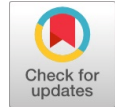# A Novel Algorithm for Soil Image Segmentation using Color and Region Based System

Prathik A, Anuradha J, Uma K

Abstract: *In the recent research era data mining is a very essential research domain. The data mining techniques are used to extract significant knowledge in agriculture management. These techniques are time consuming and less expensive than the statistical techniques. Many Researchers develop efficient techniques to improve the productivity of agriculture. This paper developed a new segment method to segment the soil region from other information. This research introduced Color and Region Based segment method to separate the soil region from its background. To evaluate the proposed segmentation the five metrics are used dice coefficient, jaccard index, Sensitivity, Specificity and Precision. The new approach produced 98% accuracy, 98% Sensitivity and 98% Specificity.*

Keywords : Data Mining, Segmentation, Classification, Jaccard Index, Sensitivity, Specificity.

I. INTRODUCTION

It is known for years that agriculture is the backbone of Indian economy. Fruit cropping gives the primary yield in the agriculture sector. Most the Indian population are employed in agriculture. Agriculture plays a vital role in contributing 18 per cent of total Gross GDP to the economy. Also, it offers about 60 per cent of employment opportunities for the population. However, this sector has to be strengthened by predicting disease in advance and ensuring good yield to the market. Data mining and image processing play a vital role in agriculture. Data mining is used to analyze large data set to produce useful information and patterns in the large data set. Various data mining techniques are used in agriculture and biological research. It can be used to predict the pattern and texture of soil in large data set and also analysis the large soil profile. So mining techniques is very useful to classifying soil texture. Soil texture is important to classification. Classification of soil is very useful to predict nature of soil, properties of soil and yield percentage. It also controls the drainage, water holding ability,strength of soil and tillage. But before classification image segmentation is an important task. It is the processing of partitioning the image into different region. This research develops new segmentation technique to improve the accuracy than the existing segmentation methods. The proposed segmentation technique in this research serves as a preprocessor to

classification. Each previous image processing step influences the execution of the next. Therefore, a good segmentation is required to obtain a good classification result. The new segmentation method is used to eliminate the background information.

II. IMAGE SEGMENTATION

In geosciences soil image processing is very important to micro scale image to large scale image. The soil components are minerals. Soil is composition with the contrast, size, shape, sharpness, frequency, and spatial arrangement of primary particles and voids. Extraction of these properties is used to identify the soil texture and differentiate. Segmentation is used to separate the each property. It is play a vital role in geosciences for texture analysis. It divide the image in to different regions, each region is homogeneous and associated with respect to some property, such as motion ,color, gray value, texture. In computer era more variety of algorithms are available to segmentation. Among these algorithm the proposed system provide a good result.

This section also discuss the segmentation and its types that used by various researcher to segment the image. Some of them are region based, thresholding based, clustering based, and classifier. This section also discuss the segmentation and its types that used by various researcher to segment the image There five types of segmentations they are: Threshold based, Edge based method, Region based, clustering based, partial equation based.

A.Thresholding Method:

It is a direct and simplest method. This method divided the pixels of the image based on their intensity value.It works based on rule of converting gray-scale image into a binary image by using a threshold value of image. It is suitable method when the images having lighter substance than background. The choice of this method is based manual or automatic. It needs prior understanding or information of features of images.[7]

B. Edge Based Approach:

It observes the rapid changes in the intensity values in the edges. Because single intensity value does not gives the edge details of image. It locate the edges wherever the first derived of intensity value is bigger than threshold or the zero crossing second derivative. In this segmentation first of all predict the all the edge information for whole image then all are connected together to form a segmented region. There are two categories: histogram based and gradient based methods. [8, 9]

Manuscript published on 30 August 2019.
*Correspondence Author(s)
Prathik.A, Research Associate, School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India
Anuradha.J, Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India,
Uma.K, Associate Professor, School of Advanced Science and Engineering, Vellore Institute of Technology, Vellore, India

