Temporal machine learning for small islands LULC and vineyard mapping using Sentinel-2

Ismael Scicluna

Supervisor: Mr Daren Scerri

June, 2022

A dissertation submitted to the Institute of Information and Communication Technology in partial fulfillment of the requirements for the degree of B.Sc. (Hons.) Software Development

Authorship Statement

This dissertation is based on the results of research carried out by myself, is my own composition, and has not been previously presented for any other certified or uncertified qualification.

The research was carried out under the supervision of Mr Daren Scerri.

Ismael Scicluna

June 6, 2022

Copyright Statement

Ismael Scicluna

June 6, 2022

Dedication

I would like to dedicate this dissertation to my late grandpa Alfred.

Acknowledgements

I would like to express my gratitude to Mr.Daren Scerri for his mentorship and contribution which without him , this research wouldn't have been possible.

I would also like to thank my loving family and friends for the constant support.

Abstract

Remote Sensing (RS) is the process of detecting and monitoring products of geographical areas from long distances, such as those produced by satellites and aircraft. This study makes use of Satellite Image Time Series (SITS) for the classification of land use and land cover (LULC) and crop classification. A large custom dataset was built making use of 9 Sentinel-2 images, 10 spectral bands and over 400,000 data points. while freely available and with a high revisit rate Sentinel-2 MSI imagery is medium resolution (10m), which is a challenge in a small island scenario. Multi-temporal Random Forst (RF) and Temporal Convolutional Neural Networks (TCNNs) were then employed to classify land. Consequently a temporal NDVI stack was used to better identify vineyards. The proposed system was then evaluated with a 2 stage experimental design; with stage 1 evaluating all LULC classes and stage 2 to classify and differentiate vineyards from other crops. Stage 1 experiment results show that the TCNN solution trained on the 10 band dataset achieved an accuracy of 98.39% which clearly outperformed the temporal RF (80.6%) solution using the same dataset. In stage 2 experiment, TCNN solution making use of NDVI SITS datasets proved to be more accurate than either TCNN with 10 bands datasets and temporal RF. A TCNN trained on a robust dataset has the potential to assist to a near accuracy LULC reporting which is critical to small islands like Malta. Moreover, the same technique can be employed to classify crops like vineyards. It is suggested that a first model is used to extract an agricultural land mask using the highly accurate LULC TCNN solution implemented in Stage 1 of this study. This mask can then be fed into another classifier using a dataset of only crop types and a temporal NDVI stack as data to address better the crop classification challenge.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The current state of Remote Sensing(RS) usage on the Maltese lands is currently very limited, while the supply of sentinel-2 imagery of the Maltese lands is available with a very updated platform providing images bi-weekly. Given the current works that are already out there on Maltese LULC, it is difficult to use the works to do accurate studies since the work done is not updated and not as accurate.

Modern day classifiers aim to solve this problem in conjunction with the mid-resolution sentinel-2 imagery . These provide pixel-based classification on an image or multiple images. The continuous struggle for innovation in the technology sector is always beneficial to such studies as classifiers in this day and age are getting much more accurate and benefiting from scientists to the everyday worker. Such classifiers improve the work being performed not only on imagery but in anything that has a pattern which can be noted, quantified and analyzed.

The most popular classifiers used for image classification are the Random Forest(RF), Convolutional Neural Network(CNN) and the Support Vector Machine(SVM) where the said classifiers analyze an image and classify the image in different patterns, then based on what the patterns analyzed taught the classifier, the classifier then is able to identify such patterns and give you the result you need based on the model trained when analyzing the image. This would mean that an LULC map would be available to anyone who needs using the aforementioned classifiers.

1.1 Purpose statement

The purpose of this research is to increase our islands awareness and increase the knowledge of the authorities concerned while also increasing the accuracy of urban planning to become more effective given the state of rapid development that is taking place in the Maltese islands. One might also instigate more research to help farmers calculate their average crop yield by using the amount of greenery from a satellite image.

1.2 Hypothesis

Given the above purpose statement, the hypothesis of this study is that by using satellite products that monitor the visual spectrum wavelengths, it is possible to identify different crops. Following on this hypothesis, three research questions were outlined :

1. - How are crops currently monitored in the Maltese Islands?
2. - Can Multi-temporal classification using neural networks be used to classify different crops in a given area?
3. - How can cropland be segmented from other land?

1.3 Motivation & Significance

The motivation for this research relies on the state of current research on the Maltese lands. Such a study provides the authorities and interested parties with a more precise

LULC map for these lands. It can be a process which can be done repeatedly to further increase the accuracy and always having an updated LULC.

Moreover, the algorithm used might also be useful for other studies in the RS sector. The data collected and processed in conjunction with the classifiers used can be used for areas like crop yield predictions and tracking of phenological cycles.

1.4 Research Outline

In this study, following the introduction, a literature review was conducted(Chapter 2) where the foundation for concepts used throughout this research such as machine learning and remote sensing. Multiple papers were also studied thoroughly to see the process in which one uses different classifiers and machine learning techniques to further improve the classification process of areas of land. That would then help in identifying what to do and what not to do when conducting the prototype.

In the methodology(Chapter 3), the tools used and the prototype pipeline was presented and explained. The steps performed to process the custom dataset and train the models were outlined step by step and the python scripts used to rectify the dataset for the classifier were shown.

The findings, analysis and discussion(Chapter 4) section outlines a description of all the results that were extrapolated from the prototype while also including multiple results other than the accuracy to highlight the research.

Finally, the conclusion(Chapter 5) includes a summary of the study performed and the results while also including some suggestions for the future to further enhance this research.

Chapter 2

Literature Review

2.1 Land Use and Land Cover Significance

With the current population growth and the need for more agricultural sites to keep up with our needs, crop classification is becoming a more important topic everyday. This also increases the need for reporting the state of land to different entities including agronomists, agricultural agencies and companies which utilize the crops for business use. (Ji et al. 2018, Vuolo et al. 2018, Paris et al. 2020, Song et al. 2017, Pelletier et al. 2019, Fan et al. 2020). Given that the gap between the current land cover available is years away from the present, it is hard to report the development and all the areas which have changed over the years, with more detailed land use and land cover(LULC), it will be easier to see the difference from year to year and the rate of development for the country of Malta. Once a dataset is available with updated data, this can help with urban planning which is a process in which one is focused on developing a plan for the use of land to help keep a good level of vegetation.

The environmental aspect to this would be for the amount of cropland that is currently being used and what amount from that is being used for vineyards. The difference can also be clearly seen between the built up areas and the cropland in the corine land cover below. According to the European Environment Agency(EAA) the current land cover is as shown in the table below.

Land cover type	Area (km^2)	%
Agricultural areas	161.5	51.2
Urban areas	70.4	22.3
Forested areas	2.1	0.7
Coastal wetlands	0.3	0.1
Natural vegetation	57.8	18.3
Industrial and commercial units, mineral extraction, airports, port areas, dump sites, green urban areas and sports and recreational facilities	23.31	7.4

Figure 2.1: Maltese Land Cover Percentage 2018

When researching the current land statistics that Malta has, one finds the website provided by the Maltese planning Authority provides a corine land cover that was last done in 2018 with very little detail and a gap of 6 years between every land use classification. This image doesn't show an accurate depiction of the Maltese lands today as in the previous years development struck with a huge margin which accelerated the growth of built-up areas thus a more frequent update would be needed rather than one every 4 years.

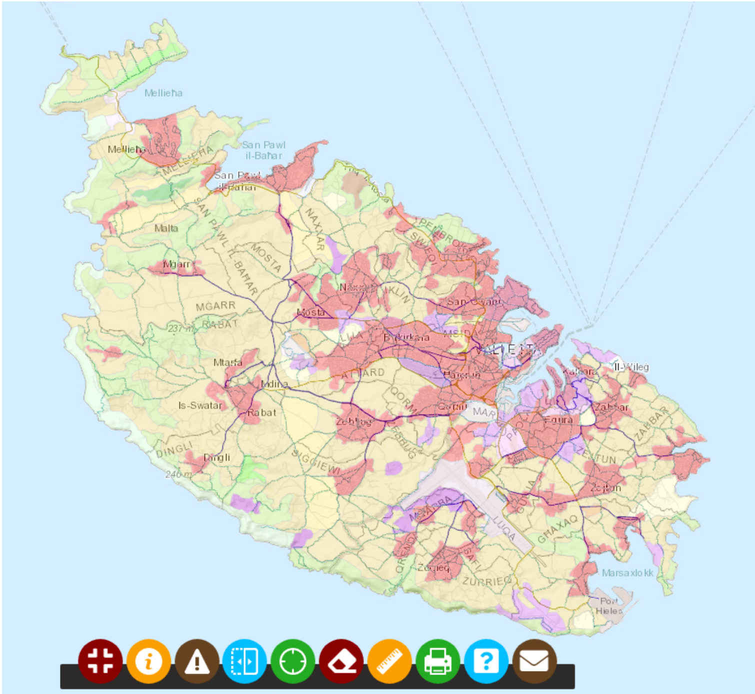


Figure 2.2: Maltese Land Cover Data 2018

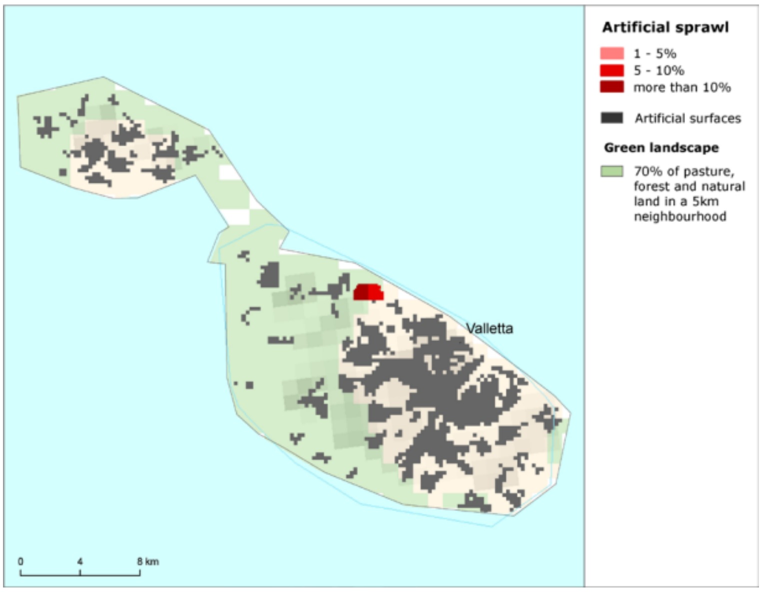


Figure 2.3: Maltese Land Cover Data 2012

2.2 Remote Sensing

With Regards to dataset, it is now possible to do such research using images from Sentinel-2 which have a $10m^2$ resolution that is globally available Mazzia et al. (2020). The Sentinel-2 is beneficial to such studies because of its high resolution and it also has a high re-visit rate with images of the Maltese islands every 3 days or less. One of the main objectives of Sentinel-2 was to introduce observation data which would provide next generation products to collect LULC and the detection of change in land.

