Accuracy assessment of supervised classification methods for extraction land use maps using remote sensing and GIS techniques

Zahraa Abbas1, Dr. Hussein Sabah Jaber2
1-2 College of Engineering, University of Baghdad, Iraq
E-mail: husseinsabah00@gmail.com

Abstract. Remotely sensed imagery identifying as the best type of data has information throughout the world. The imagery has an importance information, since it can show up-date-to-day information, and provide a truly information. Different kinds of classifiers have been used to perform that. However, there is no once test for Land cover and Land use in Hilla city. The study aims to create land use classification by making a comparison between different algorithms in Hilla city, Babylon, Iraq. The WorldView-2 imagery is used to perform the pre-processing, analysing of our comparison. Under the steps of pre-processing, the several corrections were made and performed on the imagery. For processing stages, two approaches were used; (1) Support Vector Machine and (2) Maximum Likelihood. The result reveals, that the Support Vector Machine method has the most significant of overall accuracy equal to 94.48% with kappa coefficient equal to 0.90, and these values much better and higher than those of Maximum Likelihood algorithm in estimating and extracting of Land cover/Land use. Therefore, this algorithm has been suggested to be applied as an optimal classifier for extraction of land use maps due to its higher accuracy and better consistency within the study area.

1. Introduction
Remote sensing and Geographic information system have capability of detecting and/or monitoring the features of earth’s surface using satellite images have different radiometric, spectral, spatial, and temporal resolution. These technologies resolutions have several advantages in order to minimize time and cost to extract Land cover and Land use (LULC) [1, 2]. In addition, the remote sensing provides important and different kinds of remotely data source to extract LULC information. Remote sensing data are widely used and applied to perform classification of LULC throughout the world [2]. Remote sensing data has the ability of updating information about all the features that locate on earth’s surface [3]. There are different satellite systems provides different imagery has different spatial resolutions applied for LULC detection such as WORLDVIEW-2, Quick Bird, Landsat, Spot and so on [1-3]. The classification of remote sensing imagery is an important method in order to determine the LULC information [2]. Classification approaches divided into two categories: pixel base and/or object oriented base classifiers [3]. The Maximum Likelihood approach (MLC) widely applied for LULC extraction, by creating decisions of surface regarding of the covariance and mean of each single class [4]. However, the non-parametric technique Support Vector Machine (SVM) classifier, perform the classification by no assumptions and separate the features on each class [5, 6]. SVM is a non-parametric algorithm, which contains a series of learning algorithms are conducted for classes classification and regression [5-7]. The SVM classification was proposed by Vapnik and
Chervonenkis [8] and after that discussed in more detail by Vapnik [9]. Add to that, the SVM was widely applied in different remote sensing projects and applications [10, 11]. Huang et al. [12] applied the SVM classifier for a LULC classification by using Landsat thematic mapper (TM) imagery. The SVM classification has a good accuracy and superior to the accuracy of using different methods such as (1) maximum likelihood and (2) a decision tree [1-3]. Yousefi et al. [13] has tested nine different supervised classification approaches and they are (neural network, Spectral angle mapper, maximum likelihood, SVM, mahalanobis distance, binary code, minimum distance, spectral information divergence and parallelepiped) for LULC mapping in province of Mazandaran, Iran and they used the imagery of Landsat ETM+. The results showed that the SVM classifier has the best accuracy than the other classifiers. The main objective of this study is to perform LULC classification using MLC and SVM approaches in Hilla city, Babylon, Iraq, and make comparison to find out the perfect algorithm to classify the study area.

2. Materials and methods
The detection technique to extract the thematic map of LULC was tested and referenced by previous work [12]. In this research, WorldView-2 satellite system imagery of 2017 was used to perform the pre-processing and processing steps. In this research, geometric, radiometric, and atmospheric corrections were applied to remove image noise. Then, layer stacking and image sub-setting were conducted.

The next step was to applied different kinds of supervised classification algorithms (MLC and SVM) that were applied to the collected imagery to perform imagery classification to map the LULC theme of Al-Hilla city, Babylon, Iraq. The accuracy assessment and statistical analysis of the expected results were conducted to get an accurate and suitable approach to classify the study area of this research. Figure 1 reveals the flowchart of our methodology.

