Applications of Machine Learning for Remote Sensing Satellites

Pankti Hedau
Department of Electronics & Communication
Institute of Technology, Nirma University
Ahmedabad, Gujarat, India
21bec044@nirmauni.ac.in

Markand Joshi
Department of Electronics & Communication
Institute of Technology, Nirma University
Ahmedabad, Gujarat, India
21bec065@nirmauni.ac.in

***Abstract*—The incorporation of machine learning (ML) for remote sensing technologies has completely changed several sectors, offering enhanced capabilities for data processing and interpretation. This paper explores the applications of machine learning in remote sensing satellites, focusing on four key areas: precipitation prediction, vegetation classification, mineral exploration, and oil spill detection. Through the use of advanced algorithms such as neural networks, decision trees, and support vector machines, ML models can predict precipitation patterns, classify vegetation types, identify mineral anomalies, and detect oil spills with high accuracy and efficiency. These advancements provide significant benefits, including improved environmental monitoring, better resource management, and more effective disaster response.**

Keywords— Machine Learning, Remote Sensing, Principal Component Analysis, k- Nearest Neighbour, Support Vector Machine, Neural Networks, Random Forest Classification, Clustering and Regression Analysis

I. INTRODUCTION:

The advent of satellite technology has provided unparalleled access to data and insights about our planet and beyond. Remote sensing is the practice of deriving information about the earth's land and water surfaces using images acquired from an overhead perspective, using electromagnetic radiation in one or more regions of the electromagnetic spectrum, reflected or emitted from the earth's surface. Remote sensing satellites, equipped with various sensors, capture vast amounts of data that can be utilized for a range of applications, including environmental monitoring, agricultural assessment, and natural resource exploration. However, the sheer volume and complexity of the data present significant challenges in terms of processing and analysis.

Machine learning (ML) has emerged as a transformative technology that enables efficient and effective analysis of remote sensing data. By leveraging ML algorithms such as neural networks, decision trees, and support vector machines, complex patterns and relationships can be identified within the data, leading to actionable insights. This has led to substantial advancements in fields such as precipitation prediction, vegetation classification, mineral exploration, and oil spill detection.

Here, we examine the state-of-the-art applications of machine learning in remote sensing satellites, with a focus on the aforementioned areas. We discuss the methodologies and algorithms utilized in each application, as well as the benefits and challenges associated with their implementation. By exploring these key areas, we aim to provide a comprehensive overview of how machine learning is revolutionizing remote sensing and its implications for future research and practical applications. Machine learning has become an integral part of remote sensing applications, enhancing the capabilities of satellites for various purposes.

Precipitation Prediction:

Predictive Modeling: ML models can be trained to predict precipitation patterns using satellite data such as cloud cover, temperature, humidity, and atmospheric pressure. Neural networks, decision trees, and support vector machines are commonly used algorithms.

Rainfall Estimation: Satellites equipped with sensors can capture images of cloud formations, which, when processed with ML models, can estimate rainfall amounts and distribution in real-time.

Vegetation Classification:

Land Cover Mapping: ML algorithms are implemented to classify vegetation types, distinguishing between different crops, forests, grasslands, and urban areas. This aids in monitoring land use changes over time.
Health Monitoring: ML can analyze multispectral and hyperspectral data to assess vegetation health, identify areas affected by disease or pests, and monitor crop yield potential.

Mineral Exploration:

Anomaly Detection: ML models can analyze multispectral or hyperspectral data to identify geological anomalies that may indicate the presence of minerals.
Pattern Recognition: Machine learning techniques can detect patterns in remote sensing data that suggest the presence of specific minerals, guiding exploration efforts.

Oil Spill Detection:

Image Classification: ML models can classify satellite images to identify oil spills based on their unique spectral signatures and shapes.
Change Detection: ML algorithms can compare satellite images over time to detect changes in water surfaces that could indicate an oil spill.

