FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

Novel Approach for Mixed Pixel Extraction from Remote Sensing Images

Sumit Kaur[1] and Dr. R.K. Bansal[2]
[1]*Research Scholar, Gurukashi University, Talwandi Saboo, India, sumit.bhangu87@gmail.com*
[2]*Dean Research, Gurukashi University, Talwandi Saboo, India, bansalrajk2009@gmail.com*

***Correspondence:** *sumit.bhangu87@gmail.com*

ABSTRACT- In remote sensing imagery presence of mixed pixels while performing the classification process is the biggest. Detecting and classifying the target pixels like minerals and artificial objects from RSI has a huge interest in various applications. In most of the existing techniques the pure pixels (endmembers) can be detected by using seven band values of the capturing satellite. But if those values are not available, then in such situations theses existing algorithms don not able to perform. To solve this problem we proposed a novel technique which used superpixels (based on only RGB values) for preprocessing and classify the given image by using random forest classifier to classify the image into pure and mixed pixels. This proposed method performs better than the existing techniques in terms of accuracy, RMSE and computing time.

Keywords: Remote sensing images, pure pixel, mixed pixel, neural network.

ARTICLE INFORMATION
Author(s): Sumit Kaur and Dr. R. K. Bansal;
Received: 04/05/2023; **Accepted**: 19/06/2023; **Published**: 30/06/2023;
e-ISSN: XXXX-XXXX;
Paper Id: IJCSR-020206;
Citation: 10.37391/IJCSR.020206

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Remote sensing images are the images captured from space by the use of satellites to so that the information about the objects which are not possible for humans to be being in contact direct physically and in remote sensing images a hyperspectral image is the collection of hundreds of images with different wavelength channels in the same earth area [13]. In hyperspectral image there is a large number of spectral bands. This band information, gives the information about the material within the pixel. Remote sensing image processing is used in various applications like; mineral mapping, medicine, surveillance, crop mapping, environment monitoring, object detection, etc. [14]. To extract the information from the remote sensing images the main step is segmentation of the image in different regions depends upon the applications. So, for the good quality of information or high accuracy of the extracted information the segmentation process has to be performed with high accuracy. But the biggest problem in segmentation is the presence of mixed pixels in the image. Mixed pixels are the pixels that show the belongingness to more than one class. This reduces the accuracy of segmentation. The performance of segmentation can be improved by increasing the resolution of the images but only to some extent the problem will be solved. To improve the quality of information /segmentation the mixed pixel decomposition problem need to solve. Mixed pixel decomposition is the process of assigning labels to the mixed pixel to the appropriate class (only one single spectral signature) to which it belongs. This process is also known as spectral unmixing. This process has two phases first; identification of endmember signatures second is to extraction of quantity of different end –members forming a single pixel (mixed pixel) for all the pixels available in an image. In hyperspectral images mixed pixel decomposing in easy as compare to single spectrum remote sensed images because, in hyperspectral images are captured with different spectral band values and this much information of a single pixel will make it easy to classify but on the other hand classification of pixels having single spectral value is difficult.

Mixed pixel decomposition is used in multiple applications, in recent area the research in this field has growing on. There are many techniques and algorithms are proposed and used for mixed pixel classification. All the available techniques have their own pros and cons, but he main issues arise from the lecture is the extraction of mixed pixels or identification of mixed pixels from the image (dataset). Most of the researcher used band values for extraction of mixed pixels from the image. The band value dataset for the entire images is not available easily. So, to solve this problem a new approach is proposed to extract mixed pixels from any satellite image without the band value information by using fuzzy theory and neural network.

In this paper the type of mixed pixel is presented in *section I*. in section II the different methods used for extraction of mixed pixels are given. The proposed method for mixed pixel identification is given in *section III*. *Section IV* shows the results and conclusion. The future scope of the proposed approach is given in last in *section V*.

1.1 Type of pixel

An image is a collection of pixels organized in the form of matrix. There are multiple satellites are available capturing a remote sensing image at high quality. A new technology used to capture the minerals, terrestrial vegetation, and man-made materials, etc. Each pixel has a spectral signature which

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

represents the class of the pixel to which it belongs. This information of pixels helps to segment or classify the image according to the applications. But the main problem in segmentation is mixed pixel. Remote sensed image has two types of pixels. One is pure pixel these are the pixel that shows only one spectral signature and the other is mixed pixel that show more than one spectral signature's properties which make the pixel to belongs to more than one class. While segmenting the image these mixed pixel decreases the quality of segmentation process and this problem is known as unmixing problem. Basically, there are four types of features used in image processing; low level visual features, local features, Local-global features and biologically inspired features. But the assumption of these features is that all the pixels of an image are pure pixel, which is not true in remote sensed because there is always the presence of mixed pixels in images. There are four types of mixed pixels in remotely sensing images [3]:

i. Small subpixel objects: the size of the object is smaller than the size of the pixel.
ii. Boundary pixels: the sizes of two or more land-cover classes on the ground may be larger than the sizes of the pixel, but parts of their boundaries lie in a single pixel.
iii. Intergrade pixels: a pixel allocates a space for a transition from a cluster of one class to a cluster of another class.
iv. Linear subpixels: the length of a land-cover class may be longer than a pixel but its width is thinner, and the land-cover class runs through a pixel.

