DOI: 10.18372/2310-5461.62.18708
UDC 623.454.3:528.854(623.746.519)

Ievgen Saprykin
Scientific Centre for Aerospace Research of the Earth,
NAS of Ukraine, Kyiv, Ukraine
orcid.org/0009-0004-2347-3756
e-mail: ye.sapr@gmail.com

OPTICAL DEEP LEARNING LANDMINE DETECTION BASED ON LIMITED DATASET OF AERIAL IMAGERY

Introduction

The problem of airborne landmine detection consists of two very unequal parts – one is optical detection of visible landmines, another is detection of hidden landmines using combination of sensors using different physical principles. In this paper I will consider only the first one to reasonably limit the scope of investigation. Despite the problem of hidden mines, optical detection is rather actual as a separate task, because installation of hidden mines is laborious and requires human involvement, unlike installation of landmines on the surface that is simple and can be done either by variety of technical equipment or by using cluster ammunition. Detection of hidden landmines, and particularly, using deep learning for that, is much more challenging task, mostly related to fusion of data from different sensors – magnetometers, infrared cameras, ground penetrating radars, etc., which is also discussed in many recent publications, but is not even close as actively developed as optical object detection. So it makes sense to design appropriate solution for optical detection before diving deeply into hidden mine detection.

Let's try to define the role and place of optical landmine detection in global picture of humanitarian demining process. Nowadays, a realistic goal can be set to create a kind of country-level geoinformational system of landmine contamination, where the territory can be divided into several hypothetical classes: unexplored, optically observed using unmanned aerial vehicle (UAV), professionally explored using UAV-based hidden mines detection methods, manually explored by professional team, etc. It makes sense to foresee an opportunity for non-professionals to participate in data acquisition. Because of shortage of professional deminers, Ukrainian farmers are known for their self-organized demining efforts, so it is reasonable at least to provide standardized tools for optical detection and data storing based on cheap commercial drones (that farmers may be familiar with in the context of agriculture) and free open-source data processing solutions. We also can expect teams of civil volunteers to participate at some stage.

When a large amount of data is centrally available, it becomes possible to have an application in smartphone that warns user about dangerous closeness of landmines just like some applications warn drivers about traffic jams. Taking into account the sensitivity of information, end user can be provided with information only about the closest neighborhood, while authorized professional teams may have extended access mode.

So, taking into consideration all mentioned above, I come to idea that architecture of optical detection system has to be as simple and as standard as possible.

Analysis of recent research and publications

To define the place of optical UAV landmine detection among the wide area of object detection and computer vision tasks, I have to mention that it is relatively new subdivision of *UAV remote sensing* tasks, the variety of which is described in recent publications [1], that have a lot in common, but each is specific in some way. These tasks differ in size of investigated objects, environmental conditions of sensing, dynamics of detected objects' behavior, requirements for on-board processing, etc.

First of all, it is worth mentioning that there were successful efforts of solving the problem of optical detection without neural networks and deep learning. Such a method is described in paper [2] by participants of the current research. In those studies, optical landmine detection is considered as a problem of statistical detection of a prescribed signature over a random background with Gaussian distribution. In fact, *machine learning* was used, though it was not *deep learning*. The conceptual difference is that no information about the shape of landmine was used, with exception of round form of sliding window, which was not a distinctive feature of the method. So we can expect better results using neural networks and deep learning, either in higher performance rate, or lower requirements to resolution, but only in the case of enough training data availability. Currently, I have to admit that we can't tell precisely how much training data is enough in the case of aerial landmine detection. Most likely the answer to this question will be determined experimentally, rather than in strict mathematical way.

Last several years advanced groups of researchers published their results of convolutional neural network(CNN) usage in UAV landmine detection, some of them generously sharing both code and data with public [3]. The inspiring contribution of such efforts is essential for new researchers, however, the progress in NN development is so rapid nowadays, that some statements become questionable, like that about Faster R-CNN showing better accuracy than YOLO for small object detection. Such statements need to be checked regularly as the progress in NN development goes on.

Of course, a problem of training data shortage arises, and is sometimes solved in rather sophisticated ways, like 3Ddesign and 3D-printing [4]. Ukrainian researchers usually use more general approaches like data augmentation [5].

Besides using commercial drones with RGB cameras for landmine detection, there are some researches that use different approaches. Currently, new researches in the area of landmine detection are not limited to RGB images. Multispectral imagery is also used, engaging similar network architectures, YOLO, in particular, with some advances to adopt them for different number of layers in the imagery [6, 7]. In the case of scaling the technology to Ukrainian extents it should be proven that introduction of more expensive cameras is economically feasible.

