

Choose Best SVM Kernels for Hyperspectral Satellite Image

Jalal Ibrahim Faraj

Ph.D. In Physics/Department of Physics/College of Science/Baghdad University

Address: Al-Karama Preparatory Boys School/ Kut/Wasit/Iraq

E-mail: jalalibrahimskb355@gmail.com

Bassima M. Mashkour

MSc. In Physics/Department of Physics/College of Science/Wasit University

Address: Al-Zahra Secondary School for Girls/ Kut/Wasit/Iraq

E-mail: etts341@gmail.com

Article Info

Page Number: 753-767

Publication Issue:

Vol. 71 No. 4 (2022)

Article History

Article Received: 25 March 2022

Revised: 30 April 2022

Accepted: 15 June 2022

Publication: 19 August 2022

Abstract

The use of a remote sensing technology has gained more attention due to an increasing need to collect data for environmental changes. Satellite image classification is a relatively recent type of remote sensing that uses satellite imagery to indicate several virtual environment characteristics. The support vector machine (SVM) method with various kernel functions (i.e., RBF, a sigmoid and a polynomial) were employed to recognize and classify the hyperspectral satellite image. The SVM with RBF kernel function achieves the best classification accuracy of 87.3% based on the overall Classification.

Keywords: Support vector machine(SVM), Hyperspectral image, Preprocessing of the image, Kernels of SVM.

Introduction

Remotely sensed images have a wide range of uses. A significant application of remotely sensed data is to build the classification map of the identifiable or classes of the land cover kinds in the scene or meaningful features [1]. One of the data mining techniques used to classify the object into the predefined group is Classification. They are the decision-making tasks most frequently used by human activity. A classification problem occurs when an object must be assigned to a class or predefined group depending on the number of observed attributes connected to that object. The Classification has a significant role in satellite image classification and remote sensing [2]. Image classification is a composite process that various factors can influence. A vast number of classification methods can be established in the literature; they have been commonly characterized as either unsupervised or supervised techniques. The supervised methods often require prior knowledge in choosing the exact region of interest, inadequately identifying "ROI" or the number of existing areas. It usually yields a preliminary classification result, while the unsupervised techniques must select the number of regions in the processed image. In this work, we deal with the support vector

machine techniques since several essential advantages of using these techniques are interpretation, direct geometric, elegant mathematical tractability, and high accuracy.

Additionally, the support vector machine does not involve a huge number of training data sets to avert the over suitable issues. The kernel function is essential in developing a classification model using the support vector machine technique. It supports mapping the dataset to an upper dimensional space to interpret the classification model better. Despite that, various kernel functions can be utilized, such as radial basis function, linear, polynomial, and sigmoid. This work will compare many kernel functions used at the support vector machine technique to achieve the most excellent result performance in the Classification hyperspectral satellite image [3].

Study Area

The study area is a Baghdad city. It is capital of the Republic of Iraq and is the nation's center of commercial and administration activities. With a population of 7,837,963 (Iraq Census 2015), it is the second most populous Arab city after Cairo, Egypt. It has over seventy neighborhoods and nine administrative districts. These districts initially represent the main sectors of Baghdad city. Baghdad is situated in the center of Iraq on a both sides of the a Tigris River, where a geographic coordinates: a Longitude ($44^{\circ}15'55''$) to ($44^{\circ}17'38''$) E, a Latitude ($33^{\circ}25'46''$) to ($33^{\circ}24'21''$) N. Baghdad is the a main and ahighest populated a city in Iraq. a Baghdad is located in the area of an elevatioin a between (31-39 m) over the sea level. So, nor the natural boundaries exist that bounds the city's the aerial an extension. Tigris River a passé in the town, separating into two a parts; Rusafa (Eastern part) and Karkh (Western region). The area is bounded from the east by the Diyala River, which joins the Tigris River southeast of Baghdad. The Tigris River recharges the Army Canal, which is 24 km long, empties into the Diyala River in the city's southern region. [4]. Figure (1) shows an area of study, "Baghdad city."

