Supervised Classification Model Using Google Earth Engine Development Environment for Wasit Governorate

Sajjad H Hasan¹, Amjed N M AL-Hameedawi², H S Ismael³

¹,²,³Civil Engineering Department, University of Technology, Iraq

¹bce.19.63@grad.uotechnology.edu.iq

Abstract. As a result of the advancements that have occurred in the technical field of geomatics, particularly after the development of developmental programming environments, they have become the most important machine for conducting image analyses of satellite data, creating and modifying spatial analysis tools, and performing large data analyses at a fast rate without the need for high-end specifications on the personal computer. This study has several objectives, including the definition and popularization of the use of the power of Google Earth Engine (GEE) in the speed of conducting spatial analyzes, which cite by conducting a classification at the level of a governorate and obtaining results with speed and relatively good quality. By using the Google Earth Engine (GEE) platform and through Javascript programming language, a classification of the land cover of Wasit Governorate, Iraq was created under the supervision of a satellite image (Landsat 8) by creating a training sample, Google Maps' High Resolution basemap imagery was used to create this map to identify classes of landcover (water, bare soil, vegetation, and urban). Each source pixel is assigned to one of the previously mentioned classes. Then to create a land cover map of the region using the Statistical Machine Intelligence and Learning Engine (SMILE) classifier from the JAVA library, which is used by Google Earth Engine (GEE) to implement these algorithms. The result is an array of pixels (raster data). The pixel value represents the class that was previously determined by the samples.

Keywords: GEE, Landcover, SMIL, JAVA, JavaScript, GIS, Wasit

1. Introduction and Literature Review

In recent years, due to the change in the nature of the land cover because of urban expansion, the hydrological change through the change in the area of water bodies and vegetation cover. Among the most
significant of these is climate change. This has a significant impact on planning, Subin, Z M, Riley et. al, 2012 [2], Holgerson, M A et. al 2016 [3].

As a result of climate change and increasing human activity, the surface temperature has risen, precipitation has decreased, and desertification has grown, Amjed N M AL-Hameedawi, 2020 [6]. Geographic Information System (GIS) has continued to develop and may be a useful tool in mapping, investigating, and imaging flooded regions and dry docks, Amjed N M AL-Hameedawi, 2018 [7], among other applications. Remote sensing data are a very important and reliable source of information for mapping and terrain comprehension as well as disaster management and civil engineering infrastructure design, I. A. Alwan, et. al, 2020 [11]. In order to achieve this, the study will investigate how to generate raster data for Land-cover classification maps using an automatic method based on a programming language (JavaScript) on the Google Earth Engine Platform (there are numerous sources of row data that can be used for classification maps, one of which data can be satellite imagery, O Z Jasim, et. al, 2019 [8]).

Images obtained through remote sensing techniques are a major source of row data used for mapping and monitoring natural and artificial features on the earth's surface, A Z Khalaf, et. al, 2018 [9]. Having classified data (Classification is the process by which things are linked to a common class or type, T H Shihab, et. al, 2020 [10]) of land-cover provides us with an opportunity to benefit from the information contained in the satellite images, which allows us to distinguish with reasonably acceptable accuracy between earth covers within the study area, Amjed N Al-Hameedawi & Manfred F B, 2014 [5]. A cloud computing platform intended to store and analyze massive data volumes (on the order of petabytes) for the sake of analysis and ultimately decision-making, Kumar, et. al, 2018 [1], the Google Earth Engine (GEE) Landsat data sets were stored and connected to Google's cloud computing engine for open source usage when the Landsat series became available free of charge in 2008. Among the data layers in the present collection are those from other satellites, as well as vector data sets derived from Geographic Information Systems (GIS), social and demographic data sets, meteorological data, digital elevation models, and climatic data layers.

A programming framework for displaying and analyzing geographical data for academic purposes, government initiatives, and non-profit company and government users, Google Earth Engine (GEE) is available for download. The Global Earth Explorer (GEE) allows us to interact with the Earth's surface and examine data in the form of terrains, satellite pictures, 3D buildings, and conduct analyses on them (water surface coverage, snow cover, change detection analysis and time-lapse, for example). There is a public data repository of satellite images on GEE that dates back about forty years and is accessible on a worldwide basis, Verpoorter, et. al, 2014 [4].

