Intelligent Landmine Detection with Unmanned Aerial Vehicle Mounted Thermal Camera

Victor Sineglazov[1], Kyrylo Lesohorskyi[2]

[1] Department of Aeronavigation, Electronics and Telecommunication, National University "Kyiv Aviation Institute", Kyiv, Ukraine
[2] Department of Artificial Intelligence, IASA , National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Abstract

This work is devoted to the development of landmine detection intelligent system with the usage of unmanned aerial vehicle mounted thermal camera. The problem is considered under the framework of object detection. The proposed framework is based on the robust pre-processing pipeline, with a lightweight neural network performing feature extraction, classification and bounding box detection tasks. Pre-processing pipeline includes normalization, texture extraction, and noise reduction algorithms to minimize the impact of defects in the images on the accuracy of the neural network. The neural network was trained on a custom-collected dataset of various landmines with a low-altitude flyby, with captured images being used to train the neural network. The proposed method shows perfect recall (1.0), adequate precision (0.909), high Rand index (0.98), and intersection over union(0.963) metrics.

Keywords

Object detection, landmine detection, thermal imagery, convolutional neural network

1. Introduction

Even though the usage of landmines were greatly reduced by Ottawa treaty, landmine pollution is still an acute problem around the world. It is estimated that over 60 countries are still contaminated by various types of landmines and unexploded ordnance, according to the 2023 Landmine Monitor report. Most common hazards are landmines, improvised explosive devices, and artillery shells that did not explode on impact, collectively referred to as explosive ordnance (EO). It is estimated that over 4700 civilians were killed or injured in 2022 by explosive ordnance, according to the Landmine Monitor report.

Ukraine is one of the most heavily landmine-polluted countries in the world, with various estimates stating that up to a third of its territory is polluted by EO. Removal of EO is paramount for the restoration of economic activity, which can only be achieved via the process of landmine removal. The process of landmine removal is tedious, high risk, and is complicated by high rate of false positives due to various debris, present on the minefields. As such, having a detailed map of the minefield with the most likely areas where explosive ordnance is present is extremely useful for engineers that will be performing the landmine removal operation. Drones, in particular unmanned aerial vehicles (UAVs) are particularly useful, as they are able to perform a safe and fast scan of the area.

However, the process of collecting images is not the only problem, as covering 1 square kilometer at a useful resolution requires approximately 60,000 images. An expert takes, on average, 3 minutes to verify the image for presence of EO, or 3000 man hours to process 1 square kilometer. In the context of all landmine contaminated territory of Ukraine, it is estimated that over 500 million man hours are required to manually process the images. Artificial intelligence, specifically computer vision algorithms, can greatly speed up the process and make it possible to create detailed maps of landmine polluted area for the following landmine removal operation.

[1]CMIS-2025: Seventh International Workshop on Computer Modeling and Intelligent Systems, May 5, 2025, Zaporizhzhia, Ukraine

svm@nau.edu.ua (V. Sineglazov); lesogor.kirill@gmail.com (K. Lesohorskyi);
0000-0002-3297-9060 (V. Sineglazov); 0000-0003-2773-7398 (K. Lesohorskyi);

© 2024 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

CEUR-WS.org/Vol-3988/paper17.pdf

2. Literature review

2.1. .Remote Sensing for Landmine Detection

Modern mine detection methods are based on the use of a combination of sensors and mobile platforms to quickly collect information about a mined area. The most common types of sensors used for mine detection are ground penetrating radar, electromagnetic sensors, hyperspectral cameras and infrared cameras. Most of these sensors have a number of disadvantages that limit the possibility of their use with unmanned aerial vehicles - weight, price, the requirement to be directly close to the ground (which can lead to detonation of the explosive device), however, the development of infrared camera technologies has made it possible to create lightweight, compact and relatively inexpensive sensors that can be used in combination with artificial intelligence to detect mines. Infrared cameras are used to detect shallowly buried metal mines as well as non-metallic mines [1].

The presence of a buried mine is determined based on the difference in thermal characteristics between the buried objects and the surrounding soil, since a buried mine affects the thermal conductivity within the soil, resulting in a temperature difference between the buried object and the soil. [2]. This temperature contrast is measured using a thermographic camera that detects radiation in the infrared region of the electromagnetic spectrum and appears as pseudocolor in thermal images [3].

However, detecting mines in thermal images is difficult due to the temporal behavior of soil temperature distribution during the day and night, as well as the presence of other buried objects [4].

Given the difficulty of object detection in thermal imaging images, there is a need to develop suitable image processing-based decision tools for accurate landmine detection. Various researchers have proposed various methods to improve the detection of buried mines in thermal infrared images. Infrared thermal imaging can work with passive (natural) or active (man-made) heat sources. However, they are influenced by weather conditions and soil moisture [5]. Since the thermal differences between bare soil and the soil surface above buried mines are quite small, a circular symmetrical spatial filter is applied to enhance these differences [6].

Visibility of buried targets using an infrared and charge camera has been found to be difficult during sunrise and sunset [7]. Ederra proposed mathematical morphological tools for denoising and segmentation of individual images [8]. Since the raw thermography sensor image is unlikely to provide satisfactory information due to interference from solar radiation, soil conditions, humidity, etc., the complex steps of infrared thermography processing, including data acquisition, data preprocessing, anomaly detection, and evaluation of the thermal and geometric properties of the detected anomalies, are explained using appropriate techniques [9]. Image processing techniques such as Karhunen-Loeve transform (KLT), Kittler and Young transform have been used to reduce the data size and computation time in thermal image based mine detection systems [10]. KLT and watershed segmentation were proposed for landmine detection applications [11]. The concept of spectral differentiation and detection algorithm, based on the principles of pattern recognition, were developed [12]. The dynamic behavior of the scene due to time variation and cooling of solar illumination during landmine detection and its impact on the images are analyzed using image processing tools [13]. A 3D finite difference thermal model was presented and validated for detecting landmines in outdoor minefield datasets [11].

