# Improving the Quality of Remote Sensing Images using Semi-Supervised Fuzzy Clustering

Trung Nguyen Tu[#1], TuanTran Manh[*2]

*Department of Information Technology, ThuyLoi University, Hanoi, Vietnam*

***Abstract:*** *Improving the quality of images is one of necessary stages in image processing. Multispectral images are the images with a large size and high spatial resolution. Therefore, the analysis of multispectral images is complex. The results of pre-processing progress are very important. These results affect directly to the results of image analysis progress. There were some researches related to the improvement of image quality such as noise reduction, contrast increment, histogram techniques, etc. In this paper, we propose an novel method in order to improve the quality of multispectral images by contrast increasing approach using semi-supervised clustering..*

Keywords - *Contrast enhancement, remote sensing images, Fuzzy logic, Semi-FCM.*

## I. INTRODUCTION

Remote sensing images store information about an object, an area. However, remote sensing images often have large sizes and diverse spatial resolutions. The higher spatial resolution, the more detailed images. It may also contain noise. To eliminate noise and enhance image quality, we need to use methods to improve image quality. Improving quality is a necessary step in image processing to improve some of the image characteristics such as noise filtering, smoothing, enhancing contrast, adjusting gray levels, and floating edges. Enhancing image quality includes two different stages: image enhancement and image restoration. Noise in remote sensing images includes common noises like color images and characteristic noises such as fog, cloud, etc. In image processing problems, pre-processing is important to improve input image quality. This process includes restoration and image enhancement steps. Image restoration works to eliminate or minimize the effects of the environment on the images to reduce image distortion and bring the image to its original state. Image enhancing highlights its characteristics to support the next steps more efficient. According to [7], image enhancement techniques are divided into three main groups. Firstly, the group of image processing techniques based on the histogram, which is typically characterized by the histogram balance and the histogram specification. Second, the group of image processing techniques based on fuzzy logic approach with some of the techniques presented in Part III of this paper. Third, advanced image processing techniques based on optimization. In this, Aman Tusia et al. [10] performed a performance analysis of type 2 fuzzy systems for image enhancement using cuckoo optimization algorithms.

Many traditional contrast enhancement methods use a global approach to enhance all photo brightness levels. However, it is often difficult to enhance all the objects that appear in the satellite images, because local and detailed contrast information can be lost in bright and dark areas. In [5, 6], the authors combined fuzzy logic [12] and gray level correction formulas to enhance the contrast of medical images. The method of fuzzy image enhancing [2] considers membership matrices and gray level adjustment expressions to enhance contrast. However, this approach still uses a global approach, so it has not solved the problem of traditional methods. In [6], the authors proposed a method to enhance images based on local approaches. In [9], the authors propose a number of preprocessing techniques for using cluster data. Ngo Thanh Long et al. [4] used fuzzy type 2 in combination with fuzzy clustering in the remote sensing image segmentation. In addition, Le Hoang Son et al. [8] used semi-supervised fuzzy clustering in dental X-ray image segmentation.

This paper presents a model for improving remote sensing image quality using semi-supervised fuzzy clustering method. This study has implications in finding an effective algorithm to improve the quality of remote sensing images, thereby supporting the process of remote sensing image processing.

## II. THE RELATED KNOWLEDGES

### *A. A Short Overview Of Remote Sensing*

According to [14], remote sensing is science branch remote gathering the information on the Earth surface, including sensing and taking energy released, processing, analyzing data and applying the information after analysis. Besides, most of receiving systems and remote sensing images processing include seven-step process like figure 1.

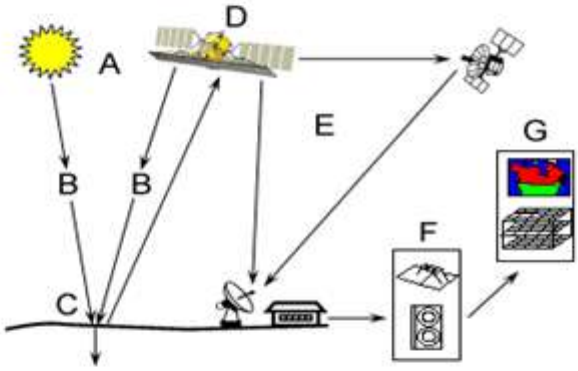


Fig 1. Process of gathering and processing remote sensing images [14].

In figure 1, A is energy source or bright source, B is radiance and atmosphere, C is interactive with destination object, D is energy gathered by sensor, E is energy transmission, reception and processing, F is interpretation and analysis, G is application.