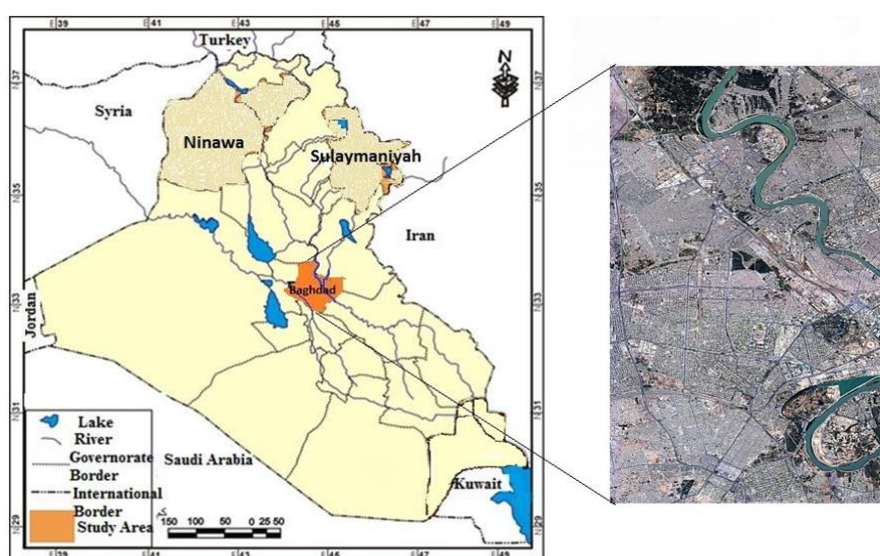


Figure 1. Area of Study "Baghdad City."

Materials and Methods

Hyperspectral Satellites Imaging Systems

"hyper" in the section "hyperspectral" refers to "too many" in "over", indicates massive the a number of the a wavelength bands. The (HSI) Hyperspectral an imaging provided possibility of the a precise and the an exhaustive information abstraction than the a potential with other type of the a remote sensing data. An increased capability that an improves the chance of noticing the concerning the materials and the supplies additional information needed for classifying and recognizing these materials. Some boundaries in the hyperspectral images are the image distortion result for spherical of Earth, give the in the characters such the distances, scale and directions. These a distortions of the image may be shadows, such as brightness of the light of a particular area and cloud covering the specific area to be the studied [5]. The Hyperspectral sensors a record reflected electromagnetized energy for a Earth's surface the across an electromagnetic the spectrum the extending for visibled wavelength a region pass mid-infrared and near-infrared region ($0.355\mu\text{m}$ to $2.5\mu\text{m}$) in the ttens to hundredse of the narrow (10nm) the contiguous abands [6]. Such the narrow bandwidths a result in the almost continuous and the detailed spectral a response for each the pixel, provide accurate and the precise information about the constituents. It is the advantage over the multispectral imaging. High spectral resolution for the hyperspectral sensor the allows us to the capture minor a deviations in spectral response for materials, thus the aiding in their the identification. Figure (2) depicts a typical Hyperspectral data cube and the spectrum of a single pixel.

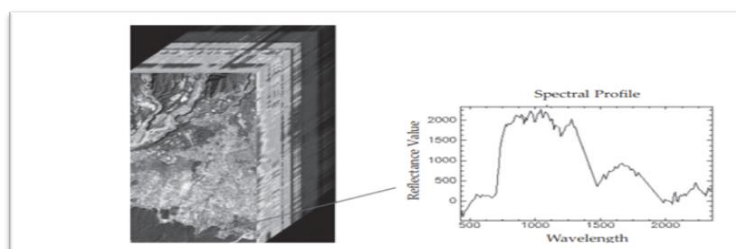


Figure 2. Hyperion Image Cube and Reflectance Spectrum.

EO-1 Hyperion Sensor”

The first space-borne Hyperspectral sensor for Earth Observation studies is the Hyperion an instrument on NASA's Earth an Observation-1 (EO1) spacecraft, which was launched on the November 21, 2000 as a part of a NASA's New the Millennium Program. It travels 705 kilometers above the Earth in a synchronous solar orbit. Hyperion is a high-resolution spectral (push-broom scanner). It includes 242 bands with a spectral series of 0.355 to 2.5 m and a sample interval of (10nm). Spatial resolution for a swath awidth of (7.7 km) is 30m and covers an area of ($7.7 \times 100 \text{ km}^2$)per an image with high a radiometric resolution (16 bit). The Hyperion sensor contains two spectrometers that operate in distinct spectral regions. The first operates in the Visible and Near Infrared zone (0.355 m to 1m) and has 70 a bands, while the second an operates in the Shortwave an Infraredthe region (0.9 to 2.5m) and has 172 bars. The overlap zone (0.9 to 1m) between the two spectrometers enables for cross-

calibration between the two spectrometers. Figure (3) depicts the Hyperion Sensor onboard EO-1.

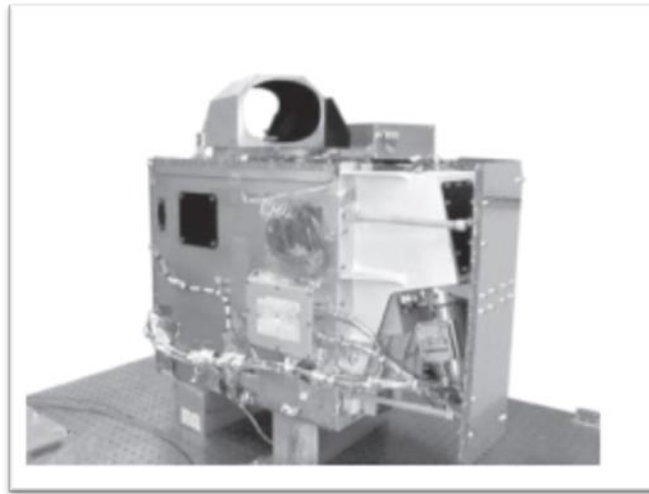


Figure 3. Hyperion Sensor[7]