2.2. .Artificial Intelligence for Landmine Detection

The operating principle of IR radiation is based on the fact that different objects can have different thermal characteristics [12], i.e. thermal conductivity and heat capacity. Mines can be thought of as an unnatural volume for heat flow within the soil. This may cause a specific spatiotemporal thermal pattern on the soil surface, which can be detected using IR imaging systems [13]. According to [14], IR-based detection systems mainly depend on the condition of the soil surface, the nature of the soil, climatic changes, the characteristics of buried objects, their position and finally the thermal excitation. When all these factors are handled properly, IR thermography is a noteworthy detection tool for locating buried objects.

If these space-time thermal patterns are due to mines, it is called a volumetric effect. On the other hand, if they occur due to disturbed soil, it is called surface effect [15]. We have experienced that the surface effect can only be detected for a short period of time after planting. During this period, the thermal contrast is quite visible [20]. An IR system can detect these anomalies as evidence of mines [16].

According [2], IR images do not require too much pre-processing and this system can work with passive (natural) or active (man-made) heat sources. However, it can be affected by weather conditions and soil moisture. Soil moisture has a positive effect on the thermal signature of a non-metallic mine and increases the detection speed; on the other hand, it reduces the detection rate of metal mines due to the shift of thermal characteristics with humidity [17].

Deeply buried objects cannot be detected using IR sensors [18]. The maximum detection limit for mines using IR radiation is about 10 cm [19]. [15] visualizes buried landmines under three different soil surface conditions. According to their conclusion, mines buried at moderate depths in the soil do not create a direct signature.

Similar studies have been published on mine detection using IR sensors. In [20], authors monitored areas containing buried anti-tank mines and analyzed changes in surface temperature over a diurnal cycle to compare different soil textures and soil moisture. According to their analysis, it is possible to predict the cyclical behavior of the thermal signatures of mines, with the exception of soil with silty loam. Authors of [21] used 24-hour time series of IR images in their studies. They used the Karhunen-Love transform (KLT) to reduce the data size and applied three different methods to segment the mines. They enhanced the image/images using gray scale morphology. The watershed marker algorithm is then applied to the data for segmentation using these three methods. In [22] authors have presented landmine detection using KLT and watershed segmentation. In their research, they propose a series of night images from 20:00 to 01:00 with a time interval of 30 minutes. According to them, images taken in the morning and afternoon contain redundant information.

Therefore, they used a series of night images and KLT, which reduces the number of images and therefore the time required to process the data.Authors of [23] worked on a 3D thermal model for mine detection problems. In [24] a 3D thermal model to study the effects of mines on bare soil is presented. They worked with mines with low or no metal content. They simulated the thermal behavior of soil with known boundary conditions. After this, they proposed an iterative method for data classification. This iterative method gives the nature and depth of the objects. In [25], a thermal radiometric model is presented. They used the finite element method to describe thermal phenomena. They used a 25cm anti-tank mine stimulator and a virtual sensor believed to be an LWIR camera operating at a wavelength of around 10μm. Additionally, they incorporated surface roughness into their thermal and radiometric models to account for surface self-shading due to soil surface topology. According to the authors, the surface temperature above the mine is lower at dawn, and the surface is hotter during the day. Finally, at night, the soil layer above the mine is colder. Additionally, they introduce the concept of spectral differentiation and developed a detection algorithm based on pattern recognition principles in another study [26]. They used a weighted difference between visible and IR images from the same scene to remove reflected radiation from the warm atmosphere to reduce interference caused by reflected light. According to the authors, there is a trade-off between reducing interference and increasing the mine signature [27]

In [28], authors investigated how a thin outer metal casing and an air gap left over buried anti-personnel and anti-tank mines affected IR images. They used the finite element method (FEM) to describe thermal phenomena. They modeled buried anti-tank mines with and without a thin metal outer casing, as well as surface/buried anti-personnel mines. To analyze the effect of the top air gap, they also simulated an anti-personnel mine with a top air gap. The simulated mines had the thermal properties of TNT in the model. According to their results, the thin metal outer shell has a significant impact on the temperature distribution due to the noticeable difference in thermal conductivity between the metal shell and TNT.

The upper air gap has a more noticeable effect on the temperature change in depth over a given time cycle due to the low thermal conductivity of the air gap compared to the soil. In addition to this, their results show that surface mines create greater temperature extremes than buried mines.

