Unexploded Ordnance Detection on UAV Thermal Images by Using YOLOv7

MILAN BAJIĆ,[1] BOŽIDAR POTOČNIK[2]

[1] Zagreb University of Applied Sciences, Department of IT and Computer Sciences, Zagreb, Croatia
mbajic@tvz.hr

[2] University of Maribor, Faculty of Electrical Engineering and Computer Science, Institute of Computer Science, Maribor, Slovenia
bozidar.potocnik@um.si

Abstract A few promising solutions for thermal imaging Unexploded Ordnance (UXO) detection were proposed after the start of the military conflict in Ukraine in 2014. Our research focuses on improving the accuracy of UXO detection in thermal images. The current state-of-the-art UXO detection method is based on the YOLOv5 Convolutional Neural Network (CNN). We accessed the effectiveness of UXO detection by using the state-of-the-art object detector YOLOv7 in this article. Two YOLOv7 models were re-implemented, fine-tuned using a grid-search approach and trained on a UXOTi_NPA public dataset of 720x480 pixel thermal images. The results showed that the models were able to identify UXOs from 11 different classes with more than 90% probability and a Mean Average Precision (mAP) of 86.8% to 89.7%, depending on the model's complexity. The metrics are just slightly behind the YOLOv5 results. Such CNN, thus, enables accurate automatic UXO detection, which is crucial to address one of the least explored and life-threatening problems worldwide.

Keywords: unmanned aerial vehicle; unexploded ordnance; thermal imaging; UXOTi_NPA dataset; convolutional neural networks; deep learning; YOLO

University of Maribor Press

DOI https://doi.org/10.18690/um.feri.4.2023.12
ISBN 978-961-286-720-1

DETEKCIJA NEEKSPLODIRANIH UBOJNIH SREDSTEV NA TERMALNIH SLIKAH UAV S POMOČJO YOLOv7

MILAN BAJIĆ,[1] BOŽIDAR POTOČNIK[2]

[1] Zagrebška univerza uporabnih znanosti, Oddelek za informatiko in računalništvo, zagreb, Hrvaška
mbajic@tvz.hr

[2] Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Inštitut za računalništvo, Maribor, Slovenija
bozidar.potocnik@um.si

Sinopsis Po začetku vojaškega spopada v Ukrajini leta 2014 je bilo predlaganih nekaj obetavnih rešitev za odkrivanje neeksplodiranih ubojnih sredstev (UXO) v termalnih slikah. Naša raziskava se osredotoča na izboljšanje natančnosti detektiranja UXO v termalnih slikah. Trenutno najsodobnejša metoda detektiranja UXO temelji na konvolucijski nevronski mreži (CNN) YOLOv5. V tem članku smo ovrednotili učinkovitost detektiranja UXO s pomočjo najsodobnejšega detektorja objektov YOLOv7. Reimplementirali smo dva modela YOLOv7, ju fino uglasili s pristopom iskanja po mreži in ju naučili na javni zbirki UXOTi_NPA, ki vsebuje termalne slike velikost 720x480 pikslov. Rezultati so pokazali, da sta modela zmožna identificirati UXO iz 11 različnih razredov z več kot 90-odstotno verjetnostjo in s srednjo povprečno točnostjo (mAP) od 86,8 % do 89,7 %, odvisno od kompleksnosti modela. Rezultati le malo zaostajajo za rezultati YOLOv5. Takšna CNN torej omogoča natančno avtomatsko detekcijo UXO, kar je ključnega pomena za reševanje enega najmanj raziskanih in življenjsko nevarnih problemov na svetu.

Ključne besede:
EU projekt AIM@VET, računalniški vid, umetna inteligenca, poklicno usposabljanje, izobraževanje

https://doi.org/10.18690/um.feri.4.2023.12 DOI
978-961-286-720-1 ISBN

Univerzitetna založba
Univerze v Mariboru

1 Introduction

Thermal imaging is a non-contact method that converts the radiation pattern of an object into a visible image, known as a thermal image or thermogram. All objects above absolute zero temperature emit infrared radiation and this energy is detected by air and space-borne Thermal Infrared (TIR) remote sensors (Kuenzer & Dech, 2013). These systems do not require an external source of infrared radiation, making it suitable for field use.

