Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, No. 1, April : 2023

CLASSIFICATION OF SIDDIPET URBAN-HYPERSPECTRAL IMAGES USING UNSUPERVISED LEARNING TECHNIQUES

G. Rajitha, Research Scholar,
A. Geetha Reddy Professor
[1,2]Dept. of Geography, University College of Science, Osmania University, Hyderabad
[1]rajithaa68@gmail.com ; [2]geetareddy47@yahoo.com

Abstract:

Siddipet Urban Hyperspectral Image (HSI) raw data collected from USGS from 2016 to 2021. A specialized sensor device collects and analyses electromagnetic spectrum data due to Siddipet urban hyperspectral imaging (HSI) pre-processing. Its statistics offer an excess of knowledge. This information may be applied to various Land Use and Land Cover applications to solve multiple issues. (LULC). Urban Boundaries, Agriculture, Water Bodies, Built-up Areas, and Open Areas of HIS and its feature extraction are examples of this application. Each pixel in a digital picture is categorized using hyperspectral imaging. The present investigation uses unsupervised hyperspectral image classification techniques to produce a classified hyperspectral image. The algorithms such as K-Means and K-nearest Neighbour (KNN) have been used. For the preliminary processing of Siddipet Urban Hyperspectral images in Telangana State, India, we use two software's (ERDAS and ENVI).

This study examined the accuracy of the K-NN and K-Means algorithms to determine how well they performed. The K-Means algorithm is shown to be less accurate than unsupervised K-Means techniques. The total accuracy of the classification process using the K-Means method is 89.7696%, compared to the overall accuracy of the K-NN algorithm, which is 87.26%. Additionally, the processing time increased as the number of repetitions to obtain the classed image rose.

1. Introduction

Remotely detected HSI has to be subjected to a classification process known as hyperspectral image (HSI) analysis. Several classification techniques have been used for categorizing hyperspectral data in the past two decades. These algorithms use both supervised and unsupervised techniques. A common classification method used in data mining techniques is K-Nearest Neighbour (KNN). Because of its straightforward implementation, understandable theory, and superior classification performance, it is frequently employed in various applications. However, KNN will raise the classification error rate if the distribution of the training samples is unequal or the sample sizes for the various classes are significantly dissimilar [1]. In this research, an enhanced KNN classification method is adopted and used for the object-oriented classification of high-resolution remote sensing images, learning from clipping KNN. First, imagine objects are gathered using image segmentation and used as sample points. Second, the initial and upgraded KNNs are shown and utilized to classify those sample points.

A high-resolution remote-sensing image may be more accurately classified using the upgraded KNN method. Remote sensing measures surface-level item attributes using information from aeroplanes and satellites. Different wavelength intervals were reflected from or emitted by objects on the Earth's surface by the onboard sensors. Remote sensing applications include spotting earthquakes, faulting, volcanic activity, landslides, flooding, wildfires, and the damages caused by each. The development of high-resolution remote sensing imaging satellites or aeroplanes to measure the quantities of energy is now a foregone conclusion due to the technology of object-oriented categorization [2]. In the present research, segmentation and feature extraction are finished using SPOT5 image data and recognition software. The research then uses an enhanced KNN algorithm to classify imagined objects produced by segmentation, aiming at the flaw in the conventional KNN approach and learning from the concept of clipping KNN. These algorithms all have their limitations. This study uses two individual unsupervised classification techniques to categorize extra raw data from

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, No. 1, April : 2023
2016 to 2021. The major goal is to use unsupervised learning techniques to evaluate the performance of all four classifiers, including border areas. Different datasets are seen by remote sensing, which is used to observe land use and cover. For hyperspectral datasets of urban hyperspectral images from Siddipet, spectral feature extraction is also part of this study.

This essay is divided into four sections: a literature review in part 2, methodology in section 3, findings discussion in section 4, and a conclusion in section 5.

2. Literature Survey

Because it requires the least computing time and primarily relies on training data, past research on minimal distance classification demonstrates that it is highly recommended in all image classification applications. The objective of the current study was to compare and assess several classification methods for identifying objects in images. The best classification method was selected after extensive research and comparison of the K-NN and K-Means algorithms.

The extraction of information or data from images is known as image classification [3]. The primary function of image classification is to find, identify, and categorize an object's characteristics in an image following the class type. Based on the closest neighbor to the unknown classes from the learned classes, the K-NN classification method finds the unknown item of a class in an image [4]. The most used classification method is K-NN, which is utilized in applications for ranking models, text categorization, determining patterns, event recognition, and object recognition. K-NN specifies the number of nearest neighbors to be considered when defining a class of sample data points using the NN rule, in which the nearest neighbor is determined from the value of k.

K-NN offers accurate findings across bigger regions quickly and transparently. Its key advantages are the K-NN algorithm's simplicity and the absence of parametric assumptions. Because it requires the least computing time and primarily depends on training data, past research on minimal distance classification demonstrates that it is highly recommended in all image classification applications. The major objective was to provide a replacement process for the region- and object-based image classification [5]. The approach uses high-resolution satellite data to extract useful area-wide spatial information for city planning and management [6]. The mixed pixel issue that plagued most pixel-based approaches was greatly minimized by the satellite image classification method employing colour and texture data [7]: spectral-based image classification, quadratic discriminant analysis and spatially based on approaches where a pixel's categorization is independent of its neighbours. An explanation of how feedforward neural networks are used in categorizing satellite images, where they are occasionally explored. The various approaches for classifying images were compared using satellite photos [7]; as a result, it was discovered that the maximum likelihood method was more useful and trustworthy for their particular type of satellite image classification. A slightly different viewpoint by extracting the various cadastral features, such as houses and roads, using particle swarm optimization [8] and mapping the land cover utilizing two k-means clustering methods and swarm computing methodologies.

