ARCH 43115: Summative Assessment

Automatic detection of archaeological sites (mounds) in Iraq

Introduction

The paper explores the use of satellite imagery (multitemporal – from various years and multisensory – various imagery types, such as optical (the multispectral type) and obtained from SAR – Synthetic Aperture Radar) within the field of machine learning, to automatically detect a specific category of archaeological sites in Iraq – the mounds (or *tells*), in the Kut area.

In Iraq, they are hills created by human activity throughout millennia (Wilkinson 2003: 100), although they also exist in other areas such as parts of Europe (ibid.) and of Asia, where automatic detection of mounds was already explored, for example regarding mounded sites of around 2500 years BC in Pakistan and India (Orengo et al. 2020). Returning to Iraq, various studies have explored the automatic detection of mounds, starting as early as the 2000s, when Ur et al. (2006), using digital elevation model data from SRTM (Shuttle Radar Topography Mission) and from ASTER mission, and also multispectral satellite imagery (from the LANDSAT, automatically detected mounds in the Khabur region (northern Iraq), through a machine learning approach – the Random Forest (RF) classifier. Similar methods were applied to automatically detect mounds in other parts of Iraq, such as in the south (the Wasit area), where they also used another type of satellite imagery, Top-of-Atmosphere (TOA) type (Tapete and Cigna 2022). However, they did not seem to use the RF method to create categories in the study area to recognise mounds (ibid.).

However, why do automatic detection? A suggestion can be found in GarciaMolsosa et al. (2021), where both deep learning and machine learning methods were used to automatically detect archaeological features, from historical maps instead of satellite imagery in the Middle East and South Asia. For South Asia, a very large number of maps with presumably very many archaeological features on them to be identified provided a response to this question – because of the difficulties associated with working with large quantities of data For example, to automatically detect features on maps means that the recording of such features on the said maps is much faster than when doing manual recording.

Methodology

For this paper, I have adapted the methodology used by Orengo et al. (2020) to automatically detect mounds in Pakistan and India – mentioned above. I have used a SAR type of imagery from Sentinel-1 mission, from 2014 to 2020. This type of imagery was used in both its form of collection – when the satellite creating the images moves to the south of Earth (descending) and when the satellite moves to the north of Earth (ascending) in the following modes of signal transmission and of signal receiving by the satellite – vertical/vertical (VV) and vertical/horizontal (VH) in the interferometric wide mode of collecting swaths of the Earth's surface on the imagery. I was interested on imagery taken over two seasons – summer and winter, combined into a single composite image with Sentinel-2 mission multispectral imagery from the same years (2014 to 2020), while also removing cloudy imagery for a better visibility of the sites.

After filtering the imagery to the area of interest around the city of Kut, roughly 2907 square kilometres, I proceeded with the machine learning approach. I used the Random Forest (RF) classifier to do an automatic classification of the area's surfaces, to prepare for the automatic detection of specific features in the study area, that is, mounds, which can be roughly explained as running a Google Earth Engine (GEE) script on known mounds, for the classifier to recognise the shapes for subsequently identifying the mounds in the landscape.

For this, I selected two collections of training data, which were parts of the database of mounds found by me manually. The training data consisted of 24 mounds, from 100 mounds mapped manually by me in the whole area. While such a number of mounds (100) may be perceived as small for the entire area, it can be suggested that when working with such a large area, I missed some of the mounds, on one hand and on the other hand, I also tried to map mounds that had representative shapes for the classifier to be trained on. The study area comprises both marshes and land, and most mounds mapped by me were in the western part of the study area (not in the marsh), where I found 24 mounds that seemed to have representative shapes on which to train the classifier.

Although it was suggested that 24 features as training data might not represent a significant sample for the entire area – personal communication: William Deadman, it seemed that more training data (around 50 elements) exceeds the capacity of the program. This issue may be corrected, but this situation appeared now in the first tries, so this can be considered in the future.

After doing the classification, since the entire study area was very large, I decided to run the classifier on a smaller area within my study area to test for automatic identification of mounds, while also checking the quality of identification – how many mounds the script identified and of these, which were correctly identified as mounds and which mounds were not identified as such but ought to have been. The smaller area, which comprised mounds from the database, but not from the training data, covers around 61 to 62 square kilometres, and included some pairs of mounds and some isolated mounds. When classifying both the larger area and the smaller area, the mounds were filtered to a 98% probability threshold of correct identification as mounds and the "focalMode" technique to ignore isolated pixels (a larger number, 5 instead of 3, for example – personal communication: William Deadman - removes smaller clusters). Since the size of the classification raster layer of the entire area was too large to be exported (and probably would have taken hours to export), I decided to only export the smaller area to QGIS to compare between the two types of mound identification – automatic and manual, and to check, based on this comparison, if the automatically detected mounds were the same as the ones manually detected.

Results and discussion

It seems that 111 clusters of pixels (and isolated pixels) representing mounds were detected by the GEE script (I counted them after exporting the layer with the detected mounds to QGIS. Also, although I could export the manually mapped mounds as vector layers, I could not export the GEE-detected mounds as such, so I exported it as an image, and created a new vector layer to manually mark the mounds on this layer.)