Some researchers demonstrate completely different approach to hardware, constructing smart terrestrial robots that use YOLO as network architecture [8]. The approach is definitely worth attention, but obviously requires different datasets than UAV platforms for training networks.

Some datasets of landmine imagery are shared with public, either on researchers' website [16], or on Roboflow open-source deep learning platform [17], but they are either dedicated to limited set of mine types, or contain ordinary photos, not taken from drone.

Formulation of the problem

In landmine detection the problem of datasets availability is painful for novice researchers. For obvious reasons of sensitive information protection, landmines filmed from drones are seldom in open-source datasets. In most cases, researchers need some cooperation with military authorities to get necessary training data. I can summarize that nowadays, despite rapid development of neural networks and deep learning methods of landmine detection, there is still shortage of publications that describe researches made on data obtained from drones, and that few researches make accent on the necessary and sufficient amount of data.

The purpose of the work is to show that the problem of limited amount of training data can be effectively overcome by data augmentation and iterational process of training optical landmine detector is possible starting from very limited amount of training data. Also the purpose is to demonstrate that the problem can be solved using open-source tools and libraries for neural networks training, object detection and dataset preparation.

Materials and Methods

The accent in this paper is made on utilizing tools and methods familiar to broad audience and affordable with limited funds. UAVs used are ordinary commercial light quadcopters, freely available at the market at the moment of research. They are DJI Phantom and Bebop Parrot series equipped with the FC300C, Zenmuse XT2, Zenmuse Z30, FLIR One Pro, etc. cameras. All the photos were taken at test sites in Ukraine in process of broader research that included experiments with multispectral and infrared cameras [2], but didn't include experiments with CNN, so getting common RGB photos of landmines was not the main concern. However, the available resulting number of photos seemed to be enough to start creating datasets and training CNN.

Different landmines, both antipersonnel and antitank, were installed at test sites, mostly on the grass surface, because deep grass is probably the most challenging environment for detection in Ukraine. Several photos of each mine were taken from different altitude, to vary visibility from fully visible in detail to recognizable only by shape. Images were then processed without scaling or using any color, contrast or brightness adjustment. The only changes were augmentation by rotation and flipping, and converting to grayscale to train separate model.

The keystone of such research is network architecture. During the last few years researchers compared the efficiency of different CNN architectures for landmine detection task, and came to some conclusions. For example, scientists from Binghamton University in their famous research [3] concluded that Faster R-CNN is more accurate than YOLO for detecting "butterfly" antipersonnel mines. But taking into account the rapid progress in the area, such conclusions are relevant for rather a short period, and need to be revised on each iteration of deep learning ecosystem development.

So, in situation where serious benchmark research is needed for qualified estimation of relative effectiveness of different detectors, but results stay actual for a relatively short period of time, it seems reasonable to make an assumption about the most effective detector based on global tendencies in research area.

Many publications show that in 2023 the default choice of CNN for detection is YOLO [4-8], because of its relative simplicity and outstanding performance [9,10]. YOLO uses the approach when detection is considered regression problem, determining simultaneously the coordinates of object's bounding box and class of the object [9]. The balance between the accuracy of bounding box and appropriate class detection varies from task to task. Speaking about landmine detection, I can state that both precise bounding box and landmine model are incomparably less important than the very fact of landmine presence. In fact, it could be satisfactory to show the presence of mine with single point coordinates and no information about class, but 100% confidence. So, modern detectors like YOLO seem even overcomplicated in some sense. But highly developed ecosystem simplifies the usage of these sophisticated systems and nowadays it is unavoidable to experiment with them whenever we deal with object detection.

One of the benefits of YOLO is availability of open-source libraries that allow performing training, validation and detection effortlessly, without even knowing the details of programming. In our opinion, such a library of first choice is open-source library Ultralytics YOLO, having its major version YOLOv8 at the moment of research [11]. It is powerful Python-based library that has a lot of possibilities of fine tuning the detector. Among them are: predefined set of network models that differ in internal number of parameters, input arguments that allow prioritizing different parameters of detection, such as accuracy of bounding box or classification confidence, built-in augmentation capabilities, etc. A friendly community of developers providing excellent support is also worth mentioning.