2. METHODS TO EXTRACT MIXED PIXELS

Endmembers are the member or the pixel in the image having single signature spectra of the materials that are present in the captured scene. Then these endmembers are used for extraction and classification of abundance fraction (mixed pixels) by using mixed pixel decomposing algorithms [17]. In spectral band i the pixel values are assumed to be the mixture of the spectral signatures of all the components with in the pixel which can be represented as:

$$m = ux + vy + wz \qquad (1)$$

Where m is known spectral signature of mixed pixel, x, y and z are the known as spectral signature of the three endmember signatures with in the mixed pixel, and what is the proportions of these all endmember with in that mixed pixel is represented by u, v and w. our interest is in identify the value of u, y and z. Basically we use Least Square Error (LSE) method to find the unknown values of u, v and z. we can also use Hough Transform(HT) for this as LSE sensitive to outliers but HT perform better in presence of outliers[2].

There are two categories of techniques/methods that are used to identify endmembers from remote sensed images, first are those that are designed under the assumption of pure pixel. These algorithms assume that there are at least one pure pixel is present per endmembers. Means that for each endmember i, there is at least one pixel having p_i being a vertex of the simplex encompassed by the data. Second techniques are those that are designed without the assumption of pure pixel that not a single pure pixel is available. Following are the techniques used for extraction of mixed pixels.

2.1 PPI

Pixel Purity Index is supervised in nature that most successful technique [1] used for endmember extraction as it is also available in Environment for Visualizing Images (ENVI) software. PPI is based on the convex set of geometry and take spectral pixels as vectors in N –dimensional space [4]. In PPI algorithm the first step is to reduce the dimensionality by using Minimum Noise Fraction (MNF) is applied on original data cube. After this random N-dimensional vector known as skewers are generated by calculating pixel purity score for each point [6]. Every data point with its position is put on each skewer, and ones falling at the extremes of each skewer are counted. Those pixels that count above a decided threshold are declared pure.

2.2 N-Finder

N-Finder algorithm is basically finding the pixels that can construct simplex with the maximum volume. In this method the first step is to reduce the dimensionality of the original data by using MNF. Then the randomly selected pixels are declared as endmembers. To calculate the trail volume Z, will be defined as:

$$Z = \begin{pmatrix} 1,1,\ldots..1 \\ z_1, z_2,\ldots.z_Z \end{pmatrix} \qquad (2)$$

Here z_i are the endmeber column vectors, Z represent the total number of endmembers which is used to compute the simplex volume. The endmember formed the volume of simplex is proportional to determinants of Z [10];

$$V(Z) = \frac{1}{(Z-1)!} abs(|Z|) \qquad (3)$$

2.3 IEA

Iterative Error Analysis (IEA) is widely used endmember extraction algorithm. This is algorithm is a collection of constrained unmixing operations. Mean spectrum is selected from the data to set the initial vector then the series of constrained unmixing operation is perform on this data vector, then error image is generated by the errors at each pixel once the unmixing operations are performed. The user then selects a desired number of endmembers E, a number of pixels R, and an angle price Θ. R is the range of pixels with the most important number of errors, selected from the error image. The spectral vector corresponding to the pixel with the only largest error is found. A subset of R consisting of all those pixels that fall at intervals associate angle Θ of the most error vector is then calculated, and these pixels are averaged to turn out the new endmember vector. This process is continuing till E endmembers are elite [5].

2.4 CCA

This method is primarily based on the very fact that some physical quantities, such as radiance and reflectance, are plus.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

The vectors formed by distinct radiance/reflectance spectra will be expressed as linear mixtures of plus parts, which lie within a plus, convex region. The objective of CCA is to seek out the boundary points for that region. To implement this concept, the method finds the eigenvectors of the sample spectral matrix of the image, and selects those eigenvectors corresponding to the E largest eigenvalues (where is that the a priori number of endmembers to model). The method then appearance for the boundaries of the convexo-concave cone, where the linear mixtures of these eigenvectors turn out vectors that area unit strictly plus, by using the following expression:

2.5 AMEE

The Automated Morphological Endmember Extraction (AMEE) rule is associate endmember extraction algorithm that makes synchronal use of special and spectral information via multi-channel morphological process [8]. The input to AMEE is the full image data cube, with no need of spatial property reduction. Let r denote the input data cube and r (x, y) denote the pixel vector at special location (x, y). Similarly, let K be a kernel defined in the special domain of the image (the x–y plane).

This kernel, usually known as structuring component (SE) in mathematical morphology nomenclature, is translated over the image. The SE acts as a probe for extracting or suppressing specific structures of the image objects, according to the dimensions and shape of the SE. Having the above definitions in mind, AMEE method is based mostly on the appliance of multichannel erosion and dilation operations to the information.

$$(r \otimes K)(x,y) = arg\{min_{(s,t)\in K}[\sum_s \sum_t SAM(r(x,y), r(x+s, y+t))]\} \quad (4)$$

$$(r \otimes K)(x,y) = arg\{max_{(s,t)\in K}[\sum_s \sum_t SAM(r(x,y), r(x-s, y-t))]\} \quad (5)$$

where SAM is spectral angular mapper.