Explosive remnants of war have been part of daily life since World War I, with 64 countries (Roberts & Williams, 1995) contaminated by landmines. There are countries with databases of explosive devices (GICHD, 2022), while others are still developing such information. Mine clearance is a slow process, taking approximately 100 times more time to clear each mine placed. Most of these remnants are well known by their physical dimensions, weight, type of fuse, and material of the cover. They mostly lie on the surface and are unaffected by weather or vegetation.

In a project by Norwegian People's Aid (NPA) in Bosnia and Herzegovina in 2019, thermal imaging (Bajic, 2020) was tested as a promising technology for surveying hazardous areas. In this research, thermal imaging and LiDAR sensors were mounted on Unmanned Aerial Vehicles (UAVs) and experimented with for this purpose.

Object detection by using thermal images is a broad research field (Leira et al., 2021), (Dai et al., 2021), (Banuls et al., 2020), however the detection of Unexploded Ordnances (UXOs) using thermal imaging is a relatively new research field. We found just few studies on this topic, like (Nikulin et al., 2018) for the PFM-1 'Butterfly Mines' detection and study (Krause et al., 2018) focused on the detection of landmines based on thermal changes in the environment. An experiment in (Yao et al., 2019) was conducted using time series thermal images to detect buried landmines. The use of Deep Learning for UXO detection is still relatively rare in the field. A research group from Binghamton University expanded their previous work (Nikulin et al., 2018) by using Deep Learning for automated detection and mapping of PFM-1 mines (Baur et al., 2020).

Recently, in our previous work (Bajić & Potočnik, 2023) we proposed the first automated Convolutional Neural Network (CNN) based solution for UXO detection using thermal imaging data that can identify multiple classes of objects. The YOLOv5 architecture was adapted successfully for the UXO detection problem in our previous study. Besides, the UXOTi_NPA public database (Bajić & Potočnik, 2023) of 808 annotated thermal images of Unexploded Ordnance was published. YOLOv5 model detection results on the UXOTi_NPA dataset are considered as the baseline results for this dataset.

In this study, we will access the effectiveness of the newest YOLO architecture, i.e. YOLOv7, on the UXOTi_NPA public dataset. The YOLOv7 architecture will be adapted and retrained to detect UXOs from 11 different classes in highly variable thermal images. This architecture will be modified to fit the UXO detection problem, fine-tuned with a grid-search approach, and finally trained end-to-end on thermal images. The effectiveness of the adapted and retrained CNN architecture will be confirmed through experiments on the UXOTi_NPA dataset.

2 Evaluation Dataset UXOTi_NPA

The construction of the evaluation dataset UXOTi_NPA is decribed in (Bajić & Potočnik, 2023). This dataset consists of 808 thermal images of the dimension of 720 × 480 pixels with annotations. The number of objects (UXOs) per image varied between one to three. Figure 1 depicts a sample thermal image from our dataset.

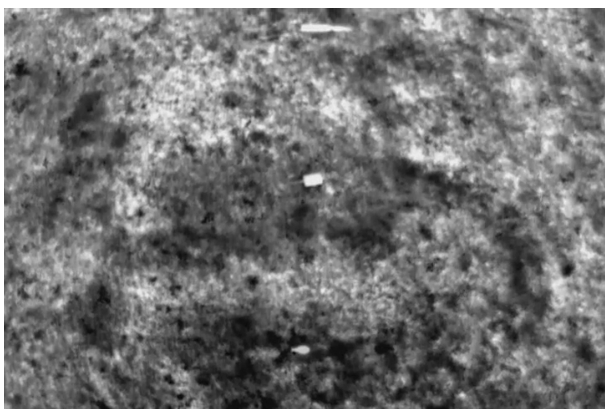
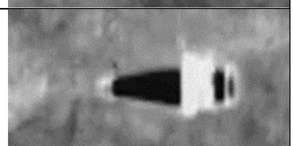
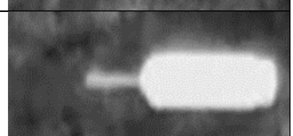
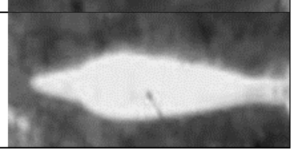
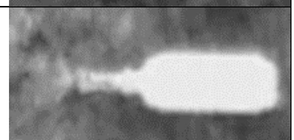
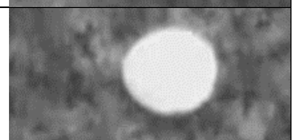


Figure 1: Sample thermal image from the UXOTi_NPA dataset.
Source: own.

