Application of General Fuzzy Min-Max Neural Network for the Clustering of Satellite Thermal Infrared Images

Barnali Goswami
Department of Computer Applications
Narula Institute of Technology
Kolkata, India
barnalig@ieee.org

Gupinath Bhandari
Department of Civil Engineering
Jadavpur University
Kolkata, India
g.bhandari@civil.jdvu.ac.in

***Abstract*— In order to estimate precipitation from thermal infrared images (TIR), the area under convective cloud cover has to be reckoned. Several pattern recognition techniques have been implemented so far. In a previous work, Fuzzy Min-Max (FMM) Neural Network was implemented on TIR images for the above purpose. There is an already proposed technique General Fuzzy Min-Max (GFMM), which is a modified and enhanced version of FMM that can be used both for supervised classification and unsupervised clustering, which is not present in FMM. In this paper, a novel application area for the clustering using GFMM Neural Network has been proffered by taking advantage of its dual quality. Previously it has been already shown that the proposed application area is appropriate for FMM; however, in the present study it appears that GFMM is even a better alternative for the identification of convective clouds from TIR images, for the further purpose of pluviometry.**

***Index Terms*—Pattern recognition, clustering, general fuzzy min-max neural networks, convective clouds, thermal infrared images.**

I. INTRODUCTION

One of the most decisive steps in the process of prediction of probable flood event, during monsoon, is the estimation of precipitation. Satellite images are useful for studying a large scale of area without visiting it physically. If the convective cloud cover can be determined from satellite thermal infrared images (TIR) then an estimation of precipitation can be done in order to determine whether a flood event is likely.

Several studies have been dedicated to identification and extraction of clouds from satellite images [1-3]. The process of identification of cloudy pixels from TIR images start with the segmentation of the images. In the previous study [1][5-9][11], *k-means* algorithm has been used for the purpose of clustering. But, *k-means* suffer from various drawbacks [15-17]; hence, fuzzy min-max (FMM) neural network [17] has been proposed [10] for the same.

General fuzzy min-max (GFMM) neural network is an augmented form of FMM [12]. In contrast to FMM, GFMM is the combination of clustering and classification tool. It can be used solely for the purpose of clustering or classification or hybrid of both. In this study, GFMM has been utilized for the clustering TIR images and the results have been analysed.

The present paper has been designed in the following format: section II defines the method of clustering with a brief detail of FMM and the description of GFMM, section III explains about the experiments, results and its analysis, followed by conclusion in section IV.

II. CLUSTERING OF TIR IMAGES

The clustering process on TIR image identifies the clusters of cloudy pixels only. Now further processing is done on these clusters for obtaining the area and other factors that may affect the precipitation condition in the study area.

There are several clustering algorithms available. In some of the previous works [1][5-9][11] *k-means* clustering algorithm has been utilized. It is a very simple algorithm meant for crisp clustering, but it suffers from several drawbacks; such as the *dead-unit* problem [15]. This means, if some of the points from the beginning are considerably far from the input data in *k-means*, then they are no more considered in the coming passes and are treated as dead units, i.e. , *k-means* is severely affected by outliers. Another problem is that the number of clusters has to be provided manually.

To solve the above mentioned problems FMM has been applied for the clustering of TIR images [10]. FMM was proposed by Simpson [16][17] and the most striking feature of this technique is that the required number of clusters is automatically calculated by the technique itself.

A. Fuzzy Min-Max Neural Network

FMM is a neuro-fuzzy technique where fuzzy logic has been merged with ART1 neural network [13]. FMM has different versions for supervised classification [16] and unsupervised clustering [17]. In a previous study [10] unsupervised clustering by FMM has been implemented. Clusters in FMM are defined by minimum and maximum points and a fuzzy membership function. Each cluster takes the form of an n-dimensional hyper-box.