The RS images were very useful in gathering the data necessary for such a study, RS itself has many application including land cover classification, atmospheric data, Marine monitoring, emergency monitoring and climate change amongst multiple other applications which can be derived from such research Gandhi et al. (2015). Such a study would help Malta and the ERA collect data regarding crops in Malta. Once the data is collected, statistics can be formed and utilized to help with EU funds allocation and agricultural planning. The bands that the Sentinel-2 satellite shows are the ones below.

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

Figure 2.4: Sentinel-2 Bands

These bands are usually all used except for the first one which is air pollution for classification of crops and LULC. There are some studies in which NDVI is used, thus only needing 2 out of the 11 bands which are the NIR(Band 8) and the Red(Band 4) but is usually used to compare the difference and how using all the bands is more beneficial rather than just using two of them. (Ji et al. 2018)

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

2A products which are images downloaded from the Sentinel-2 which is Fig. 2.5 are the images that contain the bands mentioned above. The resolution of 2A products was essential and since they are not as available as 1C products. When 2A products are not found, using Sen2cor, 1C products were converted to 2A products. Sentinel-2 has great potential for land state classification and it was key factor throughout. This research would not have been possible or feasible without Sentinel-2 due to the amount of resources needed to launch a satellite into orbit or taking high resolution images from an aircraft in the sky.

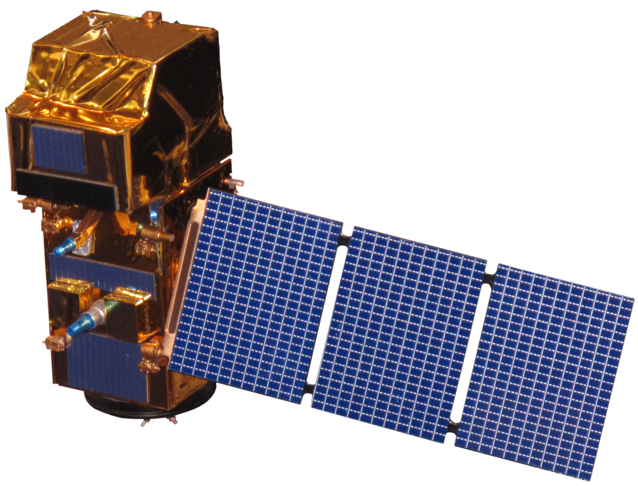


Figure 2.5: Sentinel-2

When looking at RS and how RS is used for LULC and crop classification, one must not forget to mention the classification techniques which are what makes most of the research regarding RS possible. One of the most used classification techniques is RF, this is mostly used to compare with other algorithms highlighting the improvements that new classification techniques have made in the past years. One of the most used technique to classify crops and LULC is the CNN Ji et al. (2018), Vuolo et al. (2018), Song et al. (2017), Pelletier et al. (2019), Carranza-García et al. (2019), Pena et al. (2019), it is mostly used because of it's high accuracy rate and most of the papers work on efficiently designing CNN's to increase the accuracy of such classification techniques.

2.3 AI and GIS

Computer vision is widely used and also plays a big role in this project. This technology is a method where the computer is trained to interpret visual information and classify accordingly.

Artificial intelligence is widely used in RS scenarios. Multiple researchers used Neural networks, specifically CNN to do such work. Time weighted dynamic time warping(TWDTW) also has high accuracy, upwards of 85% with a 95% confidence interval Belgiu & Csillik (2018). RF is also one of the most used algorithms especially to compare the accuracy of other algorithms to it.

Geographic Information Systems(GIS) allows us to capture, analyse and present spatial information on a map. It is widely used in marketing and environment uses.

2.4 Studies in land use and crop classification

2.4.1 Random Forest compared with Improved classifiers

In the study performed by Mazzia et al. (2020) it is stated that the use of temporal imagery and deep learning helps increase the accuracy of crop classification, it was outlined that the Overall Accuracy (OA) was increased by a huge margin of more than 15% when compared to other non-deep learning classifiers like SVM, Kernel SVM, RF and XGBoost. Apart from the deep learning, the multi-temporal images alongside the deep learning techniques in RS made it easier and more flexible to increase the amount of information gained from an image and made it easier to accurately classify land.

Following the research of Belgiu & Csillik (2018), TWDTW was evaluated to see it's use and performance when used on pixel-based and object-based classification to classify different crops. The research was performed in a total of three different areas which include the USA, Romania and Italy. Once the results were outputted, they were compared to the result of the RF and the evaluation was done. The NDVI was extracted from the images to generate a time series for the classifiers to compute on. When seeing the overall results, although similar, the object based classification performed better than the Pixel-based classification, and performed significantly better than the RF. It was also noted that the RF performed better when the phenological cycles of crops overlapped the time series. It was noted that the TWDTW has multiple applications to automated cropland mapping without the need for human intervention.

What Carranza-García et al. (2019) did was different than most studies. Rather than comparing images from the same sensor, different images were compared with a different spatial resolution and a different amount of bands. The CNN that was proposed by the research outperformed all the other classification models like the SVM and the RF by a huge margin. When seeing the results, it can easily be noticed that in two out of the three areas, the CNN outperformed the other classification techniques by more than 5% and in some cases it even outperformed the others by more than 10%. It was noted that the benefits of the CNN do not stop at the high accuracy it achieved but also its computation time and its ability to adapt to different scenarios and datasets. The computation time of the CNN was quite fast but this was due to it being performed on the Graphical Processing Unit(GPU), when the computation time was compared with all of the classification techniques using the Central Processing Unit(CPU), the performance was fairly similar between the CNN and SVM but the RF outperformed them both by a huge margin. In conclusion, the framework of the CNN proposed by this research performed very well and managed to achieve highly accurate results.

Pelletier et al. (2019) conducted a research study on the use of Temporal Convolutional Neural Network(TCNN) for crop classification and its benefits when compared to classification techniques like Recurrent Neural Network(RNN) and RF. The data processed was a Satellite Image Time series(SITS) which was a key component to evaluate this research and have the information available to be analysed and classified properly and accurately. Formosat-2 was used with 46 images that had more than 1 million pixels annotated and a total of 1419 polygons annotated. These were split into a 60/40 configuration. When analysing carefully the results, it was noted that the TCNN performed better than the RF and RNN by an average of 1-3%. It was also noted that using NDVI and other indices as a basis for the training/testing declined the performance of the TCNN as opposed to using all the data/bands available.

What Immitzer et al. (2019) noted was a method in which one can detect different tree species, what they tried to do was classify different tree species using multi-temporal RF and Sentinel-2 data in which they managed to achieve an overall accuracy of 96% which is very accurate considering the different amount of species of trees that were classified. One must also note that all the class-specific accuracies were more than 90%.

In the results of the land cover classification, where only 2 forest classes were classified amongst other land states, both of the forest classes boasted an accuracy higher than 93%. What was done afterwards was a classification in an area where there are only trees, in which all the 12 different species of trees were classified. In this study, indices and normalized differences such as NDVIHuang et al. (2020) were used thoroughly and were an important factor in the high accuracy amongst the different tree species. The indices that increased the accuracy the most were the ones based on NIR and Red bands and the ones based on Green and Red, although not the only ones used. In conclusion Immitzer et al. (2019) stated that the Sentinel-2 bands make way for a very complex and inexpensive workflow to produce such classifications.

2.4.2 Improvements to current classifiers

The study by Vuolo et al. (2018) implemented an algorithm in which they researched the accuracy variance in classifying crop types and the effects of multi-temporal data including the effect of crop growth. When looking at this research, it can be established that during the crop's peak development, it is easier to achieve high overall accuracy of data especially when compared to the development stage of the crops (March-April), the difference in the accuracy is more than 40%. Furthermore, difference can be seen between the two stages where the crop is almost fully developed(May-June) and where the crop is fully developed(July-August). This accounted for 16% difference in accuracy in favour of the fully developed crop.

According to Paris et al. (2020), although the advances of deep learning uses in such research have made quite a huge leap, the accuracy is still heavily dependant on pre-processing and classifying the crops to train models. What this study tried to achieved was to extract a training set that is made up of publicly available thematic products, one of the other problems that was addressed was to generate a time series that is in conjunction with the pheonological stages of the crops thus having a higher accuracy when using all the bands, and especially when using indices. Long Short Term Memory(LSTM) was used as a network for the crop type mapping in this study. The method that was proposed managed to achieve an accuracy of 85.87%, Which is quite comparable when seeing other state of the art methods, it showed the capability of LSTM in reducing the imbalanced classification problems that exist. The aim of this study is to pave way to make use of automated pre-processing and Time series extraction techniques to take crop classification to a country wide scale.

2.4.3 The impact of NDVI

NDVI is widely used and it is still one of the most used index although there are others Huang et al. (2020). The index has its perks when iterating through the research performed by Gandhi et al. (2015). In this paper, NDVI was used to see the vegetation change detection and how NDVI can facilitate that in conjunction with GIS and RS. As can be seen, the change detection analysis proved to be very efficient to outline the changes that happened between 2001 and 2006.