The data in the form of cubes is written as band-interleaved-line (BIL) files and saved as 16-bit signed integer radiance values in Hierarchical Data Format (HDF). The scaling factor for the SWIR bands is 80, whereas the scaling factor for the VNIR bands is 40. The actual radiance values range from 0 to about 32,767. The many forms in which USGS makes Hyperion sensor data available to users are given below[7]:

- Level 1R (L1R) The Level 1 Radiometric theproduct is only radiometrically corrected and not geometrically resampled. The L1R product is available only in HDF format. The data product consists of the HDF data file, (.L1R), a metadata file (.MET), a header file (HDR), and an auxiliary file (.AUX).
- Level 1Gs (L1Gst) The Level 1 Gs the product is radiometrically corrected, geometrically resampled, and registered to a geographic map projection. The image is terrain corrected, an orthorectified using a digital elevation models (DEM) for correcting parallax the error due to topographic relief. Each an image a band in the L1G an product is in the separate file. The L1Gst product is available in HDF v 4.1 and GeoTIFF.
- ✓ L1Gst (HDF), the Hyperion product includes a metadata file (_MTL.L1T), an HDF header file (_HDF.L1T), a Federal Geographic Data Committee (FGDC) metadata file (.fgdc), and multiple image bands (_B###.L1T).
- ✓ L1Gst (aGeoTIFF) aGeoTIFF defines an set of the public domain TIFF a tags that describe all the cartographic and geodetic information the associated with a geographic The TIFF an imagery. This Hyperionproduct includes a metadata file (_MTL_L1T.TIF), an FGDC metadata file (.fgdc), and multiple image bands (_B###_L1T.TIF).

USGS aproducts are apackaged in aHierarchical the Data aFormat (HDF), adatasets being the image data, aspectral center the wavelengths, a spectral bandwidth, the again coefficients, and the flag a mask. file naming convention utilizes an entity I.D. with an acquisition target. Table 1 describes the naming convention for Hyperion datasets; The scene parameters of the Hyperion image of the Baghdad area to be used in this thesis are listed in Table 2. The parameters of EO-1 satellite-Hyperion data are shown in Table 3.

Table 1. Data File Name Description.

EO1SPPPRRRRYYYDDDDXXXXML_GGG_VV	
EO1	Earth Observing 1 mission
S	Sensor, A=ALI, H=Hyperion
PPP	Target WRS Path of the product
RRR	Target WRS row of the product
YYY	Acquisition year of the image
DDD	Acquisition Julian day of the year
XXX	Hyperion,ALI,Atmospheric Corrector (AC),(1=sensor on,0=sensor off)
M	pointing mode, P=pointed within path/row, K=pointed outside path/row, N=Nadir
L	The scene identifier maybe 0-9 or an upper- or lower-case alpha character.
GGG	Ground/Receiving Station

Table 2. : Scene Features of Hyperion Image of Baghdad Area

Data Attribute	Attribute Value
Entity ID	EO1H1480382008133110PW_PF1_01
Scene Start Time	2005:097:07:24:10.122
Acquisition Date	2005/04/07
Scene Stop Time	2005:097:07:24:25.122
Date Entered	2005/04/18
Center Latitude	33°10'51.07"N
Sun Azimuth	132.812259
Center Longitude	44°19'20.19"E
Sun Elevation	54.860678
Satellite Inclination	98.21

Table 3: EO-1 Satellite-Hyperion Data Characteristics

Sensor altitude	705 Km
Spatial resolution	30m
Radiometric resolution	16 Bits
Swath	7.2 Km
Imaging Technology	Pushbroom

(Samples) No. of rows	1041
(lines) No. columns	3531
VNIR range	0.35-1.35 μm
SWIR	1.40-2.48 μm
NO. of Bands	242
Scaling factor (VNIR)	40
Scaling factor (SWIR)	80

Support Vector Machine (SVM) Classifiers”:

Vladimir Vapnik created the Support Vector Machine (SVM) based on the theory of structural risk minimization, to build decision functions in the input space[8,9]. Additionally, SVM involves creating one or more hyperplanes to divide the various classes. However, an ideal hyperplane must be identified. The Vapnik and Cortes [10] described an optimal hyperplane the linear decision function with maximal margin between the vectors of two classes. An ideal hyperplane is considered if examples are separated without error and if the distance between the nearest model and a hyperplane is maximal. A hyperplane can be written as:

$$W^T x + b = 0, x \in R^d \quad (1)$$

The samples that are near to the hyperplane boundaries are referred to as "support vectors". They are used to decide which hyperplane should be selected since this set of vectors is separated by the optimal hyperplane. When the data are separable, SVM is basically used as a linear decision function. We assume that the data are linearly non-separable in this paper. As a result, we should propose a nonlinear function with $\phi(\xi_i)$ a nonnegative variable which can represent the data in a high-dimensional feature space where it is linearly separable. A vector W which minimizes the functional can be used to the ideal hyperplane in a nonlinear space:

$$\phi(W, \xi) = \frac{1}{2} \|W\|^2 + C \sum_{i=1}^n \xi_i \quad (2)$$

"C: is a pre-specified value used to control the amount of regularization, and ξ is a slack variable. Furthermore, it should be noted that in this study, a multiclass recognition problem is decoupled into a two-class problem"[11]. So, we applied the one-vs-one approach suggested by Knerr et al[12]. "This method consists of $k(k-1)/2$ constructed classifiers, each of which one trains samples from two classes. The majority voting strategy was applied for the recognition decision-making".

Kernel Functions

One of the main SVM techniques is kernel functions when the samples are linearly non-separable. Thus, the kernel tricks extend the class of decision functions to the nonlinear case by mapping the samples from the input space X into a high-dimensional feature R without ever having to compute the mapping explicitly, in the hope that the samples will gain meaningful linear structure in R .

Furthermore, a kernel function may be viewed as the a measure of a similarity between samples(x_i and x_j) [13], allowing SVM the a classifiers to accomplish a separations even when the borders are quite complex.

Polynomial Kernel

A non-stationary a kernel is a Polynomial kernel. When all the training a samples are normalized, it is well suited for problems. The parameters to be standardized are the gamma slopes The polynomial a degree d and term rare constants (hence $r=0$, $d=3$) [13].

$$k(x_i, x_j) = (\sigma X_i^T X_j + r)^d, \sigma > 0 \quad (3)$$

RBF kernel

RBF (a Gaussian) a kernels are an family of the kernels where an distance measure is smoothed by a radial the function (exponential a function) [14]. Unlike a linear kernel, this kernel can the handle the a case when the a relation between the attributes is nonlinear and class labels. This is accomplished by nonlinearly mapping samples into a higher-dimensional space. The linear a kernel is a the subset of RBF [15]. Because the linear a kernel contains (C), the penalty a parameter, it performs similarly to (C, Gamma) some RBF kernel parameters.

$$k(x_{(i)}, x_{(j)}) = \exp\left[-\sigma \|X_i - X_j\|^2\right], \sigma > 0 \quad (4)$$

It is critical to locate the changeable parameter because it has a substantial impact on kernel performance and should be properly tweaked. If the exponential is overestimated, it will act almost a linearly, and higher dimensional a projection will the lose some of it's a nonlinear power. In contrast, if function is the underestimated, it will the lack regularization and decision boundaries will particularly susceptible to the noise in training data. As a result, value of the width a parameter determines (SVM) behavior.

“Sigmoid kernel”

To the satisfy a Mercer's theorem, a kernel must be a positive and the a definite. Despite its popularity, a Sigmoid kernel is not a positive or a semi-definite for parameter values. As a result, the r parameters must be carefully chosen. Otherwise, the results could be considerably

off, to the point where the SVM performs worse than random.

$$k(x_i, x_j) = \tanh(\sigma X_i^T X_j + r)^d \quad (5)$$

Consider (r) to be a shifting and scaling parameter of the input samples that governs the mapping threshold (therefore $r = 0$). The sigmoid kernel is generally inferior than linear kernels and RBF [15].

Methodology

The data set is the hyperspectral data taken by the hyperspectral image (Hyperion) sensor with 242 spectral abands ranging from 0.355 to 2.5 μm , as shown in figure (4). It was collected for Baghdad city in 2005. In processing the Hyperion image, after removing the bad spectral bands, calibration, and the atmospheric and geometric correction, about 149 bands remained. This image has a spatial resolution of about 30 m. The original image was cropped into a sub-image of size 500×730 pixels. The Data includes five classes: {tree, water, road, vacant land, building}; we will be performing support vector machine (SVM) kernels on the original image (149 bands).

Results and discussions

Hyperspectral Data Pre-processing

The investigation of Hyperion orthoimage raw data indicated that 44 of the 242 bands have zero values assigned during preprocessing.