2.6 Vertex Component Analysis

The VCA [11] is also designed to scale back expensive procedure quality suffered in MVT and CCA by substitution straightforward volume calculation with OP and growing protrusive hulls vertex by vertex till it reaches a p-vertex protrusive hull rather than substitution p-vertex protrusive hulls all at once as MVT and CCA do.

Its idea is similar to SGA within the sense that VCA conjointly grows protrusive hulls one vertex at a time consecutive in succession, but has a major distinction from that of SGA in a way to realize a vertex to grow a protrusive hull. More specifically, VCA grows convex hulls with peak orthogonal projections instead of simplexes with peak volume utilized by SGA. In other words, VCA appeals for the maximum orthogonal projection as a criterion as PPI will to grow its protrusive hulls compared to SGA that uses the peak volume of a simplex as a criterion to grow simplexes as N-FINDR will. In light of this interpretation, VCA can be thought-about a consecutive version of PPI, and SGA can be viewed as a consecutive version of N-FINDR.

Table 1: Computational Complexity of Different Algorithms

Algorithms	Numbers of flops
PPI	2EsP
VCA	$2E^2P$
N-Finder	$E^{\eta+1}P$
SGA	$P(\sum_{n=2}^{E} n^{\eta}$

Table 1 shows the computational complexity among PPI, VCA, N-finder and SGA. Where E is the total number of endmemember, P is the total number of pixels and s is the skewers.

Table 2: Comparison of Algorithms

Algorithm	PPI	N-Finder	IEA	AMEE	CCA	VCA
Type of algorithm	Supervised	Unsupervised	Unsupervised	Unsupervised	Unsupervised	Unsupervised
Category	Non-statistical	Geometric	-	-	Geometric	Geometric

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

Computational complexity	High number of iteration required	Depends on initial random pixel selection	Significant due to repeated un mixing steps	Moderate for medium sized kernel	Very high if there are large numbers of corners in the convex cone	High as the algorithm iterates until all endmembers are exhausted
Convergence property	Maximum no.of iteration	Simplex with the maximum volume found	Minimum error in unmixing	Range of kernel size	All comers of the convex cone found	projects all pixels onto a direction orthogonal to the simplex spanned
Mostly applied areas	Land cover classification, Mineral mapping	Land cover classification, Mineral mapping	Land cover classification, Mineral mapping	Land cover classification, target detection	Target and anomaly detection	Land cover classification
Parametric	Partially interactive	Non Parameterized	Non Parameterized	Parameterized	Parameterized	Parameterized
Dimensional Reduction	MNF	MNF	None	None	-	SVD & PCA
Spectral information Required?	Yes	Yes	-	Yes	Yes	Yes
Spatial information requires?	No	No	No	Yes	No	No
Support multiple Endmembers	No	No	No	-	Yes	Yes

MNF- Minimum Noise Fraction, SVD- Singular Value Decomposition, PCA- Principal Component Analysis

3. PROPOSED METHOD FOR EXTRACTION AND RESULTS

All the above said method used to extract endmember from the satellite image, but not single method from these methods is used for the identification of mixed pixels from the remote sensing images. Most of the methods that are used to classify mixed pixel use band values for the identification of mixed pixel from the pure one, but when this band value is not available then it is hard to say which pixel is mixed pixel. In this paper a new approach is proposed to classify a remote sensed image into pure pixels and mixed pixels by using neural network.

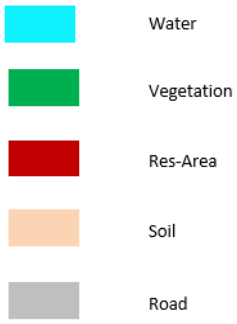
Step 1: Generate superpixel image

In this experiment different 100 satellite images have been used from airbus defence and space library. The database of proposed work is the collection of images from different satellites so that this system can be used for any resolution and without the information of band values. The Pléiades satellite image of Milano 2015, Italy shown in *figure (1)* containing five land cover classes: water, vegetation, residential area, soil, roads, which are used as the base for classification process.

(a) (b)

Figure 1: (a) The Pléiades satellite image of Milano 2015, Italy, b) ground truth image of (a)

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX



Legends

The pixel-grid illustration is Associate in Nursing "artifact" of a digital imaging method and not a natural one. Most of the image process algorithms visualize image with the utilization of pixel grid, because the underlying illustration.

Several random modelsof pictures area unit usually outlined on this regular grid. It'd be a lot of natural, and presumptively a lot of economical, to figure with meaty entities obtained from a low-level grouping method. The results of over-segmentation partitions the image into fewer range of segments referred to as superpixels. Superpixels are ordinarly outlined as catching and grouping uniform pixels within the image, that are wide utilized in image segmentation and visual perception [9][18].