When seeing the research study performed by Ji et al. (2018), one can easily determine the difference a 3D CNN makes for crop classification. The study tackles 2 different types of variables, the accuracy difference of crop classification when using a 2D CNN as opposed to a 3D CNN and the accuracy difference when using NDVI(2 bands) versus when using the original image with all the bands included. Both a 3D CNN and 2D CNN were trained and the 3D CNN increased the accuracy by an average of 3% on the 2 images from 2015 and 2016. The result was that a 3D CNN improves drastically the accuracy of crop classification while not using NDVI increased the accuracy by more than 3%. It was stated that the temporal aspect in crop classification and other similar modelling processes helps the deep learning drastically to improve accuracy.

2.4.4 The impact of pan-sharpening for crop detection

Looking at the experiment by Karakizi et al. (2016) where RS data was exploited to precisely detect vineyards and discriminate between different variety of vines, it was noted that very high resolution multi-spectral satellite imagery was acquired from the WorldView-2 (WV-2) satellite sensor. The WV-2 has a multi spectral imagery of 2m resolution and a panchromatic imagery of 0.5m spatial resolution which is significantly higher than Sentinel-2 imagery which has a spatial resolution of 10m and 20m (Mazzia et al. 2020). Being the first achievement to classify the vineyards from other crops/land, it was performed successfully with an average of 86% OA between the 4 test sites that the study was performed. From these testing sites and vineyards that were detected, vine canopy extraction using object-based Neural Network proved to be better than pixel-based SVM but not by a huge margin as both algorithms worked nicely with an average OA higher than 97%. The most challenging part was the vine variety discrimination in which Karakizi et al. (2016) used two different datasets, the multi-spectral data with 2m resolution and the pan-sharpened data with 50cm resolution, although the pan-sharpened data has a better resolution both had the same OA sitting at 61% between the 6 different varieties of vines with some different misclassification errors. It was noted that different classifiers and a fully-automated processing framework would maybe lead to better results as the resolution of the imagery doesn't have much room to be higher.

Jones et al. (2020) employed very similar spectral imagery to Karakizi et al. (2016) where in the study WV-2 images were used, both at 50cm panchromatic and 2m spatial resolution. In this study the impact of pan sharpening was highlighted and the results were very interesting. To test the impact of pan sharpening for vineyard detection, 5 different models were trained differently. M1 was trained with a panchromatic band only, M2 was trained using the panchromatic band and 8 multi spectral bands, M3 was trained using pan sharpened RGB, M4 was trained using R-RE-NIR1 as bands and the 3 of them were pan sharpened and finally M5 was trained using the panchromatic band and NDVI which was derived from the pan sharpened bands(Red and NIR). In this study, M2 seemed to perform the best out of all the models trained but M5 was very similar in terms of performance. With regards to pan sharpening for vineyard detection, it seemed that it has caused spectral distortions on the image thus reducing the ability to detect vineyards accurately. The use of NDVI had a quite similar performance to the multi-spectral model although being slightly poorer in performance. Taking into consideration this conclusion it can be stated that pan sharpening is better off not used in vineyard detection and the use of NDVI might also hinder slightly the accuracy of such a study.

2.4.5 The use of Data Fusion

Pena et al. (2019) noticed the need for data fusion in today's world of an advancing technology of RS. What was performed was a classification model on two different images and then another experiment but the deep learning was performed after the two images were fused in one to hopefully achieve a higher accuracy. The two datasets in question were collected from the Venezuelan Remote Sensing Satellite-2(VRSS-2) and the Google Earth(GE) images. Both of the datasets achieved a mid-high accuracy rate in which the VRSS-2 which is composed of RGB bands achieved an accuracy of 78.04% and the GE image which is also composed of RGB bands achieved an accuracy of 84.65%. Afterwards, the data was fused together and also added multiple bands like NIR and NDVI in the process. This managed to get a result of an OA of 93.17%, which is a significant increase in the accuracy when compared to the two individual datasets. This study was made with the intention as to increase the motivation for other researchers to take a look and study more the impacts of data fusion and what it could do for the future of RS. It proved to have a more than positive outcome between these two datasets and the possibilities of data fusion integrating lidar, radar and hyper spectral imagery.

Lanaras et al. (2018) took a different approach to deep learning using Sentinel-2 imagery, rather than using the highest resolution possible, the lower resolution bands(20m and 60m) were taken to form a uniform 10m data cube. With this method, the resolution of lower resolution imagery can be increased heavily, in-fact the impact of such a method improved the RMSE of this experiment by 50% when compared to the other methods available. It can be observed from the results they had a significant increase in resolution or 'Super-Resolution' as stated in the study. This was done with the hope of others using it and making use of the software on other multi-spectral sensors to increase the accuracy of deep-learning architectures. It was also noted that the computing times were adequate and are fast enough to be used in studies as a pre-processing step.

Aswatha et al. (2018) took Sentinel-2 data to create an unsupervised detection of surface mine sites, the research compared 3 different classification techniques on a study area in Asansol region of India. The three classification techniques were composed of spectral slopes and NDSV based one-class SVM classification with an accuracy of 82.5%, spectral slopes and stokes based one-class SVM classification with an accuracy of 21.2% and the last one was spectral slopes and stokes-NDSV based one-class SVM classification boasting an accuracy of 94.8%. The technique performed for the 3rd classifier was similar to Pena et al. (2019) where both the first 2 techniques were fused and their results proved to be quite accurate in predicting mine sites. This will hopefully be of use to help map variations in mine regions using multi-spectral imagery which can help in environmental sustainability.

Chapter 3

Methodology

This research is aimed at increasing Malta's LULC while also increasing the amount national index for vineyards through RS images that were downloaded from Sentinel-2 and classified using Artificial Intelligence. The process of classification was done by adapting the solution of Pelletier et al. (2019) to this use case and also was changed for this research to be performed on Sentinel-2, while also comparing the results to a classification using RF. The results extracted included the OA, precision and F-Score for each class classified.

3.1 Research Design

3.1.1 Qualitative

A qualitative approach focuses on the philosophy of a certain group or individual regarding a problem. This approach usually consists of a set of questions that are asked to such individuals or groups and afterwards the answers are interpreted. It helps build an inductive styled research to anchor ones meaning.

3.1.2 Quantitative

A quantitative approach is a method in which a theory is tested by means of an experiment, this experiment is quantified in variables and the relationship between the aforementioned variables is measured so that numerical results can be produced and analysed without any bias. A controlling explanation usually takes place to compare and replicate the findings.

3.1.3 Mixed Methods

When using both the methods described above, it is said to be a mixed-method approach Creswell (2013). This approach helps the study have a better overall grip on the data at hand thus the validity of such a study is stronger when using the two methods combined.

3.1.4 Justification for a quantitative research approach

For the purpose of this research, a quantitative approach was chosen. This approach was adopted due to the fact that in this research, the result is dependant on numerical attributes such as accuracy and computational speed which were essential to provide a positive result. It must also be noted that this study is not concerned with possibly subjective interpretations further increasing the reason for choosing a quantitative approach. This approach results in an extraction of multiple variables from the prototypes that were designed and a fair comparison between the two to further highlight the results obtained and their impact. The findings generated will outline the basis for RS and LULC for Malta. It can create a process in which the Maltese lands are thoroughly classified with a high accuracy and small room of error which can be performed frequently for the local authorities to have more data of land use to be used for multiple purposes including urban planning. This study will make way for other studies in which errors and possible limitations will be highlighted and can be focused on to improve such research in the future.

3.1.5 Experimental Variables

A variety of variables including control, dependant and independent were included in the study. For control variables, it was made sure that all the RS images collected all had a small to none cloud coverage to provide for a higher accuracy and less confusion during the model training as if there are clouds in the images in the same area in which a building was, the pixel of that building will be very different thus confusing the AI. The independent variable was the Time series of images provided, a total of 9 images were provided for the studies accuracy to be increased and possibly notice more the phenological cycles of crops which also leads to a higher accuracy as reviewed. The dependant variable in this case is the result obtained of the trained model and the inferencing performed on the lands of the Maltese islands which led to a classification map.

3.2 Prototype Pipeline

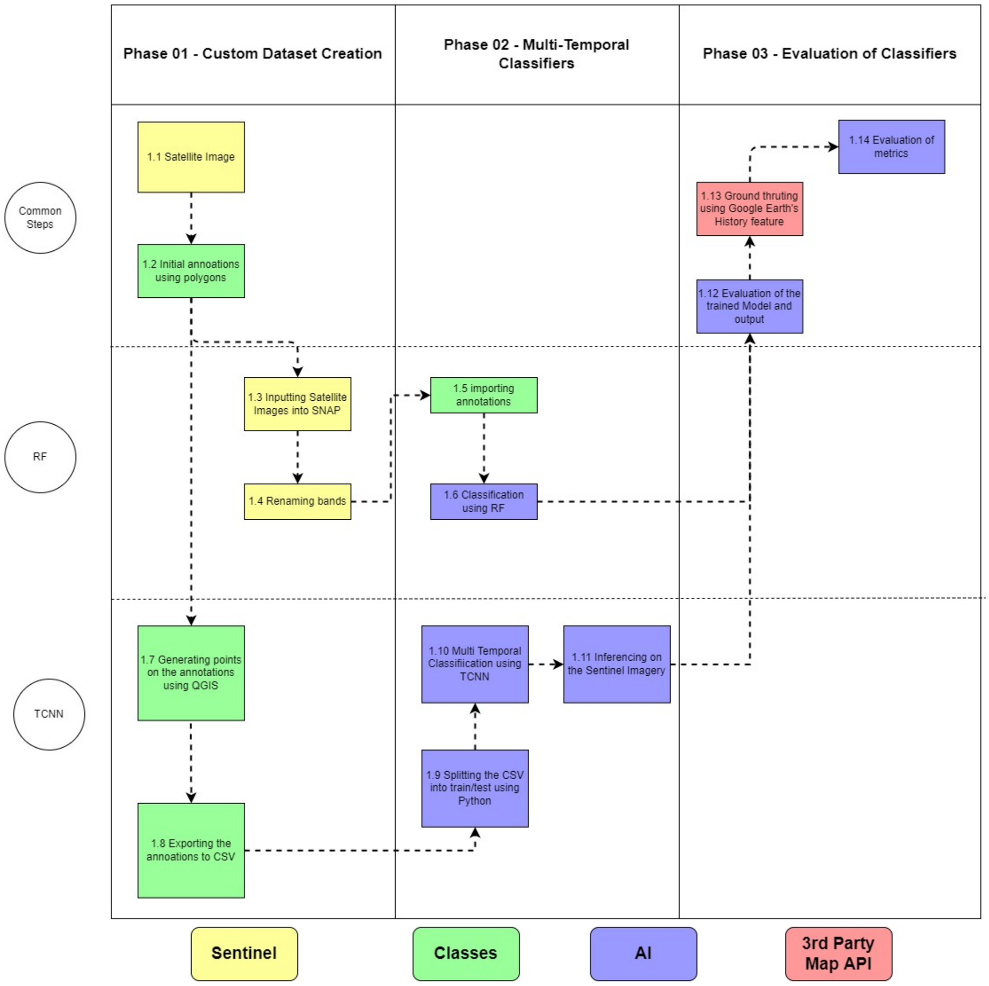


Figure 3.1: Pipeline

3.2.1 Aim of prototype

The aim of this study is to create an LULC of the Maltese islands while also creating a custom Annotation similar to the corine land cover, which highlights different types of lands and percentages of the current land use. It is especially focused on vineyard detection throughout the Maltese lands.