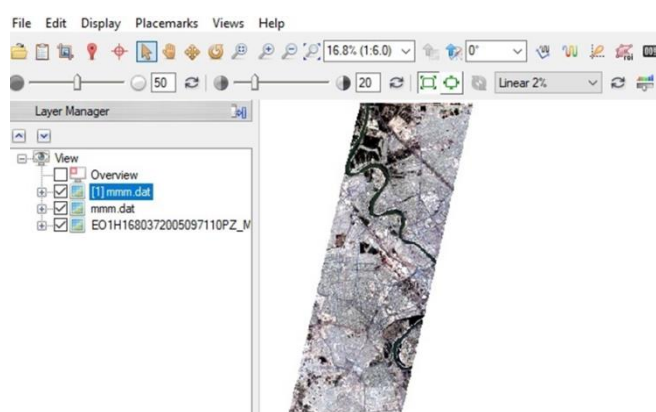


Figure 4: Area of Study (500×730), Hyperion Satellite Image (2005)

Bad Band Removal

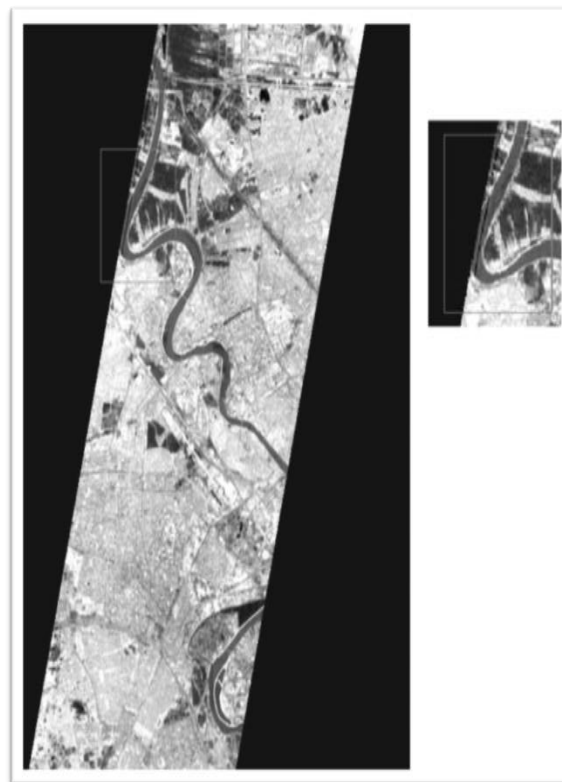
The Zero band avalues were the abands from (1 to 7), abands from (58 to 76), and abands from (225 to 242,) as shown in the table (4). The aresultantly, 198 abands were the established to an useful for the analysis. A Bands 99, 116,119-129,165-182,184-187,190-191,203,214-224 were aremoved having low aSignal to noise a value .The total of 149 a calibrated bands are the available for a further processing. Thus, we can the apply FLAASH an atmospheric correction on remaining 149 a calibrated Hyperion the imagery since an imagery is Hyperion or the image, so the processing is required. processing is the done a before main data the analysis and an extraction of the informatiion. ENVI 5.3 software was used in this thesis to perform most satellite image preprocessing stages.

Table 4: L1R Product ,Unused aBands of the Hyperion Sensor

aBands	Description
1-7	without information
58-76	without information
99 and 116	noise
119-129	noise
165-182	noise
184-187	noise
190-191 and 203	noise
214-224	noise
225-242	without information

Along Track De-striping

The result of abnormal pixels should be corrected and accounted for before further processing. Fig(5) shows the abnormal pixels containing a spatial subset for the Hyperion image. Still, Fig(6) corrected image by using Hyperion tools.

**Figure 5: Original Band: Abnormal Pixels.**

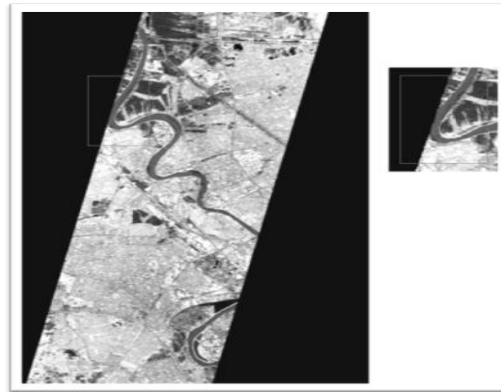


Figure 6: Band after Correction Using Hyperion Tools

Geometric Correction

All satellite pictures taken from the "<http://earthexplorer.usgs.gov/>" site were the same Universal a Transversal aMercator (WGS 1984 aUTM Zone 38N)acoordinate system using the anearest neighbor aresampling method, using the WGS 84 adatum and UTM _Zone 38N aprojection.

“Radiometric and Atmospheric Corrections”

The FLAASH atmospheric a correction was the applied to remaining 149 calibrated a Hyperion imagery. Figure 7 compares the Reflectance and the Radiance spectral profiles of the same features in separate windows before and after atmospheric corrections.

