26-osios Lietuvos jaunųjų mokslininkų konferencijos „Mokslas – Lietuvos ateitis" teminė konferencija
Proceedings of the 26th Conference for Junior Researchers "Science – Future of Lithuania"

DARNI APLINKA / SUSTAINABLE ENVIRONMENT

2023 m. balandžio 14 d., Vilnius, Lietuva
14 April 2023, Vilnius, Lithuania

ISSN 2029-7157 / eISSN 2029-7149
eISBN 978-609-476-344-1
https://doi.org/10.3846/da.2023.007

Geodezija ir kadastras / Geodesy and Cadastre
https://vilniustech.lt/346238

AI-BASED SATELLITE IMAGE CLASSIFICATION: AN ANALYSIS OF DIFFERENT ALGORITHMS ON PLANETSCOPE AND SENTINEL IMAGERY

Mina Adel Shokry FAHIM*, Jūratė SUŽIEDELYTĖ VISOCKIENĖ

Department of Geodesy and Cadastre, Environmental Engineering Faculty,
Vilnius Gediminas Technical University, Vilnius, Lithuania
**E-mail: mina.narmar.7000@gmail.com*

Received 6 March 2023; accepted 3 April 2023

Abstract. This research aims to explore the potential of machine-learning algorithms in land classification and effectively analyse and visualize the results through the development of an application. The study area and different satellite systems used for data acquisition, and various machine-learning algorithms for satellite image classification are explored. The research also delves into Geographic Information Systems (GIS) concepts and methods of classification, including different algorithms used for satellite image processing. The analysis and interpretation of the results are covered in detail, using open-source software to differentiate between Sentinel-2 and Planet da-ta sets visually. The developed application provides a comprehensive and effective visualization of the results. Overall, this study provides a comprehensive understanding of the potential of machine learning algorithms in land classification and their effectiveness in analysing and visualizing results.

Keywords: artificial intelligence, machine learning, random forest, support vector machine, artificial neural network.

Introduction

"Are we making sufficient progress on social well-being, sustainable economic development, and environmental protection? Do we have the right information available at the right time at the right place to take adequate decisions?" (Tempfli et al., 2009). Can we keep the earth under monitoring and control at all times?

While ago, before the 1800s, people obtained a creative way to be able to get images for Remote Sensing (RS). In the mid, the 1800s were taken using air balloons, and in 1897 Alfred Nobel take the first image from an abroad rocket. At the beginning of the twentieth century, photos (images) were taken using cameras tied up to a fleet of pigeons and kites (Science Education Resource Center at Carleton College, 2022).

By the beginning of the technological revolution in the 1900s, RS has become more valuable due to using aerial photography in many fields such as wartime reconnaissance, urban development, disaster damage assessment, land use surveys, agricultural status, and oil exploration until the RS upgraded to be on the global scale (Science Education Resource Center at Carleton College, 2022).

Satellite images have become more and more important through the years and the ways to classify them are significantly increased increment through the years for many applications like disaster response, law enforcement, and more sustainable life. These applications are needed for any country.

GIS have different applications, and technological advancements have significantly enhanced GIS data, specifically how it can be used and what can be achieved as a result. GIS are decision-making tool for any business or industry since it allows the collection, storage, management, and analysis of environmental, demographic, and topographic data, and land cover.

"Data, as they say, is the new oil. It is valuable and vital to unlocking the potential of an array of new technologies, such as artificial intelligence and automation" (Behr, 2021).

The research of scientist draws inspiration and builds upon the present work (Mirsanjari et al., 2021). Their study provided valuable insights into the topic under investigation, and the findings have been taken into account in the formulation of the research questions and hypotheses for this study. The results show that Compared to 1999, the most notable change in land cover in 2019

© 2023 The Author(s). Published by Vilnius Gediminas Technical University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

was the increase in the built-up area and the decrease in the forest classes. The predicted scenario for 2039 revealed a significant increase in the Built-up area, which accounted for approximately 60% of the total area. In 2019, the percentage distribution of land cover revealed that approximately 47% of the land was covered by vegetation, but if the current trend continues, this percentage will decrease to 36% by 2039.

In this research, satellite images will be obtained using *Sentinel*-2 with band combination RGB (red, green and blue), VNIR (the visible and near-infrared), and SWIR (Short-wave infrared), different algorithms will be applied, processing algorithms and tools of software using *PostGIS – PostgreSQL*, *Quantum GIS*, and *Geoserver*.