II. MACHINE LEARNING MODELS FOR REMOTE SENSING:

Here, we provide an overview of the machine learning models used in remote sensing applications and their roles in enhancing the analysis and interpretation of satellite data. These models are applied across various application areas such as precipitation prediction, vegetation classification, mineral exploration, and oil spill detection.

Principal Component Analysis (PCA)

PCA is a dimensionality reduction technique used to transform high-dimensional data into a lower-dimensional space while preserving significant patterns and variability.
In remote sensing, PCA is often used to reduce the dimensionality of multispectral and hyperspectral data, making it easier to analyze and visualize.

k-Nearest Neighbors (KNN)

KNN is a classification algorithm that predicts the label of an input based on its k nearest neighbors in the feature space.
In remote sensing, KNN is used for tasks such as land cover classification and identifying vegetation types.

Support Vector Machines (SVM)

SVM is a supervised learning model that finds the optimal hyperplane that separates classes in a dataset.
SVM is employed in remote sensing for tasks such as image classification and change detection.

Neural Networks

Neural networks are a family of models inspired by the human brain, capable of learning complex patterns in data.
In remote sensing, neural networks are used for tasks such as image classification, object detection, and segmentation.

Random Forest

Random forest is an ensemble learning method that uses multiple decision trees to make predictions.
In remote sensing, random forest is applied in land cover classification, mineral exploration, and other tasks that benefit from a robust, ensemble approach.

Clustering

Clustering algorithms group similar data points together without prior labeling.
In remote sensing, clustering is used to segment data into meaningful groups such as different types of vegetation or geological formations.

Regression Analysis

Regression analysis is used to model relationships between variables and predict continuous outcomes.
In remote sensing, regression is used for tasks such as predicting precipitation and estimating vegetation health.

These machine-learning models offer significant advantages in remote sensing applications, including improved accuracy, speed, and scalability. By leveraging these techniques, researchers and practitioners can unlock deeper insights and make more informed decisions based on satellite data.

III. PRECIPITATION PREDICTION AND VEGETATION CLASSIFICATION:

Precipitation prediction plays a vital role in weather forecasting, water resource management, and disaster mitigation. The only way to measure Precipitation (P) over any topographically complex domain is via remote sensing from space. By accurately forecasting rainfall patterns, decision-makers can effectively plan for water supply, agricultural activities, vegetation water content, soil moisture, and surface temperature prepare for potential flooding events.

Methodology:

Data Splitting:

There are two major precipitation types: stratiform and convective.
The first type is characterized by low precipitation rates (usually long steady rain at low rates) while the latter is the typical summer storm (short duration and intense rainfall). Moreover, although there is no consensus on a fixed precipitation rate threshold that clearly defines the limit between these two types, several studies consider the rate range of 4 ~ 6mm/hr as that threshold value. Therefore, we used the value of 5 mm/hr as our designated threshold. Further, and for the purposes of this study, we divided these two precipitation regimes ($P \leq 5$ and $P \geq 5$) in two subcategories for each one, as in “light- and heavy-stratiform” and “light- and heavy convective”, to achieve a better representation of rainfall characterization over East Africa. As such, we have divided our precipitation data into four bins.

Precipitation Level (P)	Precipitation Pattern	Precipitation Level
$0 < P \leq 2$	Light Stratiform	Level 0-2
$2 < P \leq 5$	Heavy Stratiform	Level 2-5
$5 < P \leq 10$	Light Convective	Level 5-10
$P > 10$	Heavy Convective	Level 10

Missing Values Interpolation:

Due to a couple temporary WindSat instrument failures, the recorded values for VWC, SM and ST are missing for those time periods. For each of these missing values at some time t, we perform a linear interpolation between two adjacent nonmissing values before and after t.

Models:

To identify the best predictive model that estimates the precipitation rate using the hydrological components, we trained five state-of-the-art machine learning models and evaluated their performances individually. We examined linear regression, nearest neighborhood regression, random forest, support vector regression and multilayer perception as the predictive model. In this section we briefly describe the mathematical derivations of these learning models.