C. Region Approach:

This method segment the image based on their properties. This method also called as homogeneous segment. There are two techniques: in the first technique divide the image into different regions and it growing based on initial pixels. Based on initial pixel region grows automatically. But in the second method pixel selection is by manual and it needs prior knowledge about the image.[10,11]

D. Clustering Based Approach:

This technique segment the image based on their pixel with related characteristics. There are two methods: Hierarchical method and Partition based method. First method is based on tree concept. In the tree root denotes database and clusters represented by internal nodes. Partition method based on optimization. [12, 13]

F.Partial Differential Equation Based Approach:

This is a fast method than the other method. There are two types: isotropic diffusion filter and convex. First method is nonlinear used to enhance the edge details. Convex method is used to remove noise in image. The result of first method is utilized to reduce the noise in image. Second approach is used to detect the edges and boundaries better than other method. [14,15]

III. EXISTING METHODS

Borselli, L et al [1] applied a new method to extract the roughness profile of soil. It consider the non-stationary profile and identify statistical profiles and index the roughness profile. The method obtain the unpredictability of numerical and roughness properties along the profile and extracts the probable mechanism random and oriented noticeable inside the sample.Soh, L. K., & Tsatsoulis, C [4] describe a new segmentation method that integrates the existing image processing technique with data mining techniques. This proposed method is to analyze and segment the natural scenes. It has three main steps: local thresholding of region, textural features for each region. Sofou, A., Evangelopoulos, G.,& Maragos, P [5] proposed equation based segmentation based on image contrast. This research is analyzing texture information of local components. It separates contrast and texture information. We propose a joint image segmentation method forfurther interpretation of soil images and feature measurements.Bushra Nazir, Md. Iqbal Qureshi [2] introduced a new segmentation technique to extract the features of soil and estimate the type of soil. Using the different segmentation techniques compare the features of soil and classify it. It is analyzing color and texture properties of soil. Charytanowicz, M., & Kulczycki, P [3] presenteda segmentation approach to predict the structure of soil pore. It contains the study of density based method. The arbitrary assumptions based on number or shape of clusters and it help to detect inherent structure. The color information removed from the grayscale images and it is used as input to gradient clustering for next segmentation processing. Pradeep Kumar Reddy, Bhavana [6] developed a local binary pattern to extract the local features and texture of soil. It is very sensitive to noise and illumination. Testing images are compared with all images with Euclidean distance. If it gives low value the test image is true image. If the space between test image and training images is more than test image is contrast as unrecognized image or match not found.

IV. PROPOSED SEGMENTATION METHOD

The proposed segmentation method based Color and Region Based Segmentation (CRBS) work, selected soil region and segmented soil alone regions for segmentation we get the accurate region based HSV region and segmented color based.

Step 1: Input image: Take soil image of T1 or T2.

Step 2:Pre-processing: Initially the image is pre-processed for noise removal where unwanted noise are removed, quality of image is enhanced here by edge sharpening. Grey conversion from RGB and reshaping of image also takes place here. Noise removal is performed using median filter.

Step 3: Multi resolution process: Aim of multi resolution image exploration is to decay image into several frequency representation in variable resolution to evaluate the area of interest. Essentially, here multi resolution steps are performed using different scale value with methods such as region growing process and canny detector. X1, x2, x3.

Step4: Region growing process: Regions which are identical in image are segmented out in this region growing process. Region of interest are identified and homogenous regions are filtered in this process. Homogenous regions are extracted at several resolution and in multi scale nature. Homogenous region in soil such as white region, grey part region can be extracted from soil using values of different scale.

Region growing process starts with a pixel or a bunch of pixels and examine the neighbor pixels. If the pixel which satisfies the criteria that pixel is added to the cluster, if it does not then it is excluded from the cluster. The cluster formation is continued until there is no more adjacent pixel can be included in the cluster. Each region is defined by one cluster.