This research aims to compare satellite image classification by different methods and machine-learning (ML) algorithms in terms of producing analytically smart visualization for the land-cover output data.

1. Satellite data and acquisition methods

The study area is located in Egypt 25 km (15 miles) from Maadi, on the southeastern edges of Cairo Governorate (Figure 1). New Cairo is one of the many new cities created in and around Cairo. The mean annual temperature is 22.1 °C.

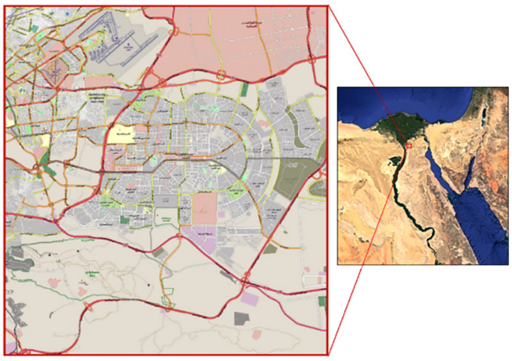

Figure 1. New Cairo (area of study)

European multi-spectral imaging mission *Sentinel*-2 is a wide-sweep, high-resolution swath. The twin satellites, which are in the same orbit, but phased at 180° are designed to have a high revisit frequency of 5 days at the Equator, according to the entire mission specification. Carry an optical payload with 13 spectral bands to sample: four bands at 10 m, six bands at 20 m, and three bands at 60 m spatial resolution. The orbital swath is 290 km wide (Sentinel Online, 2022).

The mission *Sentinel*-1 is carried out via a network of two satellites that are positioned in polar orbit. This is the first of five missions that the European Space Agency (ESA) is creating as a part of the Copernicus initiative. Each satellite is capable of performing C-band synthetic aperture radar imaging and is operational 24 hours a day, seven days a week.

The following is a list of some of the most common applications (Sentinel Online, 2022):

– Keeping an eye on the icebergs and the sea ice monitoring of ice on lands, such as glaciers, ice sheets, and ice caps, in addition to monitoring ice on rivers and lakes;
– Keeping an eye out for oil spills and ships' winds and waves of the ocean examples of these include land-use change, agriculture, deforestation, land deformation, and assistance for emergency management of catastrophes such as floods and earthquakes.

The mission of *Sentinel*-3 is to measure the topography of the sea surface, as well as the temperature of the ocean and land surfaces, as well as the colour of the ocean and land surfaces, with a high degree of accuracy and reliability in order to support ocean forecasting systems, environmental monitoring, and climate monitoring (Sentinel Online, 2022).

The mission of *Sentinel*-4 is to monitor key air quality trace gases and aerosols over Europe with a high spatial resolution and with a fast revisit time. The mission is part of the European Earth Observation Programme "Copernicus" with the ESA in cooperation with the European Environment Agency. The mission is to monitor key air quality trace gases and aerosols over Europe at high spatial resolution and with a fast (EEA) (Sentinel Online, 2022).

The project of *Sentinel*-5 is carried out as a component of the European Earth Observation Programme "Copernicus," which is managed and controlled by the European Commission (EC). The space component of the Copernicus observation infrastructure is being developed under the ESA. The primary data products being collected by this component are aerosols, O_3, NO_2, SO_2, HCHO, and CHOCHO. The focus of this component is on air quality and the interaction between atmospheric composition and climate. In addition, Sentinel-5 will provide quality parameters for CO, CH_4, and stratospheric O_3 with daily worldwide coverage for applications related to climate, air quality, ozone and surface UV (Sentinel Online, 2022).

Sentinel-5P mission is the first Copernicus mission whose only goal is to check on the state of our atmosphere. The ESA, the EC, the Netherlands Space Office (NSO), industry, data users, and scientists worked closely together to make the Copernicus Sentinel-5P satellite. The mission will be done by a single satellite that will carry the TROPOspheric Monitoring Instrument

(TROPOMI). The ESA and the Netherlands worked together to offer funds for the TROPOMI instrument. The main goal of the Copernicus *Sentinel-5P* mission is to measure the atmosphere with high spatial and temporal resolution. These measurements will be used to keep an eye on the air quality, ozone, and UV radiation, as well as to keep track of the climate and make predictions about it (Sentinel Online, 2022).

Planet Scope was established with the goal of photographing the Earth on a daily basis and making global change more visible, accessible, and actionable. Access to satellite data has been democratized beyond the conventional agriculture fence defense industries because of the work that *Planet Scope* has done over the past decade in collaboration with its customers to transform the Earth observation industry.

