**Task 8**

1. Considering my general interest in the environment, I wanted to focus my studies on addressing this sector in Malta. Knowing how small the island is and how quickly the urban growth is reducing the amount of resources we have, particularly agricultural land, is concerning. Nonetheless, it is evident that most of the land is neglected and serves no purpose. Fortunately, I have contacts with personnel of ARPA, a driving force in maximizing the benefits of the Common Agricultural Policy, who shed some light on the need of exploring ways to realize and utilize these neglected lands for beneficial purposes in the hopes of being cultivated. During the initial phase of my research, I observed that there had been little to no studies on Maltese agricultural land, particularly in terms of assessing changes through time and uncovering the possibility of non-arable land to be cultivated. This led me to validate how the usage of remote sensing in conjunction with data given by ARPA from their land inspections may reveal the aforementioned potential.

2. The initial choice of method was identifying the study area for the research to be conducted on. A small area covering the limits of Imgarr was originally the focus. The chosen algorithms to be employed were Multitemporal Random Forest, K-means, Support Vector Machine, and Temporal Convolutional Neural Networks. The initial timeframe for using the downloaded products as a dataset was anticipated to be between 2019 and 2021. The collection of ARPA's datasets to be used for the executed algorithms was also one of the initial method choices in conjunction with some of the algorithms mentioned.

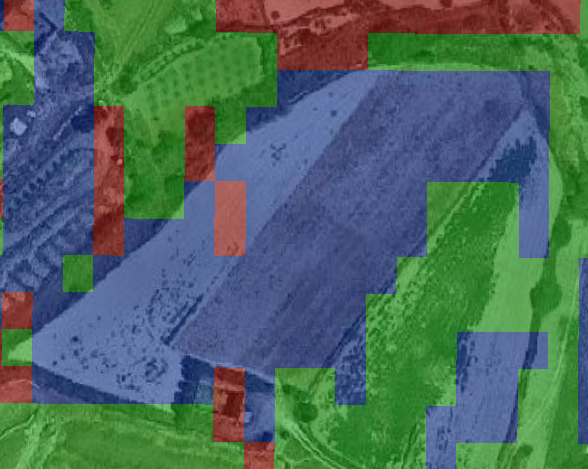
3. After obtaining and reviewing ARPA's dataset, it became clear that the targeted location of Imgarr did not have a substantial number of uncultivated land parcels. After evaluating various studies and reading through my literature review, it was decided to limit the study area to the limits of Fawwara, Siggiewi, Qrendi, and Haz-Zebbug, as the distribution of agricultural holdings by locality has a higher quota around that area, which would aid this study's purpose by having a larger dataset for better accuracy once the necessary algorithms and machine learning technique were initiated. Another set of images was downloaded for the years 2015-2017, with the justification that Sentinel data is available from 2015 onward, and a wide variety of products was assessed, from the oldest to the most recent. In this approach, the procedure is to discover changes over a longer period of time. It was also decided to conduct a monotemporal random forest classification, with the two mono-temporal images of dates 26/08/2017 and 11/09/2021 being used since ARPA established that the months of September and October precede the period when land is plowed in the winter. The majority of studies also indicate that the Random Forest classification produces the best accuracy and is preferred by the majority of researchers, followed by the K-means clustering. As a result, I chose to use them as the algorithms to be executed.

4. The objective of this research was to compare the performance of three algorithms, Random Forest classification as monotemporal and multitemporal, and K-mean clustering, when applied to agricultural land mapping using Sentinel-2 time-series data. The use of these algorithms was to confirm which gives the highest rate in accuracy, precision, and correlation in recognizing and mapping the vegetation in the determined study area. The multitemporal random forest classification was tested with two sets of datasets, as the first set produced somewhat contradictory results when compared to Google Earth Pro’s imagery. As a consequence, a minor adjustment was made to expand the dataset in all three classifications of *arable*, *non-cultivated*, and *other*, in order to generate a more accurate result. Some examples below show the changes in the results from the first and the second datasets:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **2020-2022 Random Forest Classification results**  (Green: Arable, Red: Non-Cultivated, Blue: Other) | | | | |
| Manually classifying Vineyards as Cultivated after confirming the land is still being cultivated with Google Earth Pro’s latest imagery (5/23/2021) | | | | |
| Dataset 1 | Dataset 2 | | |  |
| A picture containing text  Description automatically generated | A picture containing text, green, bedclothes  Description automatically generated | | |  |
| Result without manually annotating the vineyards as ‘Arable’ | Result after manipulating the dataset by annotating vineyards as ‘Arable’ | | | Google Earth Pro (5/23/2021) |
|  |  | | |  |
| Manually classifying trees as non-Cultivated after confirming the land is still with trees with Google Earth Pro’s latest imagery (5/23/2021) | | | | |
| Dataset 1 | | Dataset 2 | |  |
|  |  | | |  |
| Result without manually annotating trees as ‘Non-Cultivated’ | Result after manipulating the dataset by annotating trees as ‘Non-Cultivated | | | Google Earth Pro (5/23/2021) |
|  |  | | |  |
| Other changes after manipulating the annotations of the classes | | | | |
| Dataset 1 | | | Dataset 2 |  |
|  |  | | |  |
| Land marked as ‘Other’ | Land marked as ‘Cultivated’ | | | Google Earth Pro (5/23/2021) |

After establishing that the multitemporal random forest classification produced the best results with the highest accuracy of the three algorithms tested, a land cover change algorithm was used to output the metrics of the resulting images, depicting the differences between the two sets.

5. The following are some of the limitations I encountered while working on my dissertation:

* The downloaded products required to be cloud-free and glare-free in order to not interfere with the results of the algorithms. This made obtaining adequate images in alternate months problematic, and as a result, the primary objective of downloading 12 images per year had to be reduced to 11 images.
* ARPA’s dataset did not classify land which was covered in trees. As a consequence, certain lands with trees were classed as arable, while others were classified as non-cultivated. I altered the classes in the second dataset by adding new parcels and label such areas as non-cultivated.
* I fully distinguished the class 'Others', where I indicated the roads, buildings, and quarries. When the multitemporal Random Forest classification method was executed, it was discovered that a portion of land was mistaken for either a quarry or the roof of a structure, and therefore was categorized as 'other,' as seen in the example below (Blue represents the class 'Other'): Map

  Description automatically generated A close-up of some paintings

  Description automatically generated with low confidence

This led to false prediction once the Random Forest Classification was executed.

* Google Earth Pro's most recent imagery extends up to May 23, 2021, implying that ground truth data could not be determined up to the present.

6. This study falls under ecological and temporal generalization/validities. Based on the findings of my study, organizations such as the Maltese agency ARPA would benefit in providing the Commission, local entities, and the farming community with accurate and timely information. This opportunity allows them to build on the study and ensure that land is not neglected and that resources are used as efficiently as possible. The findings can also be applied for future research related to agricultural land, for other researchers in the remote sensing community to know what material and algorithms to use to obtain data on land changes throughout a period of time. A better dataset of land parcels would refine the random forest classification algorithm's output findings. When employing this technique, downloading additional satellite photos would have been more beneficial.

References:

NSO, 2022, Census of Agriculture 2020, Available at: <https://nso.gov.mt/en/News_Releases/Documents/2022/02/News2022_015.pdf>, [Accessed 8 February 2022]