Print ISSN 2285-6064, CD-ROM ISSN 2285-6072, Online ISSN 2393-5138, ISSN-L 2285-6064

EARTH OBSERVATION TECHNIQUES APPLIED FOR LAND WASTE DETECTION AND MONITORING

George BOLDEANU[1, 4], Mihaela GHEORGHE[1], Cristian MOISE[2], Iulia DANA NEGULA[2], Georgeta TUDOR[3]

[1]GMV Innovating Solutions, Calea Floreasca, District 1, Bucharest, Romania
[2]University of Agronomic Sciences and Veterinary Medicine of Bucharest, 59 Marasti Blvd, District 1, Bucharest, Romania
[3]National Institute for Research and Development in Environmental Protection, 294 Independence Embankment, District 6, Bucharest, Romania
[4]Bucharest University, Faculty of Geography, "Simion Mehedinţi" Doctoral School, 1 Nicolae Balcescu Blvd, District 1, Bucharest, Romania

Corresponding author email: george.boldeanu@gmv.com

Abstract

The dramatic increase in the amount of waste produced globally has an undeniable negative effect on the environment. The accelerated pace of urban development, the increase in consumption and the large scale of industrial activities have led to a rapid accumulation of waste in more or less proper waste dumps. All member states of the European Union are required to comply with waste management regulations, which primarily provide for the prevention of illegal dumping of any type of waste, its disposal in compliant landfills and their regular monitoring. In our ongoing project we aim to support waste management activities by proposing practical ways for Earth observation data to be used in off-site waste detection and monitoring of known landfills. Our research focuses on assessing the state of the art in earth applications techniques such as artificial intelligence/ machine learning that are currently being used for waste management and proposing approaches for building up a portfolio of scalable solutions that will support waste management not only in Romania but also at a regional, European, or global level. "This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CCCDI - UEFISCDI, project number PN-III-P2-2.1-PTE-2021-0432, within PNCDI III".

***Key words**: Deep Learning, Earth Observation, illegal waste dumps.*

INTRODUCTION

In the last 100 years global population increased 4 times, with 8 billion people reached in 2022 (World Population Prospects 2022: Summary of Results, n.d.). This drastic increase combined with changes in consumption and socio-economic patterns set the premise for appearance of illegal and unregulated dump waste sites. What a Waste 2.0 (Kaza, 2018) report estimates that until 2040 the global production of waste will be around 3.4 billion tons.

In Europe, in 2020 the waste production was of 4.808 kg/capita, from which 505 kg is classified as municipal waste (Eurostat, 2022). Due to strict regulations in waste management agreed at European Union level, Romania has an issue with illegal waste dumps resulting from various activities or from illegally imported waste. These dumps are a concern for human health and they also pose a threat for the environment with various risks associated with them, from air pollution to the spread of diseases transmitted by mosquitos (Environmental Center, 2021). Also, poorly managed wastes can cause infiltration of leachate in aquifers and the contamination of drinking water sources or rivers. Earth Observation (EO) data refers to continuously obtained data needed to improve the detection of illegal waste dumps that due to various factors can't be detected with traditional in field observations.

To perform more in-depth and least intrusive analysis of EO data, Deep Learning (DL) techniques were adopted to work with EO data, especially multi-spectral imagery. In the last 5-6 years a lot of progress was made in the field of DL and Computer Vision (CV). An approach to identify such waste dumps is proposed, which uses Sentinel 2 data with a semantic segmentation model.