Planet Scope Basemaps use cutting-edge technologies in conjunction with daily satellite photos from across the world to produce mosaics that are both aesthetically pleasing and correct from a scientific standpoint. With *Planet's* Global and Select Basemaps, you may enhance your ability to do time series analysis and data analytics (Planet, 2022).

Planet Scope Global Map combines more than 1.5 million scenes every single day in order to generate an annual *Global Basemap* that is coherent and aesthetically pleasing. The *Global Basemap* is suitable for mapping, updating charts, and visualization since it is comprehensive, cloud-free, high-accuracy, and high-resolution. Additionally, the *Global Basemap* covers all of the landmasses on Earth.

Planet Scope is a constellation of around 130 satellites that are operated by *Planet* and has the capability of imaging the whole land surface of the Earth every day (a daily collection capacity of 200 mill. km^2/day). Images captured by PlanetScope have a resolution of approximately 3 m per pixel.

Product band order:

- Band 1 = Blue;
- Band 2 = Green;
- Band 3 = Red;
- Band 4 = Near-infrared.

2. Satellite image classification algorithms

RS has been defined in a variety of ways. It can be defined as the art, science, and technology of using instrument-based techniques to observe an object, scene, or occurrence. Because the observation is done from a distance and without actual touch with the item of interest, it is called remote. Either energy detection and real-time display devices or energy recording devices can be used which come from an item or a scene and is emitted or reflected It could be light or another sort of electromagnetic radiation, force fields from remote sensing, or sonic energy (Science Education Resource Center at Carleton College, 2022).

Artificial Intelligence (AI) is the ability of a computer or a computer-controlled robot to accomplish tasks that would normally be performed by intelligent beings. The phrase is widely used to refer to a project aimed at creating systems with human-like cognitive abilities, such as the ability to reason, find meaning, generalize, and learn from past experiences. It has been proved that computers may be programmed to perform extremely complicated jobs since the development of the digital computer in the 1940s (Copeland, 2022).

ML is an artificial intelligence discipline concerned with the development of computer software that can learn on its own.

As it relates to artificial intelligence, there are several types of learning. The most basic method is trial and error. A simple computer program for solving mate-in-one chess problems, for example, might try moves at random until a mate is found. The program could then save the solution along with the position so that the next time the computer encountered the same position, it could retrieve the solution. This simple rote learning of individual items and procedures is relatively simple to implement on a computer. The problem of implementing what is known as a generalization is more difficult. Generalization is the process of applying previous experience to similar new situations (Copeland, 2022).

Artificial Neural Network (ANN) is the most frequently used machine-learning technique, which may be employed effectively in non-linear phenomena such as parameter retrieval. Land Use Land Cover (LULC) evolves with the ability to work with large amounts of data. It's now one of the most widely used non-parametric classification methods. It is not predicated on the notion that data is evenly distributed. The ANN is a black-box model with a forward structure that is trained using a back-propagation algorithm (supervised training algorithm). The ANN functions similarly to a human brain or nervous system, with nerve fibres interconnected by other axons. It can learn from examples and deliver relevant conclusions, even if the input data is incorrect, confusing, or incomplete (Talukdar et al., 2020).

Support Vector Machine (SVM) is a supervised non-parametric ML algorithm that was developed to address binary classification issues. It is based on the structural risk minimization (SRM) idea, which optimizes and separates the hyperplane and data points closest to the hyper-spectral plane's angle mapper (SAM). It uses a hyper-spectral plane to divide data points into several classes. The vectors ensure that the margin width is maximized during this operation. In different class memberships, SVM may accommodate various continuous and categorical variables, as well as linear and non-linear samples. Support vectors are the training samples or bordering samples that define the margin or hyper-plane of an SVM. The polynomial and radial basis function (RBF) kernels have been employed most frequently in remote sensing, however, RBF is the most popular methodology for LULC classification and provides superior accuracy to the other classic methods (Talukdar et al., 2020), as shown in Figure 2.

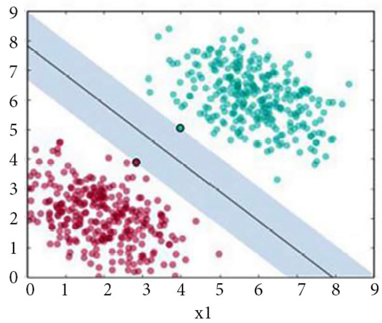

Figure 2. Support Vector Machine (Talukdar et al., 2020)

k-Nearest Neighbors (kNN) is a simple algorithm that maintains all available examples and classifies new cases based on the majority vote of their k NN. The case allocated to the class has the highest frequency among its kNN, as assessed by a distance function such as Euclidean, Manhattan, Minkowski, and Hamming distances.