Two separate techniques, the k-means algorithm and the backpropagation algorithm of an artificial neural network, were used to segment and classify satellite images. Used wavelet-based features to perform multispectral imagery classifying [6]. They observed that the wavelet transform offered an exact and comprehensive framework for characterizing and evaluating a signal. Numerous categorization algorithms are considered concerning the nature of our images. The robust k-means clustering explore was used to properly classify the picture into four environmental types [7]. The objective of the current study was to compare and assess several classification methods for identifying objects in images. The best classification method was determined after research and comparison of the k-NN and k-Means algorithms.

3. Methodology

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, No. 1, April : 2023

Hyperspectral remote sensors are used frequently in remote sensing to monitor the Earth's surface with acceptable spectral resolution. Comparing the HSI to traditional RGB images, there are typically more than three bands. Crop analysis, geological mapping, mineral exploitation, defense research, urban inquiry, military surveillance, etc., are only a few of the issues hyperspectral images address. Utilize this work or article that discusses the gathering, preparation, and exploratory data analysis of HIS data [8], [9].

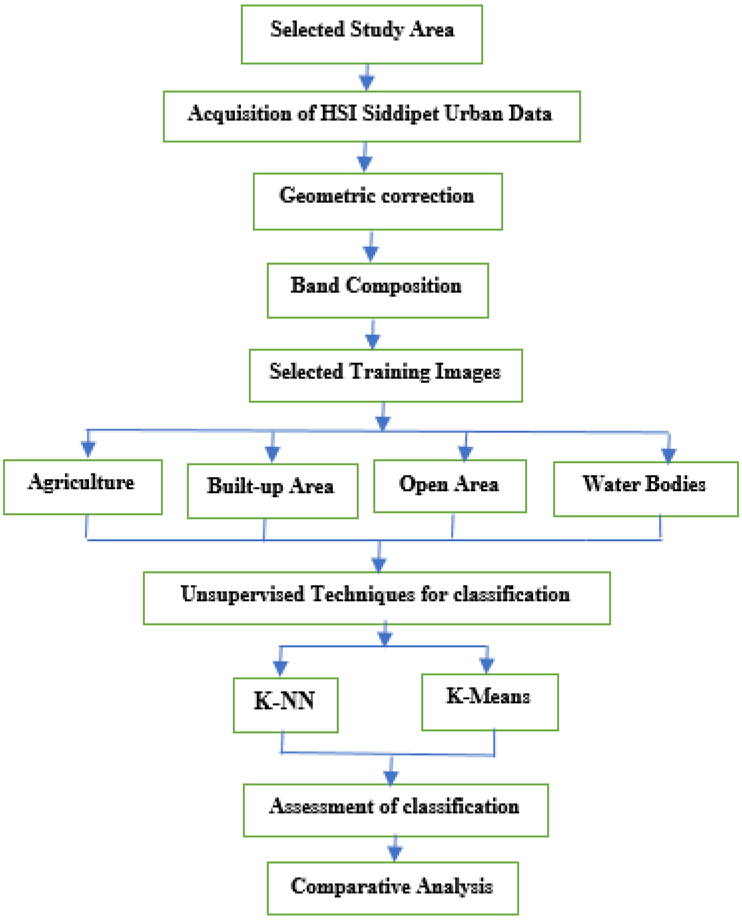
Dimensionality Reduction (DR):

By reducing the number of dimensions in the data, dimensionality reduction paves the way for classifiers to produce thorough models at a cheap computing cost. Therefore, Dimensionality Reduction (DR) has gained increased prominence to enhance pixel categorization accuracy in hyperspectral images. (HSI). The two kinds of dimension reduction are feature selection and feature extraction [3]. Selecting the dataset's feature dimensions, which contribute to machine learning tasks like classification, clustering, etc., is known as feature selection. Different techniques can be used to achieve it, including correlation and univariate analysis. Extraction of Features By choosing and combining existing features, feature extraction reduces the feature space while effectively and thoroughly characterizing the data set without sacrificing information.

Dimensionality reduction strategies are divided into Convex and Non-Convex categories based on the criteria function and convergence procedure. Several well-known dimensionality reduction methods include K-Means, K-NN, PCA, etc. For further information, see the paper "Dimensionality Reduction in Hyperspectral Images Using Python." We are now grappling with the Multi-Class Classification challenge in classification. Hyperspectral images are categorized using various classification methods, including K-Nearest Neighbors and K-Means procedures. (HSI).3.1 Hyperspectral Data of Siddipet Urban Area. The Hughes phenomenon, sometimes known as the curse of dimensionality, is the greatest obstacle to HIS categorization. By choosing the most salient features, high-difficulty feature extraction approaches are employed to lower the dimensionality. According to some statistically defined criteria, pixels with comparable spectral properties are automatically grouped into distinct clusters in unsupervised approaches [8], [9]. Furthermore, the Siddipet Urban Hyperspectral Image data may be trained using unsupervised classification algorithms without the need for prior information.