MATERIALS AND METHODS

Working with satellite images for Semantic Segmentation (SS) applications has one major deficit, which is the lack of already existing data masks for various tasks. Especially with a task with a very high specificity, like dump waste identification. Our approach in resolving this issue was to create masks from already existing Corine Land Cover (CLC) dataset that has a land cover class of waste dumps, containing 1727 training sites at European level (EEA, 2018). by detailed verification of dump waste sites from Romania, it was found that CLC database is vastly inaccurate, showing a lot of inconsistencies between ground truth and polygon masks. Also, the number of waste sites covering the national territory was very small (11 polygon with dump sites). To tackle this issue and have the possibility to control the quality of polygon masks corresponding to waste dumps, a manual identification and vectorization procedure was conducted. Although manual vectorization is a time-consuming task, it has a very high accuracy. Using QGIS and high-resolution imagery from Google Satellite a total number of 344 dump waste masks were created. Because the identification of waste dumps was intended to be done on 2 different datasets, Sentinel 2 and Sentinel 2 images enhanced through SuperResolution (S2 SR) techniques inhouse, the waste dumps were filtered based on their area extent. S2 SR data was created based on Super Resolution Generative Adversarial Network (SR-GAN) inspired algorithm that was trained with SPOT-6 imagery to enhance Sentinel 2 imagery. The SR-GAN algorithm uses a generator-discriminator dual scheme, where the generator uses a ResNet structure with several residual blocks. The resulting data SR S2 data have a resolution of 2.5 meters, displaying an upscale factor of 4x. For S2 SR all the waste dump masks were used; meanwhile, for S2 just the masks with an area larger than 0.2 ha were selected. S2 data was the best option due to its higher spectral and temporal resolutions and long mission time (>5 years). To deal with the fact that vector masks were created from high-resolution imagery and the data for semantic segmentation was of medium resolution, the masks were validated with the S2 data after creation, to achieve a high degree of correlation between the two. To create the training data, images from June till September 2022 with cloud coverage smaller than 10% were downloaded. The process for creating image tiles (or patches) was conducted using Python on overlaid S2 bands. The image tiles were created with dimensions of 120 x 120 pixels and 288 x 288 pixels for S2 data and with 126 x 126 pixels for S2 SR. Only 10 spectral channels out of the 13 available were used (the 60 m bands were filtered out due to lower spatial information), and all were resampled to 10 m resolution (Table 1). To increase the number of total training data, simple augmentations were implemented with the help of Albumentations (Buslaev et al., 2020). Augmentations are a set of data transformation to increase the training data when the data is scarce. Those augmentations consisted in one Vertical Flip, one Horizontal Flip and a Transposition of the original image, resulting in a total of 1376 images.

Table 1. S2 bands used for training

Sentinel 2 band	Wavelength (nm)
Band 1	442.2 - 442.7
Band 2	492.3 - 492.7
Band 3	558.9 - 559.8
Band 4	664.6 - 664.9
Band 5	703.8 - 704.1
Band 6	739.1 - 740.5
Band 7	779.7 - 782.8
Band 8	832.8 - 832.9
Band 8A	864.0 - 864.7
Band 9	943.2 - 945.1
Band 10	1373.5 - 1376.9
Band 11	1610.4 - 1613.7
Band 12	2185.7 - 2202.4

The accuracy of several SS models was assessed in the identification of waste dumps to select the best performing model: U-Net, ResUnet, PSPNet with different backbones or without. U-Net is a fast and precise convolutional network architecture for image segmentation tasks (Ronneberger et al., 2015). The U-Net architecture represents a Fully Convolutional Network (FCN) which is composed by two sides forming a U shaped network: the contracting path on the left side and the expansive path on the right side. The contracting path (also called the encoder) follows a convolutional network architecture with 3 x 3 convolutions, followed by ReLU activation functions and 2 x 2 max pooling operations (Figure 1). The expansive path, also called the decoder, which is the symmetric part that reconstructs the precise localization of the desired features using transposed convolutions (up convolutions), ReLU activations and a final 1 x 1 convolution to reconstruct the data to the original shape of input (same height, width, number of channels).