Step 5: White Region detection: After each cluster are formed then the white region in all clusters are separated using higher threshold value in homogenous region. For segregation of white part from the cluster, at small scale=1.

```
Step 1 : Input the soil image

    x=imread ('soil1.jpg');

Step 2 : Preprocessing of image (Image Resize, Color Conversion)

    S=zeros (size(x));

    I = imresize (I, [256,256]);

    c = makecform ('srgb2lab');

    S (273:275,423:425)


Step 3 : Multi resolution process

    ab(:,:,1) = a; ab(:,:,2)= a; ab(:,:,3)= a;

    b(c==1)=70;ab(:,:,1)=x1;

    b(c==1)=77;ab(:,:,2)=x2;

    b(c==1)=0;ab(:,:,3)=x3;

x1,x2,x3 are various resolution based on this homogeneous
regions are separated.
```

Proposed Algorithm:

3545

Figure 4.a. Image 3   b. Fuzzy C Mean   c. K-mean Clustering   d. Proposed Method

Figure 5.a. Image 4   b. Fuzzy C Mean   c. K-mean Clustering   d. Proposed Method

Figure 6.a. Image 5   b. Fuzzy C Mean   c. K-mean Clustering   d. Proposed Method
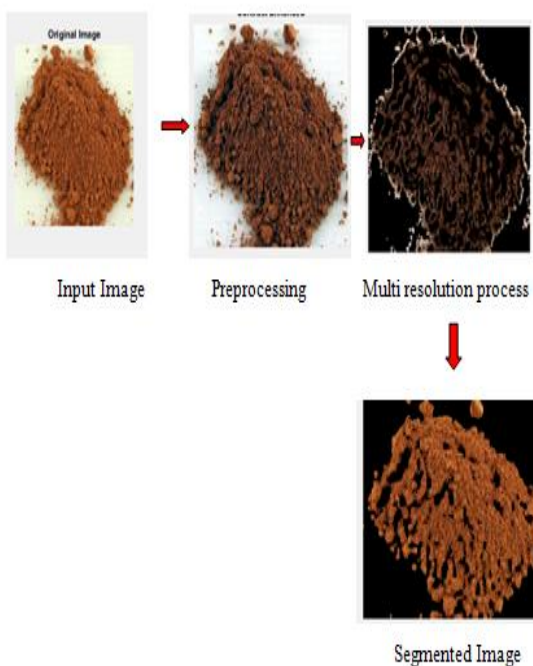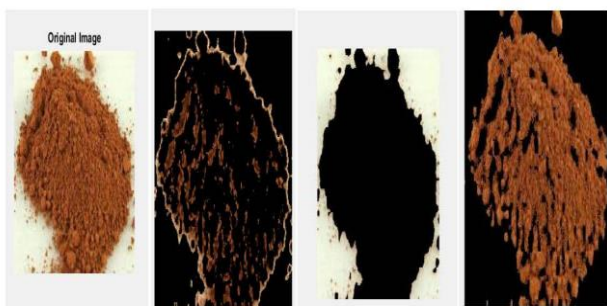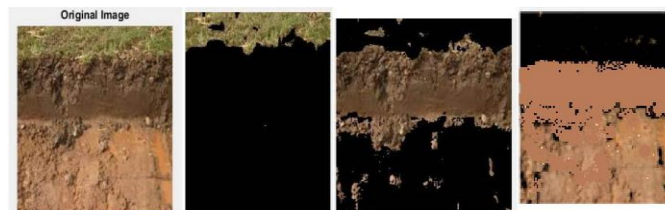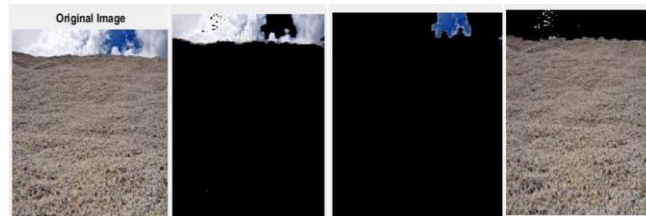

Figure 1. Proposed work Flowchart

The proposed algorithm exactly segments the soil region from the background. The proposed method exactly removes background than other existing methods. It is represented in figure d.