Target objects can belong to one of the eleven classes (also to different classes in one image). Table 1 depicts sample UXO objects for each of the eleven classes, with some additional information.

Table 1: UXO objects from each of the eleven classes. Such objects appear in images from the evaluation dataset.

Class	Name and Type	Number of Occurrences in Dataset	RGB Image	Thermal Image
0	Mortar mine M:60 mm	139		
1	Fuse M:125 mm	161		
2/3	Hand grenade	121		
3/4	AP land mine PMR čapljinka	104		
4/5	Bullet 30.2 mm	116		
5/6	Land mine PMR 2A	136		
6/7	Mortar mine M:82 mm	149		

Class	Name and Type	Number of Occurrences in Dataset	RGB Image	Thermal Image
7/8	Land mine PMR 3	150		
8/9	Land mine PROM 1	139		
9/10	Land mine PMA 3	79		
10/2	Tromblon mine TTM RP	161		

The data splitting in the UXOTi_NPA is prescribed. The training set contains 640 images, the validation set 80 and the testing set 88 images.

3 YOLOv7 Architecture

The You Only Look Once (YOLO) Deep Neural Network is a highly advanced object detector with an exceptional compromise between an accuracy and a speed of object detection. This is achieved by dividing the image into grid cells and considering each cell as a proposal to detect the object. Over time, many YOLO variants have been created with different features and capabilities, such as FastYOLO, YOLO-tiny, YOLO-lite, and YOLOv5n, which has been shown to be as effective as the more complex R-CNN or Faster R-CNN object detectors.

The most recent YOLO version is YOLOv7, i.e., a successor of the YOLOv5 version. YOLOv7 has typically higher accuracy than YOLOv5 but is slower when training on custom data due to its use of more floating-point operations. Additionally, it is slower on commonly available GPU systems, although it is faster on high-end ones.

We compared the effectiveness of YOLOv5 and YOLOv7 architectures on the UXOTi_NPA dataset in this study. Let us emphasize that YOLOv5 results mean the baseline results for our evaluation dataset. We experimented with five models of YOLOv5 in our previous study (Bajić & Potočnik, 2023). In this research, we compare these baseline results with the results of two YOLOv7 models, namely regular and tiny model as outlined in Table 2 (all presented metrics in Table are for the COCO dataset).

Training of the models was carried out on the Google Colaboratory platform, utilizing Tesla T4 GPUs. The programming code for the models was saved in Ultralytics' notebooks available on the Github page (Wang et al., 2022). The Pytorch framework, Python 3.7.15 and some supplementary libraries were employed to implement the models.

Table 2: Comparison of five mainstream YOLOv5 and two YOLOv7 detection models, trained and evaluated on the COCO dataset. The results were taken from (*YOLOv5 Models; Https://Github.Com/Ultralytics/Yolov5*) **and** (*Github YOLOv7*)**.**

Model	mAP@0.5	mAP@0.5:0.95	Parameters (in Million)
YOLOv5n	45.7%	28.0%	1.9
YOLOv5s	56.8%	37.4%	7.2
YOLOv5m	64.1%	45.4%	21.2
YOLOv5l	67.3%	49.0%	46.5
YOLOv5x	68.9%	50.7%	86.7
YOLOv7 tiny	56.7%	N/A	6.2
YOLOv7	69.7%	N/A	36.9

4 Results

This section presents the findings of our experiments. Initially, we solved the problem of classification and detection of UXOs from 11 different classes (i.e., an eleven UXO class detection problem). The five different models of the YOLOv5 version and two models of YOLOv7 version were trained separately on the UXOTi_NPA database's training set, for 300 epochs, utilizing identical hyperparameters. Subsequently, the trained CNNs were evaluated on the UXOTi_NPA testing set, with the classification and detection effectiveness measured using classic metrics, including Precision, Recall, Mean Average Precision (mAP) at a 0.5 threshold, and the average mAP within the 0.5 to 0.95 threshold range at an interval of 0.05. Table 4 presents results of our initial experiment, and, at the same time, it also includes the number of free parameters (in million) for each model.

