Intelligent hyperspectral image processing for landmine detection and classification

Viktor Sineglazov[1], Kyrylo Lesohorskyi[2] and Volodymyr Lytvynenko[3]

[1] *National Aviation University, Liubomyra Huzara Ave, 1, Kyiv, 03058, Ukraine*
[2] *National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute", Beresteiskyi Ave, 37, Kyiv, 03056, Ukraine*
[3] *Kherson National Technical University, Beryslavs'ke Hwy, 24, Kherson, Kherson Oblast, 73008, Ukraine*

Abstract

This work is devoted to a hybrid three-stage approach to hyperspectral image classification to solve the problem of remote landmine detection. A comprehensive overview of landmine pollution and its effects is given. Challenges of landmine detection and classification are highlighted. An overview of existing projects and practical applications is given. The method that utilizes a two-step pre-processing followed by a robust convolution neural network-based feature extractor and classifier is proposed. The proposed method is considered and tested in both batch and real-time scenarios. The results show state-of-the-art accuracy and viability of real-time landmine detection, however, the method is prone to high false positive incidence.

Keywords [1]

Hyperspectral imagery, intelligent systems, landmine detection, convolutional neural networks.

1. Introduction

Landmines and cluster munitions are the main obstacle to the development of societies and the return to normal life after a ceasefire. According to the latest statistics [1], 80% of landmine victims are children who have nothing to do with the war or its causes. Therefore, it is necessary to prohibit the use of this type of blind weapon. Efforts to ban landmines have already begun, and we have an international campaign to ban landmines and cluster munitions with a number of signatory countries.

One of the earliest and most widely used methods is the metal detector. Thanks to the phenomena of electrical induction, this type of detector is able to detect objects containing metal underground. The most widely used mine detection method follows the same methods developed during World War II and directly involves people. A typical explosive ordnance engineer toolkit today is very similar to those used more than 50 years ago (it consists of a metal detector and a probe). However, this method has a number of disadvantages, namely low mine detection speed, high sensitivity to noise, and a short detection radius, which leads to the need for the operator to be physically present in close proximity to explosive objects.

Modern mine detection methods focus on increasing the detection radius and expanding the types of mines they can detect. Each method is suitable for detection in specific conditions depending on the type of mine casing, explosive, and soil.

Generally, most mine detection methods consist of three main components: a sensor to capture the signature of the mine, a signal or image processing unit to organize the received data, and a decision unit to determine whether a mine exists or not. Among the sensors, the following main types are distinguished: electromagnetic, acoustic, nuclear, biological, chemical and mechanical. Electromagnetic, namely electro-optical, sensors allow detecting objects at a great distance, and do so faster than most other types of sensors. Among electromagnetic sensors, hyperspectral sensors

ProfIT AI 2024: 4th International Workshop of IT-professionals on Artificial Intelligence (ProfIT AI 2024), September 25-27, 2024, Cambridge, MA, USA

svm@nau.edu.com (V. Sineglazov); lesogor.kirill@gmail.com (K. Lesohorskyi); immun56@gmail.com (V. Lytvynenko)
0000–0002–3297–9060 (V. Sineglazov); 0000-0003-2773-7398 (K. Lesohorskyi); 0000-0002-1536-5542 (V. Lytvynenko)

are the most suitable for solving the problem of mine detection, since they have a high number of channels that allow detecting mines, while providing a good balance between dimensionality, the amount of useful information, and the speed of the intelligent system. A comparison of various electromagnetic sensors is presented in Table 1.

Table 1
Electro optical sensor types comparison

Type	Number of channels
Visible spectrum	3-5
Broadband	8-10
Multispectral	10-100
Hyperspectral	100-500
Ultraspectral	>=1000

As such, hyperspectral imagery is used for it's balance of information and redundancy. This method makes possible to measure the proportion of light reflected in hundreds of wavelengths at each image unit (pixel). Thus, we obtain a hypercube consisting of two spatial dimensions and a third dimension that contains spectral information. This method is used in the field of remote sensing for various purposes such as mapping, agriculture, astronomy, food monitoring, surveillance and others.

When light hits an object, it is either absorbed or reflected. The fraction of light that is reflected depends on the size of the molecules of the object that is reflected, the intermolecular distances, and the wavelength of the radiation. Each material, consisting of different components, can reflect light of different wavelengths differently. There are several approaches to target detection using hyperspectral imaging: some are supervised, where the spectrum of the data is known in advance; others are unsupervised, based on finding targets that are spectrally different from their surroundings. The latter type of information does not require knowledge of the target spectrum in advance. However, this type of detector is characterized by a high rate of false alarms, since rare events in the image that differ from their background will be labeled as targets.

