GEOMEMBRANE BASINS DETECTION BASED ON SATELLITE HIGH-RESOLUTION IMAGERY USING DEEP LEARNING ALGORITHMS

Mohamed Benayad*, Nouriddine Houran*, Zakaria Aamir, Mehdi Maanan, Hassan Rhinane

1 Geoscience laboratory, Department of geology, Faculty of Sciences Ain Chock, University Hassan II Casablanca, BP 5366 Maarif Casablanca Morocco.

Commission IV, WG 7

KEY WORDS: Convolutional Neural Networks. Darknet, Deep learning, Detection, Environment, Geomembrane basin, Water save, Yolo

ABSTRACT:

Agriculture is a very important economic sector in Morocco, which requires a set of tools to improve agricultural production. Among these tools, the use of geomembrane basins. The latter is of great importance in smart farming planning, management practices or even in livestock use. In this context, this study evaluates a recognition and classification of the geomembrane basin using remote sensing satellite images; based on the Yolov3 deep learning neural network. This paper first adjusts the network model to make it suitable for detecting small targets on remote sensing images, then uses the k-means algorithm to calculate the grid size of the Yolo network model suitable for geomembrane basins, then uses yolov3 to train the data that makes up the satellite remote sensing imagery. The network model for the detection of the geomembrane basins is obtained by the test phase. Finally, the geomembrane basin detection model adapted to the remote sensing image is obtained in the validation phase. Through the research and analysis of the experimental results, it can be seen that this method effectively detects the geomembrane basins in the remote sensing images and ensures the high detection accuracy of the experimental results, which gave us an accuracy of 75%.

1. INTRODUCTION

In the past several decades, it has been a great challenge to maintain the sustainable use of water resources, especially in arid and semi-arid regions, due to increasing irrigation demands and climate change (Hotchkiss et al., 2001). Since agricultural irrigation consumes the largest amount of water resources in some countries (Moreno-Camacho et al., 2019), decreasing agricultural water use becomes one of the most promising water-saving methods, in which the most important technology involves increasing water use efficiency (Wallace, 2000). In a context where water resources and the environment are of paramount importance, the water needs of agriculture must be met by water storage basins, in order to be able to maintain a sufficient low water flow, guaranteeing the fauna and flora of the sampling areas. The storage basins make it possible to withdraw water in winter in order to maintain a low flow during dry periods in summer, while providing farmers with a water resource for irrigation (Poulain & Fischer, 2015).

Deep learning has in recent years set an exciting new trend in machine learning (Rusk, 2016). The theoretical foundations of deep learning are well rooted in the classical neural network (NN) literature. But different to more traditional use of NNs, deep learning accounts for the use of many hidden neurons and layers—typically more than two—as an architectural advantage combined with new training paradigms (IEEE Xplore - Under Maintenance, n.d.-a).

A large number of remote sensing images have been generating regularly, and due to the rapid development of satellite and imaging technology, the task of object detection has gained significant attention of researchers (IEEE Xplore - Under Maintenance, n.d.-b). The objective target detection in remote sensing images is to identify object of interest and, then, predict the type and location. In other words, target detection is the process of detecting instances of semantic objects of a certain class (such as humans, vehicles airplanes, or ships) in digital images and videos. Analyzing such images contributes to social and economic aspects for decision-making as they provide a valuable source of information. For this reason, it has been applied for many applications, such as navigation, disaster management, road segmentation, agriculture survey, urban planning, geographic information system updating, intelligent monitoring, and many more. However, target object detection for each application poses some no overlapping challenges (IEEE Xplore - Under Maintenance, n.d.-c).

In 2015, YOLO (IEEE Xplore - Under Maintenance, n.d.-d) introduced an integrated detection scheme that combines candidate frame extraction, CNN learning features, and NMS optimization in order to simplify the network (Redmon & Farhadi, 2018) structure. The detection speed is nearly ten times faster than that of the R-CNN. This enables the deep learning target detection algorithm to meet the requirements of real-time detection tasks while utilizing the computing power available at the time; however, detection performance on small targets is poor. YOLOv3 (Redmon & Farhadi, 2018) utilizes the residual network based on YOLOv2, and the binary cross loss function is used as the loss function to combine the feature pyramid

network (FPN) structure. The process of target classification and bounding box regression prediction is basically the same.

The objective of this research was to design an appropriate methodology for the detection of basin geomembranes from satellite images using the Yolo model. as well as better accuracy to help detect geomembrane basins for agricultural development planning and management.