3.2 Unsupervised Techniques

The technique of extracting information from a hyperspectral image is known as image classification. (HSI). All of the pixels in a digital picture are categorized by HSI into one of many land use classifications or themes. This procedure sometimes includes classifying the land according to its different functional purposes. There are two different forms of classification, depending on how the computer and interpreter interact throughout the classification process. Data extraction from a hyperspectral image (HSI) is called image classification [10]. HSI categorizes all the pixels in the image into one of many land use classifications or themes. The process sometimes includes classifying the land according to its different functional purposes. Depending on how the computer and interpreter interact throughout the classification process, there are two different forms of categorization. The terms "Unsupervised" and "Supervised" Classification Techniques" refer to the two basic categories utilized to produce output that is categorized. Clustering image classification is another name for unsupervised classification. To extract information about land cover, it successfully divides remote sensor image data into multispectral feature space [11].



.

Figure 1: Overall Methodology of HIS of Siddipet Urban Images Classification

After that, the analyst tries to a posteriori assign or alter the spectral classes into interesting thematic information classes. (e.g., forest, agriculture). The unsupervised classification is the most of the processing is carried out autonomously by the sequencer, producing more classified. The user must determine which categories can be combined into a single land use category. In any scenario, more image processing may establish which approach is preferable in a particular circumstance. It is important to remember that maps are never totally accurate and are only crude representations of reality. Using these methods for unsupervised classification of Siddipet Urban hyperspectral images, an effort was undertaken to categorize all the different LULC. The user is not required to provide any information about the characteristics present in the photos while using unsupervised classification algorithms. Making judgments on transforming spectral groups into feature or theme classes requires consideration of the picture and classified output. For making these choices, other resources and local expertise are helpful. This work is more effective and accurate by comparing what is seen in the digital image with what was there when the image was recorded. K-NN and K-means are popular unsupervised strategies, as shown in figure1 above.

3.3 K-NN Techniques

With the rich spectrum information conveyed by its high spectral resolution properties, hyperspectral remote sensing has given remote sensing applications a considerable boost Zhang et al. With hyperspectral remote sensing, the categorization of ground objects has progressed from the coarse classification of distinct species within the same species to the fine variety of different species. It opens up a vast and alluring possibility by combining the microscopic research of plant biochemical components and the macroscopic understanding of vegetation [12]. The spatial information is a useful tool for increasing the classification accuracy of hyperspectral images. The following stages comprise the unique hyperspectral image classification approach proposed in this study based on K closest neighbor (KNN). The matching and averaging of nonlocal neighborhoods are the foundation for the suggested KNN filtering approaches. The suggested technique may utilize the nonlocal principle of actual pictures by employing KNN, giving competitive classification with quick computation without requiring complex segmentation and optimization procedures. The classification results produced by the suggested method are equivalent to earlier techniques proposed hyperspectral image classification methods, according to experiments carried out using actual Siddipet Urban Hyperspectral images.

A new technique for classifying hyperspectral images is presented [22] and based on KNN searching in an undiscovered feature space. The classification map that results is created by giving each pixel the label with the greatest likelihood. This probability optimization-based strategy is comparable to what we've done before. The suggested KNN method's main benefit is fully utilizing the nonlocal spatial information included in the hyperspectral image. The classification accuracy of such a KNN-based nonlocal filtering technique may be effectively increased without addressing a global energy optimization issue. The success of the suggested approach is demonstrated by experiments carried out on two actual urban hyperspectral data sets from Siddipet. Machine learning is an automated process, and decision trees, artificial neural networks, K-means, and K-NN are examples of machine learning techniques. A common nonparametric learning technique is the KNN classification algorithm. Because of its ease of use and excellent categorization accuracy, it has found widespread application [13]. It has long been a contentious topic in statistical pattern recognition, machine learning, and data mining.

The distance between the training sample and the sample that has to be identified determines the number of samples (K). The K value used by the K-nearest neighbor classifier is crucial. A too-small K value will not accurately represent the qualities of the example that has to be categorized. Pixel-based and object-based categorization uses the KNN classifier, which is also included as a data source. The same training samples, verification samples, and feature data sets are used to compare the classification properties of the pictures. It examines the benefits of combining object image analysis and the KNN method with multispectral remote sensing image classification [13]. The K-NN algorithm's fundamental concept is to determine how similar or distant the training sample of a given category is from the sample that must be classified. The kind of sample data to be classified is then decided according to the category to which these neighbours belong. To discover the K neighbours whose distance or similarity is closest to the sample to be classed, find them. The k-nearest neighbor samples closest to item d to be categorized are identified by searching training sets. "Nearest Neighbours" use distance metrics, such as Euclidean distance, two objectives $X_1 = (x_{11}, x_{12}, \dots, x_{1n})$ and $X_{21} = (x_{21}, x_{22}, \dots, x_{2n})$ is

$$dist(X_1, X_2) = \sqrt{\sum_{i=1}^{n}(x_{1i} - x_{2i})^2} \quad (1)$$

The choice of distance measurement is crucial, and different distances may also be utilized for calculating the weight of each class and the weight calculation formula of d belonging to C_j is

$$W(d, C_i) = \sum_{i=1}^{k} sim(d_i, d)\ell(d_i, C_j) \quad (2)$$

where $sim(d_i, d)$, is the similarity between d and the i^{th} nearest neighbor object, d_i. The class to be classed is given to the class with the biggest weight after the weights of the classes are compared. Forest mapping, resource assessments, and area monitoring have all used hyperspectral remote sensing. However, most of the user data may be collected using aerial imaging spectrometers. Despite being experimental, spaceborne imaging spectrometers have been produced successfully [13].