Spectral image processing is a non-trivial task and has a number of challenges that need to be taken into account to create an effective algorithm. The first problem is the high dimensionality of the data. Hyperspectral images cover hundreds of spectral bands, which increases the size of the data and creates the curse of dimensionality. This complicates the process of creating models and increases the cost of computations required for both training and data processing. The second problem is the limited labeled dataset. Collecting and annotating hyperspectral data is a time-consuming and expensive process, which often results in a limited amount of available training data. This limits the ability to effectively train complex models. The third problem is the coupling of spatial and spectral information. Spectral features are key for material recognition, but spatial features such as texture and context are also important for accurate classification. Efficiently integrating these two types of information is a challenge for classical computer vision algorithms. The final challenge is computational power needed for hyperspectral data processing, especially when using modern deep learning methods. This limits the ability to use complex models in real time.

This paper is devoted to developing a comprehensive landmine detection method with a novel dual-mode (batch / real-time) capability that utilizes a hyperspectral sensor in conjunction with a neural network classifier to detect the landmines. The proposed method addresses the problem of high dimensionality by applying a two-step preprocessing to reduce the dimensionality while maintaining both spectral and textual information. Proposed approach provides a flexible tool to explosive ordnance removal teams that supports initial area survey with batch mode and enables mine removal operations support with the real-time mode.

2. Literature review

2.1. Project overview

This section provides an overview of past projects using hyperspectral imagery. Note that some military studies and projects may exist in this area, but they are not considered here due to lack of information. One of the first projects to study mine detection using infrared wavelengths was conducted at Defence Research & Development Canada (DRDC). DRDC began its research in support of the Canadian Army on mine and unexploded ordnance detection in 1978 and in collaboration with Itres Research on hyperspectral imaging for mine detection in 1989. The algorithms developed during this project can be applied to pre-processed images from hyperspectral imagers. An early project proposed a hierarchical image processing algorithm to detect a sparsely distributed bright region a few pixels wide in a monochromatic image [2]. The pre-processing operation is performed to remove distortions, dropouts, overlapping regions, misregistration and any other artifacts and defects. Unsuspected regions are discarded to reduce the data size. Then, the suspected regions are segmented into homogeneous subregions and the morphological features of the subregions are extracted. Based on the extracted features, the subregions are classified. Finally, the spatial relationships between the mine-like objects are determined. The supervised method analyzes these relationships and classifies the regions as a minefield by providing a certain likelihood ratio.

Fusion of visible and SWIR bands can provide better detection results. Basic fusion of two spectral bands provides acceptable segmentation of objects from the background, regardless of the illumination conditions. In other words, choosing a set of two or three spectral bands from an image has been shown to be as effective in differentiating artificial objects from the background as using all spectral bands simultaneously [3]. Such fusion has the potential to detect mine-like objects in an image using an integrated camera with visible and SWIR sensors and more sophisticated and specialized detection algorithms. [4] describes a Defense Advanced Research Projects Agency (DARPA)-sponsored experiment to test the feasibility of detecting buried mines using midwave infrared (MWIR) (3 to 5 μm) and longwave infrared (LWIR) (8 to 12 μm) hyperspectral bands. The project focuses on detecting surface disturbances caused by buried mines. Previous experiments have shown the ability of VNIR and SWIR imagers to detect surface disturbances [5], [6], [7]. However, the problem was the high rate of false alarms caused by surrounding vegetation and rocks. According to the authors, the main rationale for detecting buried mines using spectral properties is that the surface properties are somewhat different from the properties of the subsurface soil. The impact of soil on the surface changes some of its physical and chemical properties. These experiments showed that spectral information is necessary for mine detection. The objective of the study presented in [8] is to test the design of multispectral and hyperspectral imagers that are able to achieve better detection performance while meeting the requirements and conditions of target detection. Target detection requires detection in both day and night conditions. Panchromatic or multispectral images in the VNIR and SWIR ranges provide this capability during the daytime. However, for military use, the MWIR and LWIR bands are needed for nighttime operation. Due to the high correlation of spectral bands of background materials in all background conditions, the target detection capability is high using the MWIR and LWIR bands. After testing the two bands in the MWIR and LWIR bands, the authors concluded that thermal multispectral images would provide better target detection and false alarm rate than a single-band infrared sensor. The tests showed that properly selected small bands could provide good detection, the optimal band range being between 8 and 10.5 micrometers. The usefulness of using LWIR with MWIR increased significantly compared to using MWIR alone.