To provide a more accurate result while also seeing the accuracy of the algorithms, 2 different classification techniques were used. Furthermore different models were created to find the one with the highest accuracy possible and increase the validity of this study.

3.3 Multi-temporal Custom Dataset Creation

3.3.1 Sentinel-2 MSI data

Sentinel-2 2.5 was an essential part of the study. It provided mid-resolution satellite imagery with a high re-visit rate over the Maltese lands thus making it possible to do such a study by having a big data collection where one can analyse, download and create a dataset with. The dataset was created by first going into the Copernicus hub and analysing different years to see which year provided the most imagery without or minimal cloud cover. Once it was established that the year 2020 was going to be used, the satellite images were chosen and a total of 9 images with 0% cloud coverage were found.

The image below shows the time span difference between each sentinel-2 image.

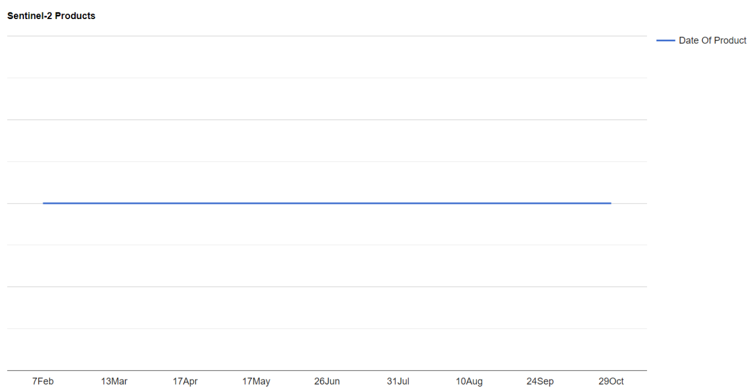


Figure 3.2: Date Of Products

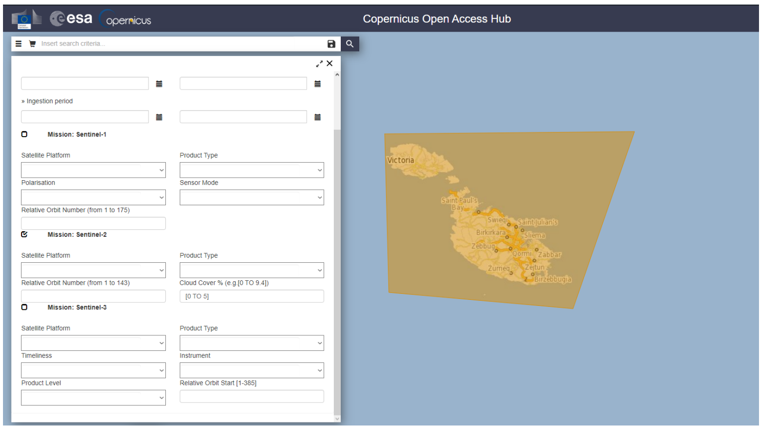


Figure 3.3: Copernicus hub

3.3.2 Custom Dataset Creation

The Dataset was then created by combining the 9 images into one using the bands needed for such a study. Bands 1,9 and 10 were removed from every image as they would add no apparent value to a classification but rather confuse it due to their poor 60m resolution as seen in 2.4. Instead of combining the images, the bands needed were combined into one image and renamed numerically after each other. Afterwards the combined multi-temporal image was cropped to include only the areas which are not sea.

3.3.3 Land classes and vector annotations

Once the images were analysed and downloaded, doing the annotations was the next step. The annotations were created using QGIS.

First one would input a satellite image for reference and the google map to check the area being annotated. This was done simultaneously with checking the area on google earth and using google earth's history function to ground truth the areas being annotated, some of the annotations which were not clear and from google earth were ground truthed by going there and verifying that they really are what the annotations are saying they are.

The annotations were done in form of polygons, every polygon contained a class and a polygon id. The class assigned to every annotation was based on the different type of land that one can find on the Maltese islands .The classes were Cropland, Forest, Grassland, Greenhouses, Other, Rocks, Settlement, Vineyards and Wetland. The polygon ID was used to identify the polygon, every class had a range of numbers in which the polygons had to abide by, this was done to be more organised while also not confusing the algorithm for the TCNN. This process was quite long as it needed to be highly accurate so the classify would be fed accurate data.

3.4 Multi-temporal Random Forest Classification

The Multi-Temporal classification using RF was performed by using SNAP.

The dataset created was imported into SNAP and the bands had to be renamed in SNAP for it not to confuse them. Once the images were in and renamed, the annotations(polygons) were imported into SNAP and it was checked to see that it overlapped the same areas. The annotations were then used to do the classification. The classification is then started using RF and once it is complete, an inference image is generated to show the result of the study including a confusion matrix with all the result scores and accuracies acquired.

3.5 Multi-temporal Convolutional Neural Network

1. The first step is to import the Custom Temporal dataset created into QGIS.

2. The annotations are then imported into QGIS and put on the overlay of the Custom Dataset.

Figure 3.4: Annotations

3. The Annotations were changed into points
   (a) Points are generated into each polygon
   (b) The points are joined by the location with the polygon attribute to retrieve essential information that was lost when generating the points
   (c) The points were merged into 1 class to be prepared for CSV export

Figure 3.5: Merged points of all polygons

(d) The CSV was extracted using the point sampling tool

(e) The CSV columns were exchanged for them to be in the format accepted by the TCNN using python

```python
import pandas as pd
df=pd.read_csv('Dataset.csv',names=['id_2','class','merged_1','merged_2',
                                    'merged_3','merged_4','merged_5','merged_6',
                                    'merged_7','merged_8','merged_9','merged_10',
                                    'merged_11','merged_12','merged_13','merged_14',
                                    'merged_15','merged_16','merged_17','merged_18',
                                    'merged_19','merged_20','merged_21','merged_22',
                                    'merged_23','merged_24','merged_25','merged_26',
                                    'merged_27','merged_28','merged_29','merged_30',
                                    'merged_31','merged_32','merged_33','merged_34',
                                    'merged_35','merged_36','merged_37','merged_38',
                                    'merged_39','merged_40','merged_41','merged_42',
                                    'merged_43','merged_44','merged_45','merged_46',
                                    'merged_47','merged_48','merged_49','merged_50',
                                    'merged_51','merged_52','merged_53','merged_54',
                                    'merged_55','merged_56','merged_57','merged_58',
                                    'merged_59','merged_60','merged_61','merged_62',
                                    'merged_63','merged_64','merged_65','merged_66',
                                    'merged_67','merged_68','merged_69','merged_70',
                                    'merged_71','merged_72','merged_73','merged_74',
                                    'merged_75','merged_76','merged_77','merged_78',
                                    'merged_79','merged_80','merged_81','merged_82',
                                    'merged_83','merged_84','merged_85','merged_86',
                                    'merged_87','merged_88','merged_89','merged_90'])
titles=list(df.columns)
titles[0],titles[1]=titles[1],titles[0]
df=df[titles]
df = df.iloc[1:]
df.to_csv("DatasetSwapped.csv", index=False,header=False)
```

Figure 3.6: Python script for swapping columns

(f) The CSV file was split into 80% train and 20% test.

```python
import pandas as pd

data = pd.read_csv('Dataset.csv', sep=",", header=None)

train = []
test = []

total_sets = len(data)//60

train_df = data[data.index % 60 < 48]
test_df = data[data.index % 60 >= 48]

train_df.to_csv("train_dataset.csv", index=False)
test_df.to_csv("test_dataset.csv", index=False)
```