Random Forest (RF) algorithm. Scientist Breiman has initiated a unique non-parametric ensemble machine-learning algorithm. The RF algorithm has been used to solve a variety of environmental issues, including water resource management and natural disaster management. It can process a wide range of data, including satellite images and numerical data. It's a decision tree-based ensemble learning system that combines huge ensemble regression and classification trees. Two factors are required to set up the RF model, which is referred to as the method's base. These parameters are

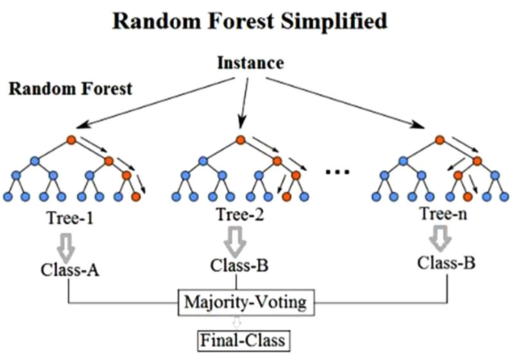

Figure 3. Random Forest (Talukdar et al., 2020)

(1) the number of trees, which "n-tree" can explain, and (2) the number of features in each split, which "m-try" can explain, as illustrated in (Figure 3) (Talukdar et al., 2020).

3. Analysis of data GIS processing

The machine has been programmed to make a variety of decisions. The training data that have been utilized here are the same as those used in the supervised classification. The machine is placed in an environment in which it may train itself continuously through trial and error.

3.1. Planet Scope platform

ML classification algorithms use a variety of methodologies to classify data into different categories or classes. Some common methodologies used in classification algorithms include RF, SVM, and kNN. As depicted in Figure 4, the diagram utilized in the research is shown.

This research utilizes various ML algorithms, RF, SVM, and kNN. The study aims to compare the results obtained from these algorithms and analyze their performance. The data obtained from the experiments will be exported and compared to determine which algorithm is most effective in the given context. This research seeks to provide valuable insights into the application of ML algorithms in practical scenarios.

This algorithm includes a goal or result variable, also known as a dependent variable, that is to be forecasted based on a predetermined set of variables (independent variables). We are able to construct a function that maps