DSTL Defense Science and Technology Laboratory Mine Defense Project. A project similar to the DRDC and DARPA projects was started in Rythen to detect mines using a VNIR imager [9]. The program was called the DSTL Mine Defense Project. Using a tripod-mounted VNIR SOC 700 hyperspectral camera, the team captured high spatial resolution images of the mines. However, the data is mainly used to explore different processing methods rather than to evaluate the PD and FAR of the sensor. To process the data, the authors used principal component analysis (PCA) for dimensionality reduction and anomaly detection for classification. The authors avoided using spectral comparisons between the target and each pixel in the image, as this would be very time-consuming due to the low target-to-background ratio. The results were still preliminary, but the authors concluded that VNIR can distinguish surface mines from the background. In India, researchers proposed a hierarchical algorithm for mine detection using infrared images, which consists of pre-processing (contrast enhancement - filtering - smoothing), segmentation, feature extraction, and ANN-based classification [10]. The authors tested the algorithm on surface mines in two soil types: black cotton and sand. During pre-processing, the image is converted to gray color. The two most important preprocessing steps are contrast enhancement and noise removal. During the tests, the authors used a small NN with 1 hidden layer and 4 neurons. The results obtained on a simple dataset are good, but it is not expected that the algorithm will work well on another field or soil type, since the data used in the training stage is not complete enough. In 2015, TELOPS, a Canadian research company specializing in infrared and hyperspectral imaging, demonstrated the feasibility of detecting buried objects using an airborne LWIR hyperspectral imager [11]. From the aircraft, they obtained thermal hyperspectral images of areas that contain previously buried artificial objects. They found that the disturbed soil directly above the buried target is warmer than the undisturbed soil area next to it [11]. Comparing the emissivity data obtained by decoupling temperature and emissivity, buried targets are displayed as part of the hottest ground region within the scene, but additional classification or information is needed to distinguish buried objects from other naturally hot regions. It is worth noting that most of the works reviewed pay much attention to the quality of electro-optical sensors and pre-processing, but the classifier itself is quite simple. The use of modern artificial intelligence technologies, namely convolutional neural networks (CNN), will have a positive effect on both the overall accuracy and specificity of the classifier. Another common problem is the limited training dataset. Semi-supervised learning methods are used to solve this problem.

2.2. Hyperspectral image processing overview

In this section, we present the detection algorithms used to detect targets in hyperspectral imagery. In addition, we present several pre-processing steps and hyperspectral data processing typically used in the pre-processing step to facilitate further detection or classification. An overview of the various processing methods used for data fusion, spectral unmixing, classification, and target detection can be found in [12].

The first step of pre-processing is normally a contrast enchantment. The process of image contrast enhancement consists of a set of methods that attempt to improve the visual appearance of an image or transform the image into a better form suitable for human or machine analysis [13]. Image contrast enhancement methods are divided into two main categories: spatial domain methods and frequency domain methods. Spatial domain methods are applied directly to the pixels of an image. In frequency domain methods, the image is processed in the frequency domain after applying the Fourier transform to the original data. Contrast enhancement is one of the most commonly used image contrast enhancement methods. In the case of mine detection, the role of contrast enhancement is to enhance the difference between the mine and the background materials [14]. The main contrast enhancement methods used are histogram equalization and morphological contrast enhancement.

Contrast enchantment is followed by filtering step. Filtering is an operation that always reduces noise or shapes blurred areas in an image to make it clearer and more suitable for further processes. In hyperspectral image filtering, several methods commonly used in image processing have been upgraded to achieve multi-channel reconstruction. There are two groups of filters: one is based on the assumption that intra-channel information is separable from inter-channel information, i.e., spectral and partial information are separable. These filters are called hybrid filters. In this case, the first step is to decorrelate the channels of the Fourier transform or PCa, and then apply a classical 2D reconstruction method such as the Wiener filter or the static wavelet transform. The other group consists of several proposed filters that do not rely on the assumption of spectral and spatial separability [15].

After filtering, segmentation is performed. In the remote sensing community, segmentation is defined as the process of finding homogeneous regions in an image followed by classification of these regions [16]. In image processing, there are many methods used for segmentation, however, not all of them are applicable to multispectral and hyperspectral images. Some methods, such as watershed algorithms, have been modified for hyperspectral image segmentation. Globally, segmentation algorithms are divided into two categories: edge-based and region-based. Edge-based methods detect the boundary using the discontinuity property. In the region-based algorithm, the pixels in a region are grouped using the similarity property. In the following, we introduce the main methods used in hyperspectral image segmentation.