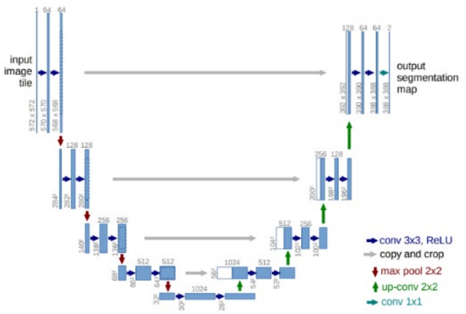


Figure 1. U-Net architecture

ResUnet is a U-net like architecture that uses residual units instead of simple neural units (convolutions followed by activation functions) to tackle the problem of vanishing gradients. Residual units are composed of 3 basic elements: BatchNormalization, ReLU activation function and convolutional layers (Figure 2). The main advantage of ResUnet is that with propagation of low level and high level information, the propagation is made without degradation, thus facilitating the design of a network with fewer parameters but comparable in performance (Zhang et al., 2017).
PSPNet or Pyramid Scene Parsing Network, is a SS model that asigns for each pixel in the image a category label using complete understanding of the scene. The main advantage of the PSPNet is that it uses dilated convolutions alongside a pyramid pooling module. The dilated convolutions are convolutions with a specified sparsity which increases the receptive field. The Pyramid Pooling Module is the central piece of the model, that captures the global context of the input image. Basically, the module upsamples and concatenates the features maps at different dimensions (1 x 1, 2 x 2, 3 x 3, 6 x 6) after which the concatenated features are convolved and lastly upsampled with a 8x bilinear upsample to create the final prediction (Zhao et al., 2016). This convolution followed by the upsampling is the decoder of the PSPNet (Figure 3).

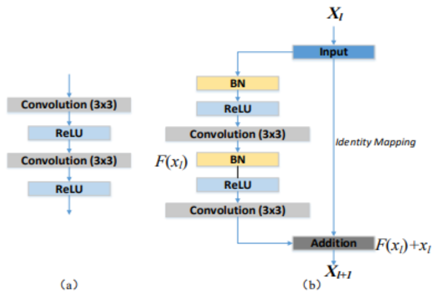


Figure 2. ResUnet residual unit

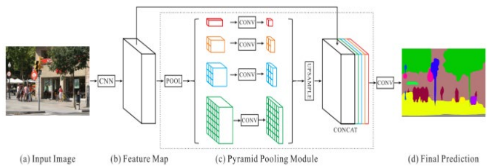
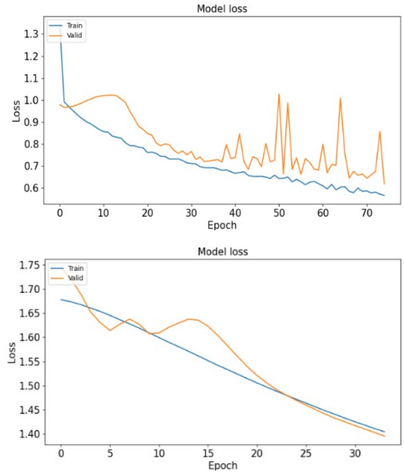


Figure 3. PSPNet architecture

For the identification of optimal hyperparameters such as: learning rate, optimizer, loss function, Optuna Framework was implemented for studying different scenarios using study/trail schema. This study/trail schema means that a study is created with all the parameters to be tested on a defined number of epochs with the desired model or models with a limited number of trials. Optuna is an automated framework that searches for the optimal hyperparameters and is not constricted by the framework where it is deployed: PyTorch, TensorFlow, Keras etc. It has an intuitive code structure and object-based orientation which makes it suitable for identifying the values to be used in finetuning the hyperparameters (Akiba et al., 2019).

RESULTS AND DISCUSIONS

Testing different approaches and sets of hyperparameters with Optuna led to the conclusion that the most important hyperparameter between the learning rate, optimizer and loss functions was the learning rate with a percentage of over 60% in most of the models tested. Testing U-Net with various backbones based on feature extractors (ResNet and DenseNet) it is observed that IOU score values are kept at a minimum (under 0.01) (Figure 4).

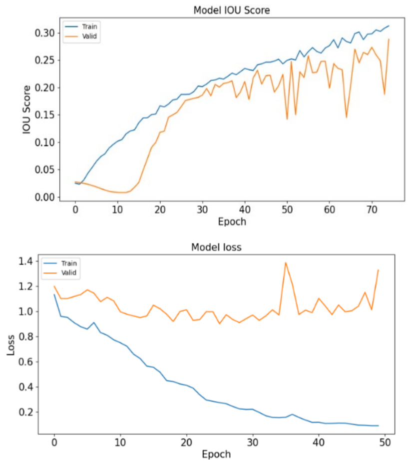


A.

B.