A standard approach to deep learning research assumes, as de-facto standard, the presence of three independent datasets – for training, validation and testing. So, limited amount of data is a serious obstacle for such a research. Intuitively, the amount of data at our disposal was enough for a single dataset. Most of the landmines represented in our photoset were filmed at two sites, less of them at three sites, and some at single, so it was problematic to create three independent datasets.

So, several antipersonnel mines were chosen, all the photos were placed in two independent groups of images. The size of images was determined by default image size used for Ultralytics YOLOv8 – 640×640. This size fitted well enough for all antipersonnel mines, the largest of them occupying nearly 2/3 of the image, while the smallest being tens of pixels large.

There are several options for creating datasets. In my research I used Label Studio open-source application [14]. It is Python-based application that runs in browser, and allows the creation of datasets in a number of different formats, including YOLO format. The application has its extended enterprise version, but minimal functionality for dataset creation is present in free version.

a *b*

c *d*

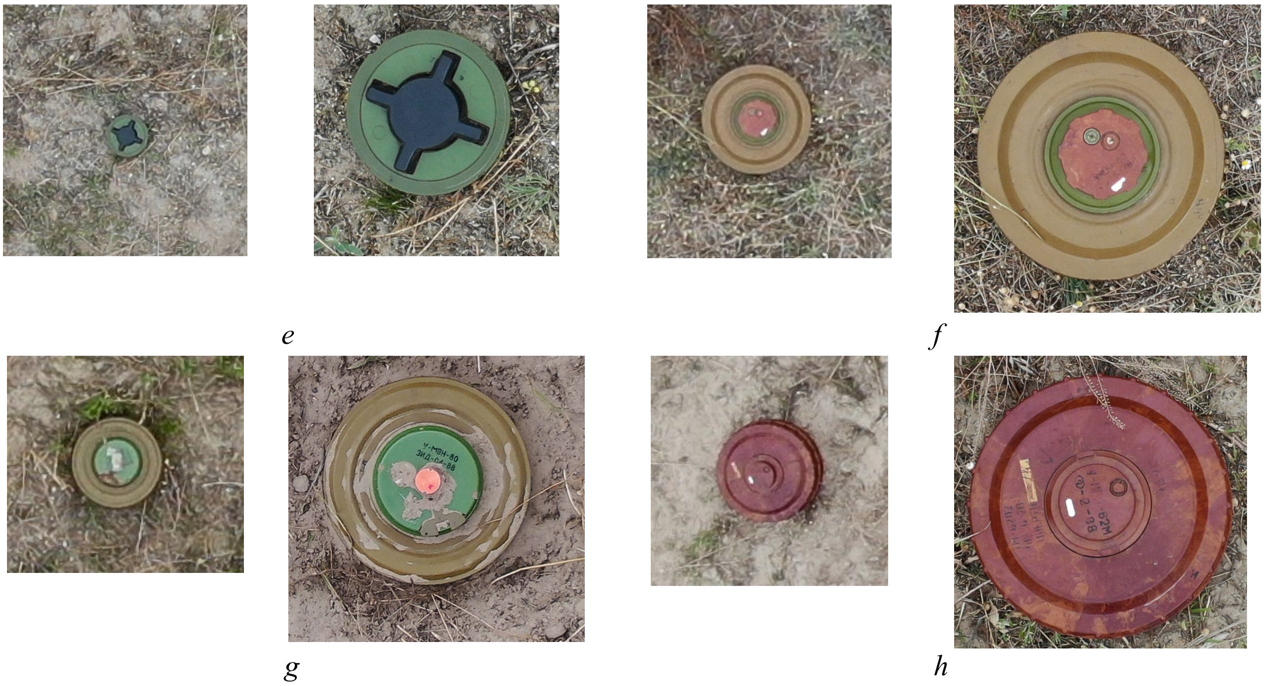


Fig. 1. Minimal and maximal size images used for training dataset:
a – MOH50 antipersonnel; *b* – MOH90 ant; ipersonnel; *c* – MOH100 antipersonnel; *d* – MOH200 antipersonnel;
e – PMN antipersonnel; *f* – TM62M antitank; *g* – TM72 antitank; *h* – TM-62П2 antitank

Augmentation of data was the important part of this research. Two options were used – built-in capabilities of Ultralytics library (default settings were used), and separate augmentation using external library. Albumentations, separate Python-based library [15] was used to increase the number of images in training dataset. This library provides a wide range of options to transform labeled images, both geometry-based and color-based. Its powerful feature is a possibility to combine several transformations easily and intuitively. Also useful is the ability to apply changes randomly to predefined percentage of dataset and to randomize some transformation parameters. In our case, I used it for random rotation, increasing dataset size x6 times.