The next step is feature extraction from the segmented image. Feature extraction is the transformation of data from a high-dimensional space to a lower-dimensional space chosen in such a way as to preserve as much of the information of interest in the data as possible. Feature extraction is used in hyperspectral image analysis to address the problem of the small number of training data samples compared to the high spectral resolution of the image and to reduce computation time. There are many feature extraction algorithms that have been introduced; some are linear while others are nonlinear. When working on mine or target detection, not all feature extraction algorithms are useful because the targets of interest are usually rare and feature extraction can highlight key features of the target. The most common feature extraction algorithms are Principal Component Transformation (PCT), Linear Discriminant Analysis (LDA), matched pursuit [17], neighborhood embedding [18], Sammon's mapping [19] and nonparametric weighted feature extraction [20], Artificial neuron Networks (ANN).

3. Problem statement

Hyperspectral image classification is one of the key tasks in remote sensing, which involves identifying the types of materials or objects on the earth's surface based on their spectral characteristics. Hyperspectral images contain information in hundreds of narrow spectral bands, which allows obtaining detailed data on objects and materials that cannot be detected by traditional multispectral images.

In this work, image classification is considered in the supervised setting. Formally, a labeled dataset $D = \{(x_1, y_1), \dots (x_n, y_n)\}$, where $x \in X, y \in Y$, X is the input space, Y is the label space is used. The goal of supervised learning is to train an approximator function $f(x, \theta) \approx P(y \mid x)$, $x \in X, y \in Y$, where f – is the approximator function, $P(y \mid x)$ – is the target marginal distribution, θ are the approximator parameters.

For hyperspectral images, each point in the input space X is a tensor (hypercube) with the dimensions of $H \times W \times C$, where W is the image width, H is the image height, C is the number of image channels. The fundamental difference from classical computer vision tasks is that the dimension C is in the range [100, 1000], which significantly increases the number of input parameters. Thus, an image in the RGB spectrum with a resolution of 256×256 pixels contains 196608 parameters, while a hyperspectral image of the same resolution using 100 channels already contains 6553600 parameters, which significantly increases the requirements for the efficiency of image coding for effective processing.

As the problem is considered in the multi class classification setting, the output (or label) space is described as a set $C = \{c_1, \dots c_k\}$, where k is the number of distinct classes. Approximator f is trained to correctly assign a label $c \in C$ for each possible input $x \in X$ in the input space, generalizing unknown samples.

4. Method

In this work we propose a three-stage approach to hyperspectral image classification that combines classical pre-processing techniques and more robust ANN-based classification methods. The three steps are:

1. Normalization;
2. Filtering and texture extraction;
3. Feature extraction and classification.

In the following subsections we describe each step in more details.

4.1. Preprocessing pipeline

The first step of the method is a preprocessing pipeline that reduces the dimensionality of the input, normalizes the data and prepares it for feature extraction and classification step. The first step of pre-processing is noramlziation. In this step, the spectral image is normalized and filtered. First, all spectral data is normalized channel by channel using min-max normalization:

$$x_{i,j,c} = \frac{x_{i,j} - \min x_c}{\max x_c - \min x_c}, \quad (1)$$

where c – is the spectrum channel, $i \in W$ – is the pixel's horizontal coordinate, $j \in H$ – is the pixel's vertical coordinate. This helps to normalize values, and by extension spectras, which simplifies further processing. Min-max normalization was chosen, as it is well suited to normalize the data for all of the channels while maintaining local patterns within each pixel's spectral bands.

The next step is filtering using the histogram method Local Binary Patterns (LBPs). This is a spatial filtering method that is used to extract spatial features, especially textures, which significantly increases the classification accuracy. LBP adjusts the intensity value of each pixel by applying a transformation function to the neighborhood function. First, a neighborhood function is selected. Often, the Moore neighborhood function is used, but other neighborhood functions can be used to increase the texture receptive field. For each pixel, a vector of texture features is calculated:

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

where g_p-is the neighborhood pixel value, P –is the neighbourhood space, $g_p \in P$, g_c – is the central pixel. After the iamge is translated into LBP encoding, they are used to build a texture's histogram. LBP's biggest advantage is high processing speed and the ability to preserve spatial patterns (and textures) for high-resolution landmine detection.

After the histogram is created, an additional filtering stage is performed. In this stage, a morphological filtering is used to remove noise and speckles form the image. Morphological filtering allows you to filter out noise and insignificant objects from the texture. Morphological image processing is a set of tools for analyzing and processing the structural features of images based on set theory. These methods allow you to highlight and enhance the spatial characteristics of objects in images, which makes 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 on the boundaries of objects. This removes insignificant noise:

$$(A \ominus \mathrm{B})(i,j) = \min_{x,y \in B} A(i+x, j+y), \quad (3)$$

where A – is the original histogram, B – is a structural element. After this, an extension stage is performed. During this stage, the objects are diluted, and effectively extended to their original size by adding extra pixels on the edge of the bright objects as follows:

$$(A \oplus \mathrm{B})(i,j) = \max_{x,y \in B} A(i-x, j-y), \quad (4)$$

where A – is the original histogram, B – is the structural element.