3.4 K-Means Techniques

K-means is one of the most straightforward unsupervised learning techniques to address the well-known clustering issue. The process uses a predetermined number of clusters (k clusters) defined a priori to categorize a given data set. To define k centroids, one for each cluster, is the main notion. These centroids should be strategically positioned since different locations lead to various outcomes. Therefore, it is preferable to situate them as far apart as possible. The following phase connects each location in a given data set to the closest centroid.

The first step is finished when an early grouping is finished and there are no remaining matters. To determine the barycentres of the clusters produced by the prior phase, we must compute k new centroids. The identical data set points must now be bound to the closest new centroid once we have these k new centroids. There has been created a loop. This loop causes the k centroids to gradually shift positions until no further changes are made, as we can see. In other words, centroids have stopped moving.

Finally, this algorithm aims to minimize an objective function, such as a squared error function, as in this equation.

$$J = \sum_{j=1}^{k} . \sum_{i=1}^{n} \left\| x_i^{(j)} - C_j \right\|^2 \qquad (3)$$

Where $\left\| x_i^{(j)} - C_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center C_j, indicates the distance of n data points from their respective cluster centers [14].

The algorithm includes the subsequent steps: 1. Insert K points into the Area that the objects in the cluster represent. These locations represent the initial set of centroids. 2. Assign each item to the group whose centroid is closest to it. 3. Recalculate the K centroids' locations after allocating all the items. 4. Continue doing Steps 2 and 3 until the centroids stop moving. As a result, the objects are divided into groups, from which the metric that has to be minimized may be calculated. Although it can be shown that the process will always come to an end, the k-means algorithm may not always locate the configuration that minimizes the global objective function. The initial, randomly chosen cluster centers considerably impact how sensitive the algorithm is. The k-means method can be used more than once to decrease this impact. The straightforward technique K-means has been used in several problem categories. K-means may categorize data in one of two ways. One technique is partitioning the data into several classes and then independently applying k-means to each class. If there are two classes, k-means will be run twice, once for each data set. Finally, we have a collection of prototypes for every class. We integrate all the prototypes when a new data point is obtained and determine which one is most similar to the new data point. Because the prototypes are made by independently clustering each data class, they are connected to a class. This prototype's class is used to determine the class of the new data point.

Utilizing R examples for each class and first-level K-means clustering independently using the training data from each class. Give each of the K and R prototypes a class designation. A new feature vector should be assigned the same class as the nearest prototype. Agriculture, built-up areas, water bodies, and open space are divided into four categories: green, red, blue, and orange. For each class, K-means is applied using five prototypes. The five prototypes selected for each class are displayed below in filled circles. The prototype closest to any new point among these 15 would be found using the categorization technique. Then, the correct class will be allocated to the new point based on the colour code of that prototype.

The black lines represent the categorization limits resulting from the k-means algorithm. These are specifically the categorization boundaries brought about by the collection of closest neighbor-based prototypes. Based on the closest neighbor rule, a decision boundary between any two prototypes is linear. Every prototype takes up some space in the available Area, and each prototype has an area surrounding it, sometimes known as the Voronoi region, surrounded by hyperplanes. The categorization border between the two classes comprises linked segments of straight lines since each class has more than one prototype [15]. Putting all the data points together and running k-means once results in the second level of categorization by k-means. Because we run k-means on the class blended data, there is no assurance that points belong to the same group or class. We count the amount of data points in each class assigned to this prototype to prototype and link to a class. The prototype is connected to the dominant class with the most data points. The process then continues when a new data point is classified. To cluster all of the training data, use k-means. Classify a new feature x into the class of the closest prototype and assign the classes to the class with the greatest count.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, No. 1, April : 2023

4. Results and Discussion

Figure 1 displays the hyperspectral (HS) image or data from Siddipet Urban data in Telangana State, India, to evaluate the performance of K-NN and K-Means methods or techniques for image classification. GPS boundaries it or Geo-Coordinates are "18.101904 and 78.852074 latitude and longitude, respectively, and its Degree, Minutes and Seconds (DMS) coordinate are 18' 6" N and 78' 51' 7" E encompassing an Area of 12 Km. Agriculture, Built-up Areas, Water Bodies, and Open Areas are the main topic aspects in this study according to the classification of HSI. HSI is purchased with funds from the IRS. For the intermediate level of the Siddipet Urban Hyperspectral images from 2016 to 2021, they pre-process the raw data using QIGS and ERDAS, as shown in figure 2. A study of confusion matrices containing true positive and negative and false positive and negative was performed using MATLAB after the classification, and the model predicted total accuracy. Therefore, compare the recall, accuracy, sensitivity, and specificity of the Siddipet Urban HSI categorization of the images from 2016 to 2021 as presented in tables 1-36 using both K-NN and K-Means. Figure 4 illustrates how the K-Means classifier was fed the identical image and training sets. A confusion matrix and overall accuracy were produced following the categorization, as illustrated in tables 1 through 36.