The j^{th} hyper-box fuzzy set is defined [17] as:

ISBN: 978-81-923777-9-7

$$B_j = \{A_h, V_j, W_j, b_j(A_h, V_j, W_j)\} \quad (1)$$

for all h = 1, 2, …, m, where $A_h = (a_{h1}, a_{h2}, \ldots, a_{hn})$ is the h^{th} input data, the minimum point for the j^{th} hyper-box is $V_j = (v_{j1}, v_{j2}, \ldots, v_{jn})$, $W_j = (w_{j1}, w_{j2}, \ldots, w_{jn})$ is maximum point for the j^{th} hyper-box with membership value $0 \leq b_j(A_h,V_j,W_j) \leq 1$. The membership function is:

$$b_j(A_h, V_j, W_j) = \frac{1}{n}\sum_{i=1}^{n}[1 - f(a_{hi} - w_{ji}, \gamma) - f(v_{ji} - a_{hi}, \gamma)] \quad (2)$$

$$\text{where } f(x,\gamma) = \begin{cases} 1 & if\ x\gamma > 1 \\ x\gamma & if\ 0 \leq x\gamma \leq 1 \\ 0 & if\ x\gamma < 0 \end{cases}$$

The sensitivity parameter γ controls the rate of decrease of membership values.

FMM is the combination of learning and clustering phases. Learning comprises of hyper-box expansion, overlap test and hyper-box contraction. Whenever a new input data is obtained, efforts are made to accommodate it in one of the existing hyper-boxes. If none of the hyper-boxes can contain the data, i.e. hyper-boxes cannot be expanded beyond the maximum permissible size then a new hyper-box is created for it. Finally the hyper-boxes are checked for any overlap. In case of any overlap, the concerned hyper-boxes are contracted to remove the overlap.

In the clustering phase, the membership value of the input data is calculated for each hyper-box (cluster). And the highest membership value determines the hyper-box to be assigned to the input data.

The benefits of using unsupervised clustering using FMM are:

- FMM itself calculates the number of clusters in the input dataset during the training phase.
- Operations are very easy and only two parameters are adjusted to get best results; i.e., size of the hyper-box (θ) and the degree of fuzzyness of the membership function.

FMM is better than *k-means* in many ways, some of the advantages are [17]:

- Unlike *k-means*, FMM calculates the number of clusters itself.
- In FMM each cluster is represented by a minimum and maximum value, i.e., there is a boundary for each cluster. But, in *k-means* a cluster is represented only by a centroid and no such boundary exists.

Clearly FMM outshines k-means and therefore has been proposed [10] for the application on satellite infrared images.

FMM has the following shortcomings [12] too:

- In FMM, classification and clustering has been treated as separate problems. But, in real-life scenario pattern recognition is the combination of supervised and unsupervised techniques.
- Fuzzy membership functions of FMM sometimes assign high membership value to those input data that may be in reality are far from the cluster.
- FMM sometimes creates more number of hyper-boxes which tends to misclassification of data.

A modification on FMM has been proposed [12] by combining classification and clustering in the same algorithm and is known as General Fuzzy Min-Max Neural Network.

B. General Fuzzy Min-Max Neural Network

GFMM has been proposed [12] as the enhanced form of FMM, where classification and clustering has been combined in the same algorithm. The training phase in GFMM combines both labelled and unlabeled data. Partial supervision in learning phase is a significant issue in the real world pattern recognition [4]. Even a small amount of labelled data can significantly increase the chances of accuracy in clustering [12]. GFMM can be used solely for the purpose of classification or clustering or the combination of both.

In FMM, it has been found that sometimes the fuzzy membership of an input data is relatively high for a cluster from which it is actually significantly far. This has been solved in GFMM by monotonically decreasing the membership value as the distance from a cluster increases. GFMM also tries to form larger hyper-boxes to avoid misclassifications.

The input in GFMM [12] is in the pair$\{X_h, d_h\}$, where, $X_h = [X_h^l\ X_h^u]$ is the h^{th} input defined by lower and upper limits, and, $d_h = \{0,1,2,\ldots,p\}$, where p represents one of the class and $d_h = 0$ means unlabeled data.