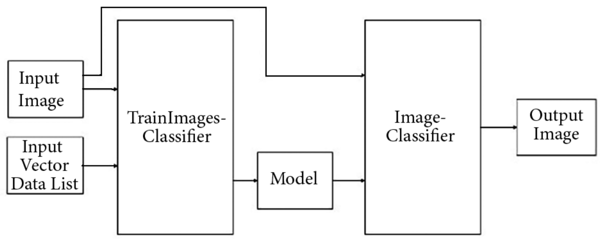

Figure 4. Machine Learning study methodology

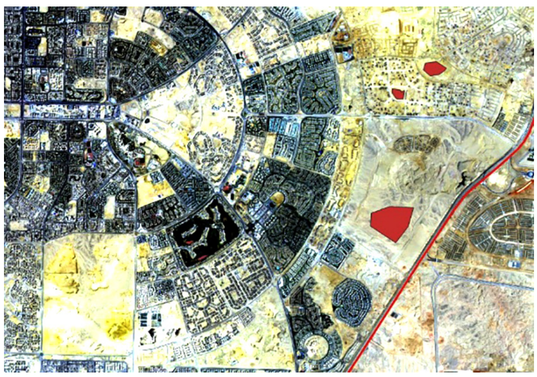

Figure 5. Training data sample

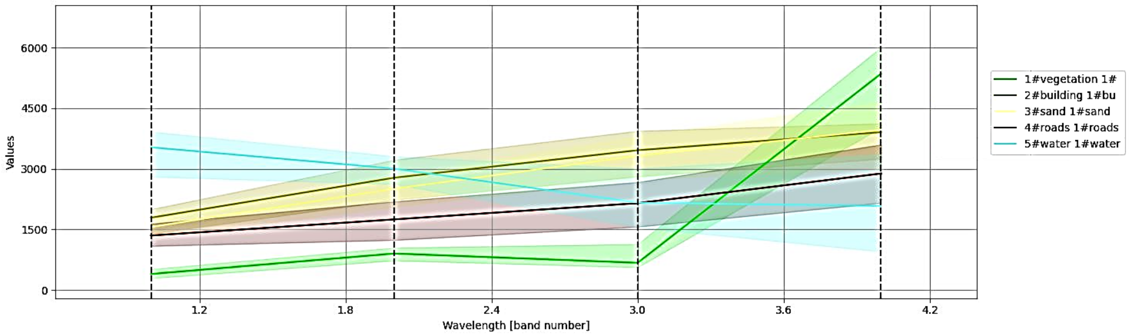

Figure 6. The signature of micro-class wavelengths

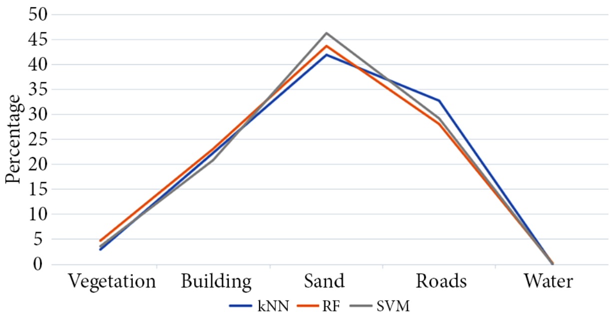

Figure 7. Planet – ML algorithms graph

inputs to the desired outputs by making use of this set of variables. The training procedure will continue until the model reaches the level of accuracy on the training data that has been determined to be satisfactory.

The sample of training data that was used in the research for the planet data set can be seen illustrated in Figure 5 above, and the wavelengths that were determined for each micro class sample of data can be seen illustrated in the Figure 6.

The training set is divided up into five different categories: vegetation, building, sand, roads, and water, in addition to an unclassified category.

Algorithms like kNN, AVM, and RF are used in ML. Since (Figure 7 and Table 1) show only a slight difference between the sand and road classes, the accuracy of ML improves with more training data, though this is also dependent on the quality of the training set.

Table 1. Planet – ML algorithms table of percentages

Type	Vegetation	Building	Sand	Roads	Water
kNN	2.94	22.35	41.86	32.80	0.047
RF	4.82	23.10	43.62	28.16	0.29
SVM	3.60	20.77	46.32	29.23	0.08

Figure 7 depicts the proportion of determined pixels out of the total number of pixels that have been classified for each class with respect to the ML for each kNN, SVM, and RF algorithm using the planet dataset, while Table 1 displays the values that have been recorded for each algorithm.

Figure 8 depicts an example of the differences between the three algorithms for the classification of planet datasets using ML.

The mean, standard deviation, and standard error for each class can be determined in regard to supervised and ML classification by computing the statistical values for each class with respect to the proportion of its pixels, as shown in the Table 2.

The comparison reveals that the mean values obtained from the *Planet* dataset record 3.48, 27.56, 44.81, 24.03, and 0.12% for the vegetation, building,

Figure 8. Planet – Machine learning algorithms RF (left) – SVM (middle) – kNN (right)

sand, roads, and water classes respectively, whimper cent standard error reveals 2.79 for the class the roads, which displays the highest error variance between the algorithms against water class which record 0.042.

Figure 9 depicts the differences between the algorithms employed on planet data, including ML and supervised classification with respect to the mean values that have been calculated.

Table 2. Statistics of classes with respect to algorithms applied on planet dataset

Type	Vegetation	Building	Sand	Roads	Water
Mean	3.48	27.56	44.81	24.02	0.12
Standard Error	0.30	2.53	1.10	2.79	0.04
Median	3.42	27.31	44.97	23.74	0.07
Standard Deviation	0.74	6.19	2.70	6.84	0.10
Sample Variance	0.54	38.31	7.28	46.86	0.01
Range	2.12	14.54	6.42	16.16	0.26
Minimum	2.70	20.77	41.86	16.65	0.039
Maximum	4.82	35.31	48.28	32.81	0.298
Sum	20.91	165.37	268.88	144.15	0.693
Confidence Level (95.0%)	0.77	6.50	2.83	7.18	0.107

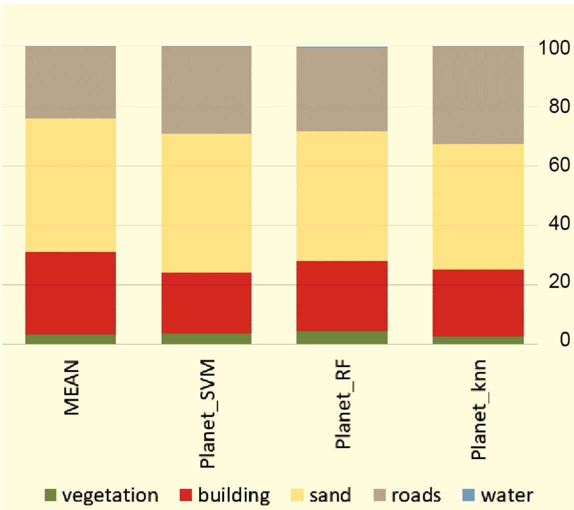

Figure 9. Differences between the algorithms for planet

3.2. Sentinel-2

The new training data set and the signature wavelengths were calculated in the Figure 10.