Table 3: Effectiveness of five models of the YOLOv5 and two models of YOLOv7, retrained on the UXOTi_NPA dataset: an eleven UXO class detection problem. The column 'Parameters' presents the number of CNN-free parameters.

Model	mAP@0.5	mAP@0.5:0.95	Parameters (in Million)
YOLOv5n	99.5%	87.0%	1.9
YOLOv5s	99.5%	88.5%	7.2
YOLOv5m	99.5%	89.9%	21.2
YOLOv5l	99.5%	90.5%	46.5
YOLOv5x	99.5%	89.7%	86.7
YOLOv7 tiny	**99.5%**	**86.8%**	**6.2**
YOLOv7	**99.5%**	**89.7%**	**36.9**

In the subsequent experiment, we combined UXOs from all 11 classes into a single class (i.e., 'a single UXO class detection problem'). The objective was to detect if any UXOs were present in a thermal image, irrespective of their type, and to identify their location. This scenario simulated a real-world situation where the primary interest of UXO removers is to ascertain whether UXOs are present in a given area before determining their type. The same YOLOv5 and YOLOv7 models as used in the first experiment were employed also in this experiment, only the output layer was modified accordingly. The training process and other parameters remained identical to those used in the previous experiment. The models were evaluated on the testing dataset of the UXOTi_NPA database, where the UXOs from all 11 classes were combined into a single testing class. The results are shown in Table 4.

Table 4: Effectiveness of five models of the YOLOv5 and two models of YOLOv7, retrained on the UXOTi_NPA dataset: a single UXO class detection problem. The column 'Parameters' presents the number of CNN-free parameters.

Model	mAP@0.5	mAP@0.5:0.95	Parameters (in Million)
YOLOv5n	99.5%	87.9%	1.8
YOLOv5s	99.5%	90.5%	7.1
YOLOv5m	99.5%	90.6%	21.1
YOLOv5l	99.5%	91.1%	46.4
YOLOv5x	99.5%	91.5%	86.6
YOLOv7 tiny	**99.5%**	**86.8%**	**6.1**
YOLOv7	**99.5%**	**90.3%**	**36.8**

Figure 2 and Figure 3 present some qualitative results. Figure 2 depicts results for the eleven UXO class detection problem, while results for a single UXO class detection problem are shown in Figure 3. The left column of both figures displays the annotated original thermal images from the UXOTi_NPA testing set. Annotations (and results) are presented in a form of bounding boxes, whereat the (detected) class is specified by an index and colour (see also Table 1 for indexes). In the case of binary detection (Figure 3), all UXOs are grouped into a single class (index 1).

On the other hand, the middle and the right columns of both figures depict the UXO detection results by using our retrained YOLOv5 and YOLOv7 models.