Figure 4. IOU score (A) and Loss function (B) for Unet with DenseNet 201 backbone

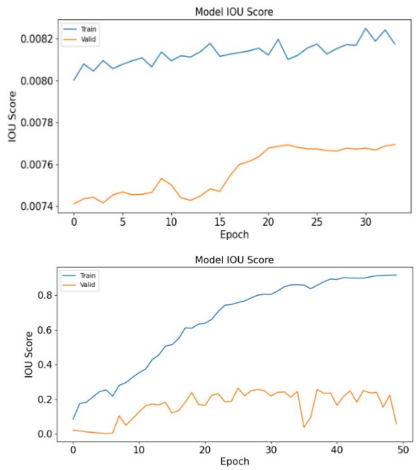
This fact can be explained by the lower learning rate that keeps the learning procedure at a minimum along the epochs. Even the network gradually improves this means that the learning procedure should be a longer one with too many epochs for training. Therefore, deeper networks (up to 60 million parameters to be trained on ResNet152) weren't tested. ResUnet gives the most promising results with values of IOU greater than 0.2 on a minimum of 50 epochs for training. The advantage of ResUnet is the propagation of information from low levels to high levels without degradation. Even if ResUnet has promising results, in Figure 5 a classic overfit scenario is presented, where the train metric kept increasing with the training time, but the validation metrics were stuck at a minimum level. This is a classical issue where learning rate plays a major role, with a value too small for an optimum learning procedure. The loss function that achieves the best results is a combination of a binary focal loss + Dice Loss, where binary focal loss is meant to discriminate between hard examples and easy examples, with a balance between positive and negative examples. Also, the Dice loss is a loss function adopted specially for semantic segmentation tasks, because is measures the similarity between two samples and works especially with imbalanced datasets, which is the case of binary semantic segmentation tasks. The combination of the two loss functions alongside a learning rate close to the optimum range leads to the best values for: validation IOU score, F1 (Dice) a score and Loss value (Table 2) (Figure 6).



A.

B.

Figure 5. IOU score (A) and Loss function (B) for ResUnet with Binary Cross-Entropy + Jaccard loss



A.

B.

Figure 6. IOU score (A) and Loss function (B) for ResUnet with Binary Focal +Dice Loss

Table 2. Results of different models and hyperparameters

Model	Optimizer	Learning rate	Loss function	Epsilon	Epochs	IOU score (max)	Loss value (min)	Val F1 score (max)
ResUnet	Adam	0.0001	Binary Cross-Entropy + Jaccard loss	0.0000001	50	0.264	0.901	0.402
ResUnet	Adam	0.001	Binary focal + Jaccard loss	1	75	0.215	0.829	0.350
ResUnet	Adam	0.001	Jaccard loss	1	50	0.163	0.838	0.278
ResUnet	Adam	0.001	Jaccard loss	1	100	0.235	0.766	0.377
ResUnet	Adam	0.001	Binary Focal + Dice Loss	1	75	0.288	0.619	0.445
ResUnet	Adam	0.001	Dice Loss	1	75	0.238	0.619	0.382
U-Net + DenseNet 201	Adam	0.0001	Binary Cross-Entropy + Jaccard loss	1	34*	0.008	1.396	0.015
U-Net + Resnet54	Adam	0.0001	Binary Cross-Entropy + Dice Loss	1	64*	0.008	1.334	0.015
U-Net + ResNet 152	Adam	0.0001	Binary Cross-Entropy + Jaccard loss	1	38*	0.008	1.144	0.015
PSPNet	Adam	0.0005	Binary Focal + Dice Loss	1	120	0.001	1.265	0.003