All the training and validation was performed using 640×640 images, but in real life various image sizes are expected, and there are at least two ways to approach this. One is using complete image as an input to standard Ultralytics library detection method, setting image size as an argument, another is slicing image into 640×640 overlapping fragments before detection. That is what SAHI approach is intended for [12,13]. It automates the pipeline of image slicing, performing detection for each fragment, and joining fragments together, including prediction boxes. It is already adopted to use with Ultralytics library, so there is no need to slice and join images manually. The size of sliced fragments is not limited to the size used in the dataset, but I kept these sizes equal for simplicity.

Standard metrics were used for model validation, such as *precision* P, *recall* R and F1 score, which are calculated as (1), (2) and (3):

$$P = \frac{TP}{FP - TP}, \tag{1}$$

where TP is the number of correctly predicted classes, FP is the number of falsely predicted classes (or false alarms)

$$R = \frac{TP}{FN - TP}, \tag{2}$$

where FN is the number of negatively predicted classes that should be positive

$$F1 = \frac{2PR}{P + R}. \tag{3}$$

It is obvious that in the context of landmine detection *recall* is the most important parameter, because it shows the ability to detect all present targets, so if it is needed to quantify landmine detector with a single parameter, recall for all classes will be the first choice.

Results

The model chosen was so-called yolov8s (small) – second in the set of available models that consists of n-nano, s-small, m-medium, l-large and x-extra-large. The training dataset was gradually increased by data augmentation from 216 images, through 648 to 1296. Validation dataset contained 90 images, and remained the same during iterations. Finally, both datasets were converted to grayscale and separate model was trained, that has shown small changes for most mines, but much worth results for MOH50. The number of iterations was taken 20. I can state that validation was performed correctly, because validation dataset is independent. No special testing dataset was provided because of data shortage. Illustrations provided are generated using images that were included neither in training, nor in validation dataset. Results of model validation are represented by standard recall-, precision-, F1-confidence curves generated by library immediately after training process. The first group of illustrations are all 640×640. Also, a comparison of usual vs SAHI approach is shown for image size 1140×1080.