The illustration of the sample of training data that was used in the study for the Sentinel-2 data set is in

Figure 10. Sentinel-2 training data sample

Figure 10, and you can see an illustration of the wavelengths that were calculated for each micro class sample of data in Figure 11.

Algorithms such as kNN, SVM, and the RF algorithm are utilized in the process of ML. Considering that the following statistics reveal just a minor distinction between the road, sand, and building classes.

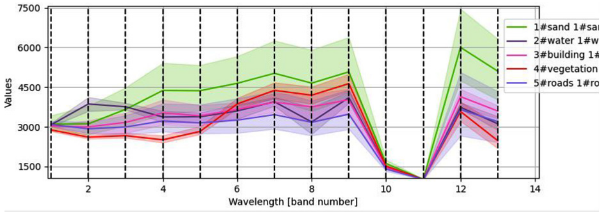

Figure 11. Sentinel-2 is the signature of micro-class wavelengths

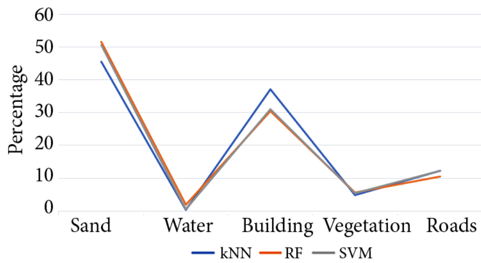

Figure 12. Sentinel-2 Machine learning algorithms

Table 3. Sentinel-2 Machine learning algorithms percentages

Type	Vegetation	Building	Sand	Roads	Water
kNN	45.51	0.394	37.062	4.83	12.20
RF	51.66	1.90	30.44	5.53	10.46
SVM	50.66	0.55	31.09	5.47	12.22

Table 3 displays the values obtained by applying the kNN, SVM, and RF algorithms on the *Sentinel*-2 dataset, while Figure 12 depicts the proportion of determined pixels relative to the total number of pixels that have been classified for each class.

An example of differences between the three ML classification algorithms for the *Sentinel*-2 dataset is shown in Figure 13.

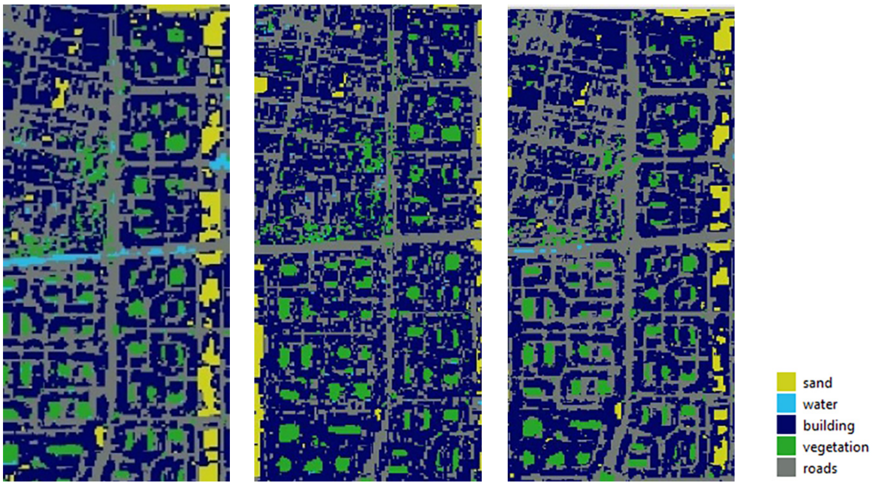

Figure 13. Sentinel-2 Machine learning algorithms RF (left) – SVM (middle) – kNN (right)

In terms of supervised and ML classification, the mean, standard deviation, and standard error for each class can be obtained by computing the statistical values for each class with respect to the proportion of its pixels, as indicated in the Table 3 for *Sentinel*-2 as well.