Annotated original	Detection by YOLOv5	Detection by YOLOv7

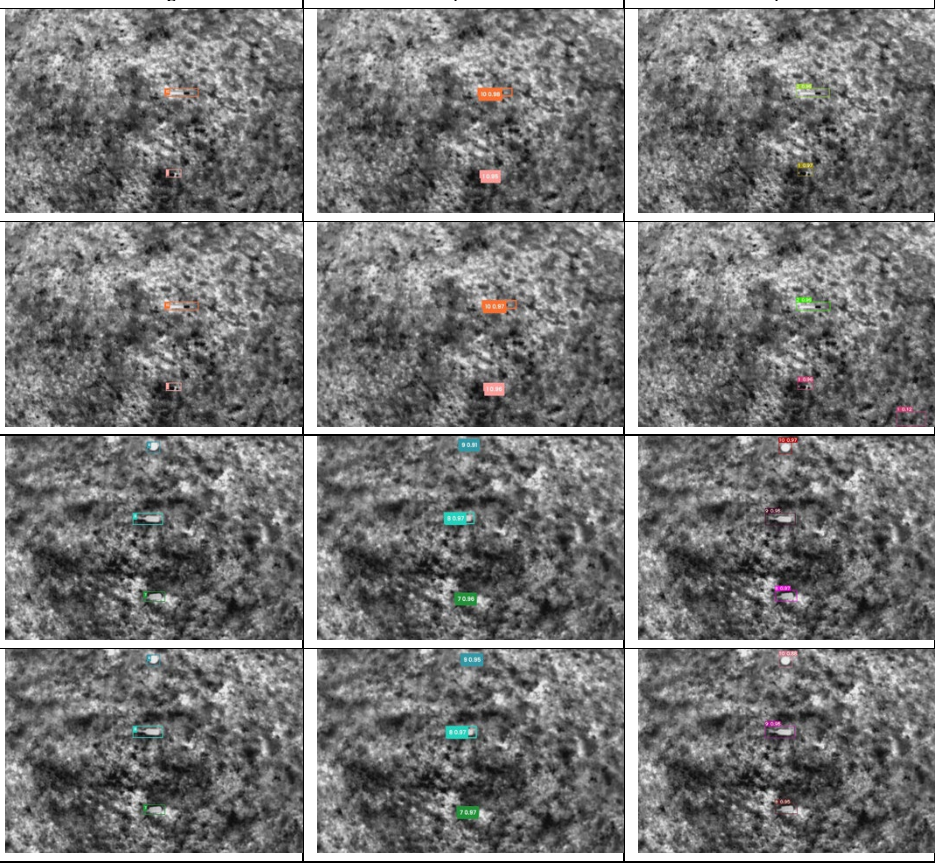


Figure 2: Qualitative results for the eleven UXO class detection problem: Expert annotations are overlaid on the thermal images (left column) and bounding boxes with UXOs, detected by using the retrained YOLOv5n and YOLOv7 tiny (rows 1 and 3, middle and right column), or YOLOv5x and YOLOv7 regular (rows 2 and 4, middle and right column). The class index and detection probability are written next to the bounding box.

Surce: own.