Figure 2.a. Image 1b. Fuzzy C Meanc. K-mean Clustering d. Proposed Method
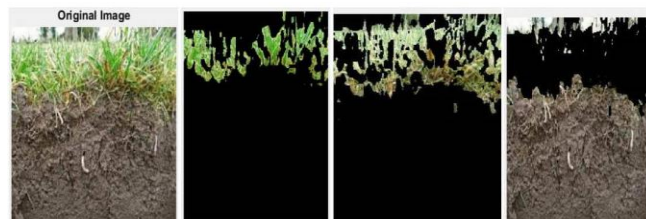

Figure 3.a. Image 2   b. Fuzzy C Mean   c. K-mean Clustering   d. Proposed Method

Different soil images are taken for testing the proposed segmentation algorithm, original images are indicated in Figure a and the segmented out image from the original image are taken for analysis is shown in Figure d. Here the proposed segmentation algorithm is compared with the Fuzzy C mean and K-Mean clustering algorithm. The Images which are processed with Fuzzy C Mean are indicated with Figure b and images which are processed with K-mean clustering are indicated in Figure c.

V. RESULT AND DISCUSSION

The To evaluate the proposed segmentation the five metrics are used dice coefficient, jaccard index, Sensitivity, Specificity and Precision. Then four parameters are used namelyTP, FP, TN and FN.These parameters are defined as follows:

True positive (TP): pixels properly segmented as foreground

False positive (FP): pixels incorrectly segmented as foreground

True negative (TN): pixels properly detected as background

False negative (FN): pixels incorrectly detected as background

Accuracy: `(TP+TN)/ (TP+FP+TN+FN)`

Dice coefficient: The overlaps among two binary images are measured by this metrics. It gives added power to instance where the images consent. This metric is used to evaluate the performance of segmentation.This is calculated by using the following equation (1).

A Novel Algorithm for Soil Image Segmentation using Color and Region Based System

$$\text{Dice} = \frac{2TP}{2TP+FP+FN} \qquad \rightarrow (1)$$

Jaccard Index: This metric is used to evaluate the similarities of the segmentation by evaluate the Ground-Truth and the segmented image. To measure the following equation (2) is used.

$$\text{Jaccard} = \frac{TP}{TP+FP+FN} \qquad \rightarrow (2)$$

Sensitivity: It is also known as True Positive Rate (TPR), estimate the percentage of positives that are recognized correctly. Equation (3) is used to calculate sensitivity.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \qquad \rightarrow (3)$$

Specificity: This is known as True Negative Rate (TNR), estimate the percentage of negatives that are recognized correctly. The equation (4) is used to calculate specificity.

$$\text{Specificity} = \frac{TN}{TN+FP} \qquad \rightarrow (4)$$

Precision: it is used to estimate level to which the same result would be created over different segmentation sections. The equation (5) is used to calculate precision.

$$\text{PPV} = \frac{TP}{TP+FP} \qquad \rightarrow (5)$$

Table 1. Evaluation of metric values

Test image	Method	DICE	JACCARD	SENSITIVITY	SPECIFICITY	PRECISION	ACCURACY
Image 1	Proposed	0.989	0.979	0.9998	0.982	99.44	97.72
	FC Mean	0.782	0.745	0.781	0.789	86.43	88.47
	K means	0.745	0.756	0.765	0.765	75.34	75.27
Image 2	Proposed	0.986	0.996	0.9583	0.965	99.73	98.45
	FC Mean	0.721	0.821	0.765	0.724	73.32	73.43
	K means	0.673	0.732	0.793	0.764	74.23	74.21
Image 3	Proposed	0.981	0.987	0.984	0.985	98.21	98.65
	FC Mean	0.781	0.693	0.751	0.698	75.02	75.43
	K means	0.752	0.765	0.696	0.675	75.11	75.21
Image 4	Proposed	0.989	0.995	0.978	0.981	98.33	98.67
	FC Mean	0.682	0.721	0.699	0.763	73.22	73.34
	K means	0.693	0.743	0.688	0.771	73.67	73.45
Image 5	Proposed	0.984	0.993	0.975	0.986	98.37	98.70
	FC Mean	0.699	0.753	0.734	0.767	74.45	74.34
	K means	0.699	0.746	0.754	0.771	73.76	74.45

Table 1 shows the evaluation performed with different parameters such as dice, jaccard, sensitivity, specificity, precision and accuracy of proposed algorithm with comparison to Fuzzy C means and K means algorithm. Here five different images are taken into for comparison.

