrics Euclidean and Manhattan incorporated with simple K-Means clustering algorithm on two real and one synthetic datasets. The results of the experiments performed showed that Euclidean approach has better outcomes than Manhattan approach on the basis of number of iterations for calculating the centroid of the datasets used during the overall clustering process.

Goyal et al. (2015) compared two similarity measures such as cosine measure and k-means algorithm measures using fuzzy similarity. The results obtained showed that fuzzy similarity required lesser time than cosine similarity, hence fuzzy similarity was considered better than cosine similarity. Zhao et al. (2015) studied a fuzzy clustering algorithm incorporated with Mahalanobis distance. The dissimilarity measure was defined by Mahalanobis distance in place of Euclidean distance to reduce the noise sensitivity. A synthetic image, a subset of Berkeley image and a generated 2D dataset were used for the experiment. A comparative study was also done on Mahalanobis distance based FCM (MFCM) and Kullback-Leibler information based FCM (KLFCM). The results showed that MFCM algorithm have much more accuracy than KLFCM as it considers the covariance and hence, can segment a divergent cluster into a single class.

5. FUZZY ALPHA-CUTS

Reznik et al. (1994) demonstrated the method of alpha-cut border mapping. . This method was implemented along with a proportional–integral–derivative controller (PID controller). The results showed that the method of alpha-cut border mapping is quicker than defuzzification of fuzzy output set. Thus, making it as good as or comparable for real-time control applications. Kainz (2007) and Ponce-Cruz et al. (2010) explained the concept of alpha-cut in vividly and described a fuzzy set being composed of crisp sets by using the concept of alpha-cuts. It was also explained that alpha-cut concept can be used to know all the elements which belong to a fuzzy set and also possess some degree of membership. Xexéo (1997) explained that the concept of alpha-cut is important as it could be used to deduce fuzzy functions from crisp sets. He also described the difference between the concepts of alpha-cut and threshold level. Dunyak et al. (1997) worked on fuzzy neural networks and showed that the previous problem of restricted weights could be solved by using alpha-cuts. The work demonstrated a method by reducing the network dimensionality at the training stage by using a subset of small size for each alpha-cut and then the network was retrained to get the interpolated values tuned. Thus, making it unconstrained and the standard neural network method could be applied without restrictions and also with reduced dimensionality of the network. Abebe et al. (2000) showed a comparative study of Monte-Carlo Simulation (MCS) and Fuzzy alpha-cut (FAC) techniques while analyzing the uncertainty in the model parameters. The paper compared the two techniques on their features, advantages and drawbacks while applying them to analyze the uncertainty in transport modelling of ground water at the Vannetin basin in France. The results showed that FAC was much faster than MCS as it took fewer simulations to get an adequate and accurate result.

Wong et al. (2001) examined the usage of Improved Multidimensional alpha-cut (IMUL) fuzzy interpolation method in mineral processing to determine the parameter d50c of hydrocyclone. The results showed that IMUL with fuzzy interpolation method could provide a good alternative for controlling on-line hydrocyclone as it resulted satisfactory human-computer interaction while incorporating fuzzy-sets. Yang et al. (2008) showed a comparative study on partition index maximization (PIM) algorithm, fuzzy c- means (FCM) clustering algorithm, FCM with alpha-cuts (FCMα), fuzzy c-regressions (FCR) and FCR with alpha-cuts (FCRα). The results showed that on implementing the concept of alpha-cuts to FCM and FCR robustness could be achieved with smaller iterations. Kreinovich (2013) extended his ideas to fuzzy mathematics and fuzzy data processing from fuzzy logic and made some important proofs for α-cuts, such as:

The membership function and α-cut representations are not same from the algorithmic point of view.

There prevails a c-membership function for which computation of α-cuts are not possible and vice-versa is also true.

In general, computation of fuzzy data processing is not possible for membership functions, but exceptions are there for α-cuts.

Other authors have shown that alpha-cuts can be used for solving various problems like;

Lee et al.(2015) showed the usage of alpha-cut as a filter in proxy caching mechanism for wireless services. This mechanism was demonstrated to monitor the traffic flow and thus guaranteeing exact and faster streaming of services while buffer caching. The results of the work showed that the given mechanism has better executing performance than other caching techniques like S-caching, I-caching and C-caching mechanisms.