Figure 3.7: Python script to split csv into test and train

4. The CSV files generated were fed to the TCNN and the model was trained.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
5	2107	208	92	1	1	1	1	1	12	42	1	116	109	20	11	3	14	10
5	2107	28	25	1	1	1	1	1	1	31	1	132	114	29	43	39	32	21
5	2107	72	27	1	1	1	1	1	1	35	1	117	95	16	5	1	8	3
5	2111	504	487	1	1	1	1	1	1	21	1	530	497	123	94	41	47	27
2	616	848	1338	1590	2253	3328	3665	4012	3270	2333	4087	1003	1460	1740	2248	3226	3514	3754
2	614	262	457	505	1125	2308	2743	2930	2143	1418	3039	547	893	1010	1560	2555	2897	2952
0	89	26	474	80	1023	3852	4951	5404	1870	954	5181	252	644	316	1070	3521	4337	4508
4	1209	3630	4300	4828	4625	4721	4793	5536	4604	3405	4888	3506	4664	5248	5169	5158	5106	5932
2	620	243	555	614	1064	2445	2815	3178	2132	1296	3151	550	878	1026	1465	2424	2675	2968
5	2112	50	1	1	1	1	1	1	11	24	1	249	96	57	64	68	55	55
2	621	267	522	590	1051	2298	2654	2988	2112	1287	3040	560	831	943	1396	2290	2599	2792
5	2117	1	1	1	1	1	1	1	20	41	1	296	121	91	77	63	74	60
7	1810	1382	1324	1872	1951	1844	1710	2112	1997	1768	1950	1502	1982	2116	2030	2140	2149	2464
5	2115	624	608	26	1	1	1	1	4	38	1	683	525	49	6	1	1	1
5	2101	286	1	1	1	1	1	1	27	49	1	486	250	85	58	39	47	44
5	2111	363	214	1	1	1	1	1	1	20	1	516	467	122	106	56	71	56
5	2105	234	85	1	1	1	1	1	6	25	1	203	107	22	35	20	33	8
8	2407	404	821	874	1434	2524	2899	2862	2600	1829	3023	663	995	1260	1868	2373	2600	2788
5	2106	117	137	1	1	1	1	1	17	68	1	187	208	62	49	18	34	29
2	623	408	844	742	1317	3285	3873	4620	1927	1061	4075	546	922	909	1447	2499	2858	3414
2	620	618	962	1094	1566	2709	3017	3430	2754	1865	3467	763	1162	1314	1832	2711	2931	3298
5	2116	495	274	1	1	1	1	1	1	28	1	317	126	37	25	11	12	13
5	2116	416	227	1	1	1	1	1	1	23	1	345	120	25	33	3	16	13
5	2107	37	6	1	1	1	1	1	11	38	1	103	108	16	11	1	17	5
2	628	537	947	998	1643	3015	3372	3728	2852	1863	3715	809	1226	1484	1917	2749	2973	3260
4	1201	2724	3086	3406	3895	3899	3873	3726	4442	4300	4042	2524	3284	3860	4837	4716	4953	4436
5	2117	32	1	1	1	1	1	1	1	8	1	297	151	115	121	101	95	86
2	618	456	730	780	1258	2471	2801	3028	2329	1445	3188	553	858	974	1555	2379	2649	2732
0	38	24	397	150	921	3129	3939	4124	1897	934	4043	228	581	271	996	3300	4069	4420
5	2109	239	121	1	1	1	1	1	43	64	1	219	174	18	18	14	23	21
0	56	124	545	134	965	4600	6131	6544	1900	893	6531	275	550	295	934	3098	3948	4192
5	2112	14	1	1	1	1	1	1	22	26	1	227	133	74	68	55	65	72
5	2115	642	582	1	1	1	1	1	1	28	1	401	255	14	1	1	1	1
5	2105	264	152	1	1	1	1	1	13	22	1	203	114	44	51	41	58	48
5	2115	621	593	10	1	1	1	1	1	44	1	320	206	2	1	1	1	1
0	43	388	590	950	1051	1274	1492	1614	2285	1714	1750	583	902	1010	1388	2202	2450	2592
5	2108	388	320	126	9	13	51	134	79	106	49	273	189	94	85	81	90	77

Figure 3.8: Dataset in CSV

3.6 Hardware and software specifications

3.6.1 Hardware

Belgiu & Csillik (2018) highlighted that deep learning is a computationally demanding task which takes a long time for it to be processed. The hardware that was used for this research was an Intel Core i7-7700K @ 4.2GHz, respectively a single NVIDIA GTX1080 with 8GB of VRAM and 48GB RAM @3200MHz.

3.6.2 Software

Sen2Cor was used to Change the L1C images downloaded to L2A so the dataset is uniform in its format. Sen2Cor provides you with a CLI interface in which you can specify the path of your local L1C satellite image and it translates it to the L2A image you need.

SNAP was used for the RF classification which was used for comparison of the results of the TCNN.

QGIS 3.16.16 was the tool most used as it was used from the cropping of the images to the extraction of annotations for the train/test dataset from the time series of images and finally to visualize the inferenced image with the classified bands having different colours to be easily identified.

Python 3.9 was used for the TCNN model training and it was also used for the inferencing to be performed on the whole time series.

3.7 Experiment Setup and Evaluation Methods

Once the Models were trained, all the models were evaluated and checked thoroughly to see the accuracy of each and how each model managed to differ the different classes using the matrix provided.

The experiment setup consisted of importing the image which included all the 9 images in it, and also including the model in the path required for the python to read from. The inferencing than started on the image with the model provided, and what the inferencing provides is a detailed image with each pixel in it classified depending on what was trained.

After the inferencing was done and the detailed image was provided. The output then outputted as a tiff image and then analysed on QGIS with the background of the google api as seen below.

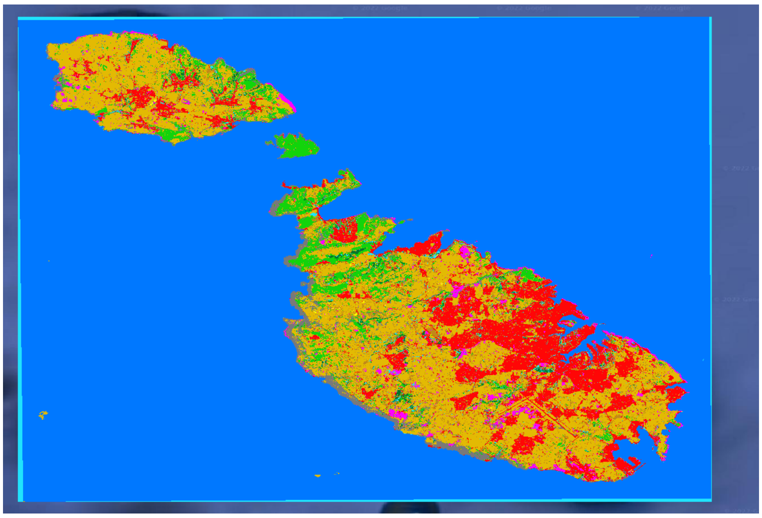


Figure 3.9: Output Image

After analysing it, a mask of Malta as seen below was used to clip the raster image to that mask thus eliminating the sea for an accurate result of the percentages of land cover of the Maltese islands.

Figure 3.10: Malta land mask

Finally after the image was clipped , the classification excluding the sea was generated as seen below.

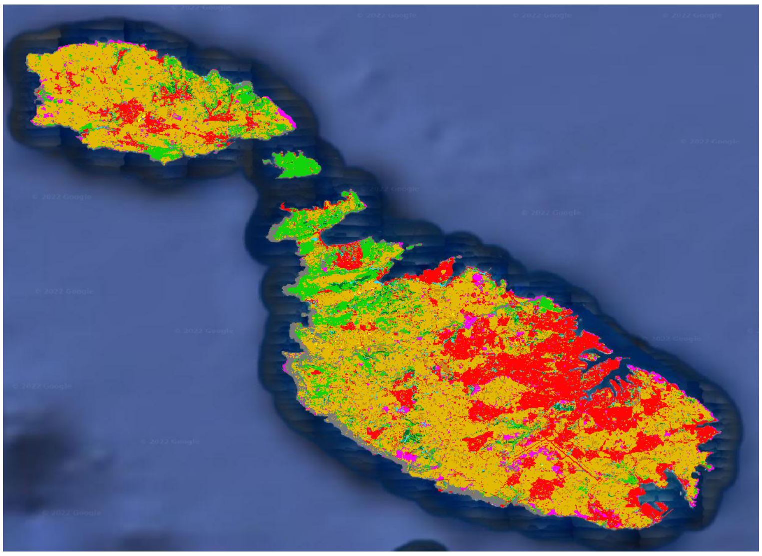


Figure 3.11: Classification without sea

3.8 Limitations of the proposed methodology

The proposed methodology has some inherent limitations. One of the most limiting factors of the RS images was the cloud coverage, in which certain months didn't have a single product which had no cloud coverage, the images collected ranged from February to October which was very similar to Vuolo et al. (2018) in which they found images ranging from March to September. For these images, Malta's good weather and almost all year summer came into play and more cloudless images were found. This will prove to be a limitation as it would be hard to track the phenological cycle of such crops to improve the classification since the whole year isn't available, although alterations can be made to the dataset to improve images according to the cycle and trying to find a balanced cycle in which products are always available.

Another limiting factor is the lack of automation that can happen, since the images need to be scanned for clouds coverage, make sure that the dates are varied enough and cropped, it limits the amount of automation that can be done. Multiple pre-processing factors took place in this research thus limiting the amount of work that can be automated and increasing the human resources needed.

The spectral resolution also was a small limiting factor but wasn't completely limiting as the accuracy could've been higher than 10m2. This resolution was still high given that it is free and publicly available, but as Karakizi et al. (2016) managed to discriminate different types of vines using a 50cm resolution image, LULC would've been close to a perfect accuracy.

3.9 Ethical Considerations

This research wasn't subject to any ethical considerations given that it didn't collect any private data, any data that can recognize a natural human being or any data in which businesses would be harmed or take any offense.

Chapter 4

Findings, analysis and discussion

4.1 Introduction

In this chapter, the results of all the methods discussed previously are to be shown and evaluated. An evaluation of all the different classifications mentioned will take place while also evaluating the dataset and how the changed in the dataset effect every iteration. The results are to be compared with each other to highlight the differences and where they excelled. The methods used will also be compared to other studies and the methods they used to also see any strength or weaknesses it might have with the metrics provided.

4.2 Model Training Evaluation

4.2.1 Random Forest Evaluation

Table 4.1: Random Forest Classifier

	Cropland	Forest	Grassland	Greenhouses	Other	Rocks	Settlement	Vineyards	Wetland	Overall
Accuracy	73.3%	71.5%	71.4%	98.2%	87.1%	72.3%	87.1%	72.6%	92.02%	80.6

The random forest classifier which was evaluated using SNAP was successful but although very accurate it was still outperformed severely by the TCNN.

The random forest worked very well when classifying the Greenhouses and wetland as they achieved an accuracy of 98.2% and 92.02% respectively. That being said the Other and Settlement class also had a good accuracy whilst the others suffered. This could be due to the band values being static as in all the classes that the RF did well are classes which do not have change during the year except for the rocks.

4.2.2 Multi-Temporal Convolutional neural network Evaluations

For the Multi-Temporal CNN, multiple iterations took place. The changes between every iteration were kept low so the difference in the changes could be highlighted.

Table 4.2: Multi-Temporal CNN Classifier

	Iteration 01	Iteration 02	Iteration 03	Iteration NDVI
Accuracy	99.50%	99.82%	98.39%	92.37%

- Iteration 01: It was the initial iteration in which the CSV's were generated and the first model trained and inferenced to see the results
- Iteration 02: The second iteration took place after seeing a lot of false positives in the first iteration and to eliminate such false positives, it was decided to train the model further with more annotations of crops and vineyards to eliminate the wrong classifications.
- Iteration 03: After eliminating plenty of false positives, there were still an amount of false positives and over training in the model. It was decided to reduce the density of points thus reducing the dataset size. What was happening was it was grabbing the same pixel more than once given that the resolution is 10m.