2. DATA AND METHODS

2.1 Data Collection

2.1.1 Images Acquisition: The images used to train the model are extracted from GoogleEarth satellite imagery with very high resolution which makes objects on them clear, visible and distinguishable. These images mined from several agricultural areas in the region of Marrakech, Morocco which consist of five principal objects that can listed as buildings, bearlands, and roads, cultivate areas and geomembrane basins.

The dataset contained 100 different images so that each image includes from one single basin to four basins with different shapes take in square, rectangle and indefinite shape, as some basins are full and others empty the thing that makes us surround all the possible states that any basin can have in order to train the model on all cases to increase its accuracy.

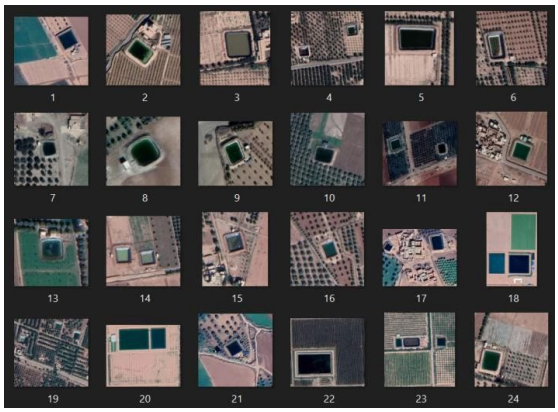

Figure 1. A section of images dataset used to train the deep learning model.

2.2 Image Compressing

To obtain good results in model training, high-resolution images are required, the thing that makes them large in size and capacity, which is not desirable in the next stages, so resorting to image compression in the way to reduce its capacity and maintain its resolution quality.

As presented in (Figure.2), The image in the right represent a high-resolution image with 55,9 MB in the size compressed to the left image with 4,60 MB instead while conserving the same resolution.

Figure 2. High-resolution compressed image

2.3 Images Labelling

Images labelling consider the most important stage to train the deep learning model, YOLO in our case, to recognize geomembrane basins done with specifying where the custom object is located on the specific image using external software. Many of them are placed for this purpose as we based on label-studio in our paper.

As shown in the (figure.3), the labeller provide us to specify where in the whole image the target object is located even the object is single or more in the same image. Most of thoselabellers offer different extensions and data types to extract the dataset after the labelling such as YOLO format that comes as jpg images accompanying with text files that represent the bounding box characterize the objects goal in the images, and other formats like JSON, XML, JPG and so on.

In this paper, more than 300 basins are bounded with the label “G.Basin” to train the model and then exported to YOLO format to pass into the training.

Figure 3. A,B: Image labelling with bounding boxes usinglabel-studio software.

2.4 Workflow

Geomembrane basin model prediction based on high-resolution satellite imagery and deep learning is divided into five major sections (Figure.4) where the first was images collection from GoogleEarth using SAS Planet software that guarantees the downloading with resolution tuning. To comes in the second section images compressing which is not important if the machine that will run the algorithms is in very high performance.

Moving to the most important step which is data labelling that need an external software or web platform to give label for each single object in the image to define them and make the model learn from them later.

These three previous steps can be assembled in the data preparation part, to take place the dataset training using deep learning algorithms, YOLO accompanied with Darknet which are considered as fast and highly accurate models in object detection.

After the dataset training, it is necessary to test the model performance and precision, so that the model testing section was taking place assuring an accuracy of 75% in geomembrane basins detection.

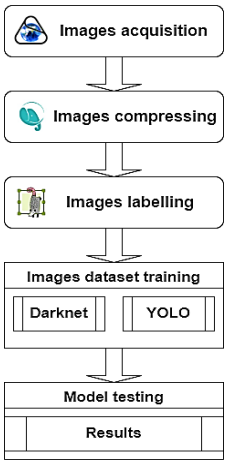

Figure 4. Workflow of geomembrane basins based on high resolution satellite images using deep learning algorithms