Thanh et al. [28] presented and validated a 3D thermal model for mine detection in open minefield datasets. They proposed a finite-difference approximation of generalized solutions of the model. In addition, they proposed methods to evaluate the thermal properties of bare soil and the air-soil interface. They validated their estimated soil parameters by comparing simulations with real

data sets. They [29] also developed a method that gives the thermal diffusivity, depth and size of buried objects. In the first stage, they presented a method that can detect landmines. This method depends on thermal differences at the soil surface caused by buried objects. In the second part, their proposed method finds the thermal diffusivity, depth and size of buried objects using an inverse problem formulation. 3D modeling was developed to simulate the passive IR signature of landmines that are buried or placed on the soil surface using (FEM). In [30], a two-step method is proposed in a review study. In the first step, they found the soil temperature using their new thermal model provided by the thermal properties of the soil and the buried object. At the second stage, the discovered objects are classified using the proposed improved inverse problem setting. They called the second step setting up an inverse problem to detect landmines. They evaluate the depth, shape of a buried object, and its thermal diffusivity using their two-step method. In [31], authors have proposed a method that can reproduce the thermal properties of outdoor conditions with reduced data size and compressed time. They generated a generalized formula for this purpose. They imaged the embedded test area for eight and six hours over a two-hour period. They used a binary reduction algorithm to detect mines.

3. Problem statement

In this work, the mine detection task will be considered as an object detection problem - in target thermal images it is necessary to generate a bounding box of pixels that correspond to mines and other objects of interest (EOs). More formally, thermal image I treated as a matrix of size $H \times W$, where H is the matrix's height, W is the matrix's width. Each element of the matrix x represents an image pixel and contains a single rational value $x \in R$, which corresponds to the "brightness" of the surface in a given pixel (the higher the temperature, the higher the brightness). The work solves the problem of constructing an object detection function that turns the input thermal image I into the list of detected objects $I' = [(((x_{1l}, y_{1l}), (x_{1r}, y_{1r})), y_1), ..., (((x_{nl}, y_{nl}), (x_{nr}, y_{nr})), y_n)]$,where each element $((x_{nl}, y_{nl}), (x_{nr}, y_{nr}))$ is a bounding box of target object and, $y \in C, C = \{1, 2, 3..., c\}$, is the class of the detected object.

To avoid the effect of the "curse of dimensionality", H and W should be limited in size. When collecting information using a drone, there are two possible projections of data - individual frames obtained by the drone and an orthomosaic (stitched frames). To limit W and H in this paper, individual frames collected by an unmanned aerial vehicle (UAV) are used. If an orthomosaic is available, it must be divided into small (depending on the configuration of the neural network) patches for further processing.

When assessing the effectiveness of the proposed solution, it is important to choose the right metrics, since it is important not only to correctly classify the mine, but also to avoid type I errors (false negatives).

The following metrics will be used in this work:

1. Precision (Mine / No Mine) – the ratio of all true results to the total number of true and false true results: $Precision = \frac{TP}{TP + FP}$;
2. Recall (Mine / No Mine) – the ratio of all true results to the total number of true and false negative results. When detecting mines, it is extremely important to avoid Type I errors, which makes recall one of the key metrics: $Recall = \frac{TP}{TP + FN}$
3. Rand index, also known as classifier accuracy. This metric evaluates the ratio of true positive and true negative predictions to all predictions: $Rand = \frac{TP + TN}{TP + TN + FP + FN}$
4. Intersection over union (IoU) is a metric that is used to measure the accuracy of bounding box predictions: $IoU = \frac{[A \cap B]}{[A \cup B]}$

When assessing the quality of the intelligent system, UN standards for humanitarian demining will be used. The UN landmine clearance standard for detection rate is 99.6% for humanitarian demining.

4. Method

The work proposes an integrated approach to information collection, data preprocessing, feature extraction, and classification. The approach is based on using a quadcopter to fly over a mined area, with further image processing using a neural network. The general scheme of the approach is presented in Fig 1.

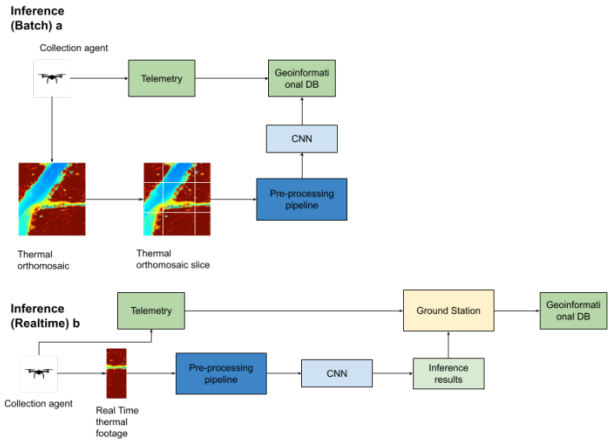

Figure 1: The proposed method's framework

4.1. . Data Collection

A quadcopter, or other type of a UAV, with a thermal camera mounted perpendicular to the ground, is used to perform a flyby over the landmine contaminated area. The flight is performed at an altitude of 10-15 meters above ground level, with a predetermined route. The height is selected depending on the area of the landmine-polluted area and weather conditions. The flyby is carried out in the afternoon, preferably in low clouds, which provides better thermal contrast between the mine and the ground.

4.2. . Data Preprocessing

The data obtained after the flight goes through a preprocessing pipeline to prepare for feature extraction and subsequent classification. The general processing pipeline is shown in Fig. 2.

The first stage of processing is normalization. When normalizing, it is important to take into account the nature of the data. Since all images collected during one flight must have the same distribution to improve generalization, normalization occurs in two stages. The first step uses linear normalization to improve the contrast of all images and bring them to the same distribution:

$$I_N = (I - Min)\frac{(newMax - newMin)}{Max - Min} + newMin\,, \quad (1)$$

where *Min* is the minimum brightness value in the original image, *Max* is the maximum brightness value in the original image, *newMax* is the new maximum value in the image, *newMin* is the new minimum value in the image.

