CLASSIFICATION OF BULK CARGO TYPES STORED OPENLY AT PORTS USING CNN

Madhu Koirala [1] Pål Gunnar Ellingsen [1] Roar Os Ådland [2]
[1]UiT The Arctic University of Norway
[2]Norwegian School of Economics

ABSTRACT

In this paper, we present a novel concept of tracking cargoes at open ports using remote sensing images and convolution neural network (CNN) to classify various dry bulk commodities. The dataset used is prepared using Sentinel-2 atmospherically corrected (Sentinel-2 L2A) images covering 12 spectral bands. There are total 4995 labeled and geo-referenced images for four different cargoes-bauxite, coal, limestone and logs. We provide benchmarks for this dataset using a CNN. The overall classification accuracy achieved was more than 90% for all cargo types. The dataset finds its applications in detecting and identifying cargoes on open ports.

Index Terms - Classification, Cargoes, CNN, Open Ports

1. INTRODUCTION

Automatic Identification System (AIS) is an automated tracking system used to track the movement of ships around the globe. However, a large proportion of AIS data contain some form of error due to improper installation and configuration of AIS equipment or intentional misleading inputs [3]. Moreover, it does not provide any information about cargoes stored on ports. The objective of this research is to use satellite images to track cargoes, estimate their volume and provide additional information to make the AIS data more reliable .

An enormous amount of remote sensing data is produced by the ESA(European Space Agency) Sentinel satellites with different missions such as monitoring land and ocean, measuring sea and land surface temperature, monitoring atmospheric composition, etc. The data is freely available, which has enabled the research community to lead their works to many different avenues such as land use land cover mapping [12], forest monitoring [6], solar panel energy optimization [11], oil storage monitoring [13] etc., providing the opportunities to solve many potential problems. However, there is a lack of research in the area of cargo transportation and optimization of ship positions using satellite data. This work focuses on classifying the cargoes stored openly on ports as a first stepping stone towards this research direction. Classification is done using the Sentinel-2 atmospherically corrected images. The ESA uses Sen2Cor [1] for atmospheric correction of Sentinel 2 images, which is done by modelling of the atmospheric constituents using libRadtran (Library for Radiative Transfer) [2] and mainly includes absorption and scattering of radiation by gas molecules and aerosol particles.

Deep learning methods are considered a very good choice for image classification tasks due to their ability to characterizing complex relationships between the feature maps and the targets [6]. While extensive studies have been conducted on classifying land covers [4], crops [1], minerals, and ores [7, 9, 8] using satellite imagery, limited research exists on the application of deep learning models for cargo classification or detection on ports. This research introduces a novel approach to cargo tracking by utilizing satellite images. To achieve this, we have developed a dataset consisting of openly stored cargoes in different ports and employed a CNN for their classification. This serves as a foundation for further exploration and advancements in the field of cargo tracking using remote sensing technology.

2. IMPLEMENTATION

We have used Sentinel-2 L2A products for the classification task. These images are acquired by the twin satellites Sentinel 2A and 2B, orbiting in the same polar orbit with a 180° shift, which has a combined temporal resolution of 5 days. Various steps involved in preparing the dataset is depicted in Figure 1. For dataset preparation, we have first located the cargo ports on Sentinel Hub EO Browser and found out the latitude and longitude of the ports. In some of the ports, it is

[1]`https://step.esa.int/main/snap-supported-plugins/sen2cor/`

[2]`https://step.esa.int/thirdparties/sen2cor/2.10.0/docs/S2-PDGS-MPC-L2A-ATBD-V2.10.0.pdf`

very difficult to locate the exact area of a berth covered by a cargo in EO Browser due to insufficient resolution. In such cases, google earth pro [2] or google maps[3] were used to locate the area accurately. After locating the coordinates, a python script was used to download the images across a temporal range from various ports. As the channels in the images are 10 m, 20 m and 60 m, the images are resampled to 10 m at the time of downloading. After the images were downloaded, each of them were visually inspected in RGB mode and images not containing the ground truth pixels were removed . Finally, they were resized to the same size maintaining their geo-reference information. A sample of each of the test images is presented in Figure 3.