Remote sensing images have features: image channel, space resolution, spectrum resolution, radiant resolution and time resolution. There are many different types of remote sensing images/satellites such as weather sensing satellites (GOES, NOAA, AVHRR…), ground observation satellites (Landsat, SPOT…)… Remote sensing images are applied in agriculture, forestry, geology, hydrography… One of concrete applications is classifying land cover, building maps on the special subject in each field.

### *B. Semi-Supervised standard fuzzy clustering*

Data clustering [1] is the process of grouping a set of similar objects in a data set into clusters so that objects belonging to the same cluster are similar, while objects belonging to different clusters are less similar.

Clustering techniques: hard clustering, fuzzy clustering, semi-supervised clustering, semi-supervised clustering. In the paper, the research team focused on fuzzy clustering algorithms and semi-supervised fuzzy clustering algorithms. The fuzzy clustering algorithm (FCM) proposed by Bezdek [1] is based on optimizing the distance of data points to the center. The semi-supervised fuzzy clustering algorithm is based on fuzzy clustering algorithms combined with additional information provided by the user. Additional information [13] for semi-supervised fuzzy clustering has three basic forms including Must-link and Cannot-link constraints; a part of data is labeled and predetermined dependency.

Yasunori et al. [11] proposed a semi-supervised fuzzy clustering algorithm with additional information that be the membership function was added to the FCM's objective function to improve efficiency in clustering. The target function [11] is then defined as follows:

$$J(U,V) = \sum_{k=1}^{N} \sum_{j=1}^{C} |u_{kj} - \bar{u}_{kj}|^m \,||X_k - V_j||^2 \to min \quad (2)$$

Where:

- m is the fuzzy number
- C is the number of clusters
- N is the number of data elements
- r is the number of dimensions of the data
- $\mu_{kj}$: membership value of $j$th cluster of $k$th pixel
- $X_k$: is the kth element
- $V_j$ : is the center of cluster j

The binding conditions are as follows:

$$\sum_{j=1}^{C} u_{kj} = 1; \qquad u_{kj} \in [0,1]; \qquad \forall k = \overline{1,N}$$
$$\overline{U} = \{\overline{u_{kj}} | \overline{u_{kj}} \in [0,1], k = \overline{1,N}, j = \overline{1,C}\}, \quad \sum_{j=1}^{C} \overline{u_{kj}} \leq 1, (\forall k = \overline{1,N}) \tag{3}$$

From condition (3) and objective function (2) we have:

$$V_j = \frac{\sum_{k=1}^{N} |u_{kj} - \bar{u}_{kj}|^m X_k}{\sum_{k=1}^{N} |u_{kj} - \bar{u}_{kj}|^m} \quad ,j = \overline{1,C} \tag{4}$$

The values of $u_{kj}$ are determined in the following two cases:

- $m > 1$ :

$$u_{kj} = \overline{u_{kj}} + \left(1 - \sum_{i=1}^{C} \overline{u_{kj}}\right) \frac{\left(\frac{1}{\|X_k - V_j\|}\right)^{\frac{2}{m-1}}}{\sum_{i=1}^{C} \left(\frac{1}{\|X_k - V_i\|}\right)^{\frac{2}{m-1}}}, k = \overline{1,N}, j = \overline{1,C}. \tag{5}$$

- $m = 1$:

$$u_{kj} = \begin{cases} \overline{u_{kj}} + 1 - \sum_{j=1}^{C} \overline{u_{kj}}, k = arg\min_i \|X_k - V_i\|^2 \\ \overline{u_{kj}}, otherwise. \end{cases}, k = \overline{1,N}, j = \overline{1,C}. \tag{6}$$

The implementation steps of the Semi-Supervised Standard Fuzzy Clustering algorithm (SSSFC) are shown in Table 2 below.