In equation (1), it is expected that, b_j, degree of membership of X_h for the hyper-box B_j, will be 1 if this input is contained in the hyper-box. Else, b_j will decrease as X_h becomes far from B_j. But, this does not happen in the previous study [16] or [17]. Hence, an alternative membership function that gradually decreases has been proposed [12]:

$$b_j = \min_{i=1\ldots n}(\min\ ([1 - f(x_{hi}^u - w_{ji}, \gamma_i)], [1 - f(v_{ji} - x_{hi}^l, \gamma_i)])) \quad (3)$$

$$\text{where } f(r,\gamma) = \begin{cases} 1\ if\ r\gamma > 1 \\ r\gamma\ if\ 0 \leq r\gamma \leq 1 \\ 0\ if\ r\gamma < 0 \end{cases}$$

$\gamma = [\gamma_1, \gamma_2, \ldots, \gamma_n]$ is the sensitivity parameter.

Initially V_j and W_j are set to 0. When the first input is obtained then V_j becomes X_h^l and W_j is X_h^u. Whenever an input X_h is supplied the hyper-box B_j is selected which has the highest membership value and the expansion criteria is verified. The expansion criterion in GFMM contains two parts that have to be satisfied:

$$(\max\ (w_{ji}, x_{hi}^u) - \min\ (v_{ji}, x_{hi}^l)) \leq \theta \quad (4)$$

ISBN: 978-81-923777-9-7

where the maximum size of a hyper box is $0 \leq \theta \leq 1$, and the second part is:

$$\text{if } d_h = 0 \Rightarrow expand\ B_j \quad \text{else}$$

$$if\ class(B_j) = \begin{cases} 0 & ,expand\ B_j\ and\ class\ (B_j) = d_h \\ d_h & ,expand\ B_j \\ else & ,consider\ another\ B_j \end{cases} \tag{5}$$

where $expand\ B_j$ for i=1,2,…,n is:

$$new\ v_{ji} = min(old\ v_{ji}, x_{hi}^{l})$$
$$new\ w_{ji} = max(old\ w_{ji}, x_{hi}^{u})$$

Equation (4) decides whether the hyper-box expansion can be carried out or not, if yes then equation (5) determines how.

If $d_h = 0$ this means the input data is unlabeled. Then the hyper-box B_j will be expanded to accommodate the data. If $d_h \neq 0$ then the class of B_j is examined. If the class is 0 then d_h is assigned as the class and B_j is expanded. If the class of the hyper-box is d_h then only the expansion is carried out. If the class of B_j is neither 0 nor d_h then it belongs to any other class so this hyper-box will be discarded and another hyper-box will be searched.

Hyper-boxes are tested for probable overlaps. If B_j is unlabeled then it is checked with all the other hyper-boxes for overlapping. If the hyper-box is labelled then it is checked with the hyper-boxes of other classes only.

Overlaps are tested for each dimension. The smallest overlap in any dimension is tested, which is done by testing the following four cases where initially $\delta^{old} = 1$.

Case 1: $v_{ji} < v_{ki} < w_{ji} < w_{ki}$
$\delta^{new} = \min(w_{ji} - v_{ki}, \delta^{old})$
Case 2: $v_{ki} < v_{ji} < w_{ki} < w_{ji}$
$\delta^{new} = \min(w_{ki} - v_{ji}, \delta^{old})$
Case 3: $v_{ji} < v_{ki} \leq w_{ki} < w_{ji}$
$\delta^{new} = \min(\min(w_{ki} - v_{ji}, w_{ji} - v_{ki}), \delta^{old})$
Case 4: $v_{ki} < v_{ji} \leq w_{ji} < w_{ki}$
$\delta^{new} = \min(\min(w_{ki} - v_{ji}, w_{ji} - v_{ki}), \delta^{old})$

If overlap for the i^{th} dimension is found then hyper-box contraction will be done for that dimension only. The contraction will be carried out according to the following cases:

Case 1: $v_{ji} < v_{ki} < w_{ji} < w_{ki}$
$v_{ki}^{new} = w_{ji}^{new} = (v_{ki}^{old} + w_{ji}^{old}) / 2$ or
$(w_{ji}^{new} = v_{ki}^{old})$
Case 2: $v_{ki} < v_{ji} < w_{ki} < w_{ji}$
$v_{ji}^{new} = w_{ki}^{new} = (v_{ji}^{old} + w_{ki}^{old}) / 2$ or
$(v_{ji}^{new} = w_{ki}^{old})$
Case 3: $v_{ji} < v_{ki} \leq w_{ki} < w_{ji}$
if $(w_{ki} - v_{ji}) < (w_{ji} - v_{ki})$ then $v_{ji}^{new} = w_{ki}^{old}$
else $w_{ji}^{new} = v_{ki}^{old}$
Case 4: $v_{ki} < v_{ji} \leq w_{ji} < w_{ki}$
if $(w_{ki} - v_{ji}) < (w_{ji} - v_{ki})$ then $w_{ki}^{new} = v_{ji}^{old}$
else $v_{ki}^{new} = w_{ji}^{old}$

Another significant difference between FMM and GFMM is that in GFMM adaptive maximum size of the hyper-box is employed. A minimum value for the size is specified and the size is decreased in every pass according to:

$$\theta^{new} = \varphi\theta^{old} \tag{6}$$

φ decreases the size of hyper-box and $0 < \varphi < 1$.

The neural network in GFMM is a three layered system. The first layer is the input layer and has $2n$ elements representing the upper and lower values of n input data. The intermediate layer corresponds to hyper-box fuzzy sets where membership values are calculated for the input data. The third layer is the set of classes. The connection between second and third layer is binary, i.e., a hyper-box belongs to one of the classes.

GFMM excels FMM in many ways [12]:

- GFMM can process both fuzzy and crisp input patterns.
- GFMM algorithm is the fusion of classification and clustering.
- It produces fewer hyper-boxes and less misclassifications.

III. EXPERIMENTS AND RESULTS

For the experimental purpose, TIR images of 13 June 2011 (0700-1130 hours) taken by Kalpana-1, were used (Figure 1). Both FMM and GFMM were applied on the images.

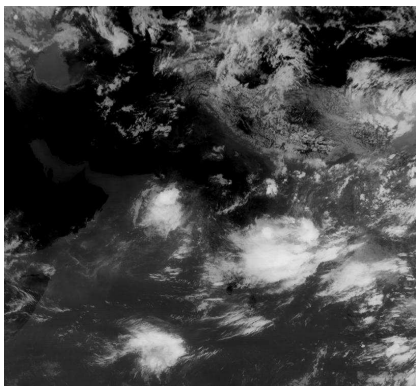

Figure 1: TIR Image taken on 13 June, 2011 at 0700 hrs

While applying FMM, the size of the hyper-box (θ) is specified at the beginning of the training and it remains

ISBN: 978-81-923777-9-7

constant throughout the process. Three values of θ (0.3, 0.1 and 0.03) were separately used in FMM. The fuzzy cluster validity index proposed by [14], *Partition Entropy* (PE), was used for the validation of the clustering. It is defined as

$$PE(c) = -\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}\, log_2\mu_{ij} \quad (7)$$

where, n is the size of the data set, c is the number of clusters and μ_{ij} is the membership value of j^{th} data in i^{th} cluster. The value of *Partition Entr*opy should be such that

$$0 < PE(c) < log_2 c$$

The PE obtained by using FMM with three different θ are shown in Table I.

GFMM was then used on the same image (13 June, 2011, 0700 hrs). Figure 2 and 3 show the extracted coldest cluster and the largest region, respectively.