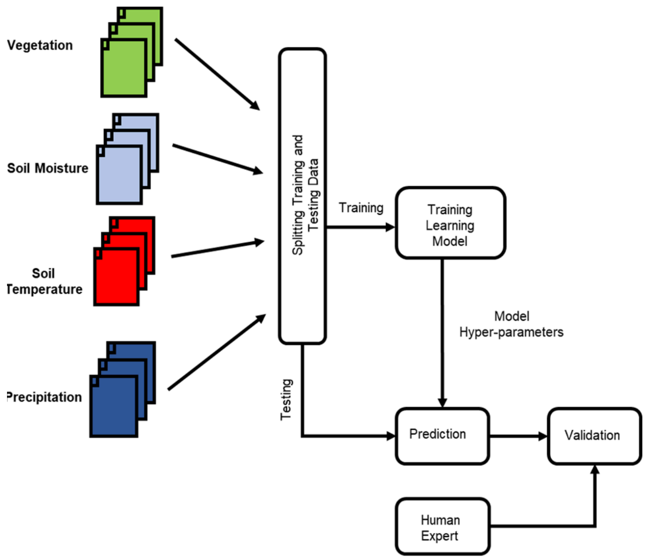

Figure: The block-diagram indicating the flow of learning methodology to predict the precipitation rate using vegetation water content, soil moisture and the soil temperature. Note that we split the white region into 80 × 80 cells and train a model for each cell.

Nearest Neighborhood

This model is the simplest predictive model, as it predicts the observed value for one testing set of feature vector as a linear combination of the observed values for the nearest feature vectors in feature space.

Linear Regression

Linear regression seeks the linear relationship between the observed values disturbed by some noise level, ϵ, and the potential predictive variables.

Random Forest

Random forest is an ensemble of learning model, constructed by a multitude of decision trees and the regressed value would be a linear combination (i.e., the mean or median) of the predicted values by each tree. Given the dataset of size K, we randomly extract N (< K) examples, fit a tree to these training models, and the final predicted value would be the linear combination of the outcome values, out of these trees. This procedure would lead to a more robust model. A single tree would be highly sensitive to the noise level, but an ensemble of the trees and taking the average of them, would decrease the variance of the model, and hence a more robust model is built.

Multi-layer Perceptron (MLP)

A class of feed-forward neural networks consisting of three fully-connected layers or more, i.e., input, hidden and output layers. MLP is a RD→RL transformer, where D and L are the input and output sizes, respectively. It would learn a non-linear transformation function like G to map the input into a space where they are linearly separable (classification mode) or they are regressed to a single value (regression mode). It also consists of an activation function which maps the weighted inputs into an output. In this work we have used Relu (Relu (x) = max (0, x)) as the activation function.

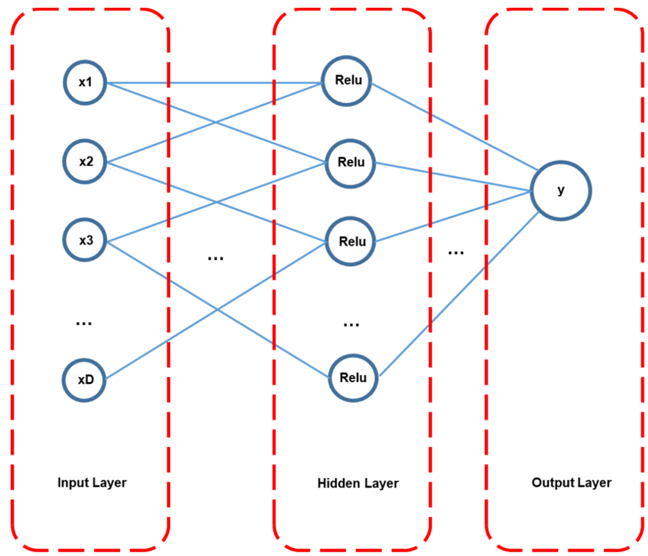

Figure: The schematic of a multilayer perception

Support Vector Regression

Support vector machine (SVM) can be used in regression mode, maintaining the variables searching for the maximal margin criterion. From a mathematical perspective, we developed a function f(x), with at most having ϵ-deviation from the target y. Here we individualize the hyperplane which maximizes the margin.