Table 3. The Semi-Supervised Standard Fuzzy Clustering algorithm

| | |
|---|---|
| Input | Data set X includes N elements, cluster number C, matrix of complementary membership $\overline{U}$, threshold$\varepsilon$, maximum number of iterations *maxStep > 0.* |
| Output | Matrix U and cluster centers V. |
| *SSSFC* | |
| 1: | t = 0 |
| 2: | Randomized initialization $V_j^{(t)}$; $(j = \overline{1,C})$ |
| 3: | Repeat |
| 4: | t = t + 1 |
| 5: | Calculate $u_{kj}(k = \overline{1,N}; j = \overline{1,C})$ by formula (5) with $m > 1$ or formula (6) with $m = 1$. |
| 6: | Calculate $V_j^{(t+1)}(j = \overline{1,C})$ by formula (4) |
| 7: | Until $\|V^{(t)} - V^{(t-1)}\| \leq \varepsilon$ or t > maxStep |

## III. IMPROVING THE QUALITY OF REMOTE SENSINGIMAGES

Improving the quality of remote sensing image is an important stage to support remote sensing image processing. Enhance image quality to support image interpretation and analysis. There are many functions in image quality enhancement such as stretching, filtering space to make patterns on the image, etc. Noise in remote sensing images includes common noise like color images and characteristic noise (fog, cloud, etc.). For common noise, we can use conventional noise reduction methods such as filtering methods. To enhance the quality of remote sensing images, contrast methods can also be used (Figure 2).

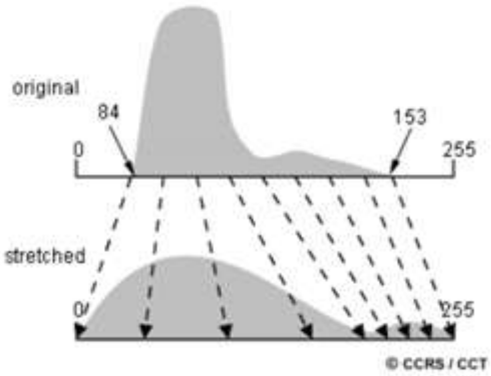


Fig. 2: Constrast Stretch.

In [6], the authors propose the algorithm of Local based Remote Sensing Image Enhancement LoRSIE using clustering FCM. Steps of this algorithm as in figure 3.

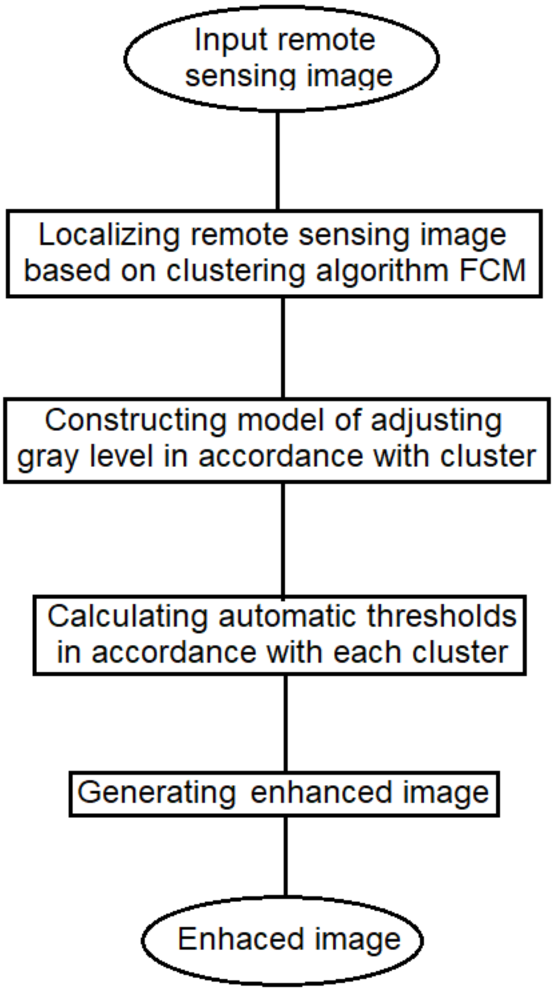


Fig. 3: The flowchart of the algorithm LoRSIE_FCM.

Following the steps in figure3, the input remote sensing image is localized based on the FCM clustering algorithm. Next, a gray level correction model was built. The gray level correction functions are listed in Table 2.

Table 2. The functions of gray level adjusting in accordance with each cluster.

| Name | Adjusting fomulaT(g) |
|---|---|
| Gray level Stretch | $255 * \frac{g-lower_{V_i}}{upper_{V_i}-lower_{V_i}}$ (1) |
| Hyperbolization | $\left(\frac{255}{e^{-1}-1}\right)\left[e^{-\mu^2(g)}-1\right]$ (2) |

Where:

$$\mu = \begin{cases} 0, & g < lower_{V_i} \\ \frac{g-lower_{V_i}}{upper_{V_i}-lower_{V_i}}, & lower_{V_i} \le g \le upper_{V_i} \\ 1, & upper_{V_i} \le g \le 255 \end{cases} \quad (3)$$

g : original gray value.