Bencherif et al. (2013) showed fuzzy alpha-cuts could be used to improve product development so that the customer requirements were met. This was done by optimizing the performance of the quality function. The optimization technique composed of the use of fuzzy sets, alpha-cut mechanisms, fuzzy ranking and clustering techniques to focus on the priorities of the customer requirements and required engineering. This resulted in higher accuracy and also exact information about the planning phase of the product.

6. ACKNOWLEDGEMENTS

This work was a part of Master of Science work at Faculty of Geo-Information Science and Earth Observation, University of Twente (Mukhopadhaya, 2016).

REFERENCES

[1] Abebe, A., Guinot, V., & Solomatine, D. (2000). Fuzzy alpha-cut vs. Monte Carlo techniques in assessing uncertainty in model parameters. *Proc. 4th International Conference on Hydroinformatics, Iowa City, USA*, (July), 1–8. http://doi.org/10.1.1.112.8621&amp

[2] Ablin, R., & Sulochana, C. H. (2013). A Survey of Hyperspectral Image Classification in Remote Sensing. *International Journal of Advanced Research in Computer and Communication Engineering*, *2*(8), 2986–3000.

[3] Asuncion, A., & Newman, D. J. (2007). UCI Machine Learning Repository. Retrieved February 1, 2016, from http://www.ics.uci.edu/~mlearn/MLRepository.html

[4] Atkinson, P. M., Cutler, M. E. J., & Lewis, H. (1997). Mapping sub-pixel proportional land cover with AVHRR imagery. *International Journal of Remote Sensing*, *18*(4), 917–935. http://doi.org/10.1080/014311697218836

[5] Atkinson, P. M., & Tatnall, A. R. (1997). Introduction: neural networks in remote sensing. *International Journal of Remote Sensing*, *18*(4), 699–709. http://doi.org/10.1080/014311697218700

[6] Bastin, L. (1997). Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities as tools for unmixing coarse pixels. *International Journal of Remote Sensing*, *18*(17), 3629–3648. http://doi.org/10.1080/014311697216847

[7] Bencherif, F., Mouss, L. H., & Benaicha, S. (2013). Fuzzy Alpha-cuts to capture customer requirements in improving product development. In *2013 11th International Symposium on Programming and Systems (ISPS)* (pp. 1–6). IEEE. http://doi.org/10.1109/ISPS.2013.6581493

[8] Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy *c*-means clustering algorithm. *Computers & Geosciences*, *10*(2–3), 191–203. http://doi.org/10.1016/0098-3004(84)90020-7

[9] Boyd, D. S., Sanchez-Hernandez, C., & Foody, G. M. (2006). Mapping a specific class for priority habitats monitoring from satellite sensor data, (August 2015). http://doi.org/10.1080/01431160600554348

[10] Brodley, C. E., & Utgoff, P. E. (1995). Multivariate decision trees. *Machine Learning*, *19*(1), 45–77. http://doi.org/10.1007/BF00994660

[11] Charulatha, B. S., Rodrigues, P., Chitralekha, T., & Rajaraman, A. (2013). A Comparative study of different distance metrics that can be used in Fuzzy Clustering Algorithms. *International Journal of Emerging Trends and Technology in Computer Science (IJETTICS)*.

[12] Das, S. (2013). Pattern Recognition using the Fuzzy c-means Technique. *International Journal of Energy, Information and Communications*, *4*(1), 1–14.

[13] De Fries, R. S., Hansen, M., Townshend, J. R. G., & Sohlberg, R. (1998). Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, *19*(16), 3141–3168. http://doi.org/10.1080/014311698214235

[14] Deer, P., Eklund, P., & Norman. (1996). A Mahalanobis Distance Fuzzy Classifer. In Narasimhan & Jam (Eds.), *4th Australian and New Zealand Conference on Intelligent Information Systems (ANZIIS '96)* (pp. 220–224).

[15] Dik, A., Bouroumi, A., & Ettouhami, A. (2014). Weighted distances for fuzzy clustering. *Applied Mathematical Sciences*, *8*(4), 147–156. http://doi.org/10.12988/ams.2014.311651

[16] Dilo, A. (2006). *Representation of and reasoning with vagueness in spatial information: A system for handling vague objects*. *Geoinformatics* (Vol. PhD).