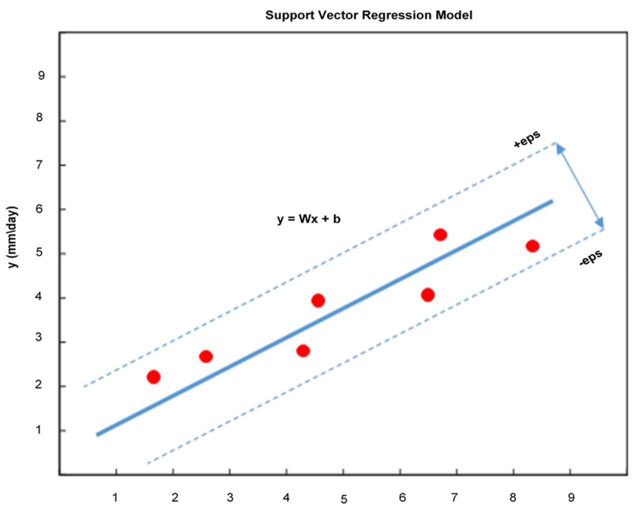

Figure: The schematic of the support vector regression model

IV. MINERAL EXPLORATION:

Mineral exploration is a critical process in identifying and assessing potential sources of valuable minerals, essential for various industries such as construction, electronics, and energy production. Machine learning offers advanced methods for analyzing remote sensing data to improve the efficiency and accuracy of mineral exploration efforts.

Methodology:

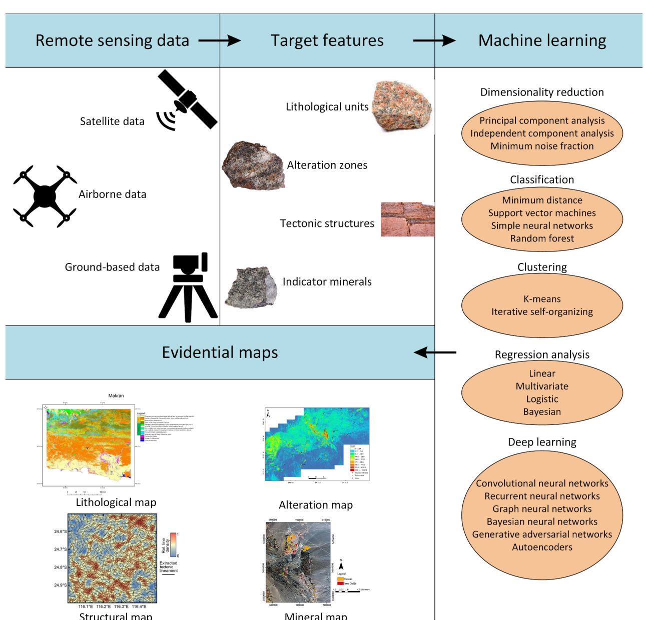

Figure: Flow Chart of ML Classification of Minerals

Remote sensing data:

Categorizing remote sensing data that are usually used for mapping geological features, particularly related to mineralization based on previous studies. The data acquiring platforms include satellites, airborne, and ground-based instruments. We summarize the characteristics of different popular remote sensing data in mineral exploration based on their platforms