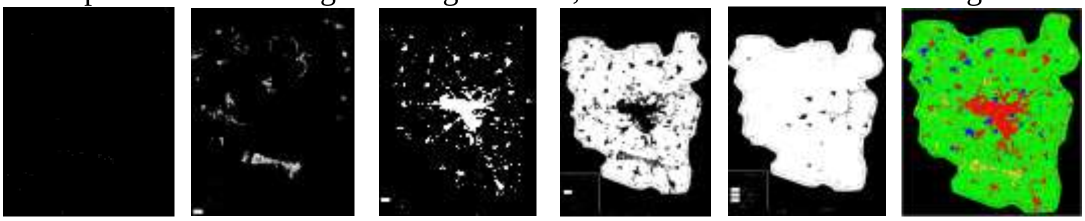
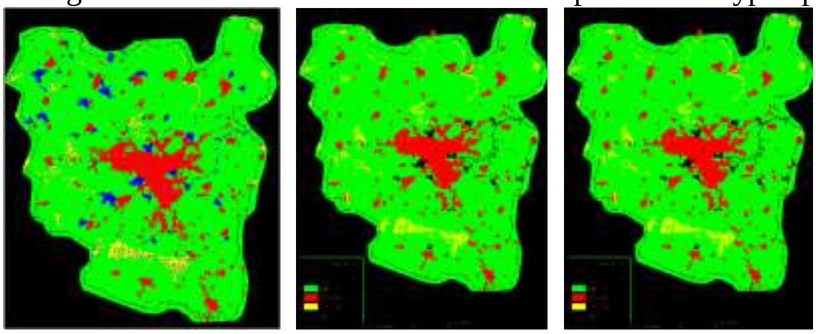
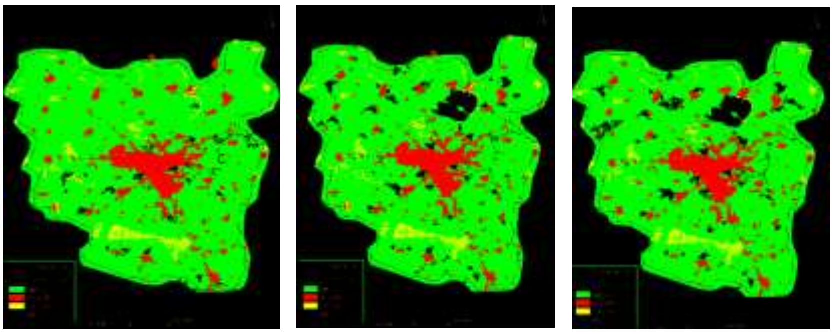


Figure 2: Raw image of the intermediate level of the Siddipet Urban Hyperspectral Images



2016 2017 2018



2019 2020 2021

Figure 3: K-NN classification of Siddipet Urban Hyperspectral Images

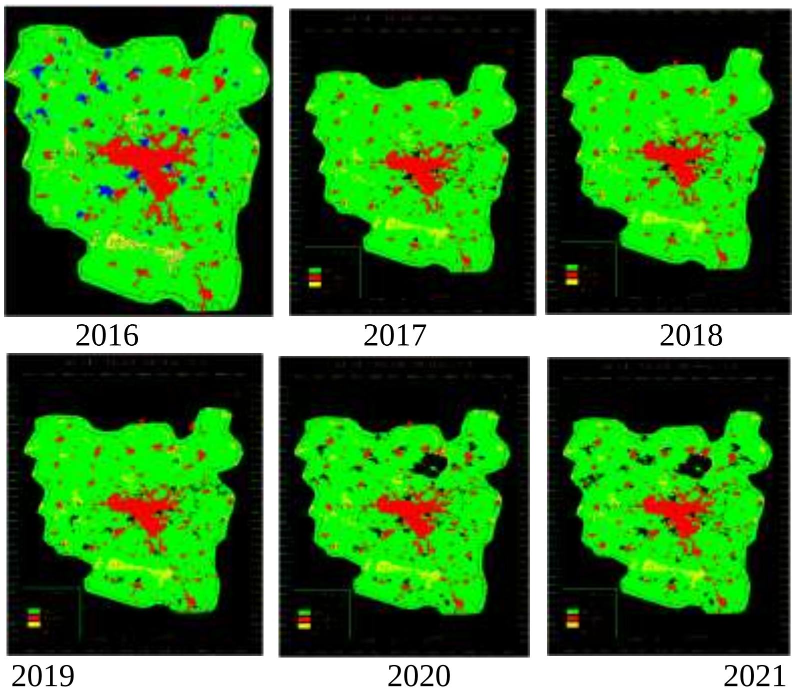


Figure 4: K-Means classification of Siddipet Urban Hyperspectral Images

Agriculture Built-up Area Water Bodies Open Area

The confusion matrix displays the results in a square matrix with the correct diagonal and incorrect classifications in the remaining positions.