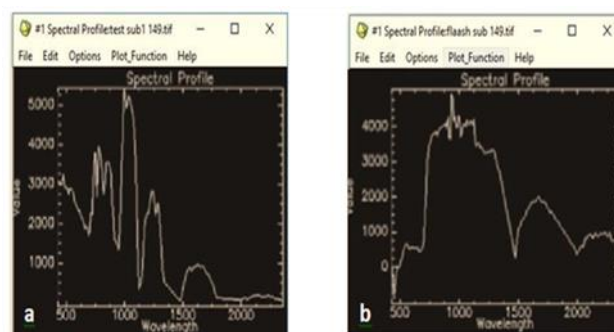


Figure 7: (Z-Profile) Spectral Profile for the Randomly Select Pixel, a) Before the Atmospheric aCorrections, b) After the Atmospheric aCorrections

From the above profile, one can observe the enhancement in the feature class after running the model. We can well keep that the dips present in shape at 550(nm) and 990 (nm) (approximated) and also at 1130(nm) are completely reduced after the atmospheric correction. Instead, we can observe the presence of several narrow contiguous peaks in the wavelength range of 900-1350 nm in the profile of the atmospherically corrected image.

Support Vector Machine (SVM) Classifier

After achieving the Hyperspectral data Preprocessing stages, we will perform support vector machine (SVM) kernels on the original image (149bands). We used support vector machines

(SVM) kernels to solve the challenge of classification of the hyperspectral remote sensing data. In this study, we used SVM-kernels classification to test the efficacy of this promising classification methodology for five classes. These the classes are a Buildings, water, roads, vacant land and Trees. SVM-kernels classification is a supervised method that needs training data for Classification. We used region interest (ROI) to create training data, as shown in table (5). After we have created a training data, we must do SVM-kernels classification a method on hyperspectral image. The Fig(8), Fig(9), and Fig(10) show classified images by the supervised Polynomial kernel, RBF kernel, and Sigmoid kernel classifier for the original image (149) bands. This Classification an achieved overall accuracy of 87.3% for RBF kernel, 61.4 for Polynomial kernel, and 68 for Sigmoid kernel, as shown in table (6).

The standard confusion matrix was used to assess image classifications accurately. The Accuracy assessment is based on the comparing classification results with the ground truth (ROIs). After getting a classified image, the post-classification and the SVM accuracy have been calculated.

Table 5: Training Data for Study Area by Using (ROIs)

No.	Class name	Training data
1	Tree (Green)	92 points
2	Water (Blue)	98 points
3	Road (Yellow)	11 points
4	Vacant land (Cyan)	14 points
5	Building (Red)	13 points

Table 6: Different Kernel SVM classification for original Data

	SVM		
	Polynomial kernel	RBF kernel	Sigmoid kernel
Overall classification %	61.4%	87.3%	68%