Types of Remote Sensing Data:
Multispectral Data: Captures information across multiple wavelengths of the electromagnetic spectrum (e.g., visible light, infrared), providing detailed insights into surface features.
Hyperspectral Data: Collects data in a higher number of narrow spectral bands, allowing for more precise identification of materials based on their spectral signatures.
Lidar Data: Uses laser technology to measure distances and create 3D representations of the Earth's surface, aiding in geological and topographical assessments.
Synthetic Aperture Radar (SAR) Data: Uses radar waves to capture images of the Earth's surface regardless of weather or lighting conditions, offering valuable data for geological and structural analysis.
Data Acquisition and Preprocessing:
Data Sources: Satellite platforms (e.g., Landsat, Sentinel) and airborne sensors provide remote sensing data for mineral exploration.
Preprocessing: The data often requires preprocessing steps such as noise reduction, normalization, and correction for atmospheric effects to improve quality and consistency.
Geological Information:
Remote sensing data provides information on the geological and topographical characteristics of an area, including rock formations, soil types, and terrain variations.
Spectral Signatures:
Each mineral has a unique spectral signature that can be identified through remote sensing data, allowing for the detection and classification of minerals.
Data Integration:
Remote sensing data can be integrated with other data sources, such as geophysical and geochemical data, to provide a comprehensive understanding of an area’s geology and potential mineral resources.
Hence, provides a foundation for understanding how remote sensing data is used in mineral exploration and the

potential benefits of leveraging machine learning for data analysis.

Target Features:

Geological Features:
Discusses specific geological formations that are indicative of mineral deposits, such as faults, folds, and fractures.
Explains how remote sensing data can be used to identify these features, which may serve as pathways for minerals.
Spectral Signatures:
Describes the unique spectral signatures of different minerals and how they can be detected in remote sensing data.
Explains how machine learning models can identify these signatures to classify and locate specific minerals.
Anomalies:
Focuses on identifying anomalies in the data that may signal the presence of mineral deposits.
Explains how machine learning algorithms can detect unexpected variations in reflectance or absorption in the data.
Pattern Recognition:
Examines how machine learning models, such as clustering and classification algorithms, are used to recognize patterns in remote sensing data that correspond to target features.
Discusses the application of these techniques in identifying areas of interest for mineral exploration.
Integration with Geological Knowledge:
Highlights the importance of integrating remote sensing data with existing geological knowledge, including geological maps and prior exploration data.
Discusses how this integration helps in confirming findings and improving the accuracy of predictions.

Machine Learning:

Mapping geological features is a fundamental step in mineralexploration. The combined use of machine learning methods and remote sensing data can be considered an easy and inexpensive approach for mapping lithological units, alteration zones, structures, and indicator minerals associated with mineral deposits. In several fields, rapid advancements in acquiring high-resolution remote sensing data have led to the explosion of big data that offers a new opportunity for data-driven discovery.

Dimensionality reduction techniques
Dimensionality reduction techniques such as PCA, ICA and MNFare multivariate statistical approaches that transform a collection of correlated input variables into uncorrelated or independent components and have been popular for processing remote sensing data.
Classification
We discuss some of the key methods used for classification problems in the scope of our applications that consider remote sensing for mineral exploration. Although some of the methods listed (such as neural networks and random forests) can also be used for regression and prediction, our focus is classification.
Here are some classification techniques that can be implemented for classification of our data:
Support vector machines
Artificial neural networks
Random forest
Clustering:
K-means clustering (KMC) can be used for separating N data points into K clusters, where each data point is categorized into a cluster with the smallest discrepancy among its value and the mean value of the cluster. In order to classify the representative data known as cluster centers, KMC can handle a dataset with high dimensionality. As a standard clustering algorithm, KMC is widely used to process hyperspectral images for recognizing objects. However, the alignment of data points and cluster centers with the standard KMC algorithm can become very complicated due to mixed pixels.
Regression analysis
Regression is a methodology for predictive analysis that can calculate the relative effect of many factors statistically and explain logically how values depend on predictor variables. Regression analysis as a branch of machine learning and statistics which can be effectively used in remote sensing for predicting target zones.