Mean values from the sentinel-2 dataset were 5.18% for the vegetation class, 33.48 for the building class, 49.75 for the sand class, 10.63 for the class of the road, and 0.96 for the water class. The standard error showed 1.29 for the class of the road, the class with the highest error variance between the algorithms, compared to the water class, which recorded 0.24.

Table 4. Statistics of classes with respect to algorithms applied on the Sentinel-2 dataset

Type	Vegetation	Building	Sand	Roads	Water
Mean	5.18	33.48	49.75	10.63	0.96
Standard Error	0.49	1.13	0.88	1.20	0.24
Median	5.15	33.05	50.41	11.57	0.75
Standard Deviation	1.19	2.77	2.16	2.94	0.58
Sample Variance	1.43	7.66	4.68	8.66	0.34
Range	3.42	6.62	6.15	8.08	1.51
Minimum	3.78	30.44	45.51	4.93	0.40
Maximum	7.19	37.06	51.66	13.01	1.91
Sum	31.09	200.87	298.50	63.75	5.78
Confidence Level (95.0%)	1.25	2.90	2.27	3.09	0.61

Differences between algorithms applied to *Sentinel*-2 data using ML and supervised classification are depicted in Figure 14.

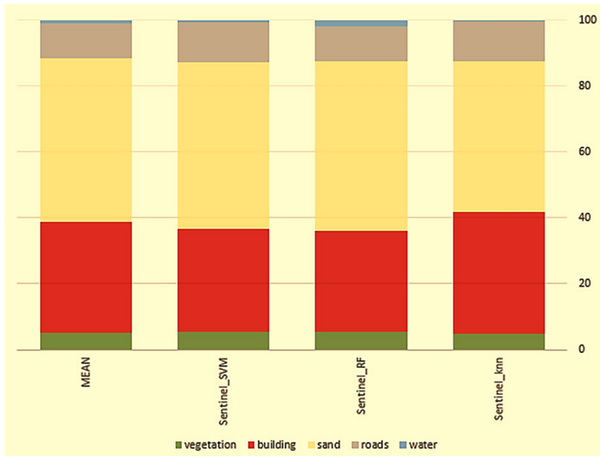

Figure 14. Differences between the algorithms for Sentinel-2

The ability to draw inferences from one's visual experiences is a crucial component of human vision, which goes beyond the simple ability to distinguish between different colours. When you are analyzing an image, you are most likely doing one of two things: direct and "spontaneous recognition" or using several clues to draw conclusions by using a reasoning process (also known as "logical inference") (Science Education Resource Center at Carleton College, 2022).

The ability of an interpreter to identify objects or features at first glance is referred to as spontaneous recognition. The circular shape of pivot irrigation systems would be instantly recognized by an agronomist. Because of previous (professional) experience, he/she would be able to do so. Similarly, the majority of people Because of scene knowledge, spontaneous recognition can directly relate what they see on an aerial photo to the terrain features of their place of residence. The quote "I see because I know" from people who see an aerial photograph for the first time refers to spontaneous recognition (Science Education Resource Center at Carleton College, 2022).

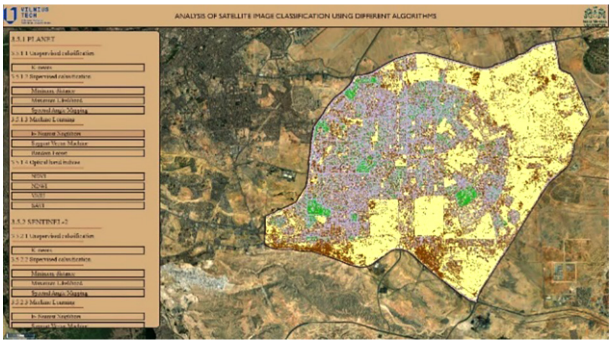

Figure 15. Application screenshot

As a consequence of this, discussing the findings with other people and propositioning their opinions will not only help to improve accuracy but will also shed light on the factors that are most relevant to the task at hand.

Develop an application that contains all of the layers' results, which will enable us to share our findings with one another and examine their perspectives on the topic.

By using HTML code that contains all of the application's texts; one for the CSS code that contains the styling for the application; and the third file will be for the javaScript since the OpenLayers library has been used.

Figure 15 is a screenshot of the application that has been created, Web Mapping Services (WMS) which accesses all layers through a GeoServer and is made available over HTTP requests, ESRI's world imagery was utilized as the base map for the application.