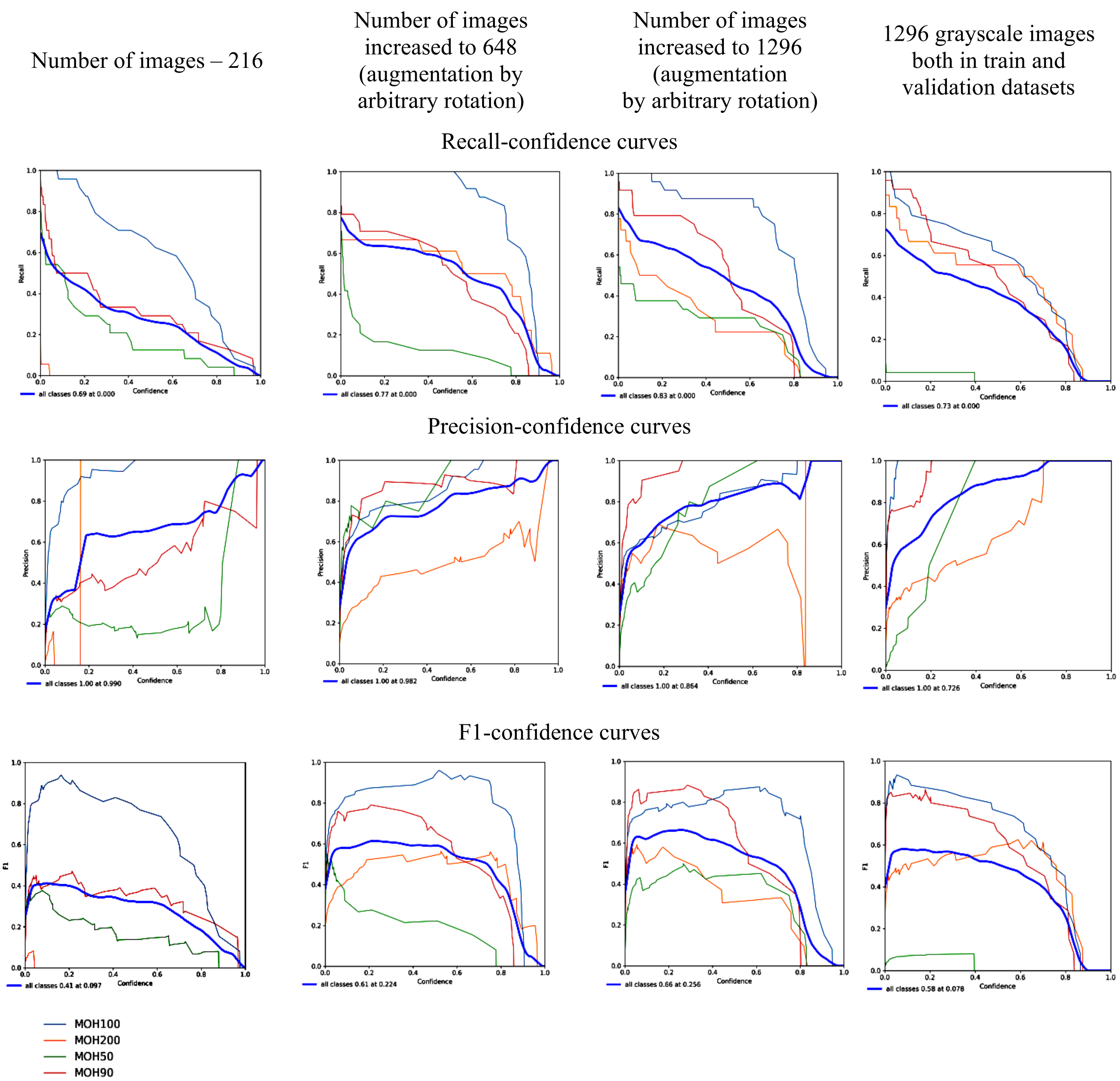


Fig. 2. Comparing results for datasets: training set contains 216, 648, 1296 images (augmentation by arbitrary rotation), and 1296 grayscale images. Validation set contains 90 images

The limited set of independent test images includes mines of arbitrary shape and size, all of them antipersonnel, placed in the grass, different from training images in color gamma and with smaller size.

a *b* *c*

Fig. 3. Testing on images different form training set in terms of color gamma and size:
a – both MOH100 and MOH50 are correctly detected; *b, c* – MOH90 is falsely detected as MOH100 and MOH200

The next figure shows changes of detection accuracy during rotation of the image.

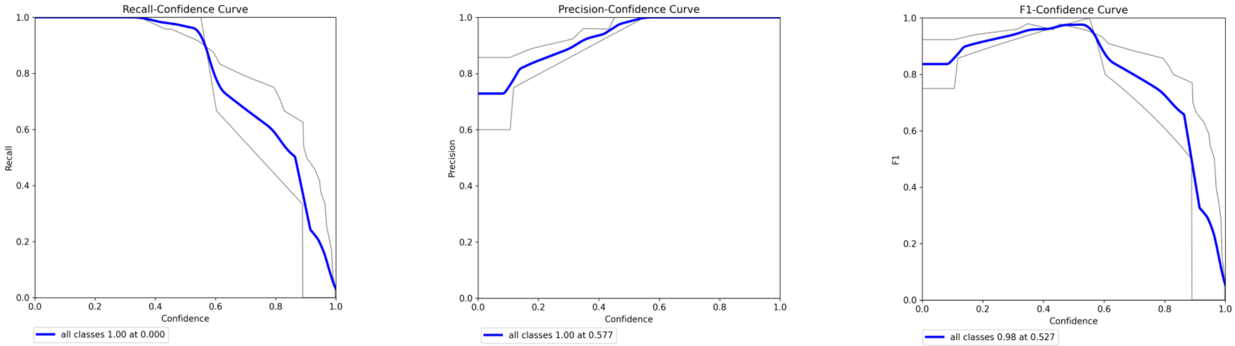
Fig. 4. Detection of MOH90 and MOH200 on the image similar to training dataset rotating the image.
The result shows that orientation matters for landmines of arbitrary shape

Landmines of round shape, both antipersonnel and antitank, are placed into separate training dataset including 72 images. Only two classes – one for PMN antipersonnel, another for a group of similar TM antitank mines, are used. Validation dataset contains 26 images of same or similar mines filmed on different background. Because of low form variability, it isn't reasonable to perform augmentation by rotation, that was done for mines of arbitrary shape.