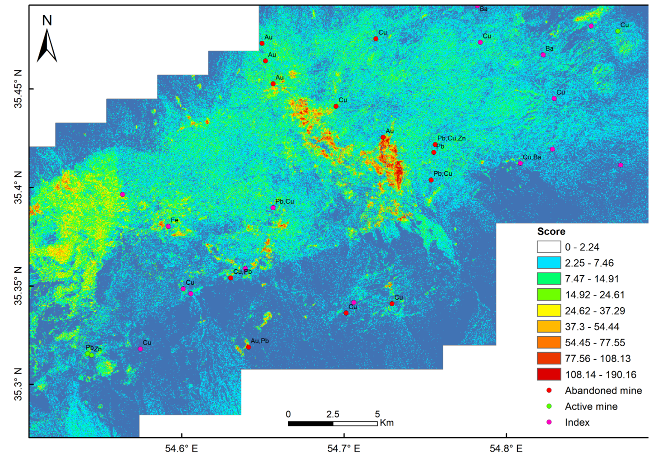

Figure: Mapping of Minerals using ML

V. OIL SPILL DETECTION:

Oil spill detection is a critical process for protecting marine environments from the adverse effects of oil contamination. Machine learning techniques applied to remote sensing data, particularly infrared (IR) images, offer a promising approach to detecting oil spills with greater accuracy and efficiency.

Methodology:

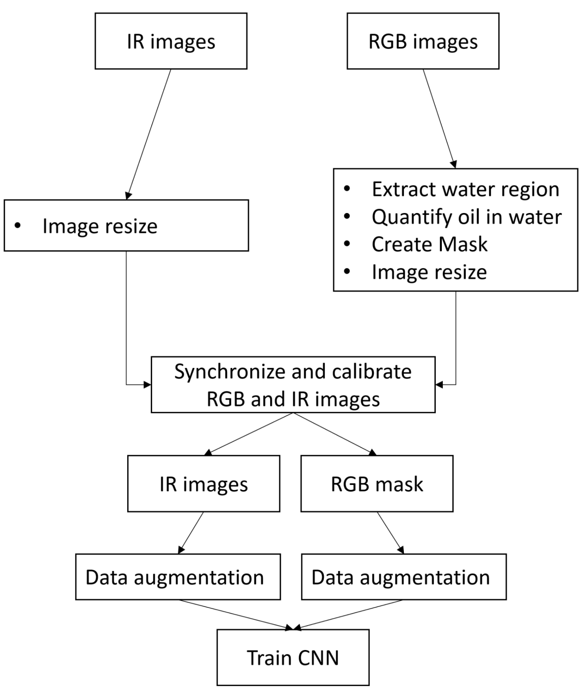

Figure: Flow chart of the training process

Training process:

To train the CNN, we use both of the RGB images and the Infrared images. From the RGB images, the oil spill is segmented, so that we have a mask representing the oil in the RGB image. We then feed both the segmented RGB images and IR images to the neural network and start the training process.

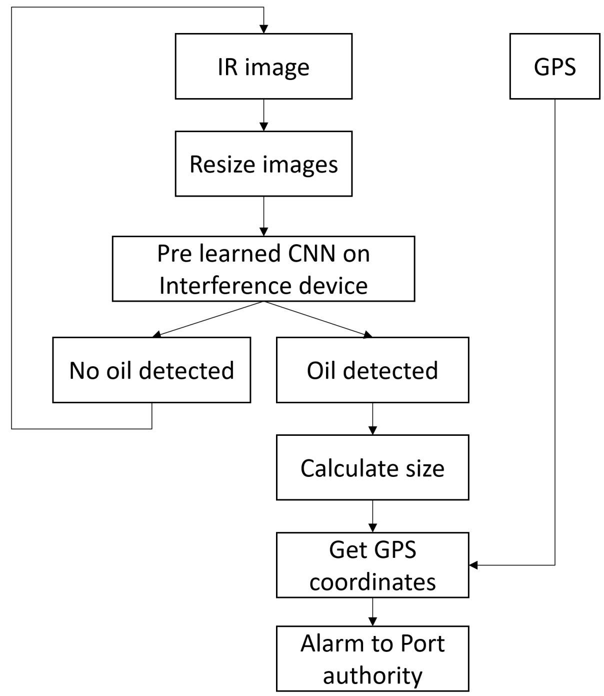

Figure: Flow chart of the operational process

Operational Process:

Once the training is done, the trained CNN can be deployed while using an interference device. Interference devices are low cost, low power computers, and are highly optimized for parallel GPU computations, ideal for CNNs. These devices make it possible to segment the images in real-time.

We start the training process with the raw IR and RGB images. For both cameras, we resize the images, from 3840 2160 pixels to 640 480, in order to reduce the training time. For the RGB images, we use thresholding algorithms to differentiate oil on the water surface. These steps are visualized in Figure shown below. After these steps are done, we end up with a binary image mask. In this mask, there are two categories, oil and no oil.

Original Image

Figure: Raw RGB image

Water Mask + Dilate + Extract Area

Figure: RGB image with the entire water/oil region extracted

Extracted Water

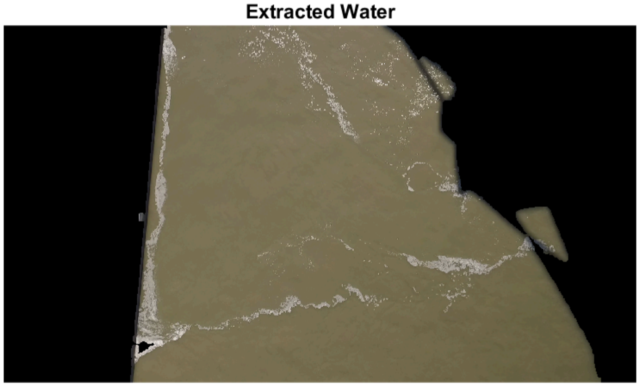

Figure: Raw RGB image with water/oil mask applied.

Oil Mask Binary

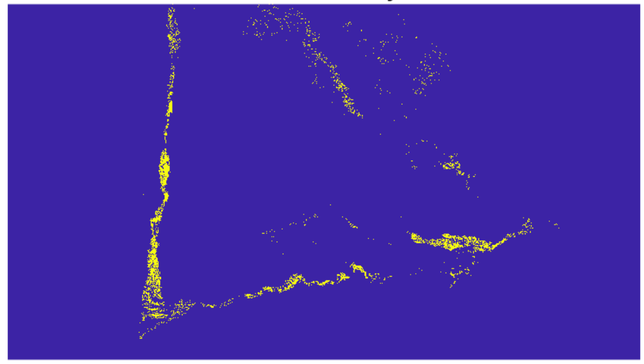

Figure: Oil extracted from above image

When these steps are finished, we end up with an IR image and a binary image with the annotated oil. Before feeding them to a neural network, the IR images on the mask are artificially augmented. Data augmentation is a common technique to improve the robustness of a neural network.

Training a Neural Network

Image segmentation is a computer vision technique, where every pixel is labeled into a predefined category The result is an image that is segmented into several categories. This creates a simplification of the image; it makes it easier to visualize certain sectors with identical labels.

Operational Process

Once we have the fully trained model, we can use this model to predict an oil spill with only the IR camera. This can be done in real time via an interference device. These interference device are small, low powered, and high performance devices that are optimized for parallel computing. The small size and low power consumption makes them the ideal choice for mounting on a UAV. Every image will be resized and segmented by the CNN on board. If there is oil detected, then the device will gather all relevant information, such as size of the spill and location. It will then send an alarm to the port authority system, where the appropriate actions can be taken.