Overall accuracy (OAA)

The diagonal elements tally the number of pixels classified correctly in each class. An overall measure of classification accuracy is defined as given below.

$$OverallAccuracy = \frac{Total\ number\ of\ correct\ \ classification}{Total\ \ number\ \ of\ \ classification}$$

After classifying Siddipet Urban Hyperspectral Images by classification using KNN and K-Means methods, the following confusion matrices are generated and given in tables 1-36.

1. Table 1-18: Model predictive correctly or Confusion matrix of Siddipet Urban Hyper Spectral Image (SUHI) classification with K-NN algorithm from 2016-2021.
2. Table 19-36: Model predictive correctly or Confusion matrix of SUHI classification with K-Means algorithm for from 2016-2021

Tables 1 to 18 have Confusion Matrixes or Model Predicted Analysis of K-NN for 2016 -2021.

Table :1

KNN-2016					
	Actual Classes				
Predicted Classes	Classes	**Agriculture**	**Built-up Area**	**Water Bodies**	**Open Space**
	Agriculture	**10198.0**	448.0	1874.0	13852.0
	Built-up Area	548.0	**8484.0**	1867.0	13792.0
	Water Bodies	504.0	440.0	**35137.0**	13709.0
	Open Space	547.0	449.0	1834.0	**262486**

Table :2

KNN-2016				
	Actual Classes			
	Agriculture	**Built-up Area**	**Water Bodies**	**Open Space**
TP	10198.00	8484.00	35137.0	262486.0
FP	16174.00	16207.0	14653.0	2830.00

FN	1599.00	1337.00	5575.00	41353.0
TN	338198.00	340141.0	310804.0	59500.0

Table :3

KNN-2016				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.86	0.86	0.86	0.86
Precision	0.39	0.34	0.71	0.99
Sensitivity	0.86	0.86	0.86	0.86
Specificity	0.95	0.95	0.95	0.95

Table:4

KNN-2017					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1030.0**	555.0	1800.0	14118.0
	Built-up Area	71.0	**11007.0**	1818.0	14150.0
	Water Bodies	47.0	538.0	**34900.0**	13968.0
	Open Space	54.0	606.0	1716.0	**274586.0**

Table:5

KNN-2017				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1030.00	11007.00	34900.00	274586.00
FP	16473.00	16039.00	14553.00	2376.00
FN	172.00	1699.00	5334.00	42236.00
TN	353289.00	342219.00	316177.00	51766.00

Table:6

KNN-2017				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.85	0.86	0.86	0.866
Precision	0.06	0.41	0.71	0.99
Sensitivity	0.86	0.87	0.87	0.87
Specificity	0.96	0.96	0.96	0.96

Table:7

KNN-2018					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**275067.0**	519.0	18330.	165.0
	Built-up Area	13964.0	**11204.0**	1728.0	159.0
	Water Bodies	13507.0	567.0	**35441.0**	172.0
	Open Space	13634.0	522.0	1765.0	**3213.0**

Table :8

KNN-2018				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	275067.00	11204.00	35441.00	3213.00

FP	2517.00	15851.00	14246.00	15921.00
FN	41105.00	1608.00	5326.00	496.00
TN	54771.00	344797.00	318447.00	353830.00

Table:9

KNN-2018				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.86	0.87	0.86	0.86
Precision	0.99	0.41	0.71	0.17
Sensitivity	0.87	0.87	0.87	0.87
Specificity	0.96	0.96	0.96	0.96

Table:10

KNN-2019					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1057.0**	538.0	1700.0	13349.0
	Built-up Area	47.0	**11067.0**	1653.0	13436.0
	Water Bodies	54.0	507.0	**35367.0**	13300.0
	Open Space	53.0	572.0	1773.0	**274936.0**

Table:11

KNN-2019				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1057.00	11067.00	35367.00	274936.00
FP	15587.00	15136.00	13861.00	2398.00
FN	154.00	1617.00	5126.00	40085.00
TN	352611.00	341589.00	315055.00	51990.00

Table:12

KNN-2019				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.87	0.87	0.87	0.87
Precision	0.60	0.42	0.72	0.99
Sensitivity	0.87	0.87	0.87	0.87
Specificity	0.96	0.96	0.96	0.96

Table:13

KNN-2020					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1068.0**	539.0	1649.0	12354.0
	Built-up Area	45.0	**11225.0**	1669.0	12487.0
	Water Bodies	53.0	535.0	**35247.0**	12463.0
	Open Space	53.0	520.0	1654.0	**265710.0**

Table :14

KNN-2020

	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1068.0	11225.00	35247.00	265710.00
FP	14542.00	14201.00	13051.00	2227.00
FN	151.00	1594.00	4972.00	37304.00
TN	341510.00	330251.00	304001.00	52030.00

Table :15

KNN-2020				
	Actual Classes			
	Agriculture	Built -up Area	Water Bodies	Open Space
Recall	0.87	0.87	0.87	0.87
Precision	0.07	0.44	0.73	0.99
Sensitivity	0.88	0.88	0.88	0.88
Specificity	0.96	0.96	0.96	0.96