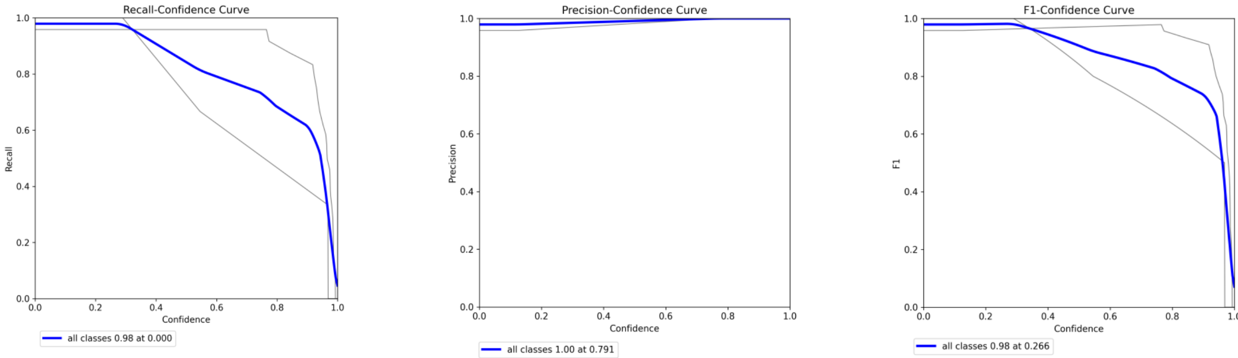
PMN antipersonnel TM62M antitank TM72 antitank

Fig. 5. Detection of round-shaped landmines

For such a limited datset it seems reasonable to use ultralight net modification – YOLOv8n (nano). There is some difference from YOLOv8s (small) model, but this difference can be considered non-critical.



YOLOv8n (nano) model result



YOLOv8s (small) model result

Fig. 6. Validation results of round-shaped landmines dataset

Also, SAHI approach was tested, showing good results for images of arbitrary size. It is recommended to use it in the case of significant difference of dataset images size and size of the images used in detection.

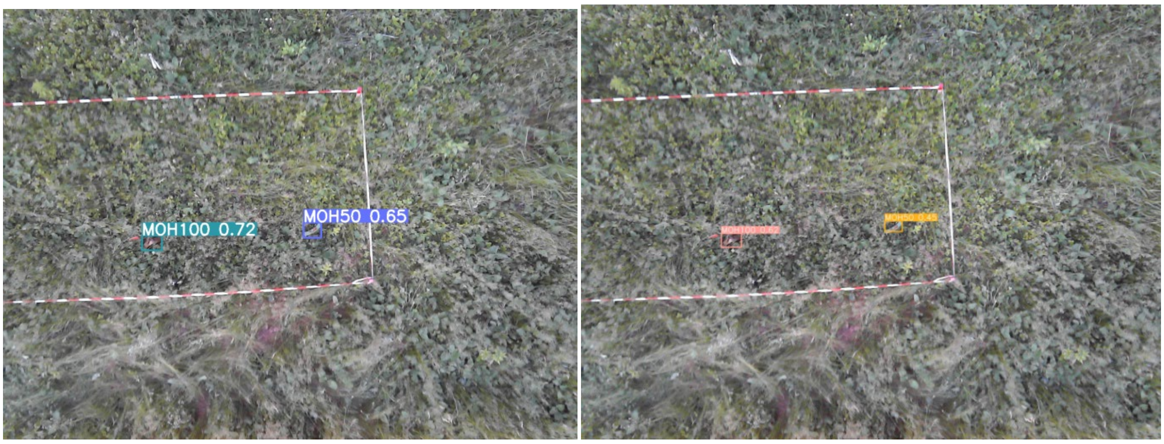


Fig. 7. Illustration of SAHI approach (on the left) on 1140x1080 images. SAHI shows some increase of confidence

Discussion

The problem of limited amount of data results in difficulties in presentation of research results. It is de-facto standard to provide independent datasets for validation and testing, and model performance indexes should be obtained from it. However, if each type of mine is represented in the dataset with one instance filmed at different angle, distance and background, can we speak about "independent" testing dataset? All we can do is to play with augmentation. To make the case more similar to other object detection problems, we can group similar landmine types to classes, for example, all antitank round-shaped landmines are candidates for such grouping, or pairs of antipersonnel landmines, such as MOH50, MOH90 and MOH100, MOH200 respectively. It can bring necessary generalization, that network lacks processing these types separately. After that, if our assumption is correct, detector should be able to detect similar mine of any class, not represented in dataset directly.

