Land-Use Land-Cover Classification by Machine Learning Approaches

Sertac İnce
Metin Abadan

Computer Engineering Department AYBU
Ankara, Turkey
nc.sertac@gmail.com
metinabadan@gmail.com

***Abstract*—Nowadays we obtain various information and inferences about almost the entire surface of the earth with satellite images. The study focuses on the development of models that can accurately classify and detect features of interest in satellite imagery, such as land use, vegetation, and infrastructure. The research includes a thorough review of relevant literature and the implementation of several experimental machine learning algorithms. The performance of the models is evaluated using a variety of metrics, including accuracy, precision, and recall. The results of the study demonstrate the potential of using satellite imagery in combination with machine learning for a wide range of applications, such as land use mapping, urban planning, and natural resource management. In this study, Land Cover Map images of Gibraltar and its surroundings obtained from Sentinel-2 satellite were used. A dataset was created with the color space values obtained from these images.**

Keywords—land cover classification; satellite imagery; classification; sentinel-2.

I. Introduction

Monitoring and managing environmental changes on Earth is important for both a sustainable environment and efficient use of natural resources. Because of their unique strengths, remote sensing images are the most widely used source for detecting these global and local changes. Remote sensing applications offer significant advantages in cost, time, and workspace size compared to terrestrial research. Thanks to remote sensing technology, important studies of large and inaccessible areas can be carried out in a short period of time. Classification of remote sensing images is a method often used in research conducted in this context. There are many classification methods and pathways in the literature. Furthermore, many studies have been conducted to analyze the performance of these methods.

In this study, machine learning models such as KNN, SVM, Bayes, Decision Trees, Random Forest, Gradient Boosting were used. In addition, the best parameters of these models were tried to be found. Various interpretations were made according to the results.

II. Study area and the data

A. Study Area

The study area is Gibraltar. It is a British Overseas Territory and headland, on Spain's south coast. It's dominated by the Rock of Gibraltar, a 426m-high limestone ridge. Gibraltar is a heavily fortified British air and naval base that guards the Strait of Gibraltar, which is the only entrance to the Mediterranean Sea from the Atlantic Ocean.

B. Optical Images

The Sentinel-2 satellites each carry a single multi-spectral instrument (MSI) with 13 spectral channels in the visible/near infrared (VNIR) and short-wave infrared spectral range (SWIR). Within the 13 bands, the 10-meter spatial resolution allows for continued collaboration with the SPOT-5 and Landsat-8 missions, with the core focus being land classification.

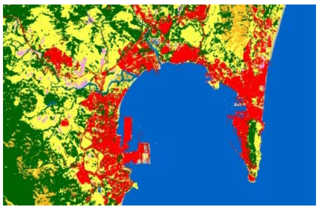

Figure 1 Cover Map

Figure 2 RGB Map

Figure 3 False Color Map

C. Ground Truth

The ground truth data includes eleven classes: Tree Cover, Shrubland, Grassland, Cropland, Built-up, Bare / sparse vegetation, Snow and ice, Permanent water bodies, Herbaceous wetland, Mangroves, Moss and lichen.

However, 9 of these 11 classes are used in the train data: Tree Cover, Shrubland, Grassland, Cropland, Built-up, Bare / sparse vegetation, Permanent water bodies, Herbaceous wetland.