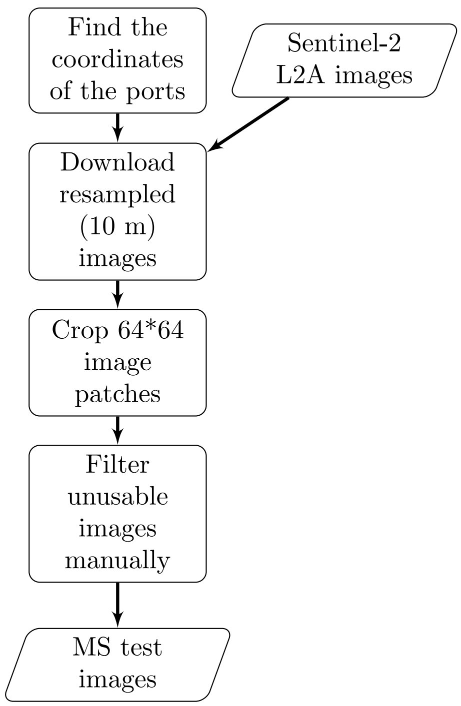


Figure 1: Dataset Preparation Steps

The dataset includes around 5000 MS(multi spectral) images, consisting of four different ores-bauxite, coal, limestone and logs from various ports scattered across different geographic locations. The ports used to download images are Tauranga, Alcoa, Kamsar, Rhodes, Trombetas, Abbot Point, Ploce, Puttalam, Qasim, Adelaide, Fujairah, Klein Point, Garcia Hernandez, Angeles, Chalmers, Marsden Point, Olympia, and Wellington. The precise location can be determined from the images as they all contain geo-reference information. These MS images have visible, near-infrared, and shortwave infrared regions of the electromagnetic spectrum.

For training the network, algorithm of SpectrumNet [10], has been used, an implementation based on the open-source code from Kaggle[4]. SpectrumNet is a convolutional neural network that expands the capability of traditional CNNs by accommodating more than just the RGB channels in images. Unlike conventional CNNs that only process RGB images, SpectrumNet can handle images with multiple channels. By leveraging the architecture of SqueezeNet [5], SpectrumNet achieves high accuracy while reducing the number of trainable parameters. As we are using 10 channels instead of 3 channels, a network with reduced trainable parameters is an optimal choice.

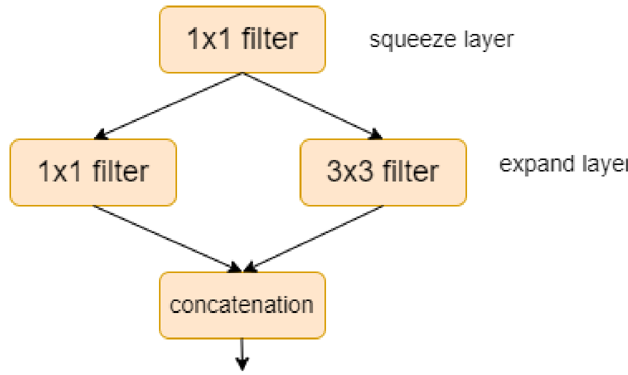


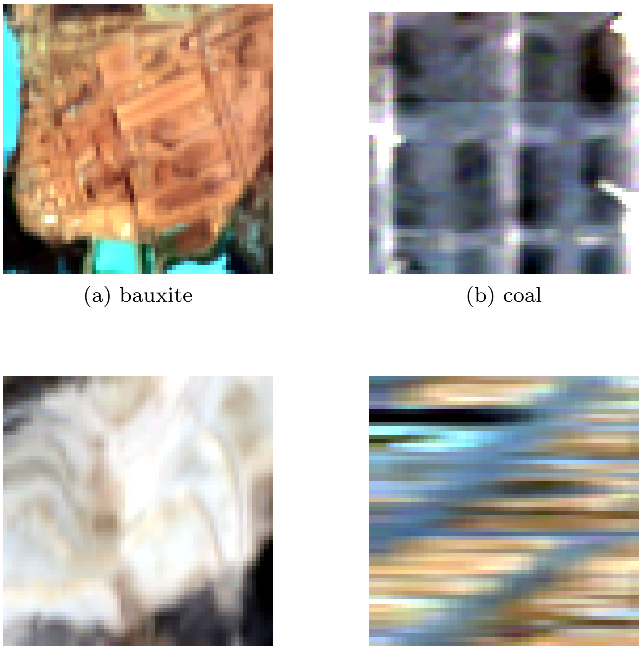
Figure 2: Spectral Module

The core building block of SpectrumNet, known as the spectral module (referred to as the fire module in the SqueezeNet paper), is illustrated in Figure 2. The squeeze layer combines and compresses the information, reducing the dimensions. On the other hand, the expand layer enhances the information by learning more intricate and diverse features. Multiple instances of this spectral module are stacked together, along with pooling operations, to construct the complete SpectrumNet architecture.