![Figure 1. The adopted Flowchart of the study area methodology.](image-url)

2.1. Study Area Location and Descriptions
Hilla city has coordinates of 32°29′N 44°26′E, with elevation 112 ft equal to 34 m. It is around 100 km (62 miles) to the south of Baghdad. The Hilla city relates to the governorate of Babylon and it is the
capital of Babil province, and it is located in central Iraq. Its boundaries area surrounding with other governorates, they are Baghdad, Kerbala, Anbar, Najaf, Wassit, and Qadissiya. It has Euphrates River, it is one of biggest rivers in Iraq, and it intersects its governorate and divided into the AL- Hindiyah and Al- Hilla branches south of the Musayib town, to create a network of canals runs through the lands of governorate to use as an irrigation system and supply the regions, fields and farms with water. Hilla city has a population of around 970, 000 people regarding registration of 2018. Figure 2 shows Location of the Hilla city.

![Hilla City Location Map](image)

**Figure 2.** The study area location map.

2.2. *The Used Date*

A World View-2 Satellite system is one of a large number of systems that provide an imagery which has high-resolution multispectral and panchromatic bands. The WorldView-2 Satellite has one panchromatic band, with eight multispectral bands, four of these bands have standard colors (red (R), green (G), blue (B) and near-infrared1 (NIR1) with another four new bands and they are; yellow, red edge, coastal, and near-infrared 2. The WorldView-2 Satellite images have a full-color for different kinds of applications such as exploration, enhanced spectral analysis, disaster relief, mapping, monitoring applications, defense, intelligence applications, land-use/land cover planning, simulation environments and visualization . The system WorldView-2 can act like a paintbrush, sweeping from back to forth in order to capture large areas of multispectral images in one single pass. WorldView-2 satellite system captures around 1 million km2 every single day. Table 1 and Figures 3 and 4 show the WorldView-2 satellite system and its bands.
**Figure 3.** The WorldView-2 satellite system bands.

**Figure 4.** The WorldView-2 image of the study area.

**Table 1.** The specifications of WorldView-2 satellite system.

| Specification                          | Details                                                                 |
|----------------------------------------|-------------------------------------------------------------------------|
| Date of launch                         | 8/10/2009                                                               |
| Altitude of orbit                      | 770 kilometers                                                          |
| Orbit type                             | sun synchronous                                                         |
| Orbit period                           | 100 mints.                                                             |
| Sensor bands                           | Panchromatic band and 8 multi-spectral bands                           |
|                                        | ground sample distance panchromatic= 0.46 m GSD with nadir and 0.52 m GSD with 20° off-nadir |
|                                        | multispectral bands = 1.84 m GSD with Nadir, 2.4 m GSD with 20° off-nadir |
| Sensor resolution GSD                  |                                                                         |
| Swath width                            | 16.4 Km at nadir                                                        |
| Attitude determination and control     | 3-Axis Stabilized                                                      |
| GPS position accuracy                  | < 500 m                                                                |
| Revisit frequency                      | 1.1 days at 1 meter GSD or less 3.7 days at 20° off-nadir or less (0.52 meter GSD) |
| Geolocation accuracy                   | demonstrated <3.5 m ce90 without GCP                                   |

(Source: [https://www.satimagingcorp.com/satellite-sensors/worldview-2/](https://www.satimagingcorp.com/satellite-sensors/worldview-2/))
3. Imagery Processing and Analysis

3.1. Atmospheric, Radiometric and Geometric Corrections
Atmospheric, radiometric and geometric corrections are required for the remove the imagery noise to perform different applications of remote sensing such as; image classification [11]. The reason is that atmospheric correction of a date image would often mean to remove a values from imagery pixels from the spectral bands of the satellite images [13]. However, the geometric correction of WorldView-2 to a satellite image is conducted by collecting ground control points (GCPs) from the fieldwork method using Envi 5.2 software to perform the geometric correction.

3.2. Fieldwork
The fieldwork conducted by using Handheld GPS type Garmin 78s, it is one of the most popular devices have been used to collect the locations of points locate in the study area with different kind of remote sensing applications. The accuracy of this device is for horizontal < 10m, and for vertical = 0.05 m/s steady-state. Figure 5 shows the used Garmin GPS.

![Handheld GPS type Garmin 78s](image)

Figure 5. Handheld GPS type Garmin 78s,

10 GCPs were collected from fieldwork to perform the geometric correction and to choose a suitable projection to the used image which is the UTM, in Zone 38N with WGS 84 datum and table 2 includes the collected GCPs from fieldwork, the geometric correction performed using Envi software as indicates in Figure 6.