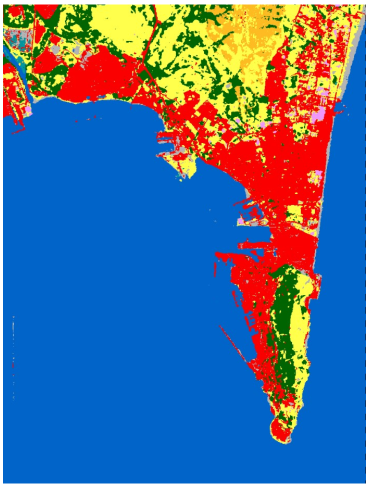

Figure 4 Ground Truth Data

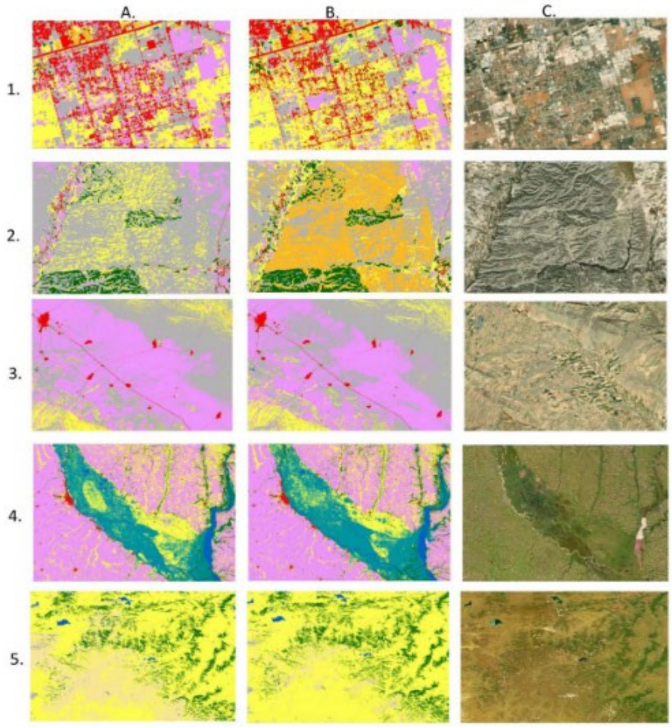

Figure 5 Comparison of WorldCover 2020 v100 (A) and WorldCover 2021 v200 (B) and VHR images available from the Google Earth (C).

III. METHOD

D. Processing of the Data

The CGLS-LC (Copernicus Global Land Service- Land Cover) validation dataset is based on probability sampling to allow a design-based inference of map accuracies. The criterion of statistical probability sampling with known and non-zero inclusion probabilities was followed. The validation dataset is based on stratified random sampling, employing a global stratification. There are two different datasets created 1 year apart.

The WorldCover 2020 v100 product reaches an overall accuracy of 74.4%.

The WorldCover 2021 v200 reaches an overall accuracy of 76.7%.

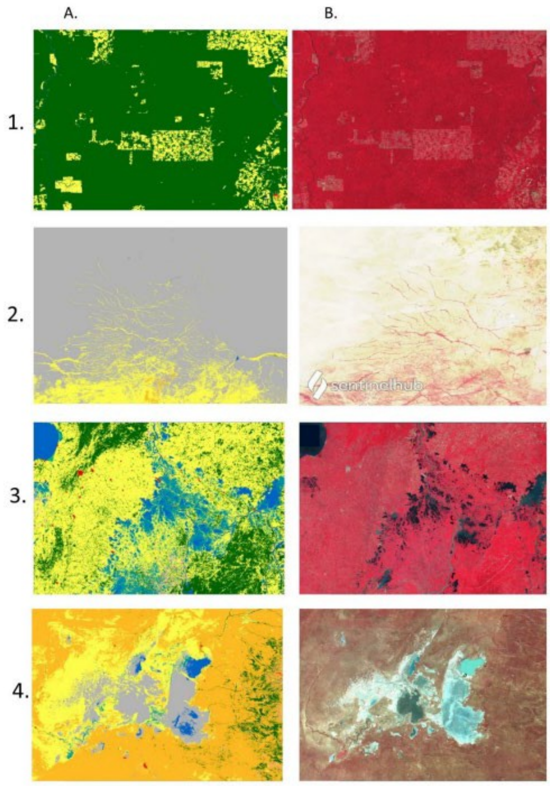

Figure 6 Comparison of the WorldCover 2021 v200 (A) and Sentinel 2 data (B)

E. *Classification*

First, the 0's in the data were cleaned and a cleaner data was created. Then, Bayes, KNN, Decision Trees, Random Forest, Gradient Boosting and SVM algorithms were used with their default parameters, respectively, without applying any normalization process with this simple processed data. Afterwards, a performance comparison was made with these results. The Grid Search method was used to find the best parameters of each algorithm. Another training was performed for each model with the best parameters and a performance comparison was made with the results. Thereupon, the data was normalized, and a training was carried out again. All these tests were performed with a subset of the provided train data. According to the obtained performance values, all the train data were used with the best parameters and the test data was estimated. Thereupon, a feature expansion operation was performed by calculating NDWI and NDVI values from the available data. Since the labels of the test data, which is approximately 4.77 times larger than the Train data, are on Kaggle, the performance comparison of this output was made on Kaggle. According to these values, parameter changes and data cleaning were made to improve the model. In summary, a larger data estimation was performed than the small data. Post-processing was carried out by visualizing the result extracted from the big data and comparing the same region over ESA and making various changes on the data of the regions that are difficult to predict.