Morfological filter is created by composing erosion and dilation functions, which removes noise from the histogram, which makes object's boundaries sharper:

$$A \circ \mathrm{B} = (A \ominus B) \oplus B. \tag{5}$$

This filter is then applied to the texture histogram generated by applying formula (2) after the data normalization and is then used as a basis for feature extraction and classification. This filter works great for landmine detection as it is fast and creates a great contrast between target objects (landmine) and surrounding pixels, while removing noise around the edges. The disadvantage of using this approach is that in high-density minefields this method can merge several mines that are nearby into a single entity, increasing the area that has to be manually surveyed.”

4.2. Feature extraction and classification

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 images with a small number of channels[29]. Due to the high dimensionality of hyperspectral images, known architectures are poorly suited for processing such images in their original form. However, after pre-processing, the dimensionality and complexity of the data is significantly reduced, which allows to synthesize [30] a simpler architecture [31]. The proposed network consists of the following types of layers:

1. convolution layer is the basic building block of a CNN, where the convolution operation occurs. Filters (kernels) slide over the entire image, calculating the dot product between the filter and a portion of the input image, creating feature maps;
2. pooling layer performs the operation of reducing the dimensionality 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 improves stability and learning speed;
4. dropout layer is used to prevent overfitting by randomly "switching off" some neurons during training;
5. fully connected layer has it’s neurons connected to all neurons of the previous layer, which enables combining features and making a final decision;

ReLu is used as the activation function, and Cross-Categorical Entropy Loss is used to train the neural network. Network architecture is presented in Figure 1.

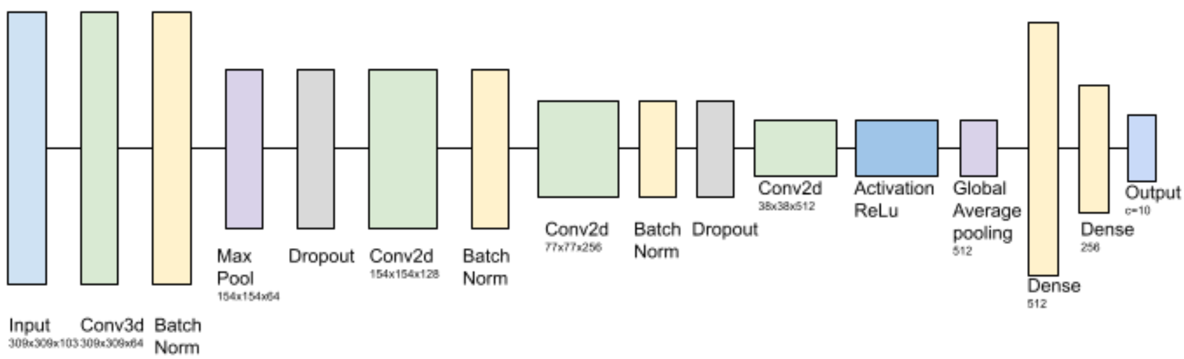

Figure 1: Proposed CNN architecture. The data is first shrinked in the 3d convolutional layer to decrease number of channels, then in is passed through two bottleneck sections, then passed into the classification head

4.3. Method application

The proposed method can be applied in two modes – real-time and batch processing of a large-area hyperspectral cube. The batch mode provides higher overall speed and accuracy due to access to larger computing resources, but requires the agent to completely scan the search area. The real-time mode provides information much faster, but increases energy costs on the part of the agent, which can negatively affect the overall speed of zone processing, since the agent will spend more time recharging. Schematically, batch processing is shown in Fig. 2 (a), and processing in batch mode is shown in Fig. 2 (b). The training phase is common for both approaches. For training, hyperspectral images of the search areas are collected and then labeled. For labeling, both binary classification (mine present – no mine) and multi-class classification can be used to determine a specific mine type. After labeling, the hyperspectral images are normalized, filtered and divided into small sections using sliding window algorithms with a small intersection between each section. After splitting, the dataset is used to train the model using supervised learning.

In batch mode, the agent performs a full scan of the area of interest, obtaining a hyperspectral image of a large area. After that, the data is also normalized, filtered and split into several sections using a sliding window algorithm. Each of the sub-sections is processed using a convolutional neural network. The classification results are combined with the telemetry collected by the agent to create a geoinformation database.