| No. | Longitude (E) | Latitude (N) | Features                  |
|-----|---------------|--------------|--------------------------|
| 1   | 44°25′15.03″  | 32°29′55.17″ | Bridge                   |
| 2   | 44°26′12.82″  | 32°29′24.17″ | Round about              |
| 3   | 44°27′26.65″  | 32°29′47.85″ | Al-Bakerly inter section |
| 4   | 44°25′09.14″  | 32°26′47.22″ | Nader inter section      |
| 5   | 44°25′31.78″  | 32°28′03.38″ | Mother square            |
| 6   | 44°26′22.21″  | 32°29′02.56″ | Province bridge          |
| 7   | 44°23′41.21″  | 32°26′40.16″ | Road inter section       |
| 8   | 44°24′48.45″  | 32°26′02.35″ | Factory gate             |
| 9   | 44°26′21.60″  | 32°26′14.92″ | Small bridge             |
| 10  | 44°24′01.47″  | 32°28′37.69″ | Al-Asatetha road inter section |
In addition, the authors use the Nearest Neighbour approach is used for resampling uncorrected pixels. Finally, a root mean square error (RMSE) of images is used to obtain to check out their accuracy and we got less than 0.4 pixels, and this value is acceptable [14]. Figure 7 shows the imagery after perform all corrections and it became a noise-free image.

Figure 6. Geometric correction of WorldView-2 satellite image by using GCPs.

Figure 7. The imagery after performing all correction and calibration.

3.3. Image Classification
Image classification was carried out by using the MLC and SVM algorithms. In the following subsections, brief explanations of the two algorithms are provided.

3.3.1. Selecting Training and Testing Sites
After removing all imagery noise by radiometric and geometric correction and conduct image clipping and layer stacking, the WorldView-2 satellite image becomes ready for further digital image
processing. The study area was classified into four major classes to perform different classification techniques based on the most located features into the study area to obtain the thematic maps of land cover and land cover. Four classes are (Water bodies, Soil, Urbanization area and Vegetation area) collected training sites were conducted by using Envi software be selected polygons of the region of interest (ROIs) for each class, then used these ROIs into classification to generate the thematic map of Land use and land cover for the study area. For this stage was about collecting the testing sites to study area, these sites very important to determine the accuracy assessment for each classification algorithms and to check the validation, producer and user accuracy of classifications. Usually testing sites are ground truth samples, and the better collecting from the field of the study area to get pure ground truth samples. Collecting the testing sites for this study was conducted by using office work from the imagery. Figure 8 (a &b) shows the number of pixels selected to be as training and testing sites for each class of the study area.

3.3.2. Maximum Likelihood Approach
A maximum-likelihood algorithm is one of the wide classifiers has been used for supervised image classification in different remote sensing application [14]. Erdas [15] mentioned that this approach works by computing weighted distances and/or likelihood D relate to an unknown vector X belong to known classes. However, Mc is regarding Bayesian equation [15]:

\[
D = \ln(ac) - [0.5\ln(|Covc|)] - 0.5(X-Mc)^T(Covc^{-1})(X-Mc)
\]

(1)

Where, C is a particular class, ac refers to the percent probability for any candidate pixel of class C, when Covc refers to pixels covariance matrix in class c. However, |Covc| is a Covc determinant, (Covc^{-1}) inverse of Covc, the T refers to transposition function [15]. After that the results is revealed and the validation is done by using the confusion matrix. The results of applied MLC and the producer and user accuracies are presented in Table 3 below for each class. Figure 9 WorldView-2 satellite image after perform MLC classification.

Table 3. Producer and user accuracies of MLC classification

| Class         | Prod. Acc. (%) | User Acc. (%) |
|---------------|----------------|--------------|
| Water Bodies  | 99.49          | 79.81        |
| Vegetation    | 82.70          | 99.57        |
| Urbanization  | 89.32          | 50.57        |
| Soil Area     | 74.89          | 95.10        |
3.3.3. Support Vector Machine Algorithm
SVM has been widely applied for imagery classification in the remote sensing applications [12]. The SVM aims to reach the optimal separating of hyper plane and/or hyper plane that locates in high and/or dimensional space, in order to locate an optimal boundary between classes. The approach this method was proposed by the authors Vapnik and Chervonenkis [8] and later discussed by the author Vapnik [9]. The success of this method regarding to how it is good to process its trains [16, 17]. With Envi. V-5.2 done the SVM classification, this classifier based on also pixels of imagery that mean conducted segmentation first for whole image [17]. The kernel type is radial based function, there were more than 50 sample selected for each class in training sites. After that the results is revealed and the validation is done by using the confusion matrix. The results of applied SVM and the producer and user accuracies are presented in Table 4 for each class. Figure 10 shows kernel as a radial based function, Figure 11 reveals WorldView-2 satellite image after perform SVM classification.
Figure 10. Kernel type as a radial based function of SVM classification.