Figure 7 shows the comparison of proposed and existing methods performed on Image 1 with metrics such as dice, jaccard, sensitivity and specificity
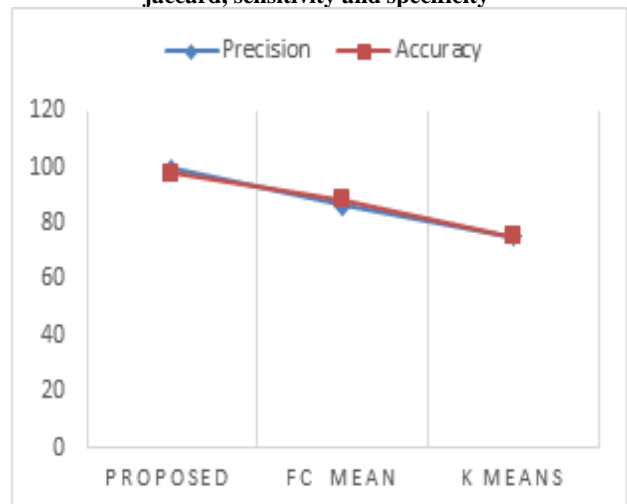

Figure 8 shows the precision and accuracy of proposed method with existing methods performed on Image 1
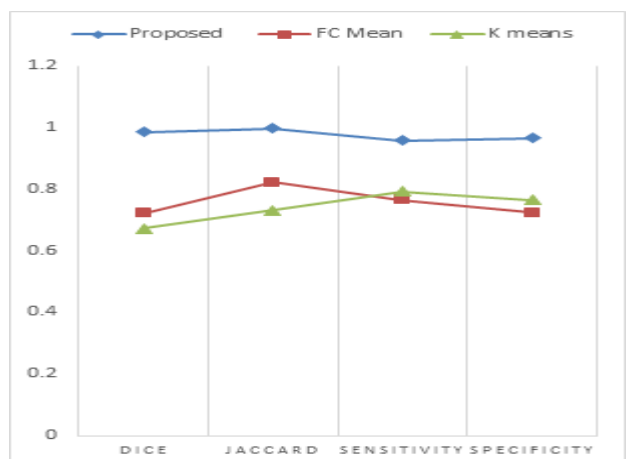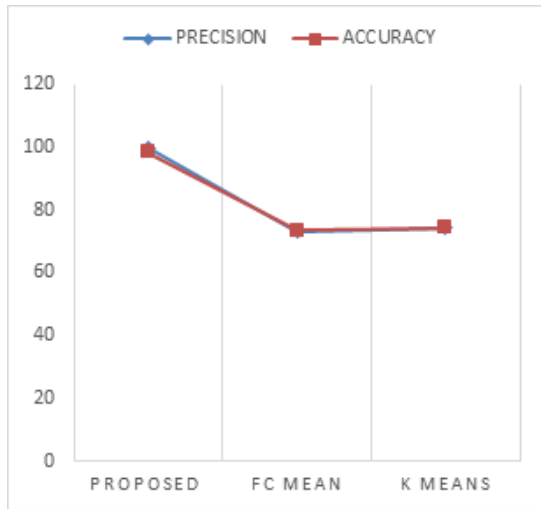

Figure 9 shows the comparison of proposed and existing methods performed on Image 2 with metrics such as dice, jaccard, sensitivity and specificity