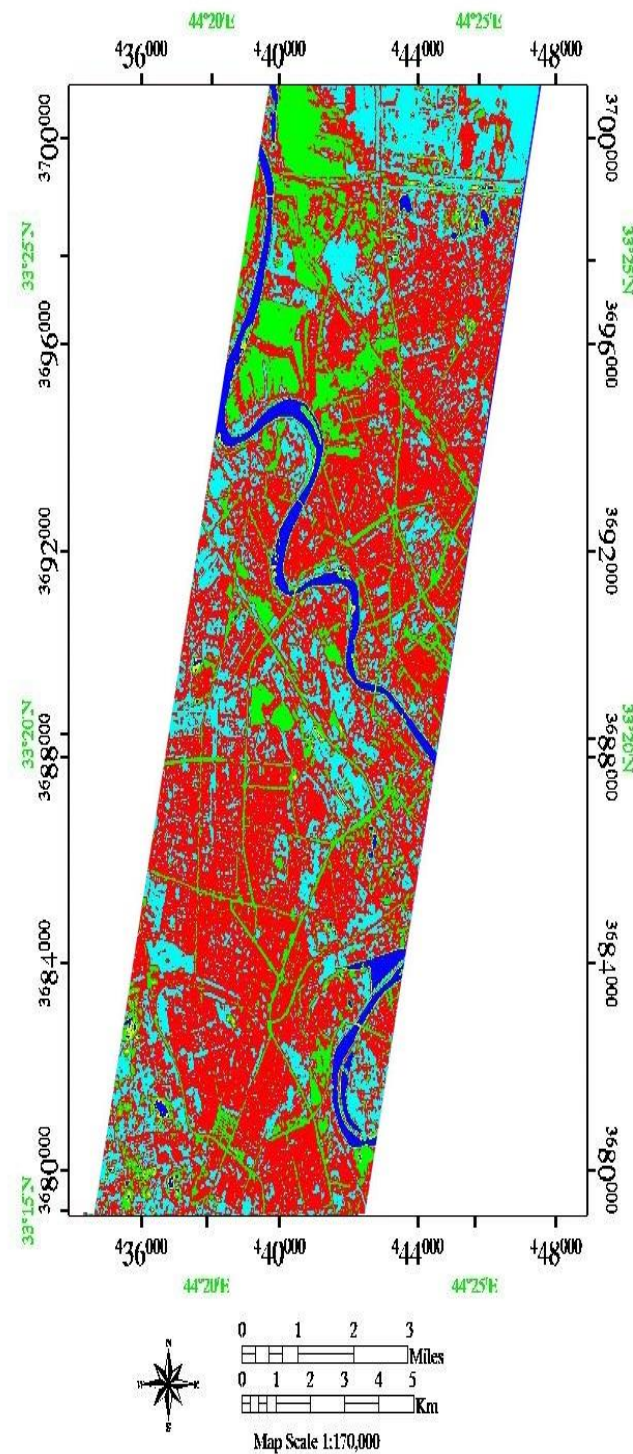


Figure 8: Polynomial kernel-SVM for Original Image with 149 Bands

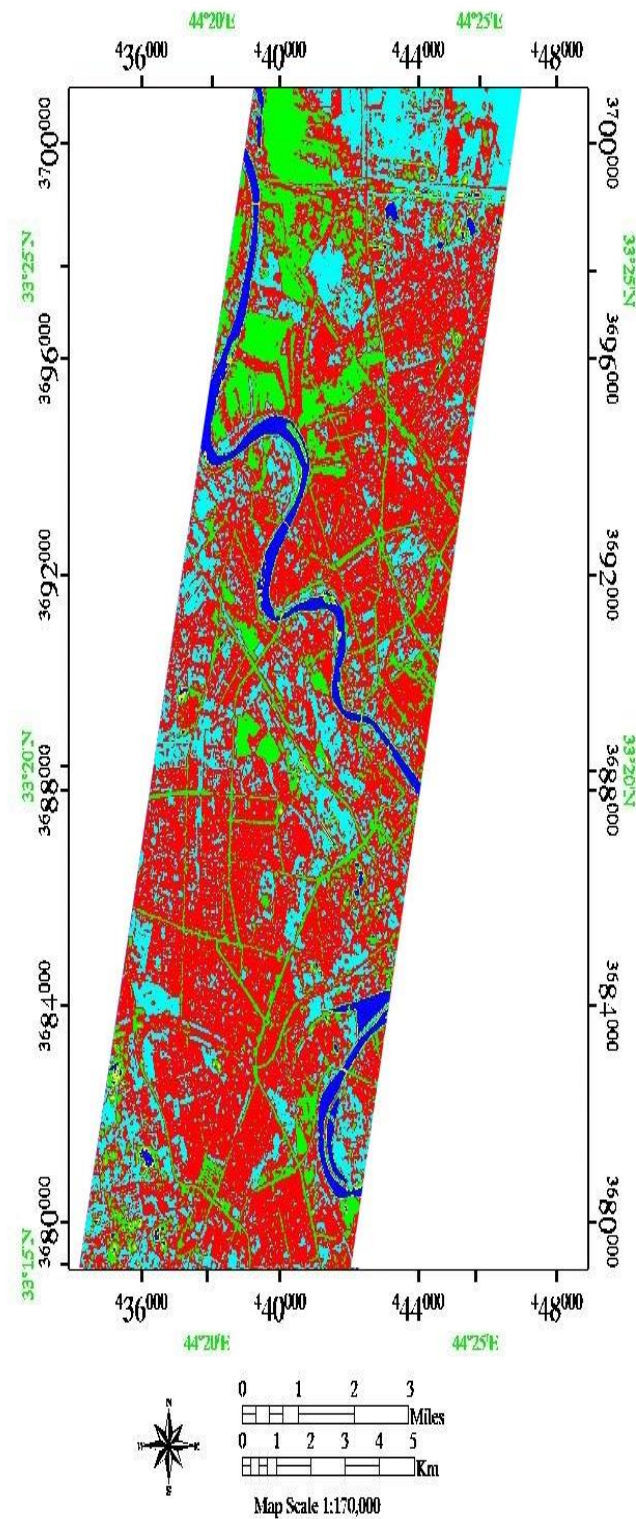


Figure 9: RBF kernel -SVM for Original Image with 149 Bands

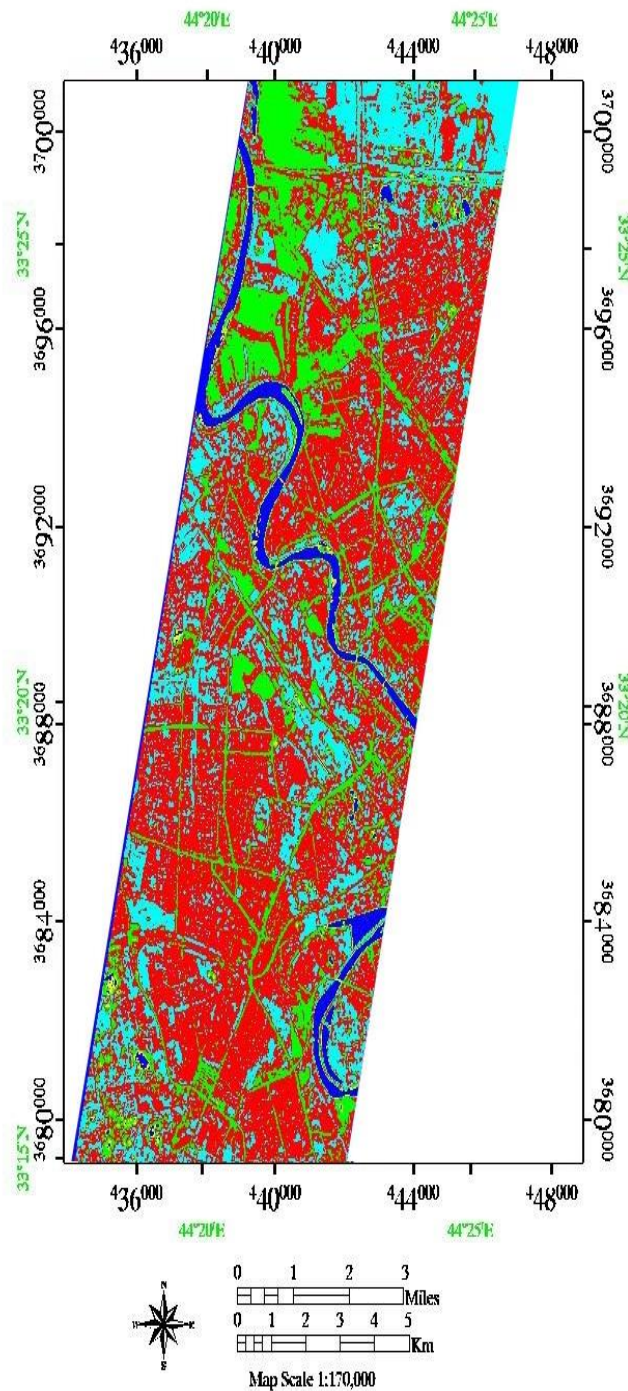


Figure 10: Sigmoid kernel-SVM for Original Image with 149 Bands

Conclusions

The paper has presented various SVM kernels that may be used for the hyperspectral satellite images. How to efficiently find out that a kernel is ideal for a given learning task is still an unsolved problem. The experiments showed that the RBF kernel performs very well on the hyperspectral satellite image, with an overall Classification of 87.3%.

References

1. Perumal, K. and Bhaskaran, R. "Supervised Classification Performance Multispectral Images", Journal of Computing ,2, pp:124-129. 2010.
2. Anand, U., Santosh, K. S. and Vipin, G.S." Impact of features on classification accuracy of IRS LISS-III images using artificial neural network", International Journal of Application Innovation in Engineering and Management, 3, pp:311-317. 2014.
3. Zhang, Y.; Wu, L. "Classification of fruits using computer vision and a multiclass support vector machine", 12, 12489–12505. 2012.
4. Bushra Q. Al-Abudi¹, Mohammed S. Mahdi and Yasser Chasab Bukheet "Study of Land Cover Changes of Baghdad Using Multi-Temporal Landsat Satellite Images" Iraqi Journal of Science, Vol. 57, No.3B, pp:2146-2160., 2016. [5] Sahar A. El_Rahman "Hyperspectral Image Classification Using Unsupervised Algorithms" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 4, 2016.