The network has been trained using k-fold cross validation using 5 folds which results in 80 : 20 train to validation ratio. The bands that have been used are RGB B04 (665.0 nm), B03 (559.0 nm), B02 (492.1 nm), vegetation red edge bands B05 (704.1 nm), B06 (740.5 nm), B07 (782.8 nm) , near infra red bands B08 (833.0 nm), B8A (864.0 nm) and short wave infra red bands B11 (1610.4 nm), B12 (2185.7 nm). Bands B01 (coastal aerosol) and B09 (water vapour) are not used as they do not contribute to the attributes of the images.

[3] https://maps.google.com

[4] https://www.kaggle.com/code/apollo2506/eurosat-allbands-classification



(a) bauxite (b) coal

(c) lime (d) timber

Figure 3: Cargo test image samples

3. RESULTS AND DISCUSSION

This section shows various results obtained from training and testing of the images. Training has been done for 5 folds and an accuracy of greater than 90% has been achieved for each fold. Figure 4 shows training and validation accuracy for the last fold of training. Both accuracy of training and validation have increased rapidly during initial epochs and gradually in the later epochs. The validation accuracy is quite low compared to training accuracy in the initial epochs, but improves with further training. This might be due to overfitting of the data in the initial epochs causing a low accuracy on validation data. As the epochs increased, the network must have learnt more meaningful features of the images, and produced a high accuracy on the validation set.

In Figure 5, accuracy of predictions on a test set is shown in a confusion matrix, which is calculated from true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions. An accuracy of more than 90% has been achieved for all cargo types on the test set also.

Table 2 presents precision, recall and F1 score for each type of cargo. The accuracy of results unequivocally demonstrate that classification of cargoes using Sentinel-2 L2A images is feasible.

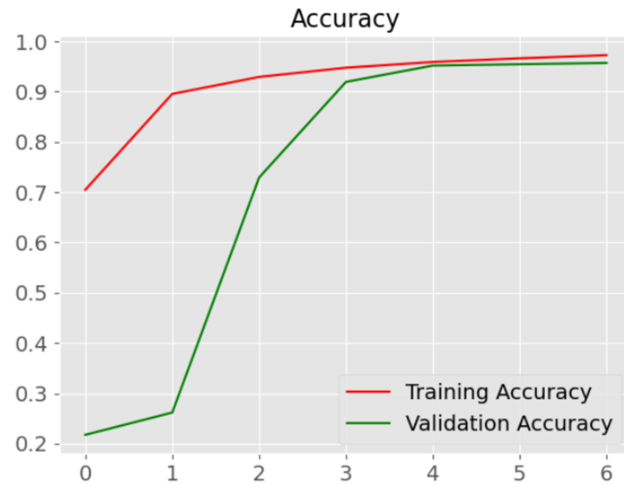


Figure 4: Training and Validation Accuracy

Table 1: Confusion Matrix Formation

	Predicted	
True	TP	FP
	FN	TN

The details of how a confusion matrix is formed is given in table 1. The diagonal of the matrix represents the correctly predicted values, which include both true positives and true negatives. True positives are instances where the model correctly identifies positive cases, while true negatives are instances where the model correctly identifies negative cases. FP indicates negative cases that are incorrectly predicted as positive, and FN indicates positive cases that are incorrectly predicted as negative.

Training of the network has been done for 7 epochs, at a constant learning rate, resulting in higher than 90% accuracy in each fold. As a good accuracy was achieved within 7 epochs, further training was not done to prevent overfitting. The high accuracy in each fold proves the robustness of the data prepared. However, as it is very difficult to generalize a dataset for all purposes, it must be noted that the prepared dataset has produced good results for classification of the images given small patches of cargoes from different ports. The training process has used 10 bands in a certain order and further training and testing with different combination and order of bands would be worthwhile for deeper understanding of the contribution of each band and is a part of future research. As the filtration of images has been done manually, it might have introduced some potential biases, because it relies on the knowledge and expertise of the checker. The classification results demonstrate the promising potential of this approach in address-