Table:16

KNN-2021					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1081.00**	447.0	1616.0	116010.00
	Built-up Area	49.00	**10644.00**	1600.00	11787.00
	Water Bodies	54.00	460.00	**36691.00**	11828.00
	Open Space	34.00	484.00	1638.00	**264892.00**

Table:17

KNN-2021				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1081.00	10644.00	36691.00	264892.00
FP	13664.00	13436.00	12342.00	2156.00
FN	137.00	1391.00	4854.00	35216.00
TN	340024.00	329435.00	301019.00	52642.00

Table:18

KNN-2021				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.88	0.88	0.88	0.88
Precision	0.07	0.44	0.75	0.99
Sensitivity	0.89	0.88	0.88	0.88
Specificity	0.96	0.96	0.96	0.96

Tables 1 to 18 have Confusion Matrixes or Model Predicted Analyses of K-Means for 2016 -2021

Table:19

K-Mens 2016					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**10345.0**	384.0	1749.0	12583.0

	Built-up Area	466.00	**8603.00**	1679.00	12446.00
	Water Bodies	480.0	483.00	**35657.00**	12556.0
	Open Space	506.0	404.0	1627.0	**266254.0**

Table :20

K-Means 2016				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	10345.00	8603.00	35657.00	266254.00
FP	14716.00	14591.00	13466.00	2537.00
FN	1452.00	1218.00	5055.00	37585.00
TN	339656.00	341757.00	311991.00	59793.00

Table:21

K-Means 2016				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.87	0.87	0.87	0.87
Precision	0.41	0.37	0.73	0.99
Sensitivity	0.88	0.88	0.88	0.88
Specificity	0.96	0.96	0.96	0.96

Table:22

K-Means 2017					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1073.00**	489.00	1477.0	11756.0
	Built-up Area	35.0	**11286.0**	1502.0	11725.0
	Water Bodies	41.0	462.0	**35799.0**	11656.0
	Open Space	53.0	469.0	1456.0	**281685.0**

Table:23

K-Means 2017				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1073.0	11286.0	35799.00	281685.0
FP	13722.0	13262.0	12159.0	1978.0
FN	129.0	1420.0	4435.0	35137.0
TN	356040.0	344996.0	318571.0	52164.00

Table:24

K-Means 2017				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.89	0.88	0.88	0.88
Precision	0.07	0.46	0.75	0.99
Sensitivity	0.89	0.89	0.89	0.89
Specificity	0.96	0.96	0.96	0.96

Table:25

K-Means 2018

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, No. 1, April : 2023

	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**279231.0**	519.0	1566.0	146.0
	Built-up Area	12412.0	**11312.0**	1548.0	144.0
	Water Bodies	12385.0	499.0	**36077.0**	154.0
	Open Space	12144.0	482.0	1576.0	**3265.0**

Table:26

K-Means2018				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	279231.0	11312.0	36077.00	3265.00
FP	2231.00	14104.00	13038.00	14202.0
FN	36941.0	1500.0	4690.0	444.0
TN	55057.0	346544.0	319655.0	355549.0

Table:27

K-Means-2018				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.88	0.88	0.88	0.88
Precision	0.99	.45	0.73	0.19
Sensitivity	0.88	0.88	0.88	0.88
Specificity	0.96	0.96	0.96	0.96

Table:28

K-Means 2019					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1053.0**	561.0	1676.0	13120.0
	Built-up Area	59.0	**11110.0**	1697.0	12852.0
	Water Bodies	58.0	483.0	**35454.0**	12966.0
	Open Space	41.0	530.0	1666.0	**276083.0**

Table:29

K-Means 2019				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1053.0	11110.0	35454.0	276083.0
FP	15357.0	14608.0	13507.0	2237.0
FN	158.0	1574.0	5039.0	38938.0
TN	352841.0	342117.0	315409.0	52151.0

Table:30

K-Means 2019				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.86	0.87	0.87	0.87
Precision	0.06	0.43	0.72	0.99
Sensitivity	0.87	0.88	0.88	0.88
Specificity	0.96	0.96	0.96	0.96

Table:31

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, No. 1, April : 2023

K-Means 2020					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1087.0**	524.0	1602.0	11888.0
	Built-up Area	50.0	**11313.0**	1511.0	11925.0
	Water Bodies	36.0	521.0	**35541.0**	11763.0
	Open Space	46.0	461.0	1565.0	**267438.0**

Table:32

K-Means 2020				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1087.0	11313.0	35541.0	267438.0
FP	14014.0	13486.0	12320.0	2072.0
FN	132.00	1506.0	4678.0	35576.00
TN	342038.00	330966.0	304732.00	52185.0

Table:33

K-Means 2020				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.89	0.88	0.88	0.88
Precision	0.07	0.46	0.74	0.99
Sensitivity	0.89	0.88	0.88	0.88
Specificity	0.96	0.96	0.96	0.96

Table:34

K-Means2021					
	Actual Classes				
Predicted Classes	Classes	Agriculture	Built-up Area	Water Bodies	Open Space
	Agriculture	**1084.0**	424.0	1536.0	11101.0
	Built-up Area	38.0	**10715.0**	1509.0	11154.0
	Water Bodies	51.0	460.0	**36944.0**	10952.0
	Open Space	45.0	436.0	1556.0	**266901.0**