5. T. Lillesand, R. W. Kiefer, and J. W. Chipman, "Remote sensing and image interpretation", 5th ed. New York: Wiley, 2004.
6. P. Varshney, "Advanced image processing techniques for remotely sensed hyperspectral data". Berlin/ ; New York: Springer, 2004.
7. V.N. Vapnik, "Statistical Learning Theory". John Wiley and sons Inc., New York, pp736, 1998.
8. V.N. Vapnik, "An Overview of Statistical Learning Theory". IEEE Transactions on Neural Networks, vol 10, no.5,1999.
9. C. Cortes, V.N. Vapnik, "Support-Vector Networks". Machine Learning, springer, NewYork, 20, 3, 273-297, 1995.
10. C. Chang, C. Lin, "LIBSVM: a library for support vector machines".National Taiwan University,2011.
11. S. Knerr, L. Personnaz, G. Dreyfus, "Single-layer learning revisited: a stepwise procedure for building and training a neural network". Neurocomputing: Algorithms, Architectures and Applications. Springer-Verlag, F68, pp41-50. 1990.
12. B. Schölkopf, "The kernel trick for distances". Technical report, Microsoft Research. 2000.
13. H.Suo, M. Li, P. Lu, Y. Yan, "Using SVM as Back-End Classifier for Language Identification". EURASIP J Audio, Speech, and Music Processing , doi:10.1155/2008/674859.2008.
14. S.S. Keerthi, C.J. Lin, "Asymptotic behaviours of support vector machines with Gaussian kernel". Neural Computation. MIT Press, Cambridge, vol 15, no7, pp 1667-1689.2003.