- Iteration NDVI: This was done to compare the output and strength of the NDVI with the all the bands, using the NDVI really helped in eliminating most of the false positives but apart from that, the other classes suffered in their accuracy. A combination of both NDVI and a normal all-band classification would be the perfect mix especially for vineyards.

Table 4.3: Confusion Matrix of Iteration 03

Iteration 03	Cropland	Forest	Grassland	Greenhouses	Other	Sea	Rocks	Settlement	Vineyards	Wetland
Cropland	10372	0	4	0	0	0	23	4	292	1
Forest	1	1722	32	0	0	0	0	1	2	70
Grassland	300	22	27119	0	15	0	247	2	62	7
Greenhouses	0	0	0	231	0	0	0	0	0	0
Other	35	0	0	0	6262	8	147	139	0	0
Sea	0	0	0	0	0	42736	4	0	0	0
Rocks	19	0	2	0	0	4	7193	7	2	0
Settlement	0	0	0	0	31	2	24	2134	13	1
Vineyards	25	0	1	0	10	0	0	7	3503	0
Wetland	27	4	24	0	1	3	20	2	14	821

The confusion matrix for iteration 03 produced some interesting results, the cropland got confused mainly with grassland which does make sense when looking from a top-down view, the forest, grassland, greenhouses, other, sea and wetland were quite accurate and didn't have many faults. The Rocks got confused mainly with grassland and other, being that rocks and grassland are somewhat similar.The vineyards were confused with cropland but to that extent it is plausible since it is also a type of cropland. Further Temporal imagery or more annotations of vineyards would help rule out this type of error since if more temporal imagery is increased, the cycle of the vineyard can be highlighted further.

Table 4.4: Confusion Matrix of iteration NDVI

Iteration NDVI	Cropland	Forest	Grassland	Greenhouses	Other	Sea	Rocks	Settlement	Vineyards	Wetland
Cropland	9732	0	332	7	79	0	238	5	297	6
Forest	4	1569	243	0	0	0	4	0	1	7
Grassland	17	77	27503	0	0	0	159	0	18	0
Greenhouses	2	0	0	197	21	4	7	0	0	0
Other	58	0	72	3	5725	29	526	178	0	0
Sea	0	0	0	0	84	42572	4	77	2	1
Rocks	139	2	2350	3	666	71	3896	86	9	5
Settlement	12	0	17	0	1018	130	281	733	5	9
Vineyards	62	2	37	0	1	0	25	0	3415	4
Wetland	47	139	132	0	10	23	25	16	25	499

The Iteration for the NDVI had somewhat similar results to those mentioned in the previous iteration. The main difference is in the rocks, settlement and wetland which are all non-vegetation classes. This might be due to the fact that the NDVI works mainly on greenery on an index of 0 to 1 and since the aforementioned classes have almost 0 greenery, they might have all been with an NDVI value of 0 hence the algorithm gets confused.

Table 4.5: Iteration 03 with multiple metrics

Iteration 03	Cropland	Forest	Grassland	Greenhouses	Other	Sea	Rocks	Settlement	Vineyards	Wetland	Overall
Accuracy	96.97%	94.20%	97.64%	100.00%	95.01%	99.99%	99.53%	96.78%	98.79%	89.63%	96.85%
Precision	96.22%	98.51%	99.77%	100.00%	99.10%	99.96%	93.93%	92.94%	90.10%	91.22%	N/A
F-Score	93.72%	86.76%	94.09%	89.34%	80.66%	99.50%	62.88%	44.42%	93.33%	68.97%	N/A

Table 4.6: NDVI With multiple Metrics

Iteration NDVI	Cropland	Forest	Grassland	Greenhouses	Other	Sea	Rocks	Settlement	Vineyards	Wetland	Overall
Accuracy	90.99%	85.83%	99.02%	85.28%	86.86%	99.61%	53.91%	33.24%	96.31%	54.48%	92.37%
Precision	96.61%	87.70%	89.63%	93.81%	75.29%	99.40%	75.43%	66.94%	90.54%	93.97%	N/A
F-Score	93.72%	86.76%	94.09%	89.34%	80.66%	99.50%	62.88%	44.42%	93.33%	68.97%	N/A

The Precision is the amount of true positive results divided by all the results possible and the F-Score is derived from precision and recall of a class. Highlighted above are both the F-score and precision. The precision for vineyards ,although slightly similar, is better with the NDVI and that was what the NDVI iteration was meant for. With the NDVI iteration the amount of vineyards detected was less but the ones detected were more precise and most of them were surely vineyards. A combination of both the NDVI iteration and the 3rd iteration would be beneficial although this can have a lot of hurdles to implement due to judging with classification has priority when there is an overlap.

Table 4.7: Multi-Temporal CNN Classifiers With Classes

Accuracy	Cropland	Forest	Grassland	Greenhouses	Other	Sea	Rocks	Settlement	Vineyards	Wetland	Overall
Iteration 01	97.75%	99.59%	99.90%	87.35%	99.80%	99.98%	99.47%	96.55%	99.53%	89.42	96.93%
Iteration 02	99.12%	99.13%	99.97%	99.77%	99.81%	100%	99.29%	99.80%	99.67%	99.76%	99.63%
Iteration 03	96.97%	94.20%	97.64%	100%	95.00%	99.99%	99.53%	96.78%	98.79%	89.63%	96.85%
Iteration NDVI	90.99%	85.83%	99.02%	85.28%	86.86%	99.61%	53.91%	33.24%	96.31%	54.48%	92.37%

These first 3 iterations all did really well in almost all the classes. Although iteration 03 seemed to be the worst of them all, when doing the inferencing, the vineyards were outputted much better. The other classes practically stayed the same and there wasn't a huge change as this iteration was mostly focused in increasing the accuracy of the vineyards. The 4th iteration with the NDVI was quite different in terms of results, it can be seen that it was very accurate in terms of crops although not as accurate as the first 3 iterations which had much more data to go on because of the NDVI's use of only 2 bands out of 10. Anything that is not cropland really suffered in terms of accuracy in the NDVI iteration, considering NDVI is targeted for live vegetation. The results in the classes like rocks, settlement and wetland reflect it perfectly.

Table 4.8: Percentage of land use of iteration03 TCNN

Class	Pixel Count	m^2	km^2	Percentage of Land Cover
Cropland	1467869	147201401.4	147.2014014	46.58%
Forest	47258	4739144.861	4.739144861	1.50%
Grassland	316514	31740778.21	31.74077821	10.04%
Greenhouses	5467	548243.7884	0.548243788	0.17%
Other	114674	11499782	11.499782	3.64%
Rocks	193319	19386490.02	19.38649002	6.14%
Settlement	677024	67893580.14	67.89358014	21.49%
Vineyards	281182	28197601.05	28.19760105	8.92%
Wetland	47661	4779558.661	4.779558661	1.51%
Sum	3150968	315986580.1	315.9865801	100%

This table shows the amount of pixels in every class classified , while also the total area covered by each class. The total area of Malta is $316km^2$ and according to the results in iteration 03, the sum came out to the same total, this helped prove the integrity of the study. The percentage of land cover was also outputted, with reference to 2.1, the total urban area/settlement in this study came out to a total of 21.49% whilst according to the Corinne land cover it is 22.3%. What might have happened is that the TCNN classified industrial and commercial units and airports as settlements which from a top-down view, an industrial unit is the same as a building.

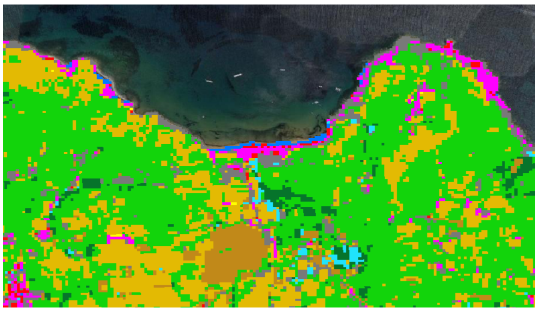
The forested areas in the TCNN amounted to 1.50% whilst in the Corinne land cover, it amounted to 0.7%. This might be due to the recent afforestation projects in which Infrastructure Malta are adding more than 25,000 trees per year.

The cropland/Agricultural areas are also quite similar in terms of percentage, although it must be mentioned that this study defined vineyards separately rather than including it with the cropland as the Corinne land cover did.

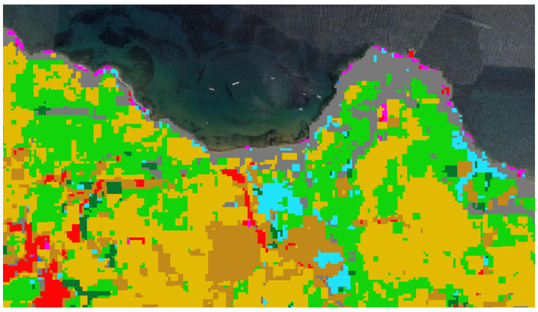
Table 4.9: Percentage of land use of iteration NDVI TCNN

Class	Pixel count	m^2	km^2	Percentage of Land Cover
Cropland	1012650	1.02E+08	101.5509554	32.44%
Forest	64629	6481150	6.481150149	2.07%
Grassland	961577	96429233	96.42923327	30.80%
Greenhouses	3575	358509.5	0.35850952	0.11%
Other	379015	38008527	38.0085275	12.14%
Rock	444645	44590060	44.59006031	14.24%
Settlement	127796	12815687	12.81568745	4.09%
Wetland	97702	9797789	9.797789411	3.13%
Vineyard	30255	3034044	3.034043506	0.97%
Sum	3121844	3.13E+08	313.0659566	100.00%

The highlighted differences in the NDVI resulted in a lower identification of cropland and vineyards, this eliminated the false positives created in iteration 03. It worked very well on cropland but suffered severely when it came to static pixels like buildings in which the percentage of Land cover in settlement and other classes had a low accuracy. This highlighted the reason for using all the bands rather than only using the 2 bands to work out the NDVI for buildings and other built-up areas that are non-vegetation.