However, when starting to explore how the GEE script detection fared in contrast to the manual detection, I should acknowledge that the manual detection can also have its limitations. For example, I mapped 33 mounds in the sample area. This number should be put in context regarding the presence of the mounds in that specific area. Maybe mounds were missed because of fear to map features that were not mounds, while also the possibility exists that some non-mounds features were mapped as mounds. Regarding this, it seems that the smaller area, although having quite a few shapes similar to a mound (roundish, sometimes with the mound's edge visible), it also seems to have features that seem too flattened on the imagery, covered with vegetation and buildings or both, which may lead to not being immediately apparent that they are mounds. In the area, there are also what can be described as traces, on the ground, of features that might be mounds (or parts of them). These traces are quite numerous, but unfortunately it could be a matter of interpretation if they are connected to mounds. There is also the issue, that may be connected to the issue of visibility above, that when the mounds are not clearly defined in the landscape and they may share the colour with other features around them, which may lead to either miss to identify the mound or identify something else as mounds.

Returning to the GEE-detected mounds, it should be also taken into account that these represent some of the first tries of automatically detecting mounds using this script. The number 111 should be also put into context. Firstly, analysing this number starts from checking the probability of detecting a mound and the use of "focalMode" – a higher degree of probability means the more likely will the script detect something that it "knows" it is a mound, while a greater number as parameter for "focalMode" might mean that smaller, more difficult to detect features will appear as individual features, instead joining a larger cluster. For example, in these 111 features, some of them which might be isolated pixels, given the quite low value of "focalMode", although a benefit of this low value seems to be in identifying isolated mounds when they exist as such in the landscape. Also, in at least a case of clustering of pixels, the GEE-detected mounds seem to form a large cluster made by several manually mapped mounds, which do not cover it entirely. Others follow the outline and position of the manually mapped mounds. However, there are cases, as with the cluster mentioned above, when a larger GEE-detected mound covers several manually mapped mounds, although this does not seem to be the norm. Also, there are quite numerous situations in which manually detected mounds seem to be located close to, but not overlapping GEE-detected mounds. Finally, there are isolated pairs/small groups of pixels, which sometimes detect smaller mounds. In numbers (noticed by me), although this is a rough estimation) that is:

- Overlap between pixels that seem to cover a single, quite small mound (where the manually mapped mound are generally within the GEEdetected mounds) – 4 mounds
- Overlap between pixels that seem to cover a single, quite small mound (where the manually mapped mound cover part of the GEE-detected mounds, but without complete overlap) – 4 mounds
- Larger GEE-detected mounds that cover several manually detected mounds: 3 larger such mounds covering respectively 2, 4, and 2 (at this last can be said that is almost 2, while for the one covering 4 manually detected mounds, there is one manually detected mound almost invisible under the large cluster, which would mean it covers 5 manually detected mounds)

- Another category is when the manually detected mounds are close to the GEE-detected mounds, which can be split into several sub-categories, regarding closeness:
    1) "glued" one to another: 4 situations, in which one GEE-detected mound is very close to another manually detected mound (one of which includes also the case of a GEE-detected cluster covering other manually detected mounds)
    2) some distance between them: 3 situations in which a mound detected by a method is quite close to another detected by another method
- There are also some manually detected mounds that are too far away from the GEE-detected ones – 4 situations, implying that the GEE-script did not pick them as mounds
- Finally, there are the remaining GEE-detected mounds which do not seem to be connected to manually-detected mounds

Conclusion

This paper explored the creation of a GEE script to do automatic detection of archaeological sites (in particular, mounds) in Iraq, around the city of Kut. However, although it is difficult to comment on how well the script performed (if the GEE-detected mounds fit the manually detected mounds), there are several future actions that will help improve the script's performance:

It may be that some features from the training data did not appear in the manually detected-mounds dataset in the smaller test area. Also, probably more than 100 mounds should have been mapped manually in the study area, which is large. Although it may be suggested that I have missed mound features in the sample area due to the fear of mapping the wrong feature, more care will be needed in choosing an area which has more mounds, although in my case, aside from the areas from which the training data came from, the rest of study area comprises dispersed mound features (and then I come back to mapping more mounds in the large area), with large gaps between them.

Before turning to the testing of the script itself, and arguably related to it, it can be suggested that maybe before advanced testing, the script can run on data similar to the training data provided – that is, in an area having shapes similar to these on which the script was trained. What I think is a necessary step in the future to be taken is to check the validity of the GEE-predicted mounds through an analysis of false positive/negative (incorrectly identified as mounds/missed identifying the correct feature), either in QGIS or GEE, which was not carried out for this script and which is also related to the manually mapped mounds. This analysis can establish percentage of correctly and incorrectly identified mounds both by the script and by manual detection in a certain area (personal communication William Deadman).

Therefore, in order to improve the script's performance, probably more mounds with known location should be mapped in this area, then choose a smaller area within it which has the most mounds and run the script on it, after which an analysis of correct identification of mounds by both methods should be taken into account.

References

Garcia-Molsosa, A., Orengo, H. A., Lawrence, D., Philip, G., Hopper, K., & Petrie, C. 2021. Potential of deep learning segmentation for the extraction of archaeological features from historical map series. Archaeological Prospection, 28(2): 187-199.

Menze, B.H.; Ur, J.A.; Sherratt, A.G. 2006. Detection of ancient settlement mounds. Photogrammetric Engineering & Remote Sensing, 3 / March 2006: 321327(7)

Orengo H.A. Conesa F.C., Garcia-Molsosa A., Lobo A., Green A.S., Madella M. and Petrie C.A. Automated detection of archaeological mounds using machinelearning classification of multisensor and multitemporal satellite data. PNAS 117 (31) 18240-18250

Tapete D. and Cigna F. 2022. Detection, Morphometric Analysis and Digital Surveying of Archaeological Mounds in Southern Iraq with CartoSat-1 and COSMO-SkyMed DEMs. Land 2022, 11(9): 1406.

Wilkinson. T.J. 2003. Archaeological landscapes of the Near East. The University of Arizona Press