2. Study Area

Wasit is a governorate in eastern Iraq, southeast of Baghdad and bordering Iran, Wasit lies along the Tigris River, halfway between Baghdad and Basra, exactly at geographic location has shown below. Table (1).

| Vertices | Longitude (dd) | Latitude(dd) |
|----------|----------------|--------------|
| A        | 47.0477370°    | 33.60353832° |
| B        | 47.04786584°   | 31.89901876° |
| C        | 44.48796555°   | 31.89898233° |
| D        | 44.48800847°   | 33.60353832° |
Al-Kut is the provincial capital and the largest city in the province. The Tigris River runs through the province of Wasit, Iraq. It is located about 172 kilometres southeast of Baghdad [19].

Wasit covers an area of about 17,153 square kilometres. Wasit bounded to the north of Maysan governorate as figure (1) shown below:

![Figure 1: Show Wasit Governorate with districts](image)

**Fig. (1).** Shows Wasit Governorate with districts [13].
3. Materials

A model is a result materials are prepared that, if correctly built, may be used to represent an accurate replica of the original [14]. As a result, satellite data is the most important and fundamental raw material for developing spatial databases in general, and particularly when dealing with raster data, as is the case in this research. The categorization of satellite pictures with spectral beams into four major categories, built-up area, water, bare soil, and vegetation, based on the visible spectrum of the Landsat 8 satellites [15], serves as the basic foundation for this study [16].

Supervised classification is one of the most essential classical machine learning techniques in remote sensing, and it is also the most difficult to master. Applications vary from the creation of Land Use/Land Cover maps to the detection of changes in land cover. Google Earth Engine is the only platform that is well-suited for doing supervised classification at scale. Because of the interactive nature of Earth Engine development, it is possible to iteratively build supervised classification processes by including a variety of different data into a single model. This lesson discusses the fundamentals of supervised classification process and accuracy evaluation, among other things. In remote sensing image data processing, the most often employed method is supervised classification. The basic idea is to divide the spectral domain into areas corresponding to different ground cover classifications [17].

3.1. Methods

Essentially, the model starts by defining the research area using a polygon feature on the one of the spatial digitization softwares, like Google Earth, and this was determined, where the polygon feature includes the entire study area of Wasit Governorate (outside its border) in the KMZ file, and the method ends with a simple summary. Then, using the Arcgis software, this formula was transformed into a shapefiles format, which consists of six files, four of which are required to be submitted to the GEE platform (shp, dbf, prj, shx).

The effective execution of terminating and extracting geographical activities was the impetus for the development of a GIS programming paradigm [18].

As of right now, the feature dataset used to delineate the boundaries of the research region has been uploaded and stored to a Google user account on these platforms (GEE).

This technique needs no previous training data and is very successful at generating high-quality classification samples anywhere in the globe, regardless of where the data is collected. Every source pixel should be classified into one of the four categories are built-up, bare, water or vegetation. Using the drawing tools in the code editor, build four new feature collections, each of which contains polygons representing pixels from the corresponding class. There is a property known as landcover that is associated with each feature collection. The value of this property is one of the following: 1,2,3,4 which indicates whether a feature collection represents an built-up area, bare ground, water or vegetation. Subsequently, using these training sets, train a Random Forest classifier to construct a model, which is then applied to all of the pixels in the picture to produce a four-class image. Afterward,

To begin, import the study region as a polygon feature to the platform utilizing the (Assets) menu on the upper left of the code-editor window's left-hand pane. On Assets tab to New Assets then, Shape files may be uploaded to the table (ESRI Shapefile or zip).
3.2. Workflow

Here's a workflow chart that shows the many steps involved in classifying the Wasit Governorate as "supervised classification.", fig(2).

Begin by importing the following characteristic of the research area and defining it as a variable in the scripting window: Fig. (2). Flow chart shows process to execute classification.
1- Call on the boundary of Wasit Governorate as a Shapefile (Study Area) which saved on user account as a variable, as shown Fig. (3):

```
Imports (6 entries)
  ▼ var Study_Area: Table users/sajjadco94/wasit
      type: FeatureCollection
      id: users/sajjadco94/wasit
      version: 1618008727599484
      columns: Object (12 properties)
      properties: Object (1 property)
```

Fig. (3). Script to define Wasit Governorate border as a variable.