[17] Dunyak, J., & Wunsch, D. (1997). Training fuzzy number neural networks with alpha-cut refinements. In *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation* (Vol. 1, pp. 189–194). IEEE. http://doi.org/10.1109/ICSMC.1997.625747

[18] Dwivedi, R., Kumar, A., & K. Ghosh, S. (2012). Study of Fuzzy based Classifier Parameter Across Spatial Resolution. *International Journal of Computer Applications*, *50*(11), 17–24. http://doi.org/10.5120/7814-1074

[19] Foody, G. M. (1996). Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data. *International Journal of Remote Sensing*, *17*(7), 1317–1340. http://doi.org/10.1080/01431169608948706

[20] Foody, G. M. (2000). Estimation of sub-pixel land cover composition in the presence of untrained classes. *Computers and Geosciences*, *26*(4), 469–478. http://doi.org/10.1016/S0098-3004(99)00125-9

[21] Foody, G. M., Mathur, A., Sanchez-Hernandez, C., & Boyd, D. S. (2006). Training set size requirements for the classification of a specific class. *Remote Sensing of Environment*, *104*(1), 1–14. http://doi.org/10.1016/j.rse.2006.03.004

[22] Friedl, M. a., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, *61*(3), 399–409. http://doi.org/10.1016/S0034-4257(97)00049-7

[23] Gopinath, R. A. (1998). Maximum likelihood modeling with Gaussian distributions for classification. *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, *2*(914), 661–664. http://doi.org/10.1109/ICASSP.1998.675351

[24] Goyal, M. M., Agrawal, N., Sarma, M. K., & Kalita, N. J. (2015). Comparison Clustering using Cosine and Fuzzy set based Similarity Measures of Text Documents. *International Conference on Computing and Communication Systems 2015 (I3CS'15)*, (x), 4. Retrieved from http://arxiv.org/abs/1505.0168

[25] Hansen, M. C., Defries, R., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *21*(6), 1331–1364. http://doi.org/10.1080/014311600210209

[26] Hansen, M. C., Dubayah, R., & DeFries, R. S. (1996). Classification trees: An alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, *17*(5), 1075–1081. http://doi.org/10.1080/01431169608949069

[27] Hore, P., Hall, L. O., & Goldgof, D. B. (2007). Creating streaming iterative soft clustering algorithms. *Annual Conference of the North American Fuzzy Information Processing Society - NAFIPS*, 484–488. http://doi.org/10.1109/NAFIPS.2007.383888

[28] Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*(4), 725–749. http://doi.org/10.1080/01431160110040323

[29] Ibrahim, M. a., Arora, M. K., & Ghosh, S. K. (2005). Estimating and accommodating uncertainty through the soft classification of remote sensing data. *International Journal of Remote Sensing*, *26*(14), 2995–3007. http://doi.org/10.1080/01431160500057806

[30] Jafar, O. A. M., & Sivakumar, R. (2013). A Comparative Study of Hard and Fuzzy Data Clustering Algorithms with Cluster Validity Indices. In *International Conference on Emerging Research in Computing, Information, Communication and Applications, ERCICA 2013* (pp. 775–782). Elsevier Publications.

[31] Ji, M. (2003). Using fuzzy sets to improve cluster labelling in unsupervised classification. *International Journal of Remote Sensing*, *24*(4), 657–671. http://doi.org/10.1080/01431160210146226

[32] Kainz, W. (2007). Fuzzy Logic and GIS. Vienna: University of Vienna.

[33] Kannan, S. R., Devi, R., Ramathilagam, S., & Takezawa, K. (2013). Effective FCM noise clustering algorithms in medical images. *Computers in Biology and Medicine*, *43*(2), 73–83. http://doi.org/10.1016/j.compbiomed.2012.10.002

[34] Kouser, K., & Sunita. (2013). A comparative study of K Means Algorithm by Different Distance Measures. *International Journal of Innovative Research in Computer and Communication Engineering*, *1*(9), 2443–2447.

[35] Kreinovich, V. (2013). Membership Functions or alpha-Cuts ? Algorithmic (Constructivist) Analysis Justifies an Interval Approach. *Applied Mathematical Science*, *7*(5), 217–288. Retrieved from http://digitalcommons.utep.edu/cgi/viewcontent.cgi?article=1711&context=cs_techrep

[36] Kumar, A., Ghosh, S. K., & Dadhwal, V. K. (2006). Sub-pixel land cover mapping: SMIC system. In *ISPRS International Symposium on Geospatial Databases for Sustainable Development, Goa, India*. Retrieved from http://www.isprs.org/proceedings/xxxvi/part4/wg-iv-2-6.pdf

[37] Lee, C., & Lee, G. K. (2015). An Alpha Cut (α-cut) Filter-Based Proxy Caching Control for Wireless Coverage Streaming Services. *Wireless Personal Communications*. http://doi.org/10.1007/s11277-015-3032-8

[38] Lee, S. H., Pedrycz, W., & Sohn, G. (2009). Design of similarity and dissimilarity measures for fuzzy sets on the basis of distance measure. *International Journal of Fuzzy Systems*, *11*(2), 67–72.