$lower_{V_i}$: upper bounds of cluster *i*th.

$upper_{V_i}$ : lower bounds of cluster *i*th.

$V_i$ : center of cluster.

The thresholds $upper_V$ and $lower_V$ is determined by selecting so that area of subregion between the graphs of $upper_V$,$lower_V$, the distributed function *d(g)* and horizontal axis equalize 95% area of region built by graph of the distributed function *d(g)* and horizontal axis (crossover bar region).

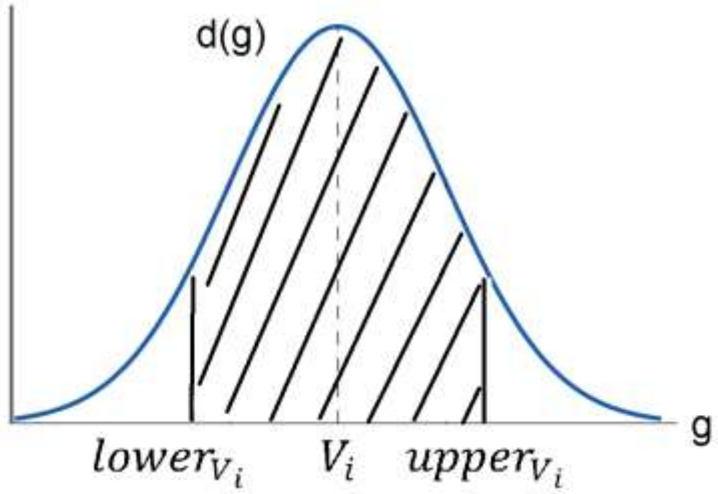


Fig 3: Distributed function and thresholds following each cluster.

Finally, the enhanced image is aggregated from the output values of the gray-scale correction function by cluster. This function has general format as follows: g |→ g' = T(g) as following:

$$(g) = \sum_{i=1}^{c} \mu_i T_i(g, lower_{V_i}, lower_{V_i}) \quad (4)$$
$$0 \le T(g) \le 255$$

## IV. THE MODEL OF QUALITY ENHANCEMENTOF REMOTE SENSING IMAGES BY SEMI-SUPERVISED FYZZY CLUSTERING

Based on the model of quality enhancement of remote sensing images LoRSIE_FCM [6] what presented above, we propose an improvement by using semi-supervised fuzzy clustering (SemiFCM) instead of FCM clustering. Figure 4 shows a flowchart of an advanced remote sensing image based semi-supervised fuzzy clustering. With this new algorithm, the input image, instead of being localized by FCM, will be localized by a semi-supervised fuzzy clustering algorithm. Then, we obtain the cluster centers and the new membership matrix from the application of the SemiFCM algorithm. Thus, instead of obtaining the function $d(g)_{FCM}$, we obtain the function $d(g)_{SemiFCM}$ the parameters $upper_{V_i}$ and $lower_{V_i}$ related to each cluster. Therefore, the output gray level values change too, leading to the change of the enhanced imge that compared to compared to enhanced image with FCM.