IV. RESULTS AND DISCUSSION

The general operation was carried out in 4 stages. First, the use of machine learning algorithms with default parameters without normalizing the data and cleaning the 0 data from training set. Second, finding the best parameters with the Grid Search method. Third, normalizing the data and using it with the best parameters and using NDWI and NDVI data. Fourth, performing a training again by cleaning various data according to the output values received.

Class - Labels	Size in Pixels	Percentage
Empty Class - (0)	1416	0.14
Tree Cover - (10)	58025	6.00
Shrubland - (20)	15711	1.62
Grassland - (30)	108276	11.19
Cropland - (40)	4348	0.44
Built-up (50)	143396	14.82
Bare / Sparse Vegetation (60)	15014	1.55
Permanent Water Bodies (80)	619074	64.01
Herbaceous Wetland (90)	1745	0.18

Figure 7 Total Sizes of the Classes

F. *Classification with Only Cleaning the 0's*

After clearing the 0 label values in the data, a subset was taken from this data. As a result of the tests made with these data, 84% accuracy with Bayes, 89% accuracy with KNN, 86% accuracy with Decision Tree, 90% accuracy with Random Forest, 90% accuracy with Gradient Boosting and 86% accuracy with SVM obtained.

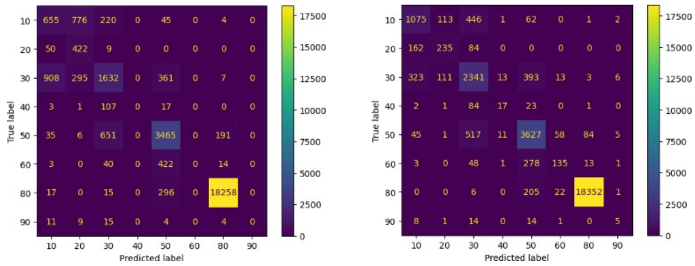

Figure 8 Confusion Matrix of the Bayes and KNN Classifiers

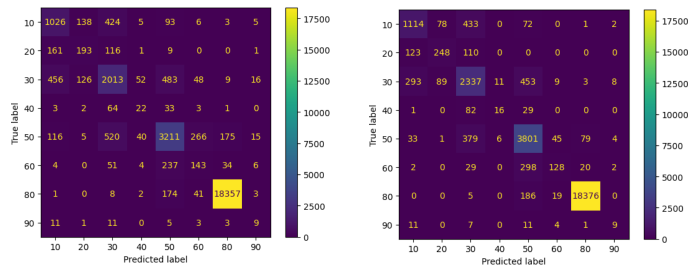

Figure 9 Confusion Matrix of the Decision Tree and Random Forest Classifiers

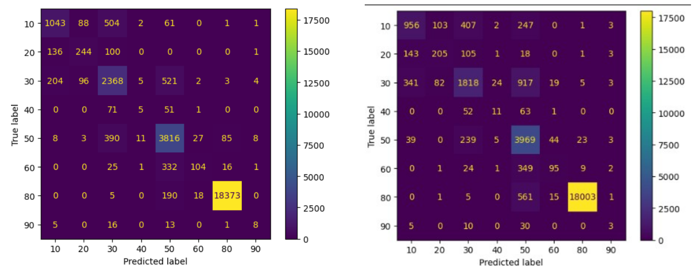

Figure 10 Confusion Matrix of the Gradient Boosting and SVM Classifiers

G. *Finding Best Parameters with Grid Search Method*

In this section, the comparisons made the best parameters found with the help of the previous subset data and the Grid Search algorithm are given. As a result of the tests 86% accuracy with Bayes, 89% accuracy with KNN, 89% accuracy with Decision Tree, 90% accuracy with Random Forest, 90% accuracy with Gradient Boosting and 91% accuracy with SVM obtained.