Table 4. Producer and user accuracies.

| Class       | Prod. Acc. (%) | User Acc. (%) |
|-------------|----------------|---------------|
| Water Bodies| 96.23          | 92.25         |
| Vegetation  | 94.28          | 97.34         |
| Urbanization| 94.81          | 69.18         |
| Soil Area   | 87.99          | 96.95         |

Figure 11. LULC thematic map with SVM classification.
4. Classification Accuracy Assessment
The final step of the satellite image classification involves the accuracy assessment stage [18]. Accuracy assessment is a quantification of estimation with aid of remotely sensed dataset to classification conditions and it is useful for evaluation of classification approach, and it is also important to determination the error that might be involves. The accuracy assessment of our classifiers is proposed in the form confusion matrix [19].In this research, to validate the our classification results with the testing sites that mention in section (3.3.1) above, with the help of collected GCPs from fieldwork (that collected by using Handheld Garmin 78s GPS) were chosen randomly for performing the accuracy assessment of each classification that made by SVM and MLC.

5. Results and discussion
The overall accuracy of the SVM and ML were 94.48% and 88.30%, respectively. However, kappa coefficients of both SVM and ML were 0.90 and 0.80, respectively. Table 5 and Figures 12 & 13 show the comparison between the Overall Accuracy and Kappa coefficient for each classification of LULC.

Table 5. Comparison between the classifications accuracy for land cover.

| No. | Approaches                | Overall Accuracy (%) | Kappa coefficient |
|-----|---------------------------|----------------------|-------------------|
| 1   | Maximum likelihood        | 88.30                | 0.80              |
| 2   | Support vector machine    | 94.48                | 0.90              |

Figure 12. The overall accuracy of two classifiers

According to the above figures, it clearly seems that the SVM classifier has the highest results of overall accuracy and the kappa coefficients than those of the MLC approach for our study area in Hilla city, Babylon, Iraq. Moreover, both Huang et al. [12] and Otukei & Blaschke [22] are evaluated various methods for classifications in LULC mapping, and they were revealed that the results of SVM approach compared to the results of ML algorithm and also to decision trees method has the most highest accuracy assessment for mapping the LULC. The author Deilmai et al. [23-27] in his research, compared between 2 classification algorithms SVM and MLC in order to extract LULC thematic map in Johor city, Malaysia., his results revealed the results of SVM classifier regarding to kappa coefficient was equal to 0.86 and it was most an accurate algorithm.
6. Conclusions

Nowadays, applying LULC mapping is a critical issue for collecting information for master city planning, detecting and monitoring the environment. Using remote sensing technology with different satellite imagery is recommended as the most perfect strategy for LULC mapping. Using images of different systems, the limitation and capability of these images can be detected. The purpose behind this research is to generate the LULC thematic maps by performing a comparison between two algorithms, by using both of SVM and MLC methods in Hilla city, Babylon, Iraq, by using imagery of WorldView-2 satellite system. After performing all necessary corrections by pre-processing, processing and analysing stages on the satellite images, two different approaches were used to classify our satellite image. The validation results reveal, that the results of SVM algorithm with overall accuracy of 94.48% and with a kappa coefficient of 0.90, it is a higher accuracy if we compare to that one that got from applied the MLC method for LULC mapping. The SVM approach is suggested to be as the best classifier to extract the maps LULC, simply because, it has higher accuracies for our study area. In this research, we confirm that the proper performance of SVM method with the extracted LULC thematic map can can be facilitatd a management in level with sustainable developing.

References

[1] H Dibs and S AL-Hedny. 2019. Detection Wetland Dehydration Extent With Multi-Temporal Remotely Sensed Data Using Remote Sensing Data Using Remote Sensing Analysis and GIS. IJIET, 2019. 10 (1)143–154.

[2] H Dibs. 2018a. Comparison of Derived Indices and Unsupervised Classification For Al-Razaza Lake Dehydration extent Using Multi-Temporal Satellite Data and Remote Sensing Analysis ARPN J Eng Appl Sci 13 (24).

[3] H Dibs, S AL-Hedny, H S Karkoosh. 2018b. Extraction Detailed Buildings 3D Model With Using High Resolution Satellite Imagery By Remote Sensing and GIS Techniques AL-Qasim Green University: A Case Study IICIET 9 (7) 1097–1108.

[4] J Richards & X Jia. Remote sensing digital image analysis. 4th edition. Heidelberg, springer, Germany, 2006. 359-88.