Conclusions

Recent advances in ML methods and AI, together with their successful implementations in a wide variety of sectors, are providing strong new tools for scientists who are actively working in the field. This is significant both for the future and how it will contribute to the reshaping of the world. highlights some useful characteristics of modern ML methods and their relevance to scientific applications.

According to the research results of the comparative analysis, this study examines and compares the different algorithms that are used for land classification. ML techniques such as kNN, SVM, and the RF algorithm show only a small difference between the sand and roads classifications for *Planet* data, despite slight differences in the values of the building, sand, and roads class for each algorithm. Additionally, they yield similar results for *Sentinel*-2 data, the technique yields distinct results for the classes of buildings, vegetation, and roads. Given that there is only a slight variation between the road, sand, and construction categories, statistical analysis suggests.

According to the study's analytics and comparison, the closest algorithms that are around the mean values and have less accumulative standard error for its classes for the sentinel data set is Support Vector Machine algorithm and Random Forest algorithm for the planet data

During the course of the research, a comparison was made between the data from *Sentinel*-2 and *Planet Scope* has shown that the classification of the land does not necessarily depend on the spatial resolution alone; rather, the spectral resolution and the number of bands also contribute significantly to the classification and help extract the necessary objective.

This research made extensive use of open source technologies, particularly for the methodologies that were applied and the open software sources that were applied. However, the image collection had to be done using both the Copernicus Open Access Hub platform, which offers access to *Sentinel*-2 data, and the *Planet Scope* platform, which required the use of research licenses that were provided as part of the Education and Research program (Planet, 2022).

References

Behr, M. (2021). *Leader insights 2021 roundup. The power of data*. https://www.digit.fyi/leader-insights-2021-roundup-the-power-of-data/

Copeland, B. J. (2022). Artificial intelligence. In *Encyclopedia Britannica*. https://www.britannica.com/technology/artificial-intelligence

Mirsanjari, M. M., Visockiene, J. S., Mohammadyari, F., & Zarandian, A. (2021). Modelling of expansion changes of Vilnius City area and impacts on landscape patterns using an artificial neural network. *Ecological Chemistry and Engineering S*, *28*(3), 429–447. https://doi.org/10.2478/eces-2021-0029

Planet. (2022). *Planet Basemaps: Comprehensive, high-frequency mosaics for analysis*. https://www.planet.com/products/basemap/

Science Education Resource Center at Carleton College. (2022). *Intro to remote sensing*. https://serc.carleton.edu/eyesinthesky2/week1/remote_sensing.html

Sentinel Online. (2022). *User guides*. https://sentinels.copernicus.eu/web/sentinel/user-guides

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y. A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations-A review. *Remote Sensing*, *12*(7), 1135. https://doi.org/10.3390/rs12071135

Tempfli, K., Huurneman, G. C., Bakker, W. H., Janssen, L. L. F., Feringa, W. F., Gieske, A. S. M., Grabmaier, K. A., Hecker, C. A., & Horn, J. A. (2009). *Principles of remote sensing: An introductory textbook*. ITC. https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesremotesensing.pdf

PALYDOVINIŲ VAIZDŲ ANALIZĖ TAIKANT ĮVAIRIUS KLASIFIKACIJOS ALGORITMUS

M. A. S. FAHIM,
J. SUŽIEDELYTĖ VISOCKIENĖ

Santrauka. Tyrimo tikslas – pritaikyti mašininio mokymo metodus ir skirtingus algoritmus žemės dangos klasifikacijai iš palydovinių vaizdų, išanalizuoti gautus rezultatus ir vizualizuoti juos taikant geografines informacines sistemas (GIS). Buvo naudoti įvairių palydovinių sistemų duomenys, o palydoviniai vaizdai suklasifikuoti remiantis įvairiais mašininio mokymo algoritmais (RF, kNN, SVM). Išnagrinėti Europos kosmoso agentūros (ESA) pagal „Copernicus" programą ir „Planet Scope" platformoje teikiami daugiaspektriai „Sentinel-2" vaizdai. Atlikta vaizdų klasifikacija naudojant skirtingus mašininio mokymo algoritmus, sukurta programa rezultatams vizualizuoti ir sugretinimui atlikti. Tyrimas parodė mašininio mokymosi algoritmų pritaikymo potencialą klasifikuojant žemę ir jų efektyvumą analizuojant bei vizualizuojant rezultatus.

Reikšminiai žodžiai: dirbtinis intelektas, mašininis mokymas, atsitiktinis miškas, atraminių vektorių mašina, dirbtinis neuroninis tinklas.