3. ALGORITHMS

3.1 Yolov3 network model structure

The YOLO series algorithms are originally target recognition methods based on regression proposed, (Redmon et al., 2016) YOLO had been developed into its third generation, i.e., YOLOv3, which has a rapid detection speed and high detection accuracy for small and dense targets. YOLOv3 uses the multi-scale prediction method to improve the defects of YOLOv2 for small target recognition, significantly improving the recognition accuracy of small targets while maintaining the rapid detection speed of YOLOv2. Therefore, YOLOv3 has a high detection accuracy and fast speed. (figure.5) shows the YOLOv3 network structure. First, YOLOv3 scale the original image to a size of 416 pixels × 416 pixels. After extraction of features with Darknet53, the original image is transformed into a feature map with a size of 13 × 13. Three feature maps are formed by combining two feature maps with sizes of 26 × 26 and 52 × 52. In other words, detection is performed on three scales, such that the feature map is transmitted to the two adjacent scales using twice the up-sampling. On each feature map, each cell predicts three bounding boxes by means of three anchor boxes, finally selecting the most suitable bounding box, which is shown in (Figure.6). For each bounding box, the network predicts its center point (XY), width and height (WH), confidence, and category. For an input image, the final output dimension is $1 \times ((13 \times 13 + 26 \times 26 + 52 \times 52) \times 3) \times (5 + k) = 1 \times 10{,}647 \times (5+ k)$, where k represents the number of categories (Ma et al., 2019).

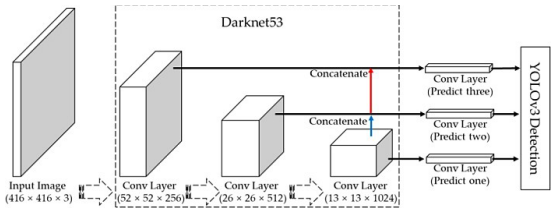

Figure 5. YOLOv3 network structure, where the blue and red lines represent two-fold up-sampling.

Figure 6. Schematic of the YOLOv3 prediction bounding box with 13 cells × 13 cells. (a) YOLOv3 detection process on 13 cells × 13 cells feature map; (b) YOLOv3 detection result.

In YOLOv3, the Darknet53 convolution network is the feature extractor, which is shown in (Figure.7). Darknet53 is mainly composed of a series of convolution layers at

dimensions of 1 × 1 and 3 × 3, with a total of 53 layers (including the last fully connected layer but excluding the residual layer). Each convolution layer is followed by a batch normalization (BN) (Ma et al., 2019)layer and LeakyReLU layer. A number of residual network modules were introduced in Darknet53, i.e., the residual layer shown in (Figure.7), which was derived from ResNet . The purpose of adding the residual layer is to solve the gradient disappearance or gradient explosion problems in the network, such that we can more easily control the propagation of the gradient and perform network training (Ma et al., 2019).

Layer	Filters size	Repeat	Output size
Image			416 × 416
Conv	32 3 × 3/1	1	416 × 416
Conv	64 3 × 3/2	1	208 × 208
Conv Conv Residual	32 1 × 1/1 64 3 × 3/1	Conv Conv Residual × 1	208 × 208 208 × 208 208 × 208
Conv	128 3 × 3/2	1	104 × 104
Conv Conv Residual	64 1 × 1/1 128 3 × 3/1	Conv Conv Residual × 2	104 × 104 104 × 104 104 × 104
Conv	256 3 × 3/2	1	52 × 52
Conv Conv Residual	128 1 × 1/1 256 3 × 3/1	Conv Conv Residual × 8	52 × 52 52 × 52 52 × 52
Conv	512 3 × 3/2	1	26 × 26
Conv Conv Residual	256 1 × 1/1 512 3 × 3/1	Conv Conv Residual × 8	26 × 26 26 × 26 26 × 26
Conv	1024 3 × 3/2	1	13 × 13
Conv Conv Residual	512 1 × 1/1 1024 3 × 3/1	Conv Conv Residual × 4	13 × 13 13 × 13 13 × 13

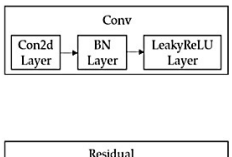

Figure 7. Structure of the Darknet53 convolutional

4. RESULTS

In this model, the performance of the deep learning predictor was tested on some new images that not trained before which attend (99,8%) as a maximum accuracy predicting geomembrane basins (figure 8) with a precision of (83, 3%), recall of (83,3 %), and mean average precision m.A.P of (80,6%) that consider satisfied results in deep learning objects detection, (figure.9).

Figure 8. A,B: Geomembrane basins detection results.