2- Define Image dataset as a variable (USGS Landsat 8 Collection 1 Tier 1 TOA Reflectance), as shown Fig. (4):

```
  ▼ var L8N: ImageCollection "USGS Landsat 8 Collection 1 Tier 1 TOA Reflectance" (12 bands)
      type: ImageCollection
      id: LANDSAT/LC08/C01/T1_TOA
      version: 1633864960978957
      bands: List (12 elements)
      properties: Object (40 properties)
```

Fig. (4). Script to define image dataset as a variable.

3- Then creating four new feature collections using the code editor's drawing tools, each containing polygons that represent pixels from the appropriate class, as shown Fig. (5):

```
  ▼ var builtup: FeatureCollection (4 elements)
      type: FeatureCollection
      columns: Object (2 properties)
      features: List (4 elements)

  ▼ var water: FeatureCollection (4 elements)
      type: FeatureCollection
      columns: Object (2 properties)
      features: List (4 elements)

  ▼ var vegetation: FeatureCollection (4 elements)
      type: FeatureCollection
      columns: Object (2 properties)
      features: List (4 elements)

  ▼ var bare_soil: FeatureCollection (5 elements)
      type: FeatureCollection
      columns: Object (2 properties)
      features: List (5 elements)
```

Fig. (5). Script to define feature collection of landcover as a variables.
4- After defining satellite sensor then image dataset have been filtered by clipping data inside boundary, filter data depending on cloud cover percent, time of data capturing, call on specific bands to visualize and made analyses and merge all feature collections into single feature contains all value of landcovers classes. As shown Fig. (6):

```javascript
var start= ee.Date('2021-06-01');
var end= ee.Date('2021-07-01');
var filter_LN8= ee.ImageCollection(LN8)
  .filterBounds(Study_Area)
  .filterDate(start,end)
  .filterMetadata('CLOUD_COVER','less_than',10)
  .median();

var final_image=filter_LN8.clip(Study_Area);
Map.addLayer(final_image, {bands: ['B4', 'B3', 'B2'], min: 0, max: 0.3}, 'True colour image');
var classNames = water.merge(bare_soil).merge(vegetation).merge(water).merge(builtup);
print(classNames);
var bands = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7'];
var training = final_image.select(bands).sampleRegions({
  collection: classNames,
  properties: ['landcover'],
  scale: 100
});
print(training);
```

Fig. (6). Script to applying filters on image dataset.

5- Applying classification by using (smile Random Forest) classifier to classify image data into 4 classes depending on training samples are determined previously, then identify color ramp to visualize the results, as shown Fig. (7):

```javascript
var classifier = ee.Classifier.smileRandomForest(4).train(
  {
    features: training,
    classProperty: 'landcover',
    inputProperties: bands
  });
var classified = final_image.select(bands).classify(classifier);
Map.centerObject(west, 9);
Map.addLayer(classified,
  {min: 0, max: 3, palette: ['00FFFF', '006400', '7FFF00', 'FF1919', '8B8B00', 'FFD700']},
  'classification');
```

Fig. (7). Script to define classifier as a variable and show the results.

6- Finally, there is a technique for calculating a Confusion Matrix, which represents the estimated accuracy of the classification, which may used to evaluate the accuracy of the classification, as shown Fig. (8):
4. Results

The results was a spatial data (raster data) consist of four classes (water, bare soil, vegetation and built up area). The result is important for planning and monitoring for this region (Wasit Governorate), especially when it comes to urban expansion and green areas, and through monitoring the operations of green areas to barren lands. It is important to study the area through classification processes on satellite images and periodically for the purposes of monitoring and facilitating decision-making regarding agricultural and urban planning in such areas. The result after running the scripts are show as Fig. (9):

The results below shows the value of pixels after manipulate classification, each pixels with a range of value had grouped and clustered into one value. Then the value of Confusion Matrix shows the accuracy of classification, as shown Fig. (10):
5. Conclusion