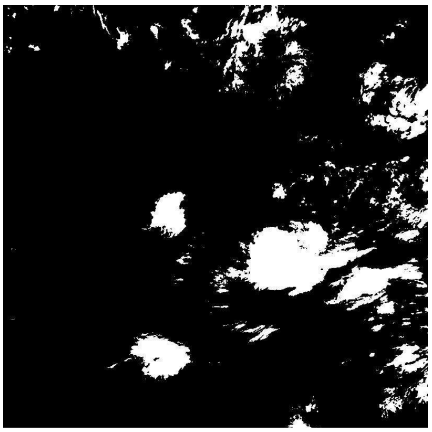

Figure 2: Extracted coldest cluster using GFMM

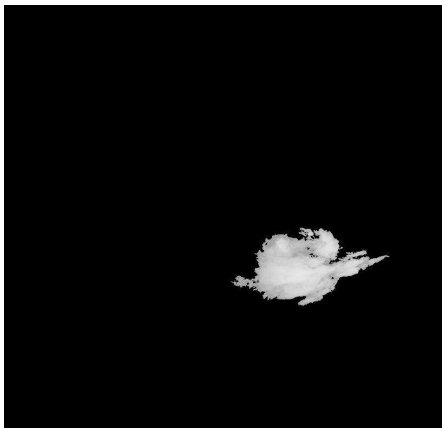

Figure 3: Largest region selected using GFMM

GFMM can start with a large value of θ and can gradually decrease it during the training using equation (6). Here θ was started with 0.3 and it was decreased by using φ as 0.9 and ended at 0.03. The PE for GFMM is also shown in Table I.

TABLE I: COMPARISON OF PARTITION ENTROPY FOR FMM AND GFMM

	FMM			**GFMM**
	θ=0.3	***θ=0.1***	***θ=0.03***	***θ=0.3 - 0.03***
No. of clusters (c)	133	336	365	7
$\log_2$ c	7.0553	8.3923	8.5118	2.8074
PE obtained	350.9567	1130.0000	1043.7000	2.7288

a. Sample image 13 June, 2011, 0700 hrs

It is clearly evident from Table I that the PEs when FMM was used is far away from the range (0 to $\log_2$c). This is due to the reason that FMM creates more clusters. This may lead to misclassification due to noise. On the other hand, GFMM created just 7 clusters and the value of PE is within the limit. The variable value of θ during the training actually created fewer hyper-boxes and gave quite appropriate result.

GFMM was applied on 10 images (13 June, 2011,0700-1130 hrs) and the obtained PE has been enumerated in Table II.

TABLE II: PARTITION ENTROPY OBTAINED FROM IMAGES USING GFMM

GFMM (13 June, 2011)			
Images (hours)	***No. of clusters (c)***	***$\log_2$c***	***PE***
0700	7	2.8074	2.7288
0730	8	3.0000	2.9527
0800	6	2.5850	2.8134
0830	7	2.8074	2.5325
0900	6	2.5850	2.4519
0930	8	3.0000	2.3565
1000	7	2.8074	3.4405
1030	7	2.8074	2.1696
1100	7	2.8074	2.1617
1130	7	2.8074	2.5891

From Table II it can be observed that GFMM's performance is quite consistent, except at 0800 and 1000 hrs where it just crossed the limit. In all the images the clustering validity criteria more or less is being satisfied. Since GFMM creates much less clusters, in comparison to FMM, there is a lesser possibility of misclassification and it's faster as well.

IV. CONCLUSION

In this study, a novel application area has been suggested for GFMM. Satellite TIR images are often used for the detection of convective clouds. Many of the existing methods use different techniques for the clustering of such images. Some of them used *k-means* as the clustering technique. Previously, it was proposed that FMM can be an alternative

ISBN: 978-81-923777-9-7

clustering tool for such application area. In this study, it has been proposed that GFMM can be even a smarter choice for the clustering of satellite TIR images in order to detect convective cloud for further estimation of precipitation. In the present study, both FMM and GFMM were applied on the same images. GFMM created less number of clusters and was found to be faster in comparison to FMM. The results were verified using a fuzzy cluster validity index, *Partition Entropy*, where GFMM gave closer values than FMM. Clearly, GFMM is a better choice than FMM for this particular application area.

ACKNOWLEDGMENT

The authors are especially indebted to *Meteorological and Oceanographic Satellite Data Archival Centre* (MOSDAC), *Indian Space Research Organization*, *Government of India*, for availing the images of Indian subcontinent from the *Indian Meteorological Satellite, Kalpana-1*, for performing the study. The authors are also indebted to both the institutions, *Jadavpur University* and *Narula Institute of Technology* for providing the opportunity to carry out the study.