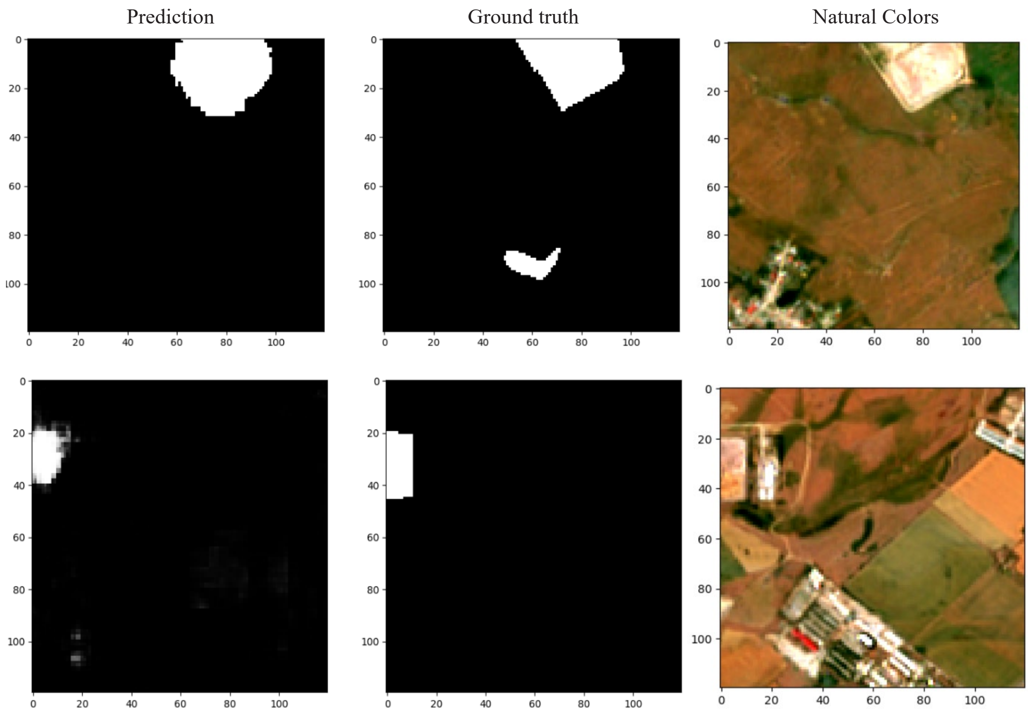


Figure 7. Inference procedure with prediction, ground truth and natural color depicted waste dump area

For the model validation task, data was derived from the initial pairs of Sentinel 2 images and dump masks. 25% percent (344 polygons) of the total data was randomly assigned for validation, Figure 7 shows an image patch in the inference step with a prediction of an administrated waste dump. The model recognizes the spectral pattern and shape of dumps that respect the norms in waste dump management and tend to have a dense material concentration. On unstructured small illegal waste dumps, the model didn't recognize the dumps from the southern part of the first image tile. This fact is due to the weak concentration of materials and therefore a weak spectral response with a lot of mixing from the surrounding area. On the second row another image tile shows how the model behaves in identifying a well-structured waste dump.

CONCLUSIONS

ResUnet is a SS model that identifies very well-structured dumps with consistent shape and spectral response. Due to the fact that the training set wasn't divided into categories of dumps, but on the criteria of area, the model was unable to differentiate between illegal waste dumps that have different properties and structure and managed waste dumps. Also, more hyperparametrization is needed for the Unet with different backbones to achieve more consistent results with higher accuracies. Even if ResUnet gives promising results, more models should be tested in the near future on a more robust training dataset that will be more rigorously created, with various approaches for limiting uncertainty. Examples of models to be tested are Mask R-CNN or DeepLabv3.

REFERENCES

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.* 2623-2631 https://doi.org/10.48550/ARXIV.1907.10902.

Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A.A. (2020). Albumentations: Fast and Flexible Image Augmentations. *Information*, *11*(2), 125. https://doi.org/10.3390/info11020125.

EEA (2018). *Corine Land Cover 2018*. Environmental European Agency. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018?tab=metadata

Environmental Center (2021). *The Hidden Damage of Landfills* https://www.colorado.edu/ecenter/2021/04/15/hidden-damage-landfills.

Eurostat (2022). *Waste Statistics*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Waste_statistics#Waste_generation_excluding_major_mineral_waste.

Kaza, S. (2018). *What a waste 2.0: A global snapshot of solid waste management to 2050*. Urban Development. © Washington, DC: World Bank. http://hdl.handle.net/10986/30317.

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, 9351, 234-241. https://doi.org/10.48550/ARXIV.1505.04597.

World Population Prospects (2022). *Summary of Results*. (n.d.).

Zhang, Z., Liu, Q., & Wang, Y. (2017). Road Extraction by Deep Residual U-Net. IEEE *Geoscience and Remote Sensing Letters.* https://doi.org/10.48550/ARXIV.1711.10684.

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2016). Pyramid Scene Parsing Network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* https://doi.org/10.48550/ARXIV.1612.01105.