By conducting the supervised classification process, starting with the process of uploading data of the administrative boundaries of the governorate and determining the samples on the basis of which the classification is made, and the accompanying recall of satellite image data, merging it and determining the time of its capture, the process is very easy and smooth despite the large volume of the summoned image data, especially satellite images with all spectral packages, as well as all this process took only a relatively short time to show the results with high accuracy without the need for a personal computer with good technical specifications, as is the case in the programs used in GIS applications (Arcgis, QGIS, Erdas Imagine, global Mapper…etc), as it only required an Internet connection and access to the GEE platform through the browser and writing remote sensing short commands using JavaScript programming language.

6. Recommendation

In terms of remote sensing image analysis, Google Earth Engine (GEE) is a really great tool. Its data library has a diverse collection of satellite remote sensing data, and it is revolutionizing our understanding and exploration of the Earth. When it comes to accessing the GEE data catalog and performing spatial analysis, the Source Code (Editor) is an each shop. GEE may be used to classify images in a supervised or unsupervised manner. Most of the classification workflows involve collecting data for training, constructing a classifier and then training it. Then the image is classified and the error is estimated using an incremental innovation data. However, attention is used when using confused matrix to evaluate the accuracy of supervised classifiers.

References

[1] Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. Remote Sens. 2018, 10, 1509. [CrossRef].
[2] Subin, Z. M., Riley, W. J. & Mironov, D. 2012 Model. Earth Syst. 4, M02001
[3] Holgerson, M. A. & Raymond, P. A. 2016 Nat. Geosci. 9, 222–226
[4] Verpoorter, C., Kutser, T., Seekell, D. A. & Tranvik, L. J. 2014 Geophys. Res. Lett. 41, 6396–6402
[5] Amjed N. Al-Hameedawi & Manfred F. B., 2014 Conference p720-723.
[6] Amjed Naser M AL-Hameedawi IOP Conference Series Materials Science and Engineering · March 2020 742 012215
[7] Amjed Naser Mohsin AL-Hameedawi , International Journal of Civil Engineering and Technology (IJCIET), Volume 9, Issue 11, November 2018, pp. 1333–1351, Article ID: IJCIET_09_11_129
[8] O.Z. Jasim, K.I. Hasoon and N.E. Sadiqe, 2019 Engineering and Technology Journal, Vol. 37, Part A, No. 04, pp. 140-147.
[9] A.Z. Khalaf, I.A-K. Alwan and T.A. Kadhum, 2018 Engineering and Technology Journal, Vol. 36, Part A, No. 1, pp. 01-09.
[10] T. H. Shihab, A.N. Al-Hameedawi and A.M. Hamza, 2020 Engineering and Technology Journal, Vol. 38, Part A, No. 04, pp. 510-514.
[11] I. A. Alwan,, Z. W. Samueel, and Q. K. Abdullah, 2020 Engineering and Technology Journal, Vol. 38, Part A, No. 11, pp. 1580-1592.
[12] "Wassit Governorate Profile". NGO Coordination Committee for Iraq (NCCI). December 2015. Retrieved 30 April 2017.
[13] General Authority of Survey , December 2007.
[14] M.S. Hussain, I.A. Alwan and T.A. Hussain, 2020 Engineering and Technology Journal, Vol. 38, Part A, No. 04, pp. 515-522, 2020.
[15] Feng, M., Sexton, J. O., Channan, S. & Townshend, 2015 J. R. Int. J. Digit. Earth 9, 113–133
[16] Yamazaki, D., Trigg, M. A. & Ikeshima, D. 2015 Remote Sens. Environ. 171, 337–351
[17] Richards J.A. (2013) Supervised Classification Techniques. In: Remote Sensing Digital Image Analysis. Springer, Berlin, Heidelberg, pp 247-318
[18] Amjed Naser Mohsin AL-Hameedawi , International Journal of Civil Engineering and Technology (IJCIET), Volume 9, Issue 11, November 2018, pp. 1333–1351
[19] https://govinfo.library.unt.edu/cpa-iraq/regions/south-central/provinces/english/wasit.html