[5] M Bray & D Han. Identification of Support Vector Machines for Runoff Modelling. J. Hydroinform. 2004, 6, 265-80.
[6] A Moghaddamnia M Gosheh M Nuraie M Mansuri D Han & E Schmitter. Performance Evaluation of LLR, SVM, CGNN and BFGSNN Models to Evaporation Estimation. Water Geosci. 2010, 9, 108-13.
[7] D Han L Chan and N Zhu. Flood Forecasting Using Support Vector Machines. J. Hydrolinform. 2007, 9, 267-76.
[8] N Vapnik & A Chervonenkis. On The Uniform Convergence of The Relative Frequencies of Events to Their Probabilities. Theory Probability. Appl. 1971, 16, 264-80.
[9] N Vapnik. An Overview of Statistical Learning Theory. IEEE Trans. Neural Network. 1999, 10, 988-99.
[10] B Dixon & N Candade. Multispectral Land Use Classification Using Neural Networks and Support Vector Machines. International Journal of Remote Sensing, 2008, 29, 1185-206.
[11] G Foody & A Mathur. Toward Intelligent Training of Supervised Image Classifications: Directing Training Data Acquisition for SVM Classification. Remote Sensing and Environment. 2004, 93, 107-17.
[12] C Huang LS Davis and JRG Townshend. An Assessment of Support Vector Machines for Land Cover Classification. Int. J. Remote Sens. 2002, 23, 725-49.
[13] S Yousefi M Tazeh S Mirzaee HR Moradi and SH Tavangar. Comparison of Different Classification Algorithms in Satellite Imagery to Produce Land Use Maps, Case study: Noor city. J. Appl. rs gis tech. Nat. Res. Sci. 2011, 2, 15-25.
[14] C Tucker D Grant & J Dykstra. NASA’s Global Orthorectified Landsat Data set. Photogramm. Eng. Rem. Sens. 2004, 70, 313-22.
[15] Erdas Inc. Erdas field guide. Erdas Inc., Atlanta, Georgia, 1999.
[16] EE Osuna & R Freud. Support Vector Machines: training and applications. Massachusett institute of technology cambridge, USA, 1977. p. 1-144.
[17] A Erener. Classification Method, Spectral Diversity, Band Combination and Accuracy Assessment Evaluation for Urban Feature Detection. Int. J. Appl. Earth Observ. Geoinform. 2013, 21, 397-408.
[18] B Manish G Chitra & N Umrikar. Image Classification Tool For Land Use/ Land Cover Analysis: A Comparative Study of Maximum Likelihood and Minimum Distance Method. Int. J. Geo. earth. environ. 2012, 6, 189-96.
[19] R Congalton. A Review of Assessing The Accuracy of Classifications of Remotely Sensed Data. Remote Sens. Environ.1991, 37, 35-46.
[20] D Lu P Mausel E Brondizio & E Moran. Change Detection Techniques. Int. J. Remote Sens. 2004, 25, 2365-407.
[21] K Sundarakumar M Harika S Aspita S Yamini & K Balakrishna. Land Use and Land Cover Change Detection and Urban Sprawl Analysis of Vijayawada City Using Multi-temporal Landsat data. Int. J. Eng. Sci. Tech. 2012, 4, 170-8.
[22] J Otukei & T Blaschke. Land Cover Change Assessment Using Decision Trees, Support Vector Machines and Maximum Likelihood Classification Algorithms. Int. J. Appl. Earth Observation. Geoformation. 2010, 12., 527-31.
[23] B Deilmay B Ahmad & H Zabihi. Comparison of Two Classification Methods (MLC & SVM) to Extract Land Use and Land Cover in Johor city, Malaysia. IOP Conf. Ser. Earth Environ. Sci. 2014, 20, 1-6.
[24] H Dibs, M O Idrees, and G B A Alsalhin. 2017a Hierarchical Classification Approach for Mapping Rubber Tree Growth Using Per-Pixel and Object-Oriented Classifiers with SPOT-5 Imagery Egyptian Journal of Remote Sensing and Space Science 20 (1) 21–30. https://doi.org/10.1016/j.ejrs.2017.01.004.
[25] H Dibs, A S Al-Janabi, C Gomes. 2017b. "Easy To Use Remote Sensing and GIS Analysis for Landslide Risk Assessment." Journal of University of Babylon 26 (1): 42 - 54.
[26] H Dibs, M O Idrees, V Saeidi and S Mans. 2016. "Automatic Keypoints Extraction from UAV Image with Refine and Improved Scale Invariant Features Transform (RI-SIFT)." International Journal of Geoinformatics 12 (3): 51-58.

[27] H Dibs, S Mansor, N Ahmed and B Pradhan 2015 Band-to-band registration model for near-equatorial Earth observation satellite images with the use of automatic control point extraction Int J Remote Sens 36 (8) 2184–2200.