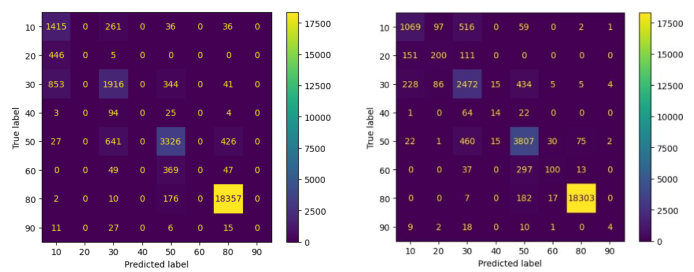

Figure 11 Confusion Matrix of the Bayes and KNN Classifiers

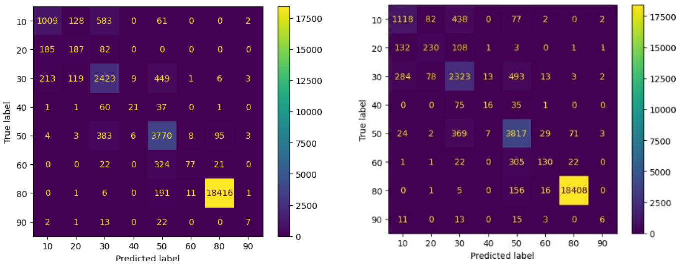

Figure 12 Confusion Matrix of the Decision Tree and Random Forest Classifiers

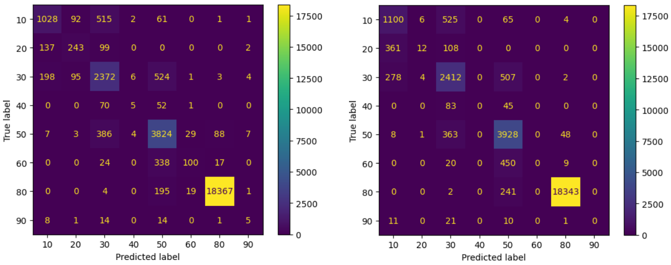

Figure 13 Confusion Matrix of the Gradient Boosting and SVM Classifiers

H. Normalizing the Data and Using It with Best Parameters and Using NDWI – NDVI Values.

In this section, the observed changes after normalization and adding NDWI and NDVI values to the data set will be explained. Since the train data now gives the same accuracy rates, this section will examine the changes on the large test data. It has been observed that the most important change is on the bayes and Decision Tree classifiers. A positive change was observed in all other classifiers, although not as much as Bayes and Decision Tree. It was seen that the main change was provided by the NDWI and NDVI data, not the normalization. After these operations, an accuracy rate increased from 28% to 38% on Kaggle for the Bayes model. For the Decision tree, this accuracy rate increased from 15% to 21%.

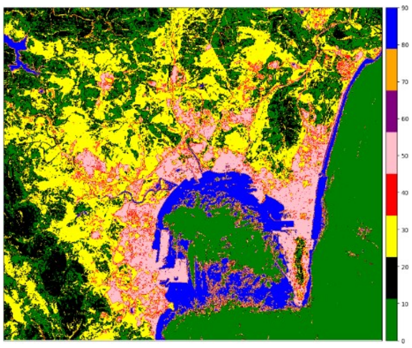
Figure 14 Before the NDWI and NDVI Data for the Bayes

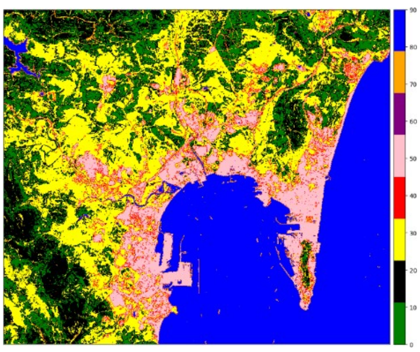
Figure 15 After the NDWI and NDVI Data for the Bayes

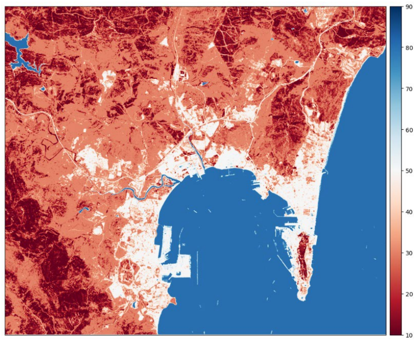
Figure 16 Before the NDWI and NDVI Data for the KNN