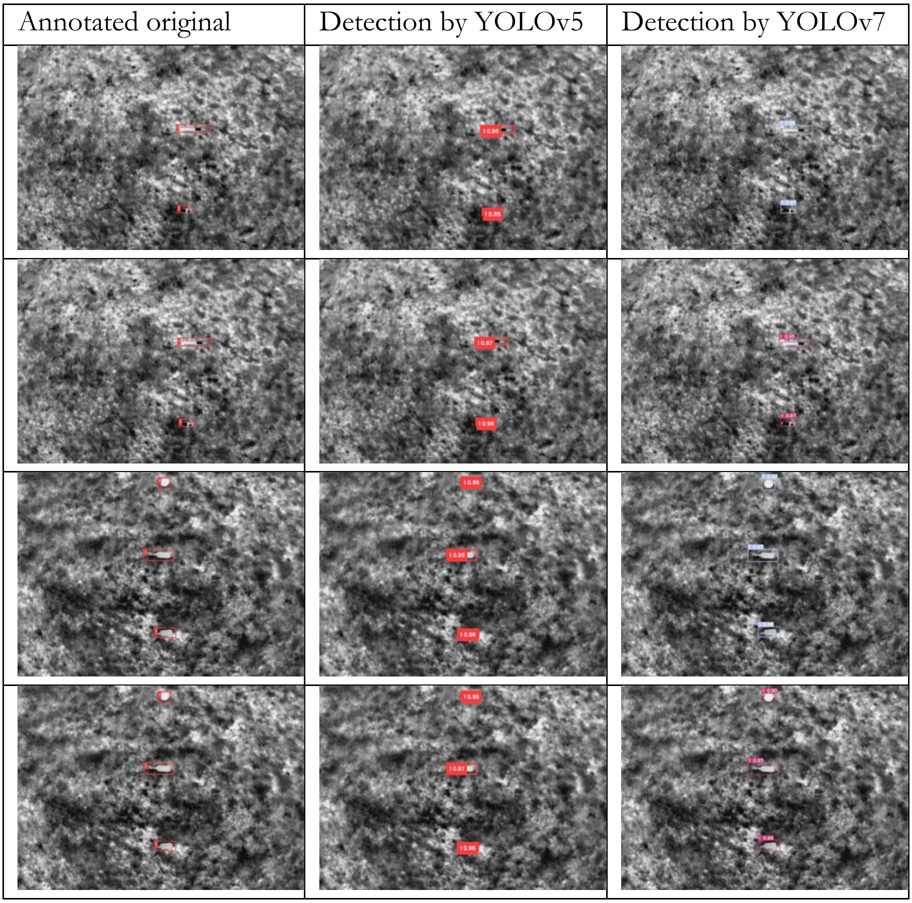


Figure 3: Qualitative results for the single UXO class detection problem: Expert annotations are overlaid on the thermal images (left column) and bounding boxes with UXOs, detected by using the retrained YOLOv5n and YOLOv7 tiny (rows 1 and 3, middle and right column) or YOLOv5x and YOLOv7 regular (rows 2 and 4, middle and right column). A single class index 1 and detection probability are written next to the bounding box.

Surce: own.

5 Discussion and Conclusion

In this study, we have experimented with a state-of-the-art computational approach utilizing Deep Learning and Convolutional Neural Networks aimed to detect Unexploded Ordnances (UXOs) from thermal imaging materials. The detection of UXOs is an essential support activity by the neutralization of explosive war remnants. With our automated solution, a terrain can be examined from the air (e.g., by using Unmanned Aerial Vehicles) and potentially dangerous areas can be identified in advance.

Our study's innovation is the utilization of thermal images for UXO detection. Thermal imaging provides valuable information about the environment and its changes, particularly if this information is not detectable in the visible spectrum. Different materials, including UXOs, leave their unique thermal signatures in the image, making it possible to identify them with high accuracy. An example is a land mine with a green metal casing placed on green grass. Such a mine is impossible to detect in the visible spectrum, but it is easily separable in the thermal spectrum (see Figure 4).

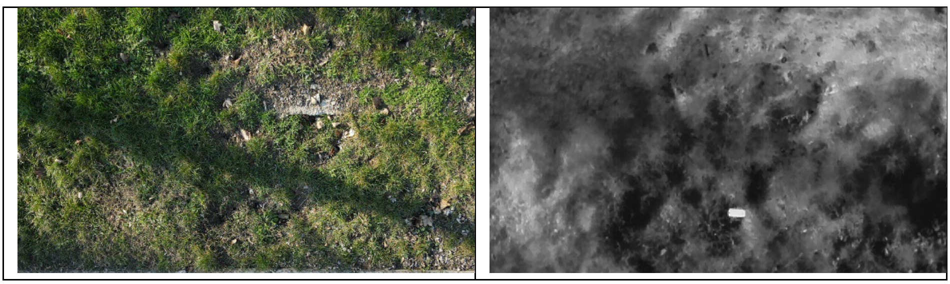


Figure 4. A land mine with a green metal casing placed on green grass: An image in the visible spectrum (left) and the same terrain in the thermal spectrum (right). Note: This example is from our private database and is not part of the UXOTi_NPA dataset.
Source: own.

We also compared the performance of YOLOv7 with that of YOLOv5 on the UXOTi_NPA public dataset in this article. YOLOv5 is an established architecture thoroughly tested on various datasets, while YOLOv7 was introduced recently (in 2022) and is still in the development phase. Based on the obtained standardised metrics mAP on the UXOTi_NPA dataset, we conclude that the YOLOv5 seems to be a better trade-off between model size and detection effectiveness than YOLOv7. The reason lies undoubtedly in the higher number of false positive detection by YOLOv7 (see Figure 2). We also see a possible cause in a small UXOTi_NPA dataset. Considering a training time on the UXOTi_NPA dataset, YOLOv7 requires more processing time and resources compared to YOLOv5, whereat such an increase was not justified by the results. A similar discussion can be found in (Durve et al., 2023).