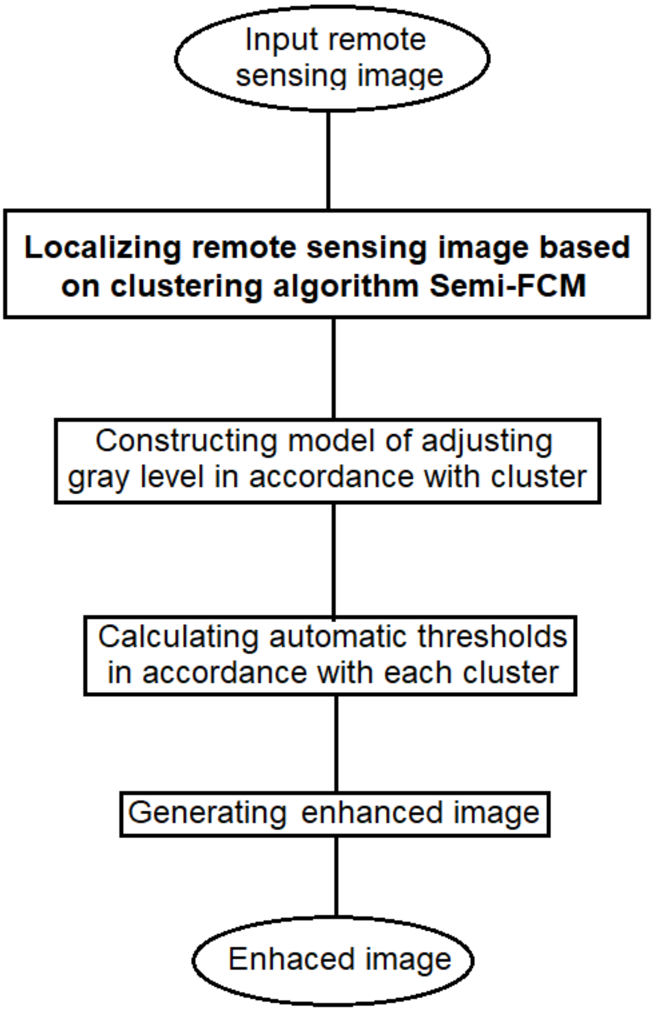


Fig. 4:The flowchart of the algorithm LoRSIE_SemiFCM.

## V. EXPERIMENTS

### A. *Experimental description*

We tested the LoRSIE_SemiFCM algorithm on the specific data set LANDSAT ETM + and compared it with the results of the LoRSIE_FCM algorithm together on this data set. In the tests, $\overline{U}$ was calculated by performing an FCM cluster. From the obtained matrix $U$, determine $\overline{U}$ as follows for each pixel, keeping the maximum value of membership values with clusters, the remaining values reassigned to zero.

Data description: Data set used for experiments are Landsat ETM+ images that are taken in Hoa Binh area in 2001, including 11 pictures about districts and 1 picture about Hoa Binh province. Landsat ETM+ image includes 7 channels [12]: Indigo, Green, Red, Near infrared, Medium infrared, Heat infrared, Medium infrared.

Measurements for evaluation: To measure the quality of the original image and the enhanced image we use the Entropy index. Shannon Entropy (or information entropy) is a method of measuring the uncertainty of information. Assuming there are n events in the sample space, the probability of each event is pi (i = 1, 2, ..., n), with pi ≥ 0, and the sum of pi equals 1. A function H is defined to measure the uncertainty of the sample space [3]. The value of H is calculated by formula (7). For image processing, n is given by the number of gray levels.

$$H = -\sum_{i=0}^{n-1} p_i \ln(p_i) \quad (7)$$

Where:
- $n$is the number of gray levels
- $p_i$ is the probability of level i in the histogram.

### *B. Experimental results*

Due to the limited scope of the paper, the authors presented the test results on four different input image samples. These testls included satellite imagery in Cao Phong, Da Bac, Kim Boi and Lac Thuy districts.

Table 3 shows the input image and the result image with image quality enhancement when applying the LoRSIE_FCM algorithm and the SemiFCM algorithm with a combination of 3 channels of Indigo, Green, Red. We see that the following images are enhanced, the brightness and contrast are better than the original image. However, if we look more carefully, we will see that the contrast of the resulting image of the algorithm using SemiFCM is a little better than the result image of the algorithm using FCM.

**Table 3.**Constrast Enhancement of Remote sensing images with FCM and SemiFCM.

| Sample | Input | Result of LoRSIE_FCM | Result of LoRSIE_SemiFCM |
|---|---|---|---|
| Cao Phong |  |  |  |
| Da Bac |  |  |  |
| Kim Boi |  |  |  |
| Lac Thuy |  |  |  |

Tables 4, 5, 6 and 7 show the Entropy index of the experimental results when enhancing the image quality of Cao Phong, Da Bac, Kim Boi and Lac Thuy. Based on this result we can clearly see that the entropy values of the result image channels using the algorithm using LoRSIE_SemiFCM are higher than the entropy values of the result image channels using the algorithm LoRSIE_FCM. This proves that quality of the image enhancement algorithm using LoRSIE_SemiFCM is better than that quality of the image enhancement algorithm LoRSIE_FCM.