VI. CONCLUSION:

Inspired by recent advances in artificial intelligence, and machine learning strategies in particular, as a powerful tool to approximate the physical-based hydrological models, evaluating the performance of the top state-of-the-art machine learning models to predict the precipitation rate using three potential hydrological predictors, i.e., vegetation water content (VWC), soil moisture (SM) and surface temperature (ST).

In this work, to enhance the prediction accuracy, the investigated variable was divided into four categories based on the precipitation rates, and a learning model was trained for each category. The Random forest and Linear Regression models outperformed the others, achieving the minimum prediction NMAE for most of P levels. Our results present the surface temperature as the main element to forecast the precipitation rate, followed by soil moisture and vegetation water content. We contribute this to the strong correlation between soil moisture and surface temperature in water-limited regions. Such a strong correlation would then lead into water fluxes in the region. Moreover, temperature fluctuation will directly affect the Earth's water cycle, via impacts on evapotranspiration, and changes in the conditions for cloud formation, and will consequently alter precipitation patterns. As such, surface temperature plays a pivotal role in determining the precipitation rate.

Remote sensing datasets have provided a new data resource to overcome problems associated with mapping geological features from field data alone. As a data-driven classification or prediction tool, neural networks have been widely applied in remote sensing data processing as well as a large number of research areas ranging from engineering and environmental science to physics and astronomy. Dimensionality reduction techniques can transform high-dimensional problems into a low-dimensional space and potentially mine special features from remote sensing data for mineral exploration. Recent advancements in deep learning methods have the potential to deal with large and complex remote sensing data with features in processing spectral and ground truth measurements against noise and uncertainties. Deep learning methods can be very e ective in identifying target features and mineral discovery using remote sensing data. Advanced deep learning methods can improve the mapping of geological target features for both small and large-scale studies as the success rate of mineral exploration in the face of increasing demand for critical metals.

For oil spill detection inside a port area while using an UAV and IR camera. Implementing this solution can increase the detection rate and decrease the overall cleaning costs of an oil spill. We tested our hypothesis and framework on a deliberately created oil spill under controlled circumstances.This allows for minimal human interaction during operation. This method is primarily used in order to detect smaller oil spills that might go unnoticed by the port authority. This early-stage detection is crucial in aiding the efficient cleaning of the oil spill. During the test environment, we had a field of view of 31.9 m by 42.1 m, and were able to detect oil spills that fell within the field of view.

VII. REFERENCES:

[1] Maxwell, A.E., Warner, T.A. and Fang, F., 2018. Implementation of machine-learning classification in remote sensing: An applied review. International journal of remote sensing, 39(9), pp.2784-2817.

[2] Stampoulis, D., et al. "Using satellite remote sensing and machine learning techniques towards precipitation prediction and vegetation classification." Journal of Environmental Informatics 37.1 (2021): 1-15.

[3] Diaz-Gonzalez, Freddy A., et al. "Machine learning and remote sensing techniques applied to estimate soil indicators–review." Ecological Indicators 135 (2022): 108517.

[4] Shirmard, Hojat, et al. "A review of machine learning in processing remote sensing data for mineral exploration." Remote Sensing of Environment 268 (2022): 112750.

[5] De Kerf, Thomas, et al. "Oil spill detection using machine learning and infrared images." Remote sensing 12.24 (2020): 4090.

[6] Yang, Yi-Jie, Suman Singha, and Roberto Mayerle. "A deep learning based oil spill detector using Sentinel-1 SAR imagery." International Journal of Remote Sensing 43.11 (2022): 4287-4314.

[7] Janati, M., M. Kolahdoozan, and H. Imanian. "Artificial neural network modeling for the management of oil slick transport in the marine environments." Pollution 6.2 (2020): 399-415.

[8] Anderson, R. C., Fralish, J. S., and Baskin, J. M. (2007). Savannas, barrens, and rock outcrop plant communities of North America, Cambridge University Press, 2007.