Figure 10 shows the precision and accuracy of proposed method with existing methods performed on Image 2
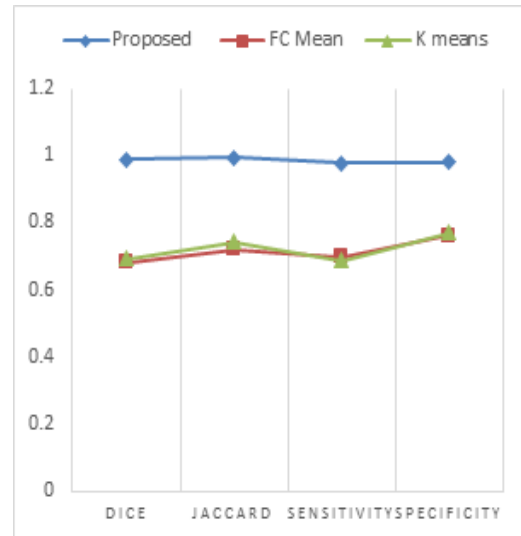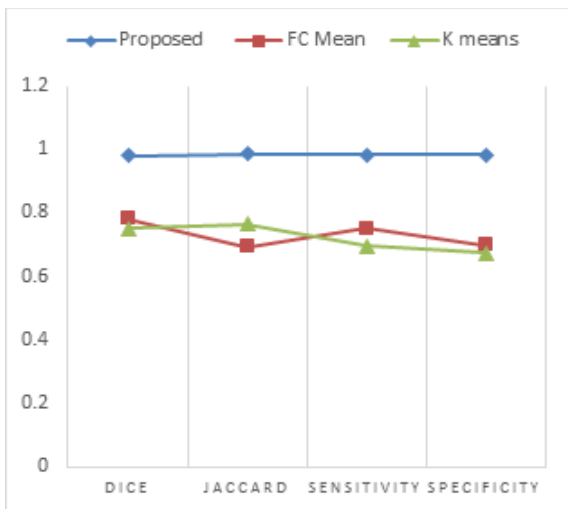

Figure 11 shows the comparison of proposed and existing methods performed on Image 3 with metrics such as dice, jaccard, sensitivity and specificity

Figure 12 shows the precision and accuracy of proposed method with existing methods performed on Image 3

Figure 13 shows the comparison of proposed and existing methods performed on Image 4 with metrics such as dice, jaccard, sensitivity and specificity
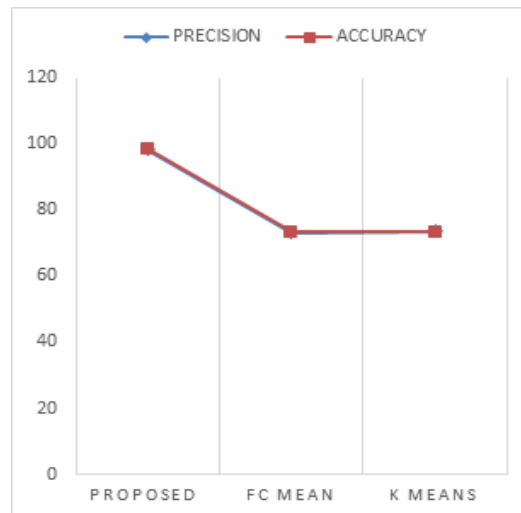