In real-time mode, classification and database creation are performed directly on the agent during the scan. This allows you to analyze a part of the area much faster, but in this case the algorithm itself is much less efficient (since normalization and filtering are performed on slices with a high percentage of intersection). Also, the overall scanning time of the area increases, since the algorithm requires high energy costs, which results in the need for additional recharges.

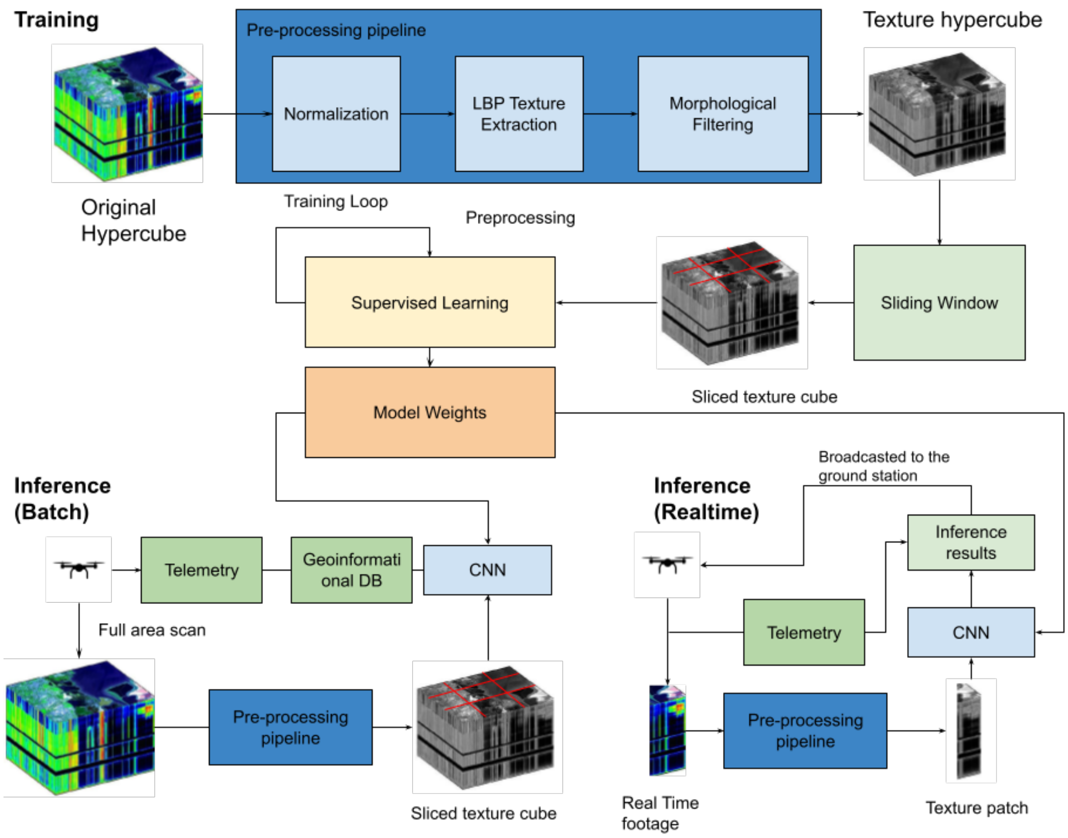

Figure 2: Proposed method – training loop (top section), batch inference mode(a, on the left), real time inference mode

5. Results and discussion

5.1. Results

The algorithm was tested on the Pavia University hyperspectral dataset [27], which was injected with synthetically generated signatures of various mines. To emulate more realistic camouflaged mines, we used a blend operator to weightedly average the spectral signature of the mines with the surrounding environment.

This dataset was collected using an aircraft-mounted hyperspectral sensor and contains images of an urban area near the University of Pavia in Italy.

The dataset has a resolution of 1.3 meters per pixel and has 103 spectral channels ranging from 430 to 860 nanometers. These channels cover the visible and near-infrared spectra, providing high spectral resolution for identifying different materials and objects. The dataset includes 9 classes: asphalt, lawn, road, trees, soil, bitumen, tiles, shadow, buildings. An additional class, mine, was introduced into the dataset in a ratio of 5-95% percent, replacing some of the existing class labels soil, lawn, road, shadow. During training, the dataset is split into a training and validation dataset in a ratio of 80-20%. During training, the Adam optimizer [28] is used, the learning rate is set to 0.01, with a subsequent decrease to 0.001. Recommendations outlined in [31, 32] are used during the classifier training. The batch size is 32. During the training of the neural network, we measure the loss, average accuracy, classifier specificity (false positive rate) and false negative rate. The training lasts for 500 epochs. The training results are presented in Table 2. Confusion matrix is presented in the Table 3. An example of the original and classified hyperspectral image is shown in Fig. 3.