And still the question is – to what extent are these images independent, in comparison with such diverse detection objects as persons, dogs or bicycles? So, the question of correct validation and testing of such models is still open. During investigation, we have to sacrifice some amount of data for validation purposes, but finally, when optimal hyperparameters are found, we have to use all the available data for the final dataset in production. After all, our final objective is not getting accurate numbers and plots, but building reliable detector that will simplify deminers' work.

Conclusion

In this study I showed that it is possible to start the development of AI-based landmine detector even based on a very limited dataset. The current paper shows only the first iteration, so results are still far from requirements for accuracy needed in humanitarian demining. In the next iterations, when more data is available, I will be able to quantify the threshold of necessary amount of data needed for reliable detector. For this purpose, activities on test sites in Ukraine are needed. Also, different strategies of dataset assembling need to be tested, such as grouping landmines of different models having similar shape and size. Also, for different environments we may require separate datasets. An important milestone should be testing detector on real minefield imagery. So, current paper presents important step in complex iterational process of optical landmine detector development.

REFERENCES

[1] Osco, L. P., Junior, J. M., Ramos, A.P.M., Jorge, L.A.C., Fatholahi, S.N., Silva, J.A., … Li, J. (2021) A review on deep learning in UAV remote sensing. *International Journal of Applied Earth Observations and Geoinformation*, 102 (2021) 102456. https://doi.org/10.1016/j.jag.2021. 102456.

[2] Popov, M. O., Stankevich, S. A., Mosov, S. P., Titarenko, O. V., Dugin, S. S., Golubov, S. I., Andreiev, A. A. (2022) Method for minefields mapping by imagery from unmanned aerial vehicle, *Advances in Military Technology*, 17(2), 211–229. https://doi.org/ 10.3849/aimt.01722.

[3] Baur, J., Steinberg, G., Nikulin, A., Chiu, K., Smet, T. (2020). Applying Deep Learning to Automate UAV-Based Detection of Scatterable Landmines. *Remote Sensing*. 12(859) https://doi.org/ 10.3390/ rs12050859.

[4] Harvey, A., LeBrun, E. (2023) Computer vision detection of explosive ordnance: a high-performance 9N235/9N210 cluster submunition detector. *The Journal of Conventional Weapons Destruction*. 27(2).

[5] Kunichik, O., Tereshchenko V. (2023) Improving the accuracy of landmine detection using data augmentation: a comprehensive study. *Artificial Intelligence*, 2023(2). https://doi.org/10.15407/ jai2023.02.042.

[6] Cho S., Ma J., Yakimenko O. A. (2023) Aerial multi-spectral AI-based detection system for unexploded ordnance, *Defence Technology*, 27, 24–37, ISSN 2214–9147. https://doi.org/10.1016/ j.dt.2022.12.002.

[7] Qiu, Z.; Guo, H.; Hu, J.; Jiang, H.; Luo, C. (2023) Joint Fusion and Detection via Deep Learning in UAV-Borne Multispectral Sensing of Scatterable Landmine. *Sensors*, 23(5693). https://doi.org/ 10.3390/s23125693.

[8] Vivoli, E., Bertini, M., Capineri, L. (2024) Deep Learning-Based Real-Time Detection of Surface Landmines Using Optical Imaging. *Remote Sensing*. 16(4):677. https://doi.org/10.3390/ rs16040677.

[9] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016) YOLO: You Only Look Once: Unified, Real-Time Object Detection. arXiv 2016, arXiv:1506.02640. https://doi.org/10.48550/ arXiv.1506.02640.

[10] Terven, J. R., Cordova-Esparza, D. M. A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS. arXiv 2023, arXiv:2304.00501. https://doi.org/ 10.48550/arXiv.2304.00501.

[11] Ultralytics (2023) Ultralytics YOLOv8-State-of-the-art Vision AI. Retrieved from https://www. ultralytics.com/yolo (access data 20.05.2024).

[12] Durai, P. (2023) Exploring SAHI: Slicing Aided Hyper Inference for Small Object Detection, Learn OpenCV, Retrieved from https://learnopencv.com/ slicing-aided-hyper-inference/ (access data 20.05.2024).

[13] Akyon, F. C., Altinuc, S. O., Temizel, A. (2022) Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection. IEEE International Conference on Image Processing Proceedings, 2022, https://doi.org/10.1109/ICIP46576.2022. 9897990 (access data 20.05.2024).