Figure 14 shows the precision and accuracy of proposed method with existing methods performed on Image 4
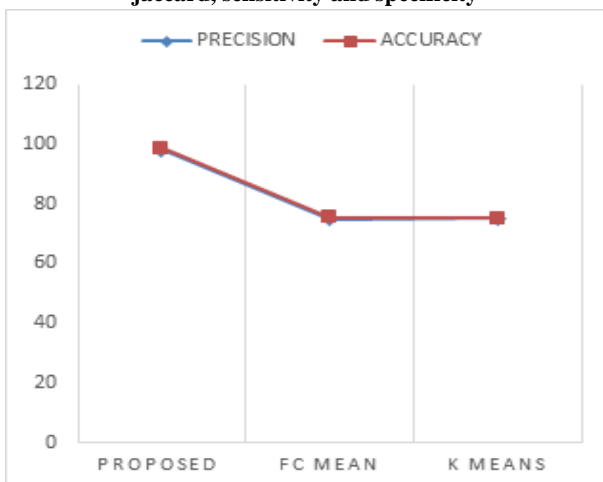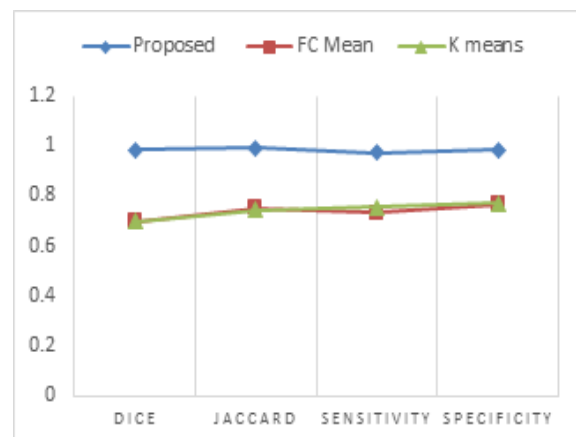

Figure 15 shows the comparison of proposed and existing methods performed on Image 5 with metrics such as dice, jaccard, sensitivity and specificity
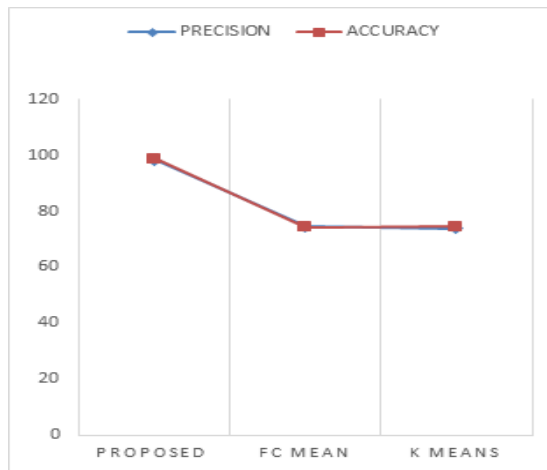
A Novel Algorithm for Soil Image Segmentation using Color and Region Based System

Figure 16 shows the precision and accuracy of proposed method with existing methods performed on Image 5

Different evaluation metrics are taken for comparison of the proposed and the existing methods, which is performed on different images. In order to evaluate the performance of segmentation, dice coefficient is taken. Similarity of segmentation is evaluated by jaccard index. Figure 7 shows the graphical representation of proposed segmentation method which proven to be effective when compared to Fuzzy C Mean and K means algorithm.

Figure 8 shows the precision and accuracy of proposed segmentation method performed on image 1. Performance of proposed method on image 2 is depicted in figure 9 and figure 10, where the accuracy of the segmentation algorithm is 98.45%. . Performance of proposed method on image 3 is depicted in figure 11 and figure 12, where the accuracy of the proposed segmentation method is 98.65%. . Performance of proposed method on image 3 is depicted in figure 13 and figure 14, where the accuracy of the segmentation method is 98.67%. . Performance of proposed method on image 5 is depicted in figure 15 and figure 16, where the accuracy of the proposed segmentation method is 98.70%

VI. CONCLUSION

In this research work a new segmentation approach is introduced to segment the soil region based on the color and region. The proposed approach evaluated by following metrics such as Dice, Jaccard, Sensitivity, Specificity and Precession. The new segmentation method gave good result than other methods. The method is tested with five sample soil images it detect the soil and remove the background and other information. The result is compared with Fuzzy C Mean (FCM), K Mean algorithm. The result of new method is as follows Dice coefficient is 98%, Jaccard Index 98%, Sensitivity 98%, Specificity 98%, Precision 98% and finally accuracy achieved 98%.