Precision	Recall	A.P.
83.3%	83.3%	80.6%

Figure 9. Model metrics

5. RESULTS CRITIC

As shown in (Figure.10), YOLO do not perform well for objects that are very close to each other (middle points of more than one objects fall into the same grid), and for small object group. Because there are only two boxes belong to one category are predicted in a grid. There are problems of bad training approximation and generalization for test image of unusual aspect ratio. Due to the problem of loss function, positioning error is one of the most important reasons that affect the detection effect, especially for large or small objects. Again, YOLO even if can detect the object trained, this model can’t detect the exact shape of the target.

Figure 10. YOLO model drawbacks.

6. DISCUSSION AND CONCLUSION

Among the most used model for object detection we found Yolo model where based on two steps which are applied to images of predefined size during learning. The first step is the detection of objects operated by convolutional neural networks. The second is a grid of the image where the class of the object is predicted if it exists (in our case it is a geomembrane basin).

One of the main problems while training these detectors is the number of geomembrane basin images needed to train the classifiers, as well as the quality of the images generally the basins of these images must be cantered and of the right size. In our case, the data used for the training is composed of a canteen of annotated images containing in all nearly 150 geomembrane basins. Model improvement depends on training data, epochs, batch size, and some other factors. The results showed that the network is able to generalize the image better and process images faster, which gave us an accuracy of 75%.

In recent decades, it has been difficult to maintain the sustainable use of water resources, especially in arid and semi- arid regions,

due to the increasing demand for irrigation and climate change. In many regions in Morocco, the use of the geomembrane basin to increase the yield of agricultural cultivation and the intelligent use of water. With these questions in mind, this research is motivated by the need to prepare a sensing model of the geomenbrane basins for high and medium resolution satellite imagery and this is based on the YoloV3 deep learning neural network. Experiments on the geomembrane basins dataset demonstrate that the YOLO model has strong applicability for satellite imagery, especially in terms of prediction speed, which gave a 75% accuracy. The main disadvantages of YOLO are its low positioning accuracy, poor learning approximation, and generalization for images of unusual aspect ratio and objects very close to each other. It needs a large number of high-quality Ground Truth labels for model training, which relies on professional interpretation experiences and a lot of manual work. Therefore, solving these problems is the direction of future research.

REFERENCES

Hotchkiss, R. H., Wingert, C. B., & Kelly, W. E. (2001). Determining Irrigation Canal Seepage with Electrical Resistivity. *Journal of Irrigation and Drainage Engineering*,*127*(1),20–26. https://doi.org/10.1061/(ASCE)0733- 9437(2001)127:1(20)

IEEE Xplore—Under Maintenance. (n.d.-a). Retrieved May 6, 2022, from https://s3-us-west-2.amazonaws.com/ieeeshutpages/xplore/xplore-ie- notice.html?

IEEE Xplore—Under Maintenance. (n.d.-b). Retrieved May 6, 2022, from https://s3-us-west-2.amazonaws.com/ieeeshutpages/xplore/xplore-ie- notice.html?

IEEE Xplore—Under Maintenance. (n.d.-c). Retrieved May 6, 2022, from https://s3-us-west-2.amazonaws.com/ieeeshutpages/xplore/xplore-ie- notice.html?

Ma, H., Liu, Y., Ren, Y., & Yu, J. (2019). Detection of Collapsed Buildings in Post-Earthquake Remote Sensing Images Based on the Improved YOLOv3. *Remote Sensing*, *12*(1), 44. https://doi.org/10.3390/rs12010044

Moreno-Camacho, C. A., Montoya-Torres, J. R., Jaegler, A., & Gondran, N. (2019). Sustainability metrics for real case applications of the supply chain network design problem: A systematic literature review. *Journal of Cleaner Production*, *231*, 600–618. https://doi.org/10.1016/j.jclepro.2019.05.278

Poulain, D., & Fischer, S. (2015). *RETOUR D’EXPÉRIENCE SUR LES DISPOSITIFS D’ÉTANCHÉITÉ PAR GÉOMEMBRANE (DEG) UTILISÉS SUR DES BASSINS D’IRRIGATION*. 8.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *ArXiv:1804.02767 [Cs]*. http://arxiv.org/abs/1804.02767

Rusk, N. (2016). Deep learning. *Nature Methods*, *13*(1), 35–35. https://doi.org/10.1038/nmeth.3707

Wallace, J. S. (2000). Increasing agricultural water use efficiency to meet future food production. *Agriculture, Ecosystems & Environment*, *82*(1), 105–119. https://doi.org/10.1016/S0167-8809(00)00220-6