**Table 4.**The Entropy values of enhanced image with Cao Phong.

| Channel | Enhancement with LoRSIE_FCM | Enhancement with SemiFCM |
|---|---|---|
| 1 | 3.44578672451698 | 3.75960630294584 |
| 2 | 3.53055329114634 | 3.87734956539595 |
| 3 | 3.57985248661498 | 3.92747171506403 |

**Table 5.** The Entropy values of enhanced image with Da Bac.

| Channel | Enhancement with LoRSIE_FCM | Enhancement with SemiFCM |
|---|---|---|
| 1 | 2.45684844447667 | 3.50167678500134 |
| 2 | 2.66391133356626 | 3.49133354235306 |
| 3 | 2.65419094174343 | 3.44077893674379 |

**Table 6.** The Entropy values of enhanced image with Kim Boi.

| Channel | Enhancement with LoRSIE_FCM | Enhancement with SemiFCM |
|---|---|---|
| 1 | 4.1858282177615 | 4.49061166054001 |
| 2 | 4.19923187498745 | 4.53125963217986 |
| 3 | 4.27514951159837 | 4.54546630160856 |

**Table 7.**The Entropy values of enhanced image with Lac Thuy.

| Channel | Enhancement with LoRSIE_FCM | Enhancement with SemiFCM |
|---|---|---|
| 1 | 2.40240944416823 | 3.11144330891485 |
| 2 | 2.61518856088777 | 3.22093197523361 |
| 3 | 2.77044329227435 | 3.24632169784426 |

## VI. CONCLUSIONS

In this paper, we focus on improving the quality of remote sensing image based on semi-fuzzy supervised cluster. The main contribution of the paper is to use semi-supervised fuzzy clustering to improve the quality of remote sensing images. In addition, the authors have also experimentally installed the proposed algorithm with remote sensing images collected in Hoa Binh area. The experimental results also show that the Entropy values of the LoRSIE_SemiFCM method have a better value than the results obtained when applying the LoRSIE_FCM algorithm. From this study, we will continue to study other parameters to improve image quality. Moreover, this research also opens the direction of development for algorithms using clustering based on advanced fuzzy sets.

## REFERENCES

[1] Bezdek, J. C. (1981). Pattern recognition with fuzzy objective function algorithms. *Kluwer Academic Publishers*.

[2] Hassanien, A. E., & Badr, A. (2003). A comparative study on digital mamography enhancement algorithms based on fuzzy theory. *Studies in informatics and control*, *12*(1), 21-32.

[3] Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical review*, *106*(4), 620-630.

[4] Ngo, L. T., Mai, D. S., & Pedrycz, W. (2015). Semi-supervising Interval Type-2 Fuzzy C-Means clustering with spatial information for multi-spectral satellite image classification and change detection. *Computers & geosciences*, *83*, 1-16.

[5] Nguyễn Tu Trung, Đặng Văn Đức, Vũ Văn Thỏa (2016), Một số cải tiến kĩ thuật phân cụm cho ảnh viễn thám, *Tạp chí Công nghệ thông tin và Truyền thông, Bộ Thông tin và Truyền thông*, số 16 (36).

[6] Nguyễn Tu Trung, Vũ Văn Thoả, Đặng Văn Đức (2015), Một phương pháp tăng cường độ tương phản ảnh viễn thám dựa trên tiếp cận cục bộ, *Chuyên san các công trình nghiên cứu, Tạp chí Công nghệ thông tin và Truyền thông, Bộ Thông tin và Truyền thông*, Kỳ 3, tập V-2, số 14 (34).

[7] Sharo, T. A., & Raimond, K. (2013). A Survey on Color Image Enhancement Techniques. *IOSR J. Eng*, *3*, 20-24.

[8] Son, L.H., Tuan, T. M. (2017). Dental segmentation from X-ray images using semi-supervised fuzzy clustering with spatial constraints. *Engineering Applications of Artificial Intelligence*, *59*, 186-195.

[9] Sudhavani, G., Srilakshmi, M., Rao, P. V., & Prasad, K. S. (2014). Comparison of fuzzy contrast enhancement techniques. *International Journal of Computer Applications*, *95*(22). pp. 0975 – 8887.

[10] Tusia, A., & KUMAR, D. N. (2014). Performance analysis of type-2 fuzzy system for image enhancement using optimization. *International Journal of Enhanced Research in Science Technology & Engineering*, *3*(7), 108-116.

[11] Yasunori, E., Yukihiro, H., Makito, Y., & Sadaaki, M. (2009, August). On semi-supervised fuzzy c-means clustering. In *Fuzzy Systems, 2009. FUZZ-IEEE 2009. IEEE International Conference on* (pp. 1119-1124). IEEE

[12] Zadeh, L.A. (1965). Fuzzy sets. *Information and Control*, *8*, 338-353.

[13] Zhang, H., & Lu, J. (2009). Semi-supervised fuzzy clustering: A kernel-based approach. *Knowledge-Based Systems*, *22*(6), 477-481.

[14] Canada Center for Remote Sensing, *Fundamentals of Remote Sensing*, http://www.ccrs.nrcan.gc.ca, 2008.