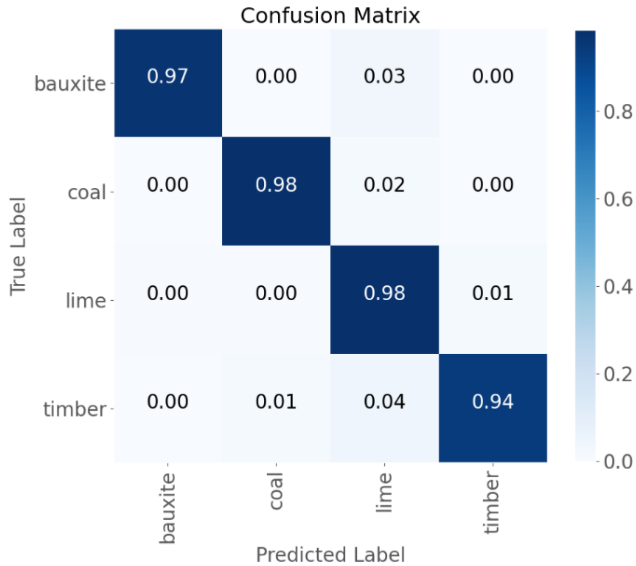


Figure 5: Confusion Matrix

ing the issue of cargo tracking. Further research can address the limitations presented by the current AIS system and bridge any existing gaps in the field.

Table 2: Accuracy Metrices

Cargo	Precision	Recall	F1 Score
Bauxite	0.99	0.96	0.98
Coal	0.98	0.98	0.98
Lime	0.91	0.97	0.94
Timber	0.98	0.94	0.96

4. CONCLUSION

The findings of the experiment indicate that it is feasible to accurately classify cargo using Sentinel-2 images with a spatial resolution of 10 m. To enhance the applicability of the dataset for practical purposes, it would be beneficial to expand the size of the dataset used in this research by including data from more ports. This would enable a broader representation of cargo classifications and facilitate more reliable generalizations of the findings. Since some of the images are initially small in size, they have been extrapolated to a larger size. This process of extrapolation might introduce potential constraints that could affect the practical usability of the data.

5. ACKNOWLEDGEMENT

This work was supported by the Norwegian Research Council [contract 326609].

6. REFERENCES

[1] Krishna Karthik Gadiraju, Bharathkumar Ramachandra, Zexi Chen, and Ranga Raju Vatsavai. Multimodal deep learning based crop classification using multispectral and multitemporal satellite imagery, 2020.

[2] Google. Google earth pro, 2022. URL: `https://www.google.com/earth/versions/`.

[3] Abbas Harati-Mokhtari, Alan Wall, Philip Brooks, and Jin Wang. Automatic identification system (ais): Data reliability and human error implications. *The Journal of Navigation*, 60(3):373–389, 2007.

[4] Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. 2017.

[5] Forrest N. Iandola, Matthew W. Moskewicz, Khalid Ashraf, Song Han, William J. Dally, and Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <1mb model size. *CoRR*, abs/1602.07360, 2016.

[6] Ma Lei, Liu Yu, Zhang Xueliang, Ye Yuanxin, Yin Gaofei, and Johnson Brian Alan. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152:166–177, 2019.

[7] Y. Liu, Z. Zhang, X. Liu, W. Lei, and X. Xia. Deep learning based mineral image classification combined with visual attention mechanism. *IEEE Access*, 9:98091–98109, 2021.

[8] Yang Liu, Zelin Zhang, Xiang Liu, Lei Wang, and Xuhui Xia. Ore image classification based on small deep learning model: Evaluation and optimization of model depth, model structure and data size. *Minerals Engineering*, 172:107020, 2021.

[9] Yuanyuan Pu, Derek B. Apel, Alicja Szmigiel, and Jie Chen. Image recognition of coal and coal gangue using a convolutional neural network and transfer learning. *Energies*, 12(9):1735, 2019.

[10] J. J. Senecal, J. W. Sheppard, and J. A. Shaw. Efficient convolutional neural networks for multi-spectral image classification. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.

[11] Shoki Shimada and Wataru Takeuchi. A machine-learning based scheme for solar pv detection using medium-resolution satellite images in vietnam, July , 2022.

[12] Naftaly Wambugu, Yiping Chen, Zhenlong Xiao, Mingqiang Wei, Saifullahi Aminu Bello, José Marcato Junior, and Jonathan Li. A hybrid deep convolutional neural network for accurate land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 103:102515, 2021.

[13] Tong Wang, Ying Li, Shengtao Yu, and Yu Liu. Estimating the volume of oil tanks based on high-resolution remote sensing images. *Remote Sensing*, 11(7):793, 2019.