(a) *NDVI Gozo area*



(b) Iteraion 03 Gozo area

The difference in Ramla bay between the NDVI and iteration 03 really got highlighted when seeing the cropland and vineyards detected. In the NDVI the cropland and vineyards got detected perfectly while the layer of rocks on the border didn't get picked up unlike in the iteration 03 image in which the rocks got picked up perfectly but had multiple false positives in the vegetation.

(a) *NDVI harbour area*

(b) iteration 03 harbour area

The NDVI highlighted most of the settlements as other, while iteration 03 highlighted the settlements as settlement. This shows the strength of using all the bands on static images. The cropland however was detected more accurately in the NDVI as can be seen in the images above.

More sample outputs can be seen in the appendix.

4.3 Experiments

4.3.1 Level 1 experiment : Vegetation Versus Non-Vegetation

Table 4.10: Vegetation Versus Non-Vegetation

Level 1	Vegetation	Non Vegetation
Accuracy	96.90%	96.82%

The accuracy between vegetation against non-vegetation was quite equal as the accuracy was quite high overall. This outlined that the TCNN didn't have a problem differentiating between the two but managed to do it quite smoothly. Since non-vegetation doesn't change as frequent, the Temporal aspect really helps a lot to differentiate the two as the pattern of change outlines the two.

4.3.2 Level 2 experiment : Different types of Vegetation

Table 4.11: Different types of vegetation compared

Level 2	Cropland	Vineyards	Forest	Grassland
Accuracy	96.97%	98.79%	94.20%	97.64%

The vineyards were of the highest accuracy and the forest class had the lowest accuracy. One major issue that was happening was that the forest and vineyards were getting confused due to their appearance from a top down view. This was an issue that its impact was reduced when increasing the vineyards annotations but it was still an apparent problem.

4.3.3 Level 3 experiment : Different types of crops

Table 4.12: Different Types of crops compared

Level 3	Cropland	Vineyards
Accuracy	96.97%	98.79%

The vineyards were more accurate in the classifier. This might be because of the annotations that were specific to them whilst the cropland were still highly accurate , it suffered due to it being generic over different types of crops.

4.4 Comparison with other classifiers and architectures

Table 4.13: Comparison with other classifiers

	TCNN With all bands	TCNN with NDVI	Belgiu & Csillik (2018)	Pena et al. (2019)	Aswatha et al. (2018)
Accuracy	96.85%	92.37%	92.62%	93.17%	94.8%

The accuracy between the NDVI and all the bands is clearly in favour of all the bands combined. When comparing the TCNN of this study to Belgiu & Csillik (2018), it clearly performed better in terms of accuracy. Both of these studies used sentinel-2 imagery hence had the same resolutions available. The difference could be due to multiple factors. A major factor in the difference in accuracy was that in the aforementioned study, only cropland was classified and inferenced, while in this research, the whole island was classified including several different classes. Another factor which surely affected the accuracy was the classifier used, as time-weighted dynamic time warping analysis isn't as accurate as the TCNN in this case.

When compared to Pena et al. (2019) research, the accuracy is similar although once again it is being done on a field and the only comparison it has to do is between similar land, that might affect the final OA.

The result outputted by Aswatha et al. (2018) was very accurate but like the previous studies, it is also done only on areas which are very similar as all the area in the study is grassland and green areas, although it performed quite well with the fused stokes-NDSV, the TCNN performed in this study including all the bands was still more powerful whilst the TCNN with the NDVI came a bit short in terms of accuracy.

Chapter 5

Conclusions and Recommendations

This study aimed to identify Malta's land and classify it into different classes as defined previously. The research aimed to do such a task by implementing multiple classifiers while also comparing with the similar previous research that are performed on Malta. This approach provides new insight into modern classification techniques used in small islands. The research that was intended was achieved by using both the TCNN and the RF. Both classifiers achieved a very high accuracy, with the TCNN outperforming the RF with an accuracy of 98.39% and the RF achieving an accuracy of 80.6%. The custom temporal dataset that was created was robust as it was composed of a temporal stack of 9 images with 10 bands and over 400,000 data points in 10 classes. Multiple studies took on the challenge to classify land cover using CNN, although considering Malta's size and it fitting into one satellite image, the whole country could be identified without any fusing of images. Given the Maltese island's small size, there is limited research knowledge in this area . The prototype pipeline developed was aimed to be used with quantitative techniques after the development and dataset creation was finished to evaluate and quantify the results of such a research. The conclusions attained will be able to help government entities and European entities to focus on Malta's environmental issues.

5.0.1 Summary

The hypothesis in this study was that by using satellite products that monitor the visual spectrum wavelengths, it is possible to identify different crops. After discussing the results, it can be said that the hypothesis stands. The initial research question was how are the current crops monitored in the Maltese islands. It was seen that the current Maltese LULC map is the Corinne map provided by the ERA which isn't as detailed and updated as one would hope given the current development Malta has gone throughout the last few years. The second research question was if multi-temporal classification using neural networks can be used to classify different crops in a given area, specifically to differentiate from different crops, vineyards were used to outline a specific crop. This research question was found to be true as the classification in the results was achieved by both the RF and the TCNN in a feasible amount of time with a computer that has mid-end specifications while also a way to create the dataset for such a study was defined which will surely be of use for further research. The final research question was how can cropland be segmented from other land. Other than using the classifier, the cropland could be segmented to be prepared for the classifier using GIS tools and annotating the land, while also using ground truthing techniques such as the google earth history feature and also going to view the land cover by sight. The inferencing results were quite positive with various high accuracies being produced while also worth mentioning that most of the false positives produced made logical sense as to why the classifier confused them in a top-down view from the sky while also being limited by the mid-end spatial resolution of the camera.

5.0.2 Suggestions for future research

Given the current focus on space-research programs, this area is in a growing state especially in the past few years. That being said with all the advancements and money being put into RS, the potential for growth in such a study is quite big. One recommendation that would surely be beneficial to further research would be to include several more annotations which include different pieces of land, for example including different types of the forest class in which some location, the trees were more dispersed thus having a different top-down view leading to inaccuracies. Finally a process in which the NDVI classification and the classification with all the bands is merged into one with an overlapping technique which benefits the most accurate classifier of them all in the said class, thus the settlement classification would be left for the classification with all the bands, while when trying to classify different croplands, the NDVI would be used due to its strength in outputting the vegetation and laying out the pattern in such a temporal study. Another possible suggestion would be to train a model to extract an agricultural land mask using the highly accurate LULC TCNN solution implemented in Stage 1 of this study. Then this mask can be used in another classifier for crop classification using the NDVI stack combined with a dataset of only crop types. This would help in identifying different crops.