The superpixel map is natural and perceptually important illustration of the input image. Therefore, compared to the standard element illustration of the image, the superpixel illustration greatly reduces the amount of image primitives and improves the representative potency. The convenience and effectiveness of superpixels to reckon the region primarily based visual options provides vital advantages for the vision tasks like visual perception [9]. This segmentation procedure employs SLIC superpixels [19] and the DBSCAN clustering algorithm. SLIC superpixel approach formulates an over-segmented image and the generated superpixels are then processed through DBSCAN method to create superpixels segmentation clusters. This strategy is easy and fast relatively. *Figure (2)* shows the different images and their conversion to super pixels and then merged super pixel images

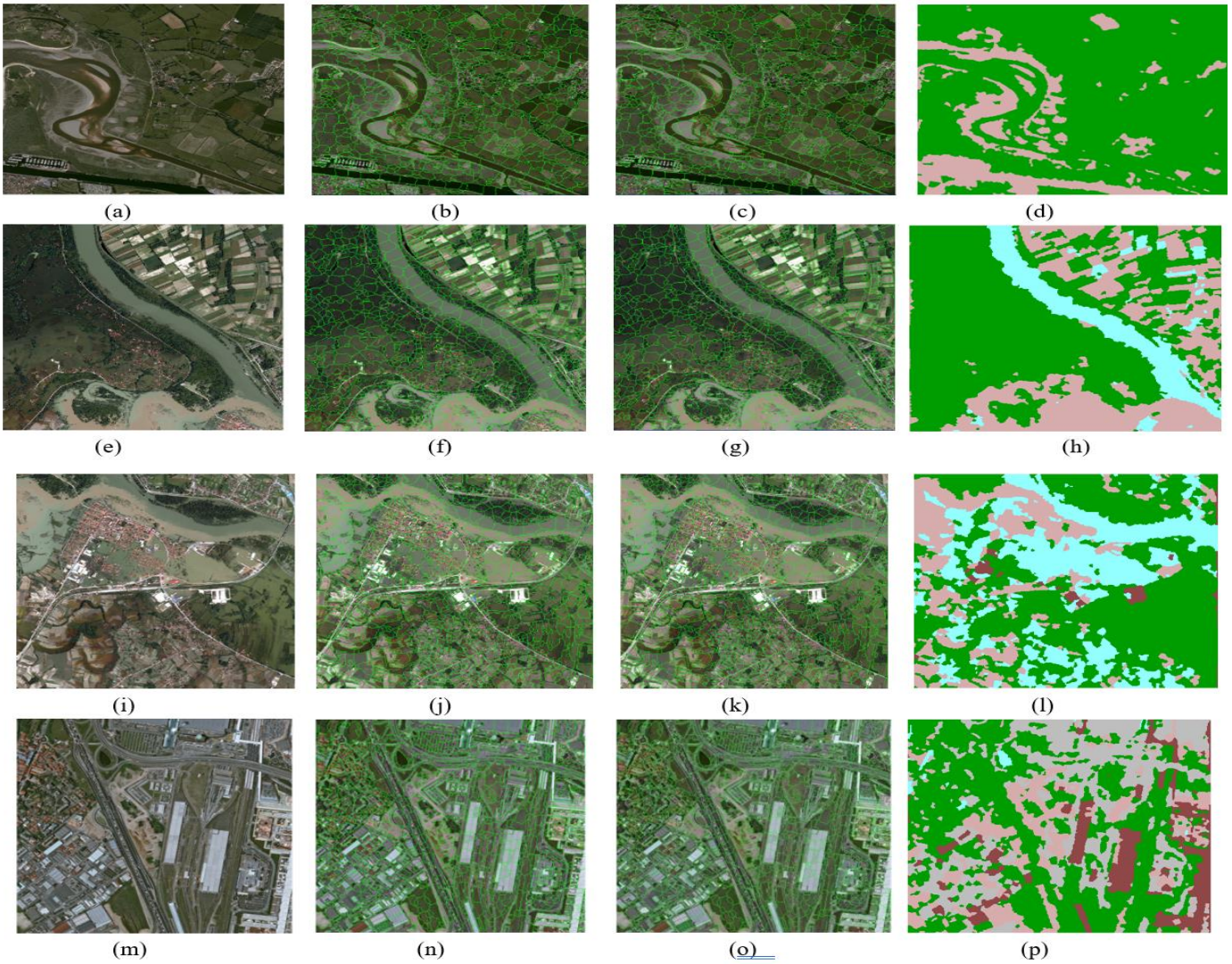


Figure 2: (a, e, i, m) remote sensed images of different resolution, (b, f, j, n) Super pixel image of RSI, (c, g, k, o) merged super pixel image, (d, h, l, p) Ground truth image of images of merged super pixel images

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

Step 2: Merge internal objects of super pixel

Super pixel clustering technique of *step 1* produce internal object in a single super pixel, which causes confusion and increases processing time and reduces accuracy. To solve this issue in this proposed method we will merge internal object of a super pixel to its main objects by following *algorithm 2*.

Algorithm 2: Merging internal objects of super pixel

$for\ i \coloneqq 1\ to\ n$ // n is the number of unique objects found after applying super pixel

$A_i = bin(O_i)$ // $bin(O_i)$ is the function to compute binary of i^{th} object.