[39] Lee, Y., & Horng, S.-J. (1996). The Chessboard Distance Transform and the Medial Axis Transform are Interchangeable. In *Parallel Processing Symposium, 1996., Proceedings of IPPS*

'96, The 10th International (pp. 424–428). Honolulu, HI: IEEE. http://doi.org/10.1109/IPPS.1996.508090

[40] Li, W., Guo, Q., & Elkan, C. (2011). A positive and unlabeled learning algorithm for one-class classification of remote-sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(2), 717–725. http://doi.org/10.1109/TGRS.2010.2058578

[41] Lippmann, R. P. (1987). An Introduction ' to Computing with Neural Nets. *Ieee Assp Magazine*, (April).

[42] Liu, H., Jeng, B., Yih, J., & Yu, Y. (2009). Fuzzy C-Means Algorithm Based on Standard Mahalanobis Distances. In *Proceedings of the 2009 International Symposium on Information Processing (ISIP'09)* (Vol. 2, pp. 422–427). Huangshan, P. R. China: ACADEMY PUBLISHER.

[43] Lucas, N. S., Shanmugam, S., & Barnsley, M. (2002). Sub-pixel habitat mapping of a costal dune ecosystem. *Applied Geography*, *22*(3), 253–270. http://doi.org/10.1016/S0143-6228(02)00007-3

[44] Mather, P., & Tso, B. (2009). *Classification methods for remotely sensed data* (Second). CRC Press - Taylor & Francis Group.

[45] Mukherjee, S., Osuna, E., & Girosi, F. (1997). Nonlinear prediction of chaotic time series using support vector machines. *IEEE Workshop on Neural Networks for Signal Processing VII*, 511. http://doi.org/10.1109/NNSP.1997.622433

[46] Mukhopadhaya, S. (2016). *Exploring measures of similarity / dissimilarity for fuzzy classifier : from data quality to distance quality*. (A. Stein & A. Kumar, Eds.). Enschede: University of Twente. Retrieved from http://www.itc.nl/library/papers_2016/msc/gfm/mukhopadhaya.pdf

[47] Okeke, F., & Karnieli, a. (2006). Methods for fuzzy classification and accuracy assessment of historical aerial photographs for vegetation change analyses. Part I: Algorithm development. *International Journal of Remote Sensing*, *27*(1), 153–176. http://doi.org/10.1080/01431160500166540

[48] Osuna, E., & Freund, R. (1997). An improved training algorithm for support vector machines. *Neural Networks for Signal Processing VII --- Proceedings of the 1997 {IEEE} Workshop*, 276–285. http://doi.org/10.1109/NNSP.1997.622408

[49] Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, *26*(5), 1007–1011. http://doi.org/10.1080/01431160512331314083

[50] Paola, J. D., & Schowengerdt, R. A. (1995). A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of Remote Sensing*, *16*(16), 3033–3058. http://doi.org/10.1080/01431169508954607

[51] Patra, S., & Bruzzone, L. (2011). A fast cluster-assumption based active-learning technique for classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(5), 1617–1626. http://doi.org/10.1109/TGRS.2010.2083673

[52] Ponce-Cruz, P., & Ramirez-Figueroa, F. D. (2010). Fuzzy Logic. In *Intelligent Control Systems with LabVIEW™* (pp. 9–46). London: Springer London. http://doi.org/10.1007/978-1-84882-684-7_2

[53] Reznik, L., & Stoica, A. (1994). Mapping alpha-cut borders: classification and PID realization. In *Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference* (Vol. 3, pp. 1604–1607). IEEE. http://doi.org/10.1109/FUZZY.1994.343935

[54] Safavian, R. S., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transaction Is on Systems, Man, and Cybernetics*, *21*(3), 660–674. http://doi.org/10.1109/21.97458

[55] Shalan, M. a., Arora, M. K., & Ghosh, S. K. (2003). An evaluation of fuzzy classifications from IRS 1C LISS III imagery: A case study. *International Journal of Remote Sensing*, *24*(15), 3179–3186. http://doi.org/10.1080/0143116031000094791

[56] Shankar, B. U., Meher, S. K., Ghosh, A., & Bruzzone, L. (2006). Remote Sensing Image Classification : A Neuro-fuzzy MCS Approach. *Imaging*, 128–139.