REFERENCES

1. Borselli, L. Segmentation of soil roughness profiles. Earth Surface Processes and Landforms, 24(1), 71-90, 1999.
2. Bushra Nazir, Md. Iqbal Qureshi. Soil image segmentation and texture analysis: a computer vision approach. International Journal of Innovative Research in Science, Engineering and Technology, Vol. 5, 2016.
3. Charytanowicz, Małgorzata, and Piotr Kulczycki. "An image analysis algorithm for soil structure identification." In Intelligent Systems' 2014, pp. 681-692. Springer, Cham, 2015.
4. Soh, Leen-Kiat, and Costas Tsatsoulis. "Segmentation of satellite imagery of natural scenes using data mining." IEEE Transactions on Geoscience and Remote Sensing 37, no. 2,1086-1099, 1999.
5. Sofou, A., Evangelopoulos, G., & Maragos, P. Soil image segmentation and texture analysis: a computer vision approach. IEEE Geoscience and Remote Sensing Letters, 2(4), 394-398, 2005
6. Pradeep Kumar Reddy, Bhavana,Soil Texture Classification And Analysis With Local Ternary Pattern (Ltp) Technique, Recent Advances In Electronics And Communication Engineering-2017.
7. Cheng HD, Lui YM, Freimanis RI. A novel approach to micro calcification detection using fuzzy logic technique. IEEE Trans Med Imaging; 17:442–450. 1998.
8. Manousakas N, Undrill PE, Cameron GG, et al. Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions. Comp Biomed Res; 31: 393–412, 1998.
9. Udupa K, Samarasekera S. Fuzzy connectedness and object definition: theory, algorithms and applications in image segmentation. Graph Models Image Process; 58:246–261, 1996
10. Bezdek JC, Hall LO, Clarke LP. Review of MR image segmentation techniques using pattern recognition. Med Phys; 20:1033–1048. 1993.
11. Schalkoff J. Pattern recognition: statistical, structural and neural approach. New York: Wiley & Sons, 1992.
12. Zijdenbos AP, Dawant BM. Brain segmentation and white matter lesion detection in MR images. Crit Rev Biomed Eng; 22:401–465, 1994
13. Pham D. L., Xu C., and Prince J. L., A Survey of Current Methods in Medical Image Segmentation. Annual Review of Biomedical Engineering, 1998.
14. Xu R., and Wunsch D. Survey of Clustering Algorithms. IEEE Transactions on Neural Networks, vol. 16, no. 3, May 2005.
15. Engr. V. C. Chijindu et.al.,Medical Image Segmentation Methodologies – A Classified Overview, African Journal of Computing & ICT. ISSN 2006-1781, Vol 5. No. 5,2012.
16. Prathik, A., J. Anuradha, and K. Uma. "Survey on Spatial Data Mining, Challenges and Its Applications." Journal of Computational and Theoretical Nanoscience 15, no. 9-10: 2769-2776.

AUTHORS PROFILE

Prathik. A is a Research Associate in School of Information Technology and Engineering, VIT University in Vellore, Tamil Nadu. He received M.Tech in the field of Information Technology-Networking from VIT University,Vellore in the year 2015.He is currently doing PhD in VIT University, Vellore, Tamil Nadu, India.His research interest includes Spatial Data Mining,communication networks, coding theory, algorithms, computation, operations research, circuit design, and communication network addressing and data base management.

Dr Anuradha. J is working as Associate Professor in School of Computer Science and Engineering, VIT University in Vellore, Tamil Nadu. She has more than 15+Years of Teaching Experience. She is currently guiding 6 PhD scholars. Her research interest includes Machine Learning, Classification, Computational Intelligence,Spatial Data Mining, and coding theory, Soft Computing, Clustering Algorithm, Rough Set Theory, Feature Selection and Algorithms.

Dr Uma. K is working as Associate Professor in School of Advanced Sciences, VIT University in Vellore, Tamil Nadu. She has more than 20+ years of teaching Experience. She is currently guiding PhD scholars. Her research interest includes Engineering, Applied and Computational Mathematics, Mathematical Analysis, Numerical Mathematics, Real and Complex Analysis, Measure Theory, Harmonic Analysis, Constructive Approximation.