In this work, we have shown that the combination of deep Convolutional Neural Networks and thermal imaging can be used advantageously to detect UXOs in a real environment. Our future work will go in several directions, primarily in an expansion of the UXOTi_NPA dataset with new objects and environments, and in an upgrade of YOLOv7 with a module for the more efficient feature analysis and fusion.

References

Bajic, M. (2020). Testing of remotely piloted aircraft systems with a thermal infrared camera to detect explosive devices at contaminated areas and validation of developed standard operational procedures. Norwegian Peoples Aid Olso Norway.

Bajić, M., & Potočnik, B. (2023). UAV Thermal Imaging for Unexploded Ordnance Detection by Using Deep Learning. Remote Sensing, 15(4), 967. https://doi.org/10.3390/rs15040967

Banuls, A., Mandow, A., Vazquez-Martin, R., Morales, J., & Garcia-Cerezo, A. (2020). Object Detection from Thermal Infrared and Visible Light Cameras in Search and Rescue Scenes. 2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), 380–386. https://doi.org/10.1109/SSRR50563.2020.9292593

Baur, J., Steinberg, G., Nikulin, A., Chiu, K., & de Smet, T. S. (2020). Applying Deep Learning to Automate UAV-Based Detection of Scatterable Landmines. Remote Sensing, 12(5), 859. https://doi.org/10.3390/rs12050859

Dai, X., Yuan, X., & Wei, X. (2021). TIRNet: Object detection in thermal infrared images for autonomous driving. Applied Intelligence, 51(3), 1244–1261. https://doi.org/10.1007/s10489-020-01882-2

Durve, M., Orsini, S., Tiribocchi, A., Montessori, A., Tucny, J.-M., Lauricella, M., Camposeo, A., Pisignano, D., & Succi, S. (2023). Benchmarking YOLOv5 and YOLOv7 models with DeepSORT for droplet tracking applications. https://doi.org/10.48550/ARXIV.2301.08189

GICHD. (2022). Explosive Ordnance Guide for Ukraine—Second Edition; https://www.gichd.org/fileadmin/GICHD-resources/rec-documents/GICHD_Ukraine_Guide_2022_Second_Edition_web.pdf. GICHD. https://www.gichd.org/fileadmin/GICHD-resources/rec-documents/GICHD_Ukraine_Guide_2022_Second_Edition_web.pdf

Github YOLOv7. (n.d.). Retrieved February 12, 2023, from https://github.com/WongKinYiu/yolov7

Krause, P., Salahat, E., & Franklin, E. (2018). Diurnal Thermal Dormant Landmine Detection Using Unmanned Aerial Vehicles. IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society, 2299–2304. https://doi.org/10.1109/IECON.2018.8591378

Kuenzer, C., & Dech, S. (Eds.). (2013). Thermal infrared remote sensing: Sensors, methods, applications. Springer. https://doi.org/10.1007/978-94-007-6639-6

Leira, F. S., Helgesen, H. H., Johansen, T. A., & Fossen, T. I. (2021). Object detection, recognition, and tracking from UAVs using a thermal camera. Journal of Field Robotics, 38(2), 242–267. https://doi.org/10.1002/rob.21985

Nikulin, A., de Smet, T., Baur, J., Frazer, W., & Abramowitz, J. (2018). Detection and Identification of Remnant PFM-1 'Butterfly Mines' with a UAV-Based Thermal-Imaging Protocol. Remote Sensing, 10(11), 1672. https://doi.org/10.3390/rs10111672

Roberts, S., & Williams, J. (1995). After the guns fall silent: The enduring legacy of landmines. Veterans of America Foundation, Washington D.C., USA.

Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. https://doi.org/10.48550/ARXIV.2207.02696

Yao, Y., Wen, M., & Wang, Y. (2019). Multi-Temporal IR Thermography For Mine Detection. 2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1–4. https://doi.org/10.1109/Multi-Temp.2019.8866906

YOLOv5 models; https://github.com/ultralytics/yolov5. (n.d.). Retrieved September 1, 2022, from https://github.com/ultralytics/yolov5