REFERENCES

[1] Achintya K. Mandal, Srimanta Pal, Arun K. De and Sibhasis Mitra, "Novel approach to identify good tracer clouds from a sequence of satellite images", IEEE Transactions on Geoscience and Remote Sensing, vol. 43, issue 4, pp. 813-818, April 2005.

[2] Adrian N. Evans, "Cloud tracking using ordinal measures and relaxation labelling", IEEE Proceedings of International Geoscience and Remote Sensing Symposium, vol. 2, pp. 1259-1261, 1999.

[3] Antonio Turiel, Jacopo Grazzini and Hussein Yahia, "Multiscale techniques for the detection of precipitation using thermal IR satellite images", IEEE Geoscience and Remote Sensing Letters, vol. 2, no. 4, pp. 447-450, October 2005

[4] A. V. Nandedkar and P. K. Biswas, "A General Reflex Fuzzy Min-Max Neural Network", Engineering Letters, vol. 14, issue. 1, pp. 1-11, 2007.

[5] Barnali Goswami and Gupinath Bhandari, "Near real-time detection of heavy rain clouds from IR image for estimation of precipitation", Proceedings of the 3rd International Conference on Water and Flood Management (ICWFM-2011), vol. I, pp. 277-281, 2011.

[6] Barnali Goswami and Gupinath Bhandari, "Cloud motion prediction using mean path adjustment method from satellite infrared images", Proceedings of CALCON11, IEEE Calcutta Section and CAS Chapter. pp. 313-316, 2011.

[7] Barnali Goswami and Gupinath Bhandari, "Automatically adjusting cloud movement prediction model from satellite infrared images", Proceedings of India Conference (INDICON), 2011 Annual IEEE, pp. 1-4, 2011.

[8] Barnali Goswami and Gupinath Bhandari, "Convective Cloud Detection and Tracking from Series of Infrared Images", Journal of Indian Society of Remote Sensing, Springer [In press], 2012.

[9] Barnali Goswami and Gupinath Bhandari, "Temperature Induced Mean Based Cloud Motion Prediction Model from Satellite Infrared Images", Proceedings of India Conference (INDICON), 2012 Annual IEEE, pp. 719-723, 2012.

[10] Barnali Goswami, Gupinath Bhandari and Sanjay Goswami, "Fuzzy Min-Max Neural Network for Satellite Infrared Image Clustering", Proceedings of 3rd International Conference on Emerging Applications of Information Technology (EAIT-2012), vol. 1, pp. 239-242, 2012.

[11] Barnali Goswami and Gupinath Bhandari, "Development of Irregular Cloud Cluster Encapsulating Structure from Satellite Infrared Images", Proceedings of 33rd Asian Conference on Remote Sensing (ACRS-2012), 2012.

[12] Bogdan Gabrys and Andrzej Bargiela, "General fuzzy min-max neural network for clustering and classification", IEEE Transactions on Neural Networks, vol. 11, no. 3, pp. 769-783, May 2000.

[13] G. Carpenter, S. Grossberg and D.B. Rosen, "Fuzzy ART: An adaptive resonance algorithm for rapid, stable classification of analog patterns", Proceedings of International Joint Conference of Neural Networks (IJCNN-91), vol. 2, pp. 411-416, 1991.

[14] James C. Bezdek, "Cluster Validity with Fuzzy Sets", Journal of Cybernetics, vol. 3, issue. 3, pp. 58-73, 1974.

[15] Kanchan Deshmukh and Ganesh Shinde, "Adaptive color image segmentation using fuzzy min-max clustering", Proceedings of Engineering Letters, pp. 57–64, 2006.

[16] Patrick K. Simpson, "Fuzzy min-max neural networks-Part1: Classification", IEEE Transactions on Neural Networks, vol. 3, no. 5, pp. 776-786, September 1992.

[17] Patrick K. Simpson, "Fuzzy min-max neural networks-Part2: Clustering", IEEE Transactions on Fuzzy Systems, vol. 1, no. 1, pp. 32-45, February 1993.

 ISBN: 978-81-923777-9-7