Appendices

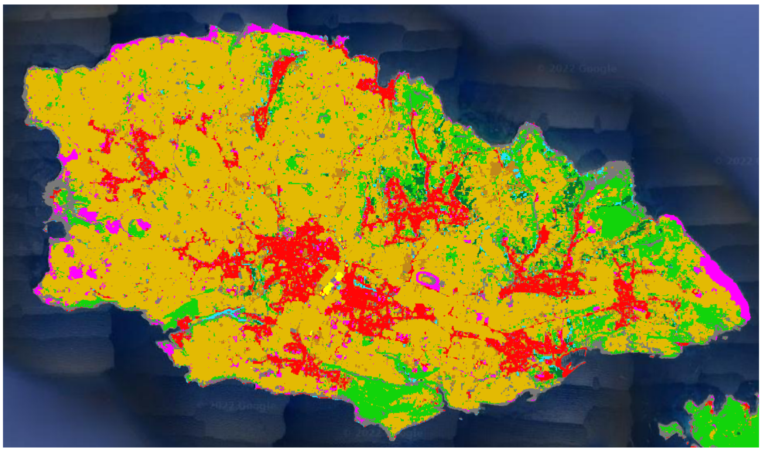


Figure 1: Gozo iteration 03

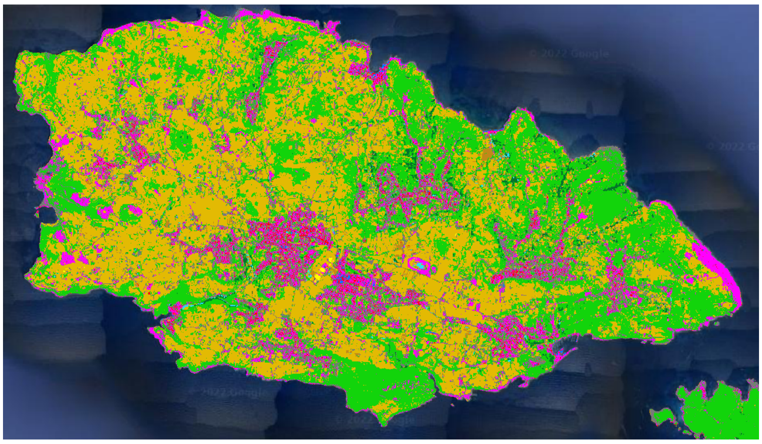


Figure 2: Gozo NDVI

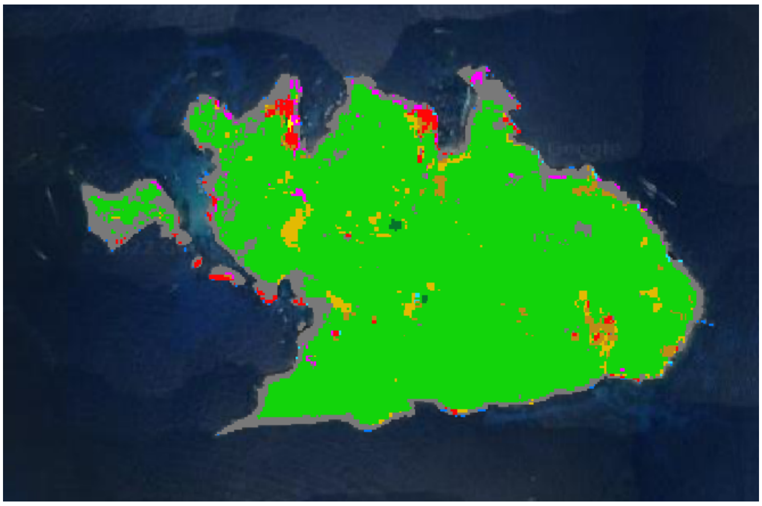


Figure 3: Kemmuna iteration 03

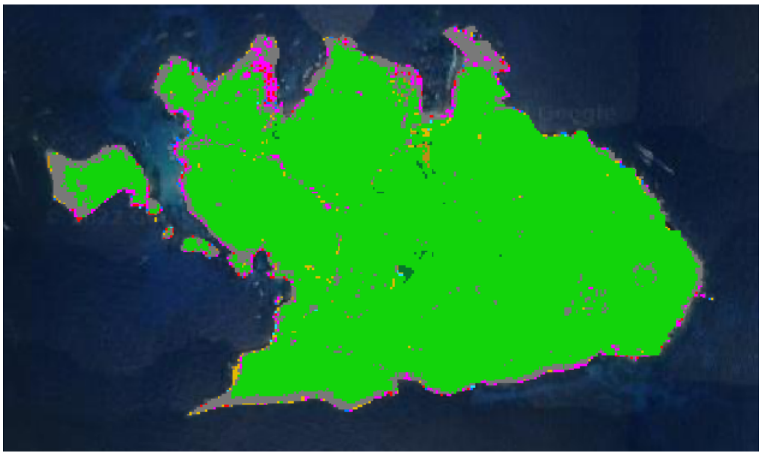


Figure 4: Kemmuna NDVI

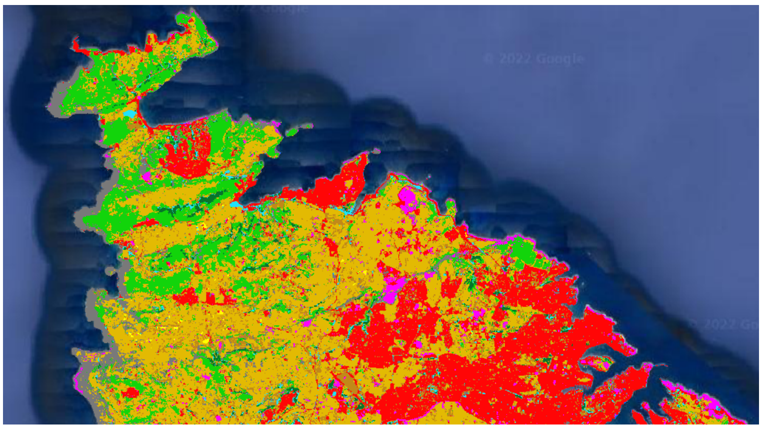


Figure 5: North iteration 03

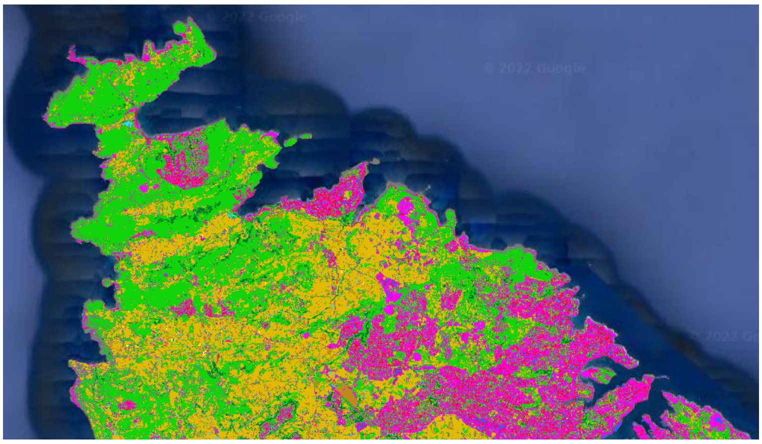


Figure 6: North NDVI

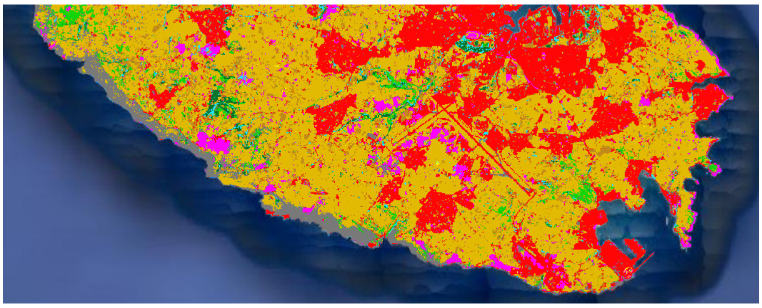


Figure 7: South iteration 03

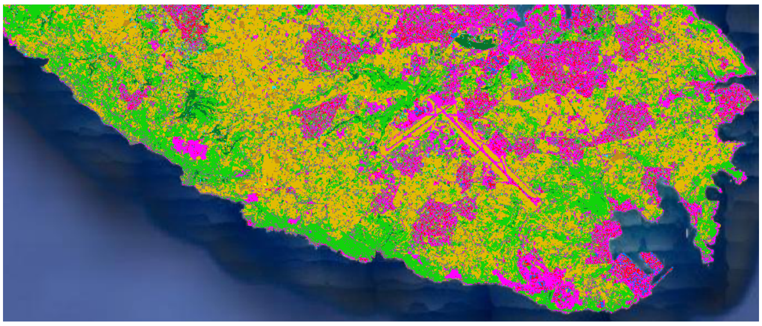


Figure 8: South NDVI

Figure 9: Legend

Bibliography

Aswatha, S., Saini, V., Mukherjee, J., Biswas, P., Aikat, S. & Misra, A. (2018), 'Unsupervised detection of surface mine sites using sentinel multi-spectral imagery and dual-polarimetric sar data', pp. 1–8.

Belgiu, M. & Csillik, O. (2018), 'Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis', *Remote Sensing of Environment* **204**, 509–523.
URL: *https://www.sciencedirect.com/science/article/pii/S0034425717304686*

Carranza-García, M., García-Gutiérrez, J. & Riquelme, J. C. (2019), 'A framework for evaluating land use and land cover classification using convolutional neural networks', *Remote Sensing* **11**(3).
URL: *https://www.mdpi.com/2072-4292/11/3/274*

Creswell, J. W. (2013), *Research design: Qualitative, quantitative, and mixed methods approaches*, Sage.

Fan, X., Vrieling, A., Muller, B. & Nelson, A. (2020), 'Winter cover crops in dutch maize fields: Variability in quality and its drivers assessed from multi-temporal sentinel-2 imagery', *International Journal of Applied Earth Observation and Geoinformation* **91**, 102139.
URL: *https://www.sciencedirect.com/science/article/pii/S0303243420300106*

Gandhi, G. M., Parthiban, S., Thummalu, N. & Christy, A. (2015), 'Ndvi: Vegetation change detection using remote sensing and gis–a case study of vellore district', *Procedia Computer Science* **57**, 1199–1210.

Huang, S., Tang, L., Hupy, J. P., Wang, Y. & Shao, G. (2020), 'A commentary review on the use of normalized difference vegetation index (ndvi) in the era of popular remote sensing', *Journal of Forestry Research* .

Immitzer, M., Neuwirth, M., Böck, S., Brenner, H., Vuolo, F. & Atzberger, C. (2019), 'Optimal input features for tree species classification in central europe based on multi-temporal sentinel-2 data', *Remote Sensing* **11**(22).
URL: *https://www.mdpi.com/2072-4292/11/22/2599*

Ji, S., Zhang, C., Xu, A., Shi, Y. & Duan, Y. (2018), '3d convolutional neural networks for crop classification with multi-temporal remote sensing images', *Remote Sensing* **10**(1).
URL: *https://www.mdpi.com/2072-4292/10/1/75*

Jones, E. G., Wong, S., Milton, A., Sclauzero, J., Whittenbury, H. & McDonnell, M. D. (2020), 'The impact of pan-sharpening and spectral resolution on vineyard segmentation through machine learning', *Remote Sensing* **12**(6).
URL: *https://www.mdpi.com/2072-4292/12/6/934*

Karakizi, C., Oikonomou, M. & Karantzalos, K. (2016), 'Vineyard detection and vine variety discrimination from very high resolution satellite data', *Remote Sensing* **8**(3).
URL: *https://www.mdpi.com/2072-4292/8/3/235*

Lanaras, C., Bioucas-Dias, J., Galliani, S. & Baltsavias, E. (2018), 'Super-resolution of sentinel-2 images: Learning a globally applicable deep neural network', *ISPRS Journal of Photogrammetry and Remote Sensing* **146**.

Mazzia, V., Khaliq, A. & Chiaberge, M. (2020), ‘Improvement in land cover and crop classification based on temporal features learning from sentinel-2 data using recurrent-convolutional neural network (r-cnn)’, *Applied Sciences* **10**(1).
URL: *https://www.mdpi.com/2076-3417/10/1/238*

Paris, C., Weikmann, G. & Bruzzone, L. (2020), ‘Monitoring of agricultural areas by using sentinel 2 image time series and deep learning techniques’, p. 18.

Pelletier, C., Webb, G. I. & Petitjean, F. (2019), ‘Temporal convolutional neural network for the classification of satellite image time series’, *Remote Sensing* **11**(5).
URL: *https://www.mdpi.com/2072-4292/11/5/523*

Pena, J., Tan, Y. & Boonpook, W. (2019), ‘Semantic segmentation based remote sensing data fusion on crops detection’, *Journal of Computer and Communications* **07**, 53–64.

Song, X.-P., Potapov, P. V., Krylov, A., King, L., Di Bella, C. M., Hudson, A., Khan, A., Adusei, B., Stehman, S. V. & Hansen, M. C. (2017), ‘National-scale soybean mapping and area estimation in the united states using medium resolution satellite imagery and field survey’, *Remote Sensing of Environment* **190**, 383–395.
URL: *https://www.sciencedirect.com/science/article/pii/S0034425717300081*

Vuolo, Neuwirth, Immitzer, Atzberger & Ng (2018), ‘How much does multi-temporal sentinel-2 data improve crop type classification?.’, *International Journal of Applied Earth Observation and Geoinformation* **72**, 122–130.