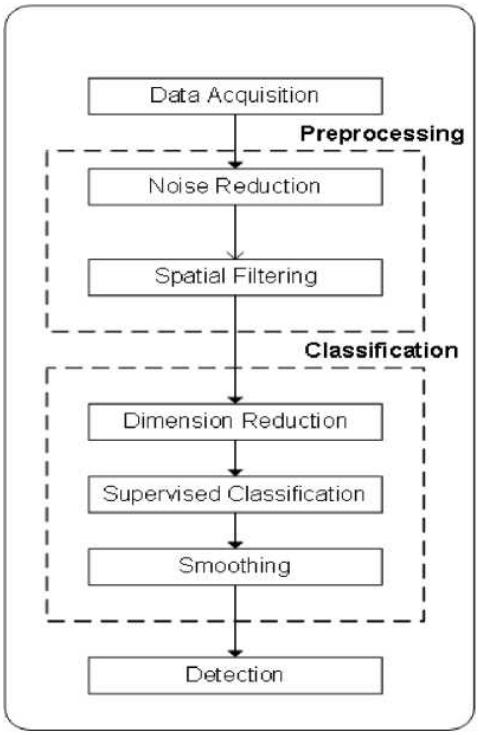

Figure 2: Data processing pipeline

Linear normalization uses classical parameters *newMax* = 255, *newMin* = 0. At the same time *Max* And *Min* the parameters are selected as the maximum and minimum values for the entire span, and not in each image separately.

After linear global normalization, local normalization is performed to increase the contrast of regions that may be unevenly illuminated by the sun. To do this, local contrast stretching is used, which is equivalent to a convolution operation using an averaging kernel:

$$I_n(x,y)=\frac{newMax*(I(x,y)-\min_I(x,y))}{\max_I(x,y)-\min_I(x,y)}, \quad (2)$$

Where *newMax* is a maximum value after normalization, $I_{x,y}$ is the pixel value *x,y* in the original image, *min(x,y)* is the minimum value for the convolution kernel in pixel x,y, *max(x,y)* is the maximum value for the convolution kernel in pixel x,y.

For local gradient stretching, it is recommended to use a convolution kernel with padding, which allows for a maintainance the original image size. Parameter values are set depending on the resolution of the input image.

After normalizing the image and stretching the gradient, the result is a high-contrast image, but it will contain noise. Stones, debris, grass, and immitators will create noise in the image, leading to a high rate of false positives. From a demining perspective, this is not a critical issue as they can be safely inspected manually, but it does add significant labor and time to demining operations. To overcome this limitation, filtering removes speckle noise and weak signals that complicate further image processing.

Filtering consists of two stages – de-texturization and morphological filtering.

Using the local binary pattern (LBP) histogram method. This is a spatial filtering method that is used to extract spatial features, especially textures, which significantly increases classification accuracy. LBP adjusts the intensity value of each pixel using a mapping function to a neighborhood function. Initially, the neighborhood function is selected. Moore's neighborhood function is often used, but other neighborhood functions can be used to increase the floorI perceive texture. For each pixel, a vector of texture characteristics is calculated:

$$LBP_p=\sum_{p=0}^{p-1} s(g_p-g_c)2^p, s(x)=\begin{cases}1 \, if \, x\geq 0\\ 0 \, otherwise\end{cases}, \quad (3)$$

where g_p is the pixel's neighbourhood value, *P* is the selected neighbourhood type, g_c is the central pixel of the nighborhood *P*.

Once the image is converted to LBP encoding, they are used to construct a texture histogram. The biggest advantage of LBP is its high processing speed and ability to store spatial patterns for high-resolution mine detection.
After the histogram is created, an additional filtering step is performed. This step uses morphological filtering to remove noise and spots that form the image. Morphological filtering filter removes noise and insignificant objects from the texture. Morphological image processing is a set of tools for analyzing and processing structural features of images based on set theory. These techniques can extract and enhance the spatial characteristics of objects in images, making them extremely useful in image processing and computer vision.
The first stage is erosion - reducing the number of objects in the image by removing pixels at the boundaries of objects. This removes minor noise:

$$(A \ominus B)(i,j) = \min_{x,y \in B} A(i+x, j+y), \tag{4}$$

where A is the original image, B is a structural element. After this, the image must be restored to avoid loss of features, for which the expansion operator is used:

$$(A \oplus B)(i,j) = \min_{x,y \in B} A(i-x, j-y), \tag{5}$$

where A is the original image, B is a structural element. The morphological filtering operation is a composition of the erosion and dilation operator, which removes noise from the image and makes it clearer.
These steps ensure that the input images are cleaned, noise is removed and textures are preserved (if possible). These steps also partially extract features (through LBP and morphological filtering), which allows for a more simple neural network architecture, decreasing the number of learnt parameters.

4.3. . Feature Extraction and Classification

This paper considers a hybrid architecture for the task of segmentation and classification. Due to the need to calibrate the sensitivity of the mine detection network, two-step segmentation is used, which allows the sensitivity and accuracy of the network to be adjusted separately from each other.
Segmentation will be performed in two stages - the first stage uses U-net with residual connections to identify areas of interest that are most likely to contain a mine. These zones are marked, expanded and fed into a convolutional neural network to classify the type of mine.
Convolutional neural networks are used to solve the problem of feature extraction and classification. This is a common approach for solving computer vision problems. There are many architectures, but most of them are designed to process complex images with a large number of features and possible classes. After pre-processing, the dimensionality and complexity of the data is significantly reduced, which makes it possible to synthesize [32, 33] a simpler architecture [34].