$AF_i = fill(A_i)$ // $fill(A_i)$ is the function to fill internal holes of objects as given in [21]

$AN_i = AF_i = \& \sim A_i$ // AN_i is the internal object which are filled by morphological operation

$[r, c] = pos(AN_i)$ // r, c are the position found to be merged with function pos of object AN_i

$for\ j \coloneqq 1\ to\ len - r$ // len-r is the variable having length of pixels to be merged

$O(r_j, c_j) = i$ // values at position r_j, c_j will be replaced by i

$end\ for$

$end\ for$

This merged object image give better and accurate clusters and also reduces the processing time for classification.

Step 3: Generate inputs and outputs for classifiers

Once we get final super pixel image then find the features on which feature vector table and targets will be generated which will be feed to classifiers for classification of pixels into pure and mixed pixels. In this experiment inputs and outputs are created by calculating probability of maximum existing object in a single super pixel and by calculating mean of red, green and blues values of a super pixel respectively.

Algorithm.3: Generating inputs and outputs

$for\ j = 1\ to\ \propto$

$r_\beta, c_\beta = pos(O_j)$

$for\ i = 1\ to\ \beta$

$r_i, c_i = pos(O_i)$

$P_i = len - r_i / len - r_\beta$

$end\ for$

$UP_j = sort\{P\}$

$P_j = \left[\frac{\sum_{k=1}^{k=\beta} I_{rk,\ ck\_R}}{\beta}, \frac{\sum_{k=1}^{k=\beta} I_{rk,\ ck\_G}}{\beta}, \frac{\sum_{k=1}^{k=\beta} I_{rk,\ ck\_B}}{\beta},\right]$

$FVT_j = P_j$

$T_j = \max(UP_j)$

$end\ for$

Where
$\propto$ is the number of object after super pixel
r_β , c_β are the position of j^{th} object
r_i , c_i are the position of i^{th} unique object in j^{th} super pixel
P_i is the probability of i^{th} object in j^{th} super pixel
P_j is the variable containg mean value of j^{th} superpixel for R, G and B
FVT_j & T_j are the feature vector table & targets for j^{th} object

Also getting the features and area of different super pixels it will help to identify the total amount of share of the existing classes in the given image

Step 4: Classification of super pixels in pure and mixed pixel super pixels

Once the endmember (distinct classes) are clearly identified from the image next step is to classify the pixels into pure and mixed ones. There are various techniques used to classify the hyper-spectal images based on seven band values, which make the classification process easier, but for remote sensed images without having values for different spectra and only having only information of three bands (Red, Green, and Blue) then the previously used technique will not be able to perform better with that accuracy.

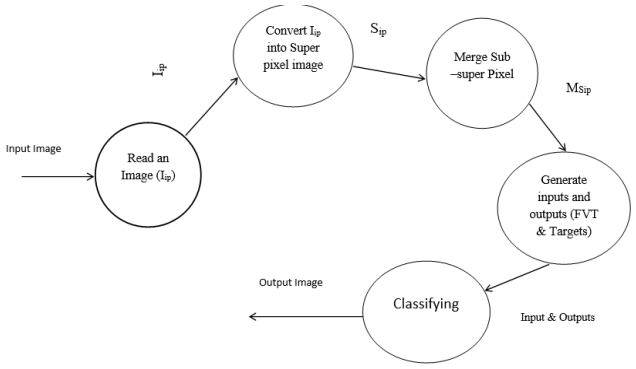


Figure 3: Data Flow Diagram of Proposed Algorithm

There are mainly two type of classification processes that are used; supervised classification and unsupervised classification. In supervised classification, the spectral signatures are developed from the specified locations of the image and theses specified locations are given the generic names 'training sites' and are defined by user. From these spectral signatures statistical characteristics of different land cover classes are generated which are used as reference to classify the pixels to their appropriate class. On the other hand, unsupervised classification, classify pixels into different classes based on the natural grouping exist in the image values. Unsupervised classification does not require expert analytic data for the reference to classify unlike the supervised process. Unsupervised classification approach is not preferred one because results are purely relying on software's knowledge of identifying pixels which is not always correct and as in image there are many type of other instances exist that change the band values that we are using in this experiment like; shadows and these can only analysis by the human. As the result of these in this experiment supervised classification is used. There are various classifiers that can be used to perform supervised classification. *Figure (3)* shows the complete data flow diagram (DFD) of the proposed method used in this experiment to extract the mixed pixels from the inputted satellite image which

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

is pre-processed and to generate super pixel image of the input image. *Table 3* shows the comparison of different classifiers.