Table:35

K-Means 2021				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
TP	1084.0	10715.0	36944.0	266901.0
FP	13061.0	12701.0	11463.0	2037.0
FN	134.0	1320.0	4601.0	33207.0
TN	340627.0	330170.0	301898.0	52761.0

Table:36

K-Means 2021				
	Actual Classes			
	Agriculture	Built-up Area	Water Bodies	Open Space
Recall	0.88	0.89	0.88	0.88
Precision	0.08	0.46	0.76	0.99
Sensitivity	0.89	0.89	0.89	0.89
Specificity	0.96	0.96	0.96	0.96

Comparing classifications accuracies obtained using K-NN and K-Means classification provides statistics on overall accuracy, Confusion Matrix, or Model predicted correctly. Comparative analysis of four classifications reveals that the K-Means classification performs better than K-NN with the predicted analysis. True Positive and Negative, False Positive and Negative, and also Recall, Precision, Sensitivity, and specificity of K-NN and K-Means for SUH Images from 2016 to 2021 of 87,87,87,88 and 88 percent of K-NN and 2016 to 2021of 88,89, 88,88,88 and 89 percent of K-Means respectively, recall, precision, sensitivity and specificity respectively in tables 1-36. The K-Means show better results for five classes except for open spaces.

5. Conclusion

K-Means and K-NN were used to classify the HSI of Siddipet Urban data, and a comparison revealed that K-Means were generally superior to K-NN. The K-Means, however, may have produced a better-categorized result. The problem with K-Means can be that it returns an index that corresponds to a cluster. Give the clustered index for each pixel in the image. The output of k-means is the cluster center. Future research will test the effect of kernel functions. In the future, supervised methods or algorithms will also be used to test classification.

6. Acknowledgement

The authors would like to thank all reviewers and sincerely thank the institute staff for the technical support of the Department of Geography, University College of Science, Osmania University Hyderabad, and for remarkable suggestions during research work. The author would like to acknowledge ICSSR, a fellowship provided by Govt of India, New Delhi, India.

Reference

1. Li, M., Zang, S., Zhang, B., Li, S., & Wu, C. (2014). A review of remote sensing image classification techniques: The role of spatio-contextual information. *European Journal of Remote Sensing*, *47*(1), 389-411.
2. Mehmood, M., Shahzad, A., Zafar, B., Shabbir, A., & Ali, N. (2022). Remote sensing image classification: A comprehensive review and applications. *Mathematical Problems in Engineering*, *2022*, 1-24.
3. Mazhar, S., Sun, G., Bilal, A., Li, Y., Farhan, M., & Awan, H. H. (2022). Digital and Geographical Feature Detection by Machine Learning Techniques Using Google Earth Engine for CPEC Traffic Management. *Wireless Communications & Mobile Computing (Online)*, *2022*.
4. Wang, J., & Biljecki, F. (2022). Unsupervised machine learning in urban studies: A systematic review of applications. *Cities*, *129*, 103925.
5. Fiorini, L., Falasca, F., Marucci, A., & Saganeiti, L. (2022). Discretization of the Urban and Non-Urban Shape: Unsupervised Machine Learning Techniques for Territorial Planning. *Applied Sciences*, *12*(20), 10439.
6. Sataer, G., Sultan, M., Emil, M. K., Yellich, J. A., Palaseanu-Lovejoy, M., Becker, R., ... & Abdelmohsen, K. (2022). Remote Sensing Application for Landslide Detection, Monitoring along Eastern Lake Michigan (Miami Park, MI). *Remote Sensing*, *14*(14), 3474.
7. Thary, A. H. (2014). Satellite Image Classification Using K-Means and SVD Techniques. *A Thesis*.
8. Lu, W., Qi, J., & Feng, H. (2022). Urban functional zone classification based on self-supervised learning: A case study in Beijing, China. *Frontiers in Environmental Science*, *10*, 2352.
9. Fatima, W., & Ejaz, I. (2021). Hyperspectral Image Classification Using Unsupervised Learning Algorithms.
10. El_Rahman, S. A. (2016). Hyperspectral image classification using unsupervised algorithms. *International Journal of Advanced Computer Science and Applications*, *7*(4).
11. Delima, A. J. P. (2020). An enhanced K-nearest neighbor predictive model through metaheuristic optimization. *International Journal of Engineering and Technology Innovation*, *10*(4), 280.

12. Usman, B. (2013). Satellite imagery land cover classification using k-means clustering algorithm computer vision for environmental information extraction. *Elixir International Journal of Computer Science and Engineering*, *63*, 18671-18675.
13. Lv, Z., Hu, Y., Zhong, H., Wu, J., Li, B., & Zhao, H. (2010). Parallel k-means clustering of remote sensing images based on mapreduce. In *Web Information Systems and Mining: International Conference, WISM 2010, Sanya, China, October 23-24, 2010. Proceedings* (pp. 162-170). Springer Berlin Heidelberg.
14. Luo, H. (2018). Shorten spatial-spectral RNN with parallel-GRU for hyperspectral image classification. *arXiv preprint arXiv:1810.12563*.
15. Cai, J., & Chen, Y. (2022). A novel unsupervised deep learning method for the generalization of urban form. *Geo-spatial Information Science*, *25*(4), 568-587.