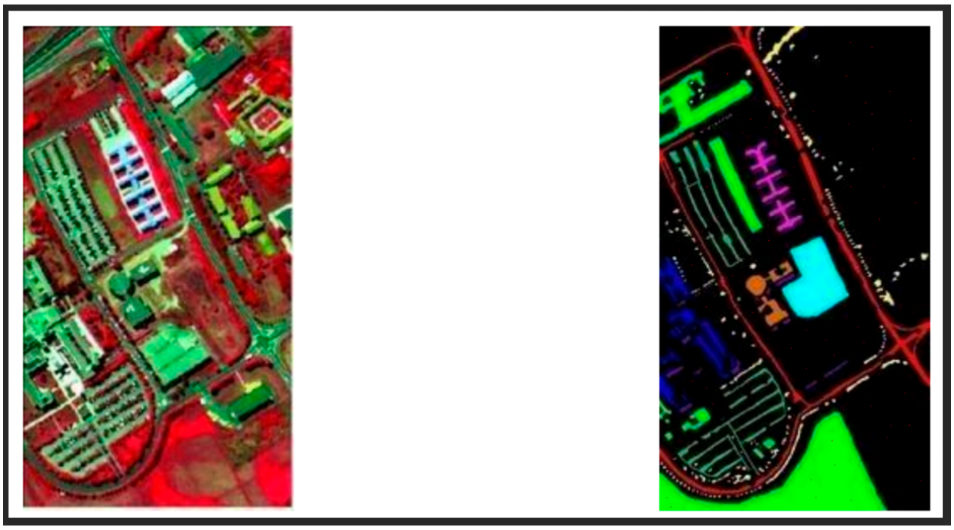

Figure 3: Results of method application in multi-class setting. Red dots represent landmines

Table 2
Model training results

Avg. Accuracy (all classes)	Precision (landmine class only)	Recall (landmine only)	Avg. Inference Time	Avg Preprocessing time (patch)
87.5%	86.33%	100%	67.8 ms (~15FPS)	21.73 ms ~(46FPS)

Table 3
Confusion matrix (for the Landmine class only)

	Positive	Negative
True	120	2261
False	19	0

5.2. Discussion

The results obtained in this experiment are promising. They indicate a viability of applying the proposed method in the real world scenario. The proposed model was able to achieve high throughput in the batch mode, and adequate inference time in real-time mode. Average accuracy of 87.5% is somewhat low, however in this setting the network was trained to classify not only landmines, but other classes as well, which brought down the overall results. Confusion matrix, however, indicate a decent classification accuracy. The most important aspect is recall, as false negative errors are extremely dangerous when dealing with explosive ordnance. The proposed method was able to achieve 100% recall, which means that no landmine was undetected. However, precision metric is 86.33%. While this is not critical on the scale the test was performed (as only 19 false positives would have been checked), real minefields can have over 1000 mines planted, which would drastically slow down mine removal process.

While results are promising, it is important to note that used dataset has severe limitations. The key limitation is the granularity of source dataset, which somewhat skews the precision and recall metrics. Higher-granularity dataset is unlikely to impact recall metric as long as decision boundary is set reasonably, however will greatly increase false positive rate near the edges of the landmines.

Another major limitation of the experiment is that synthetic dataset does not give realistic recall metrics, as they can depend on many factors, such as the type of soil, time of day, ambient temperature, landmine depth, landmine material, atmospheric pressure. A purpose-built dataset which can control these parameters are required to validate the experimental results before the method can be applied in real-world landmine removal operations.

6. Conclusion

This paper presents a hybrid method for detecting mines in hyperspectral images that combines the strengths of various existing methods. The proposed method uses an integrated approach to preprocessing and filtering the obtained data, which significantly simplifies the feature extraction and classification step. The preprocessing step made it possible to significantly simplify the architecture of the convolutional neural network of the classifier, which has a positive effect on the accuracy and speed of both training and inference. The proposed method operates both in batch mode and in real time (at ~13 FPS), which expands the practical possibilities of its application.

The main limitation of this work is the use of a sensor with a resolution of 1.3 meters per pixel, which is not sufficient to accurately detect the location of mines. Also, in the synthetic dataset, the noise was not so pronounced due to the fact that the mines were introduced using the blend operator.