Table 3: Difference between Classifiers

Classifiers	Understandability	$Pred_{acc}$	$Train_{sp}$	$Pred_{sp}$	$Perf_{smlobs}$	$Handl_{irrfeat}$	$Autolearn_{featinter}$	$Calibprob_{membership}$	Parametric	$Feat_{needscal}$
Naive Bayes	Yes	Lower	Fast	Fast	Yes	Yes	No	No	Yes	No
Random Forests	No	High	Slow	Moderate	No	Yes	Yes	Possibly	No	No
AdaBoost	No	High	Slow	Fast	No	Yes	Yes	Possibly	No	No
Neural networks	No	High	Slow	Fast	No	Yes	Yes	Possibly	No	Yes

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

$Pred_{acc}$ – predictive accuracy, $Train_{sp}$ - Training speed, $Pred_{sp}$ - Prediction speed, $Perf_{smlobs}$ - Performs well with small number of observations, $Handl_{irrfeat}$ - Handles lots of irrelevant features well , $Autolearn_{featinter}$ - Automatically learns feature interactions, $Calibprob_{membership}$ - calibrated probabilities of class membership, $Feat_{needscal}$- Features might need scaling

There are various statistical parameters used to evaluate the performance of the classifier systems. To evaluate the proposed system, we are using five parameters; True Positive (TP), False Positive (FP), Precision, Recall, F-measure, Accuracy as given in *table 4*.

Table 4: Performance Parameters

Parameter	Formula	Description
Recall	$\frac{TP}{TP+FN}$	measures the proportion of positives that are correctly identified as such
Accuracy	$\frac{TP+TN}{TP+FP+TN+FN}$	measurement system is the degree of closeness of measurements of a quantity to that quantity's true value
Precision	$\frac{TP}{TP+FP}$	proportions of positive results in statistics and diagnostic tests that are true positive results
F Measure	$\frac{2TP}{2TP+FP+FN}$	This measure test the accuracy, it's at best when value is 1 and ta its worst when value is 0

TP-true positives, FN- false negatives, TN- true negatives, FP- false positives

The values of these parameters in the experiment are shown in *table 5* for different classifiers to choose the best among them for further classification.

Table 5: Values of Evaluation Parameters

Name of Classifiers	TP_Rate	FP-Rate	Precision	Recall	F_Measure	Accuracy (%)
Conjugate gradient backpropagation with Powell-Beale restarts	0.2395	0.0586	0.6829	0.6829	0.6829	81.50
Scaled conjugate gradient backpropagation	0.2673	0.0740	0.6686	0.6686	0.6686	81.12
Gradient descent with adaptive learning rate backpropagation	0.0735	0.0098	0.4362	0.4362	0.4362	82.89
Gradient descent with momentum and adaptive learning rate backpropagation	0.1409	0.0246	0.4333	0.5200	0.4333	82.36
Random Forest	1.0000	0.0000	1.0000	1.0000	0.9990	99.94
AdaboostM1	0.9040	0.1780	0.8180	0.9040	0.8590	90.43
Navie Bayes	0.9500	0.0060	0.9630	0.0950	0.9540	95.01
Bagging	0.9990	0.0010	0.9990	0.9990	0.9990	99.89

As it shows from the *table 3* that Random Forest and Bagging classifiers perform better as compare to others. Random Forest is an improvement to bagging decision tree. The fundamental difference between random forest and Bagging is that in Random Forest best split feature from the randomly selected subset of features from the total features are used to split the each nodes in a tree, where as in bagging all the features are considered for this splitting of nodes. This random selection of subset of features to split the nodes reduces correlation between trees which further improve the variance. For the selection from these two in the experiment is clearly evident that Random forest have zero FP rate and have higher accuracy as compare to other for which in this proposed approach Random classifier will be used to classify an image pixels to super mixed and pure super pixels. *Figure (4)* show the algorithm of the Random Forest used in the proposed scheme.

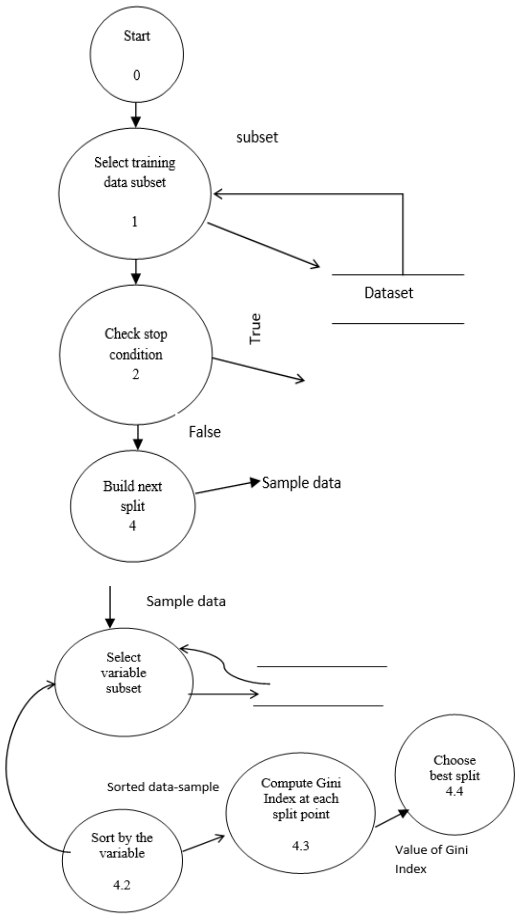


Figure 4: Data Flow Diagram of Random Forest Algorithm

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

There are also various other than accuracy benefits of Random Forest like; it can run on large amount of database efficiently, it can generate an internal unbiased estimate of the generalization error as the forest building progresses, it computes proximities between pairs of cases that can be used in clustering, locating outliers, or (by scaling) give interesting views of the data, Capabilities of the above can be extended to unlabeled data, leading to unsupervised clustering, data views and outlier detection, Offers an experimental method for detecting variable interactions[7]. After passing inputs to classifiers (Random Forest) the image will be classified into pure super pixels and mixed super pixel with only using three bands value. *Figure (5)*, shows the ground truth images of input images and after processing through proposed approach the ground truth of classified images in pure and mixed (red) super pixels.