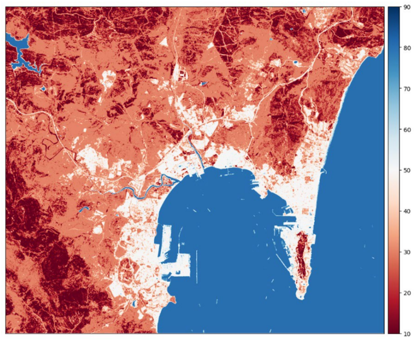
Figure 17 After the NDWI and NDVI Data for the KNN

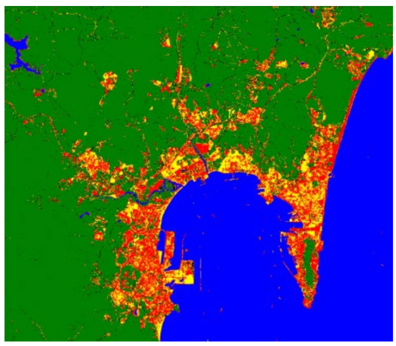
Figure 18 Before the NDWI and NDVI Data for the Decision Tree

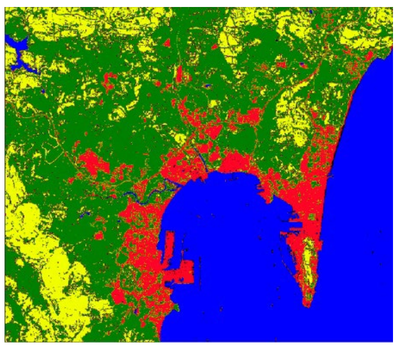
Figure 19 After the NDWI and NDVI for the Decision Tree

I. *Cleaning the Data According to the Outputs*

Changes made in this section have had little effect and have been made only on fast-running classifiers (Bayes and KNN). The procedure is to replace the scarce and low-predicted labels in the dataset with the closest data. In addition, another retrospective estimation was made for control purposes. In other words, a model is trained again with the outputs of the predicted test data and train data with it. Its change on Bayesian and KNN outputs is negligible. The reason for this is that this data is known at a low rate because it is mixed with data that is already nearby. In other words, this process has not observed much effect on this dataset. After this process, there was a decrease in accuracy for Bayes and Decision Tree, an increase of 0.37% for KNN, a 23% increase for Random Forest, and a 0.59% increase for SVM.

J. *Discussion of the Results*

In line with the results, it has been seen that making inferences of this dataset in its pure form does not give very high accuracy rates. With the new data derived from pure data, a good increase in the accuracy of the results of many models has been observed. Apart from these, it has been observed that the data that are scarce in the train data (shrubland, cropland, Bare/Sparse Vegetation) can hardly be estimated accurately in the test data.

V. Conclusions

In this study, the performance of the dataset obtained from sentinel-2 satellite images of the Gibraltar region on different machine learning applications is mentioned. The importance of data cleaning and the effect of normalization on accuracy were also investigated. In addition, it has been observed that positive changes can be made on the accuracy rate of machine learning algorithms by using feature expansion.

References

Umbria, A., Galán, M., Munoz, M. J., & Martín, R. (2004). Characterization of atmospheric particles: analysis of particles in the Campo de Gibraltar. Atmósfera, 17(4), 191-206.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: a review. *Remote Sensing*, *12*(14), 2291.

Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GIScience & Remote Sensing, 57(1), 1-20.

Castro Gomez, M. G. (2017). Joint use of Sentinel-1 and Sentinel-2 for land cover classification: A machine learning approach. Lund University GEM thesis series.

Rana, V. K., & Suryanarayana, T. M. V. (2020). Performance evaluation of MLE, RF and SVM classification algorithms for watershed scale land use/land cover mapping using sentinel 2 bands. Remote Sensing Applications: Society and Environment, 19, 100351.

Hu, B., Xu, Y., Huang, X., Cheng, Q., Ding, Q., Bai, L., & Li, Y. (2021). Improving urban land cover classification with combined use of sentinel-2 and sentinel-1 imagery. ISPRS International Journal of Geo-Information, 10(8), 533.