[57] Sinwar, D., & Kaushik, R. (2014). Study of Euclidean and Manhattan Distance Metrics using Simple K-Means Clustering. *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, *2*(5), 270–274.

[58] Swain, P. H., & Davis, S. M. (1979). *Remote Sensing: The Quantitative Approach*. *McGraw-Hill International Book Company*. New York, USA.

[59] Takahashi, A., Endo, Y., & Miyamoto, S. (2011). Fuzzy c-means clustering for data with tolerance using cosine correlation. In *2011 IEEE International Conference on Granular Computing* (Vol. 11, pp. 619–624). IEEE. http://doi.org/10.1109/GRC.2011.6122668

[60] Tyagi, U., & Kumar, A. (2015). Similarity and Dissimilarity Measures with Fuzzy Classifiers : SMIC Tool. In *9th International Conference on ASEICT*. New Delhi.

[61] Upadhyay, P., Ghosh, S. K., & Kumar, A. (2013). Ecological Informatics Moist deciduous forest identi fi cation using temporal MODIS data — A comparative study using fuzzy based classi fi ers. *Ecological Informatics*, *18*, 117–130. http://doi.org/10.1016/j.ecoinf.2013.07.002

[62] Upadhyay, P., Ghosh, S. K., & Kumar, A. (2014). A Brief Review of Fuzzy Soft Classification and Assessment of Accuracy Methods for Identification of Single Land Cover, *2*(Mlc), 1–13.

[63] Wang, F. (1990). Fuzzy Supervised Classification of Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, *28*(2), 194–201. http://doi.org/10.1109/36.46698

[64] Wang, L., Zhang, Y., & Feng, J. (2005). On the Euclidean Distances of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(8), 1334–1339.

[65] Wong, K. W., Fung, C. C., Eren, H., & Gedeon, T. (2001). Determination of parameter d50c of hydrocyclones using improved multidimensional alpha-cut based fuzzy interpolation technique. In *IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No.01CH 37188)* (Vol. 3, pp. 1602–1606). IEEE. http://doi.org/10.1109/IMTC.2001.929474

[66] Xexéo, G. (n.d.). Fuzzy logic. Rio de Janeiro: Computing Science Department and Systems and Computing Engineering Program, Federal University of Rio de Janeiro, Brazil.

[67] Xuecheng, L. (1994). Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems*, *62*(3), 377. http://doi.org/10.1016/0165-0114(94)90124-4

[68] Yang, M., Wu, K., Hsieh, J., & Yu, J. (2008). Alpha-Cut Implemented Fuzzy Clustering Algorithms and Switching Regressions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *38*(3), 588–603. http://doi.org/10.1109/TSMCB.2007.915537

[69] Ye, J. (2011). Multicriteria decision-making method based on a cosine similarity measure between trapezoidal fuzzy numbers. *International Journal of Engineering, Science and Technology*, *3*(1), 272–278. http://doi.org/10.4314/ijest.v3i1.67654

[70] Zhang, C., Zhou, Y., & Martin, T. (2008). Similarity Based Fuzzy and Possibilistic c-means Algorithm. *Proceedings of the 11th Joint Conference on Information Science*, 1–6.

[71] Zhang, J., & Foody, G. M. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing*, *19*(14), 2721–2738. http://doi.org/10.1080/014311698214479

[72] Zhang, J., & Foody, G. M. (2001). Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and artificial neural network approaches. *International Journal of Remote Sensing*, *22*(4), 615–628. http://doi.org/10.1080/01431160050505883

[73] Zhao, X., Li, Y., & Zhao, Q. (2015). Mahalanobis distance based on fuzzy clustering algorithm for image segmentation. *Digital Signal Processing*, *43*, 8–16. http://doi.org/10.1016/j.dsp.2015.04.009

[74] Zhu, A. (1997). Measuring uncertainty in class assignment for natural resource maps under fuzzy logic. *Photogrammetric Engineering & Remote Sensing*, *63*(10), 1195–1202.

[75] Zwick, R., Carlstein, E., & Budescu, D. V. (1987). Measures of similarity among fuzzy concepts: A comparative analysis. *International Journal of Approximate Reasoning*, *1*(2), 221–242. http://doi.org/10.1016/0888-613X(87)90015-6