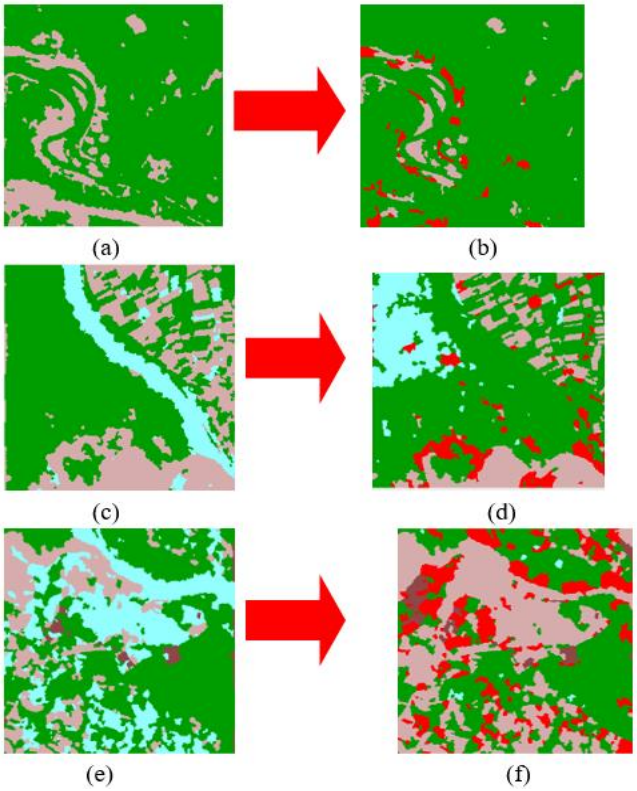


Figure 5: (a, c, e shows the ground truth images of input images, (b, d, f) final classified images (into pure pixel and mixed pixels (in red color))

To compare the proposed algorithm with existing techniques for extraction Root Mean Square Error (RMSE) is used as a parameter. Suppose X is the original image $\hat{X}$ reconstructed image of X. RMSE is calculated as following:

$$RMSE(X,\hat{X}) = \frac{1}{P}\sum_{i=1}^{P}\left(\frac{1}{B}\sum_{j=1}^{B}[y_{ji} - \hat{y}_{ji}]^2\right)^{\frac{1}{2}} \qquad (6)$$

Where y_j is pixel vector given by $y_j = [y_{1j}, y_{2j}, , y_{Bj}]$ of original image and $\hat{y}_j = [\hat{y}_{1j}, \hat{y}_{2j}, \ldots, \hat{y}_{Bj}]$ and j=1, . . , P.

Table 6: RMSE value of the algorithm

Algorithm	Average RMSE
IEA	0.1599
N-Finder	0.1068
VCA	0.1007
Proposed Algorithm	0.0085

Table 6 shows Root Mean Square Error (RMSE) value of different algorithm and it shows that the proposed algorithm in this paper have the minimum RMSE so it is better than other mention algorithm and *figure (6)* show the graphical comparison of algorithm.

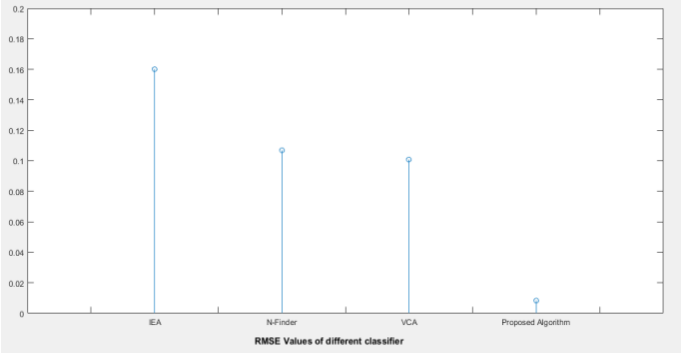


Figure 6: RMSE value of different mixed pixel extraction algorithms

While comparing with different algorithm on the other parameter like computing time, the proposed method take less time as compared to the other algorithms mention in literature survey. Computing time is the time taken by the technique to classify the given image to pure pixels and mixed pixels.