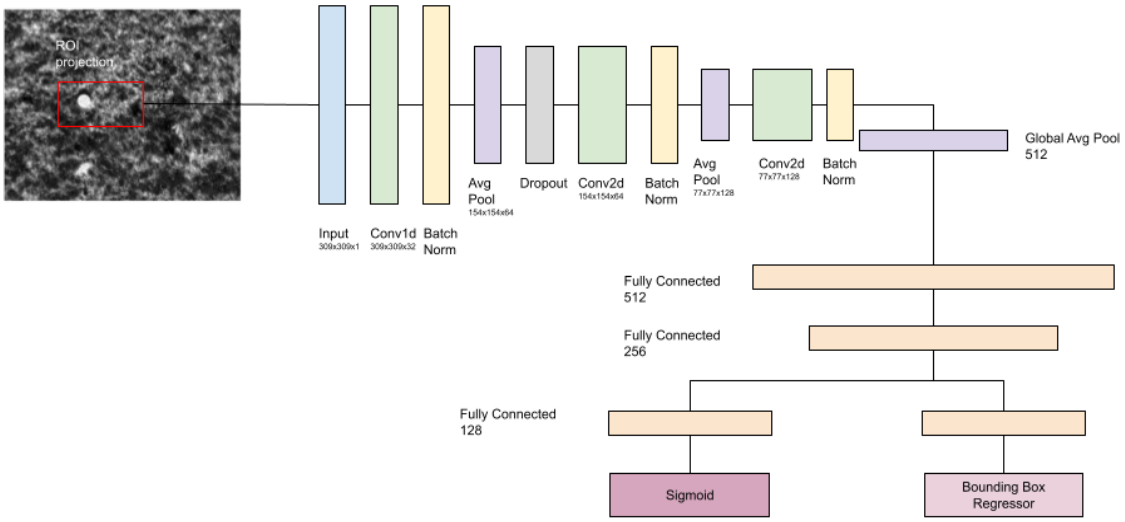

Figure 3: Proposed neural network architecture

The proposed network consists of the following types of layers:
1. Convolutional layer is the primary building block of CNN where the convolution operation occurs. Filters (kernels) slide over the entire image, calculating the dot product between the filter and part of the input image, creating feature maps;
2. The pooling layer performs dimensionality reduction of feature maps, preserving the most important features. The most common types are MaxPooling (selects the maximum value in each window) and AveragePooling (selects the average value);
3. BatchNorm layer is used to normalize feature maps, which increases stability and learning speed;
4. The exclusion layer is used to prevent overfitting by randomly "turning off" some neurons during training;
5. A fully connected layer has its neurons connected to all the neurons of the previous layer, which enables combining features, making a final classification decision;
A key feature of the proposed architecture is the relatively low depth of the convolutional network. This allows you to reduce the number of parameters, which speeds up training and processing. This was achieved through the use of a comprehensive preprocessing pipeline.
For object detection, region-based convolutional neural networks were selected as a baseline for the model. Specifically, we use fast R-CNN with a convolutional pathway outlined above. While this is not the most robust algorithms, it performs reasonably well due to the nature of the domain and robust pre-processing pipeline, which partially extracts the features, reducing the learning capacity that is expected of neural network.
To achieve object detection, ROI projection is used to extract areas of the image that are then passed through the convolution path outlined above. Global average pooling is used to build a feature vector from the image, passing through two fully connected layers. After this, a pathway branches into two – the classification pathway which utilizes sigmoid to perform binary classification task and through a bounding box regressor, which extracts the bounding box from the feature vector of the image. The selected architecture is somewhat simplistic, however it was considered to have sufficient learning capacity in the context of this problem [38].
To train the learner, a multi-task loss is used. Classification loss is based on a simple binary cross-entropy loss. Bounding box repressor needs a location-aware loss, as such, a modification of IoU is used as a loss function. The problem with using IoU itself is that if no overlap is detected between the target object and the classified image, the loss function becomes constant. To overcome this, a variety of modifications of IoU is designed to serve as a loss function. In this paper, a generalized intersection over union (GIoU) is used. The rectangle bounding box C is used to build a convex of an object that encloses both A and B:

$$GIoU = \frac{[A \cap B]}{[A \cup B]} - \frac{C \setminus (A \cup B)}{C}, \tag{6}$$

where A and B are groundtruth and predicted bounding box areas, C is the bounding box that covers both A and B.
The GIoU loss is constructed as following:

$$L_{GIoU} = 1 - GIoU. \tag{7}$$

To handle object detection, a selective search, ROI polling, and bounding box prediction modules are added to the network's architecture.

5. Experiment & Result

To test and evaluate the proposed system, an experiment was conducted to collect data, train and evaluate the proposed system. Data was collected using a DJI ZH20T thermal camera attached to a Mavic Phantom T4 quad copter. The mines were installed on the surface and also buried in the ground to a shallow depth (up to 10 cm). The study used two types of mines - anti-tank and anti-personnel, both types had a metal casing. The flight was carried out at a low altitude (5-6 meters) and medium altitude(10-11 meters). The integrity of the grass cover was damaged only in the places where the mines were installed, but otherwise the cover remained intact. Data were collected in clear, warm weather to minimize noise and maximize image quality. The temperature was shifting thought the day, which ensured high quality thermal gradient in the collected dataset.
A limited amount of debris and uneven ground was present in the collection area, which also created additional noise and false-positive spots in the thermal gradient on the ground. This creates additional challenge for the neural network, as this noise is not fully removed by the pre-processing

pipeline. Collected data set consists of 436 thermal images, with 62 images being removed due to low quality.

After data collection, the resulting frames undergo pre-processing, eliminating low-quality frames. Low-quality footage refers to footage with high levels of noise or blur. Such frames have an extremely negative impact on the quality of classification and detection of objects, reducing the accuracy of the neural network. One of the main problems is motion blur. If the desired object falls into a blurred area of the image, this significantly increases the bounding box of the desired object, negatively affecting the training of the network, so in this study such frames were removed from the original data set. In further research, it is possible to use an algorithm to restore the quality of images. An example of a blurry (low-quality image) is shown in Fig. 4

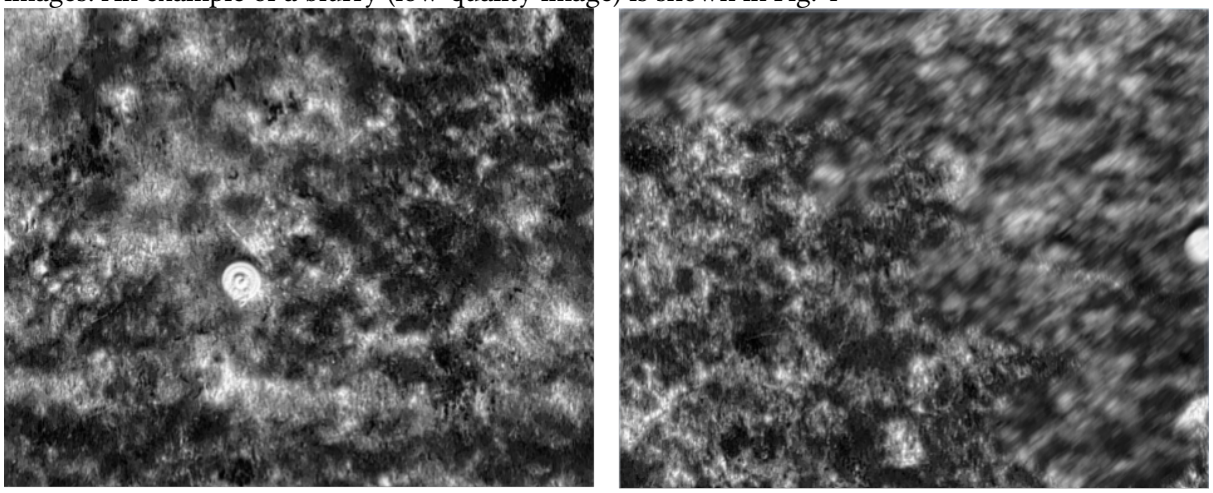

Figure 4: Normal(a) and blurred image (b) examples from the dataset. Motion blur is one of the major contributing factors.

The neural network was trained in batches of 16 images. The sample size is bolstered by applying "weak" augmentations that consist of rotations, stretching, and other augmentations from RandAugment. The preprocessing stage was performed for each batch separately to increase the generalization ability of the neural network. Adam learning algorithm with a decaying learning rate from 0.001 to 0.00001 during the training of the neural network. Default parameters were used during the initial training, with further fine-tuning during the experiment using methods outlined in [35, 36]

The results of the training are presented in the table 1, and a sample of classification is given in Fig. 5

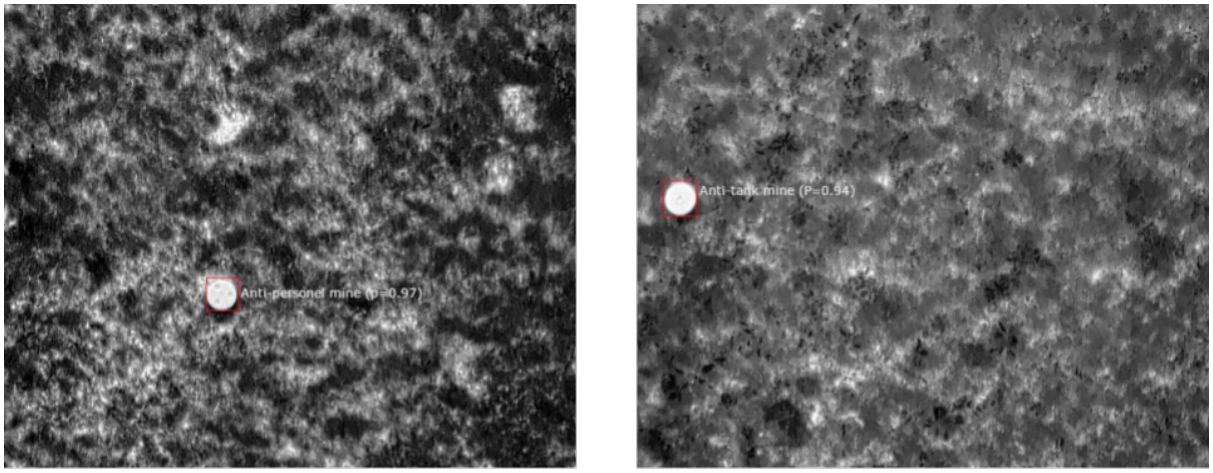

Figure 5: Object detection result. The system is unable to classify the type of the landmine, labels were added manually for clarity

Table 1
Experiment result

Metric	Value	Note
Precision	0.909	Precision is reasonably high, with several samples being mislabeled (with weak blur / noise)
Recall	1.0	All landmines were successfully detected
Rand index (OA)	0.98	Overall high, apart from few misclassifications
IoU	0.963	Bounding boxes around the mines are precise, apart from few cases where they are somewhat larger than anticipated.

Overall, the algorithm performs well, detecting both landmines installed over the ground and under the ground, with high reliability, however it should be noted that even on a small number of samples and with a large number of training iterations, the network has modest precision. While in the case of landmine detection this is not as bad as having low recall, it is still something that should be addressed, as a high number of false-positives leads to slower and more costly landmine removal operation.

Conclusion

This paper presents a comprehensive framework for landmine detection with thermal imagery cameras installed on a mobile platform (UAV). The proposed approach is based on the combination of deterministic pre-processing to pre-extract features from the images, followed by a region-based convolution neural network detector for feature extraction, classification and ROI extraction.

The proposed approach was able to achieve high recall (1.0) and moderate precision (0.92). Algorithm's average IoU is reasonably high at 0.875, with results being skewed by false positives and unclear edge for buried landmines.

Future research will be focused on addressing some of the shortcomings of the algorithm discovered in this paper. First such shortcoming is the diversity of the dataset. The research is based on the dataset collected over 1 flyby over a limited area, with limited number of landmines available and in new-perfect weather conditions. Future dataset collection should be focused around building a more challenging and diverse dataset for neural network training and evaluation. The second consideration to address is identifying optimal architecture for the neural network itself. In this research, a simple yet robust fast CNN-based region object detection. It works reasonably well, considering the nature of data and feature pre-extraction step, however exploration of other options, such as mask-based CNN could improve performance of the method. Lastly, in this paper binary object detection was used. For the proposed method to be practically valuable, it would also be beneficial to identify the type of the landmine, so future research will focus on multi-class object detection to not just detect the mine itself, but also it's type or exact model.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly in order to: Grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] R. Bello, "Literature review on landmines and detection methods," Frontiers in Science, vol. 3, no. 1, pp. 27-42, 2013.

[2] H. Kasban, O. Zahran, S. M. Elaraby, and M. El-Kordy, "A comparative study of landmine detection techniques," Sensing and Imaging: An International Journal, vol. 11, pp. 89–112, 2010. DOI 10.1007/s11220-010-0054-x

[3] T. Nguyen, D. Hao, P. Lopez, F. Cremer, and H. Sahli, "Thermal infrared identification of buried landmines," in Proceedings of the SPIE, 2005, vol. 45794, pp. 198–206.

[4] J. A. Richards and X. Jia, “The effect of the atmosphere on radiation,” in Remote Sensing Digital Image Analysis: An Introduction. Canberra: Springer, 2005, p. 28. DOI 10.1007/978-3-030-82327-6

[5] A. Linder, S. Nyberg, S. Sjokvist, and M. Uppsal, “Optical method for detection of mine fields,” Swedish Defence Research Agency, Base data report, September, 2004. DOI 10.4186/ej.2021.25.3.61

[6] S. Kaya, “Buried and surface mine detection from thermal image time series,” Degree of Master of Science in Geodetic and Geographical Information Technologies Department, Middle East Technical University.

[7] Y. H. L. Janssen, A. N. de Jong, H. Winkel, and F. J. M. van Puten, “Detection of surface laid and buried mines with IR and CCD cameras, an evaluation based on measurements,” in Proceedings of SPIE Detection and Remediation Technologies for Mines and Minelike Targets, A. C. Dubey, R. L. Barnard, C. J. Lowe, and J. E. McFee, Eds, 1996, vol. 2765, pp. 448–459. DOI 10.1117/12.241248

[8] G. Ederra, “Mathematical morphology techniques applied to anti-personnel mine detection,” MS Thesis, Department of Electronics and Information Processing, Vrije Universiteit Brussel. 1999.

[9] N. T. Thanh, D. N. Hao, and H. Sahli, “Infrared thermography for land mine detection,” in Augmented Vision Perception in Infrared—Advances in Pattern Recognition Series, R. I. Hammoud, Eds. London: Springer, 2009. DOI 10.1007/978-1-84800-277-7_1

[10] L. Kempen, M. Kaczmarec, H. Sahli, and J. Cornelis, “Dynamic infrared image sequence analysis for anti-personnel mine detection,” in Proc. IEEE Benelux Signal Processing Chapter, Signal Processing Symposium, 1998, pp. 215–218.

[11] N. T. Thanh, H. Sahli, and D. N. Hao, “Finite-difference methods and validity of a thermal model for landmine detection with soil property estimation,” IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 3, pp. 656-674, 2007. DOI 10.1109/TGRS.2006.888862

[12] A. Ajlouni and A. Sheta, “Landmind detection with IR sensors using Karhunen Loeve transformation and watershed segmentation,” in The 5th IEEE International Multi-Conference on Systems, Signals and Devices, 2008, pp. 1-6. DOI: 10.1109/SSD.2008.4632869

[13] I. K. Sendur and B. A. Baertlein, “Numerical simulation of thermal signatures of buried mines over a diurnal cycle,” in SPIE 4038, Detection and Remediation Technologies for Mines and Mine like Targets V, 2000. DOI 10.1117/12.396243

[14] Gonzalez, P., Cobano, J. A., Garcia, E., Estremera, J., & Armada, M. A., "A six-legged robot-based system for humanitarian demining missions," Mechatronics, vol. 17(8), pp. 417-430, 2007. DOI 10.1016/j.mechatronics.2007.04.014

[15] Khanafer K., Vafai K., and Baertlein, B. A., "Effects of Thin Metal Outer Case and Top Air Gap on Thermal IR Images of Buried Antitank and Antipersonnel Land Mines," IEEE Transactions on Geoscience and Remote Sensing, vol. 41, no. 1, pp. 123-135, 2003. DOI: 10.1109/TGRS.2002.807755

[16] Nguyen, T., T., Sahli, H. and Nho, H., D., "Thermal infrared technique for landmine detection: Mathematical formulation and methods," RICAM, 2011.