[14] Label Studio. Open Source Data Labeling Platform. https://labelstud.io/ (access data 21.05.2024).

[15] Albumentations: fast and flexible image augment tations. https://albumentations.ai/ (access data 21.05.2024).

[16] Earth Sciences Faculty Scholarship |Earth Siences| Bindhamton University https://orb.binghamton.edu/ geology fac/ (access data 21.05.2024).

[17] Roboflow Universe: Open Source Computer Vision Community https://universe.roboflow.com/ (access data 21.05.2024)

Саприкін Є.

ОПТИЧНЕ ВИЯВЛЕННЯ НАЗЕМНИХ МІН З ВИКОРИСТАННЯМ ГЛИБОКОГО НАВЧАННЯ НА ОСНОВІ ОБМЕЖЕНОГО НАБОРУ ДАНИХ АЕРОЗНІМАННЯ

Виявлення наземних мін є одним із найбільш інноваційних застосувань безпілотних літальних апаратів, що зобов'язане своєю появою швидкому розвитку як літальних апаратів, оснащених різними оптичними камерами та датчиками, що засновані на різних фізичних принципах, так і методів класифікації та виявлення об'єктів, включно з машинним навчанням та зокрема глибоким навчанням. Оптичне виявлення є важливою частиною цілісного процесу виявлення мін,і може виконуватися окремо або в поєднанні з обробкою даних з інших типів камер або датчиків. Розвиток глибоких згорткових нейронних мереж кардинально змінив стан речей у галузі оптичного виявлення, зробивши їх де-факто вибором номер один для більшості завдань класифікації, виявлення та сегментації об'єктів. Проте стримуючим фактором у випадку виявлення наземних мін є обмежена доступність відповідних даних для навчання, яку різні дослідники намагаються подолати різними способами. Визначення необхідного обсягу навчальних даних все ще залишається експериментальною задачею. Незважаючи на кілька років розвитку цієї галузі, все ще існує дефіцит досліджень, заснованих на реальних зображеннях наземних мін, отриманих з безпілотних літальних апаратів, тому наразі будь-які зусилля в цьому напрямку є цінними та служать джерелом натхнення для нових дослідників. У даній статті описане таке дослідження, а саме його перша ітерація, коли популярні інструменти з відкритим вихідним кодом використовуються для побудови системи виявлення, а їх ефективність оцінюється на базі обмеженої кількості даних. Показано, що проблема обмеженої кількості тренувальних даних може бути ефективно подолана нарощуванням даних, і продемонстровано ітераційний процес навчання оптичного детектора мін. Також продемонстровано ефективність використання вільно розповсюджуваних інструментів та бібліотек для тренування нейронних мереж, виявлення об'єктів та підготовки наборів даних.

Ключові слова: виявлення наземних мін, безпілотні літальні апарати, глибоке навчання, YOLO.

Saprykin Ievgen

OPTICAL DEEP LEARNING LANDMINE DETECTION BASED ON LIMITED DATASET OF AERIAL IMAGERY

Landmine detection is one of the most innovative applications of unmanned aerial vehicles that became possible due to rapid development of both aerial vehicles equipped by different optical cameras and sensors using different physical principles, and object classification and detection methods, including machine learning and especially deep learning. Optical detection is an essential part of the overall landmine detection process that can be performed either separately or in combination with data processing from other types of cameras or sensors. The development of deep convolutional neural networks has dramatically changed the landscape of optical detection by making them de-facto choice number one for the majority of object classification, detection and segmentation tasks. However, the deterrent factor in the case of landmine detection is limited availability of appropriate data for training that different researchers try to overcome in different ways. The assessment of necessary amount of training data for any particular object detection problem still remains an experimental task. Despite several years of development in this area, still there is a shortage of research based on real landmine imagery obtained from unmanned aerial vehicles, so currently any public effort in this direction is valuable and works as an inspiration for new researchers. This paper describes such a study, namely its first iteration in which popular open-source tools are used to build detection pipeline and estimation of their efficiency is done using limited amount of data. It is shown that the problem of limited amount of training data can be effectively overcome by data augmentation and iterational process of training optical landmine detector is demonstrated. The effectiveness of open-source tools and libraries for neural networks training, object detection and dataset preparation is also demonstrated.

Keywords: landmine detection, unmanned aerial vehicles, deep learning, YOLO.

Стаття надійшла до редакції 30.04.2024 р.
Прийнято до друку 12.06.2024 р.