Table 7: Computing time of the algorithms

Algorithm	Computing time
Proposed Method	0.39
VCA	1.16
PPI	1.65
SGA	5.91
N-Finder	31.26

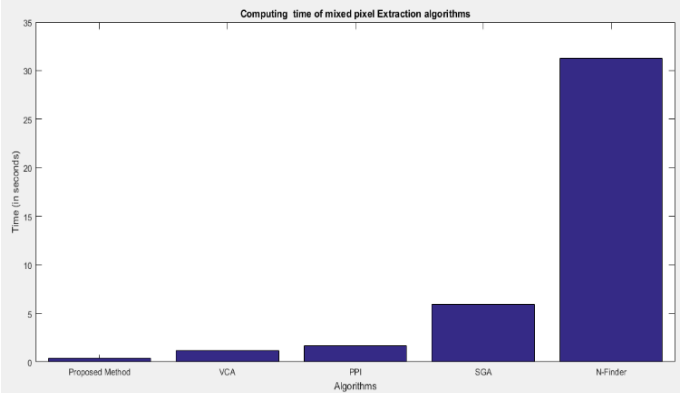


Figure 7: Computing time of different mixed pixel extraction algorithms

International Journal of Computer Science Research (IJCSR)
Research Article | Volume 2, Issue 2 | Pages 59-68 | ISSN XXXX-XXXX

4. CONCLUSION

In this paper a novel approach is proposed for the extraction of mixed pixel from a remote sensed image based on three band values. In this proposed method super pixel concept is used to make classification easy and to increase the speed. So, the image is classified into pure and mixed super pixels and this experiment also merge the internal instances of the super pixel to its parent super pixel. Merging of internal object to its parent will increase the accuracy by 5% of the classification process as compare to use a normal super pixel image. Various classifiers like, Navie Bayes, Adaboost, Bagging, Random forest and neural networks are tested to select the best from them. As per the performance evaluation parameters Random forest perform better than other. Kappa coefficient factor of random forest is 1, which is good for the classification.

REFERENCES

[1] Boardman, J.W., F.A. Kruse and R.O. Green, 1995. Mapping Target Signatures via Partial Unmixing of AVIRIS Data, in Summaries of the V JPL Airborne Earth Science Workshop

[2] Panagiota Bosdogianni, Maria Petrou, 1996. Robust Mixed pixel classification in remote sensing, International Archives of Photogrammetry and remote sensing 31, p. 85-90

[3] P. Fisher, 1997. The pixel: A snare and a delusion, Int. J. Remote Sens. 18, p 679–685

[4] Ifarraguerri, A. and C.I. Chang, 1999. Multispectral and hyperspectral image analysis with convex cones, IEEE Transaction. Geoscience. Remote Sensing 37, p.756- 770.

[5] R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, 1999. Automatic endmember extraction from hyperspectral data for mineral exploration, in Proc. 21st Can. Symp. Remote Sensing, Ottawa, ON, Canada

[6] Theiler, J., D.D. Lavenier, N.R. Harvey, S.J. Perkins and J.J. Szymanski, 2000. Using blocks of skewers for faster computation of Pixel Purity Index, SPIE Proc. 4132, p. 61-71.

[7] Liaw, A. and Wiener, M., 2002. Classification and regression by randomForest. R news, 2(3), pp.18-22.

[8] Plaza, A., Martínez, P., Pérez, R., & Plaza, J., 2002. Spatial/spectral endmember extraction by multidimensional morphological operations. IEEE Transactions on Geoscience and Remote Sensing, 40, p. 2025-2041.

[9] Mori, G., Ren, X., Efros, A. A., & Malik, J., 2004. Recovering human body configurations: Combining segmentation and recognition. In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on IEEE, 2, p. II-326.

[10] Plaza, A., Martínez, P., Pérez, R., & Plaza, J., 2004. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. IEEE transactions on geoscience and remote sensing, 42, p. 650-663

[11] Nascimento, José MP, and José M. Bioucas Dias, 2005. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, 43, p. 898-910.

[12] Chang, C. I., Wu, C. C., Liu, W., & Ouyang, Y. C., 2006. A new growing method for simplex-based endmember extraction algorithm. IEEE Transactions on Geoscience and Remote Sensing, 44, p.2804-2819.

[13] Chang, C. I. (Ed.). 2007. Hyperspectral data exploitation: theory and applications. John Wiley & Sons.

[14] Schaepman, M. E., Ustin, S. L., Plaza, A. J., Painter, T. H., Verrelst, J., & Liang, S., 2009. Earth system science related imaging spectroscopy—An assessment. Remote Sensing of Environment, 113, p. S123-S137.

[15] Song, D., & Tao, D., 2010. Biologically inspired feature manifold for scene classification. IEEE Transactions on Image Processing, 19, p. 174-184.

[16] Huang, K., Tao, D., Yuan, Y., Li, X., & Tan, T., 2011. Biologically inspired features for scene classification in video surveillance. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 41, p. 307-313.

[17] Javier Plaza, Eligius M.T. Hendrix, Inmaculada García, Gabriel Martín, Antonio Plaza, 2012. "On Endmember Identification in Hyperspectral Images Without Pure Pixels: A Comparison of Algorithms", J Math Imaging Vis, 42, p. 163–175

[18] Patz, T., & Preusser, T., 2012. Segmentation of stochastic images with a stochastic random walker method. IEEE Transactions on Image Processing, 21, p. 2424-2433

[19] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua and S. Susstrunk, 2012, SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. PAMI, 34, pp. 2274-2281

[20] Soille, P., 2013. Morphological image analysis: principles and applications. Springer Science & Business Media