[17] Lillesand, T., M., Kiefer, R., W., Chipman, J., W., Remote Sensing and Image Interpretation, John Wiley & Sons,Inc., 2007.

[18] Richards, J., A. and Jia, X., "The Effect of the Atmosphere on Radiation," in Remote Sensing Digital Image Analysis An Introduction, Canberra, Springer, 2005, p. 28.

[19] Sendur, Ibrahim K. and Baertlein, Brian A., "Techniques for improving buried mine detection in thermal IR imagery," in 3710, Detection and Remediation Technologies for Mines and Minelike Targets IV, 1999. DOI 10.1117/12.357009

[20] Paik, J., Lee, C., P. & Abidi, M., A.i, "Image Processing-Based Mine Detection Techniques:A Review," Subsurface Sensing Technologies and Applications, vol. 3, no. 3, 2002. DOI 10.1023/A:1020399314530

[21] Bruschini, C. & Gros, B., "A survey of Current Sensor Technology Research for the detection of landmines," in In the proceedings of International workshop on Sustainable Humanitarian Demining, 1997.

[22] Cremer, F., Nguyen, T. T. , Yang, L. & Sahli, H., "Stand-off Thermal IR Minefield Survey: System concept and experimental results," in Proceedings of the SPIE, Vol. 5794, pp 209 - 220., 2005. DOI 10.1117/12.626264

[23] Dam, R. L. V., Borchers, B., Hendrickx, J. M. H. & Harmon, R. S., "Effects of soil water content and texture on radar and infrared landmine sensors:implications for sensor fusion," in In the proceedings of European Demining, 2003, 2003.

[24] Bruschini, C., & Gros, B., "A Survey of research on sensor technology for landmine detection," Journal of Humanitarian Demining, no. 2.1, 1998.

[25] Khanafer, K., Vafai, K., "Thermal analysis of buried land mines over a diurnal cycle," Geoscience and Remote Sensing, vol. 40, no. 2, pp. 461-473, 2002. DOI: 10.1109/36.992811

[26] Hong, S. H., Miller, T. W., Borchers, B., Hendrickx, J. M., Lensen, H. A., Schwering, P. B., & Van Den Broek, S. P., " Land mine detection in bare soils using thermal infrared sensors," in In AeroSense 2002 (pp. 43-50). International Society for Optics and Photonics., 2002. DOI 10.1117/12.479124

[27] Nguyen, T., Hao, D., P. Lopez, F. C., and Sahli, H., "Thermal infrared identification of buried landmines," in In Proceedings of the SPIE, volume 45794, pages 198–206., 2005. DOI 10.1117/12.626263

[28] Martнnez, P.,L., Kempen, L.,v., Sahli, H., Ferrer, D., C., "Improved Thermal Analysis of Buried Landmines," Transactions on Geoscience and Remote Sensing, vol. 42, no. 9, 2004. DOI: 10.1109/TGRS.2004.831884

[29] Thanh,N., T., Sahli, H., and Hao, D., N., "Infrared Thermography for Buried Landmine Detection: Inverse Problem Setting," Transactions on Geoscience and Remote Sensing, vol. 46, no. 12, 2008. DOI: 10.1109/TGRS.2008.2000926

[30] Muscio, A., Corticelli, M., A., "Experiments of thermographic landmine detection with reduced size and compressed time," Infrared Physics & Technology, vol. 46, no. 1-2, pp. 101-107, 2004. DOI 10.1016/j.infrared.2004.03.014

[31] Lundberg, M., Gu, I., Y., H., "3D matched filter for detection of land mines using spatio-temporal thermal modeling," in SPIE 4038, Detection and Remediation Technologies for Mines and Minelike Targets V, 179, 2000. DOI 10.1117/12.396245

[32] Sineglazov, V., Kot, A. Design of Hybrid Neural Networks of the Ensemble Structure Eastern-European Journal of Enterprise Technologies, 2021, 1, pp 31–45. DOI 10.2139/ssrn.3807474

[33] Khotsianivskyi, V., Sineglazov, V. Robotic manipulator motion planning method development using neural network-based intelligent system. Machinery and Energetics, 2023, 14(4), pp 131–145. DOI 10.31548/machinery/4.2023.131

[34] Zgurovsky, M., Sineglazov, V. , Chumachenko, E. ”Classification and Analysis Topologies Known Artificial Neurons and Neural Networks”. Studies in Computational Intelligence, 2021, 904, pp 1–58. DOI 10.1007/978-3-030-48453-8_1

[35] Sineglazov, V.M., Riazanovskiy, K.D., Chumachencko, O.I. Multicriteria conditional optimization based on genetic algorithms. System Research and Information Technologies, 2020, 2020(3), pp 89–104. DOI 10.20535/SRIT.2308-8893.2020.3.07

[36] Zgurovsky, M., Sineglazov, V., Chumachenko, E. ”Classification and Analysis of Multicriteria Optimization Methods”. Studies in Computational Intelligence, 2021, 904, pp. 59–174. DOI 10.1007/978-3-030-48453-8_2