Further research can be developed in several directions. Firstly, obtaining a real landmine dataset with a higher resolution will allow testing the proposed algorithm in conditions close to real world. In this paper, we considered a multi-class classification problem where each grid segment could belong to 10 classes, but only 1 of them could be a mine. A more interesting scenario is the absence of irrelevant classes with several different types of mines. In this scenario, the sensitivity of the classifier will be overstimulated, which will lead to a high level of type I errors. Also, the classifier created using the current method has poor generalization ability with respect to the soil in which the mines are buried. When using this method in a location different from the one in which the training dataset was collected, a drop in accuracy is expected. Improving the generalization ability of the classifier without losing accuracy and increasing sensitivity to false positives is an open problem.

Based on the factors, outlined above, future research will be concentrated on acquiring a real landmine dataset to validate and calibrate the existing approach. By introducing multiple types of landmines, it will be possible to properly calibrate the decision thresholds for the classifier to decrease it’s sensitivity to noise and increase precision.

References

[1] Reynolds, P. Preston. Landmines and Sustainability: Remaking the world through global citizenship, activism, research and collaborative mine action. Development in Crisis. Routledge, 2015. 204-222.

[2] J. E. McFee ; K. L. Russell and M. R. Ito, Detection of surface-laid minefields using a hierarchical image processing algorithm, Proc. SPIE 1567, Applications of Digital Image Processing XIV, 42 (December 1, 1991); doi:10.1117/12.50802

[3] L. B. Wolff ; D. A. Socolinsky ; C. K. Eveland ; J. I. Yalcin and J. H. Holloway, Jr. Image fusion of shortwave infrared (SWIR) and visible for detection of mines, obstacles, and camouflage, Proc. SPIE 5089, Detection and Remediation Technologies for Mines and Minelike Targets VIII, 1298 (September 15, 2003).

[4] E.M. Winter, M. J. Schlangen, A.P. Bowman, M. R. Carter, C. L. Bennett, D. J. Fields, W. D. Aimonetti et al. Experiments to support the development of techniques for hyperspectral mine detection. In Aerospace/Defense Sensing and Controls, pp. 139-148. International Society for Optics and Photonics, 1996.

[5] J. E. McFee, H. T. Ripley, R. Buxton, A. M. Thriscutt, Preliminary study of detection of buried landmines using a programmable hyperspectral imager, Proc. SPIE 2765, Detection and Remediation Technologies for Mines and Minelike Targets, 476 (May 31, 1996); doi:10.1117/12.241250

[6] J. E. McFee and H. T. Ripley, Detection of buried land mines using a CASI hyperspectral imager, Proc. SPIE 3079, Detection and Remediation Technologies for Mines and Minelike Targets II, 738 (July 22, 1997)

[7] J. E. McFee, S. Achal, T. Ivanco and C. Anger, A short wave infrared hyperspectral imager for landmine detection, Proc. SPIE 5794, Detection and Remediation Technologies for Mines and Minelike Targets X, 56 (July 08, 2005);

[8] M. T. Eismann, C. R. Schwartz, J. N. Cederquist, J.A. Hackwell, and R. J. Huppi. Comparison of infrared imaging hyperspectral sensors for military target detection Applications; In SPIE#39; 1996 International Symposium on Optical Science, Engineering, and Instrumentation, pp. 91-101. International Society for Optics and Photonics, 1996.

[9] N. Playle, Detection of landmines using hyperspectral imaging. In Defense and Security Symposium, pp. 62170A-62170A. International Society for Optics and Photonics, 2006.

[10] G. Suganthi, R. Korah, Discrimination of Mine-Like Objects in Infrared Images Using Artificial Neural Network, Indian Journal Of Applied Research, Volume : 4 Issue : 12 , India, p 206-208, Dec 2014

[11] M.-A. Gagnon, P. Lagueux, J.-P. Gagnon, S. Savary, P. Tremblay, V. Farley, É. Guyot and M. Chamberland. Airborne thermal infrared hyperspectral imaging of buried objects. In SPIE Security+ Defence, pp. 96490T-96490T. International Society for Optics and Photonics, 2015.

[12] J. M Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot. Hyperspectral remote sensing data analysis and future challenges. IEEE Geoscience and Remote Sensing Magazine 1, no. 2 (2013): 6-36.

[13] V. A. Kotkar , S. S. Gharde, Review Of Various Image Contrast Enhancement Techniques, International Journal of Innovative Research in Science, Engineering and Technology Vol. 2, Issue 7, July 2013.

[14] C. P.Lee, Mine detection techniques using multiple sensors. The Project in Lieu of Thesis, Electrical and Computer Engineering The University of Tennessee at Knoxville (2000).

[15] J-M Gaucel, M. Guillaume and S. Bourennane. Adaptive-3D-Wiener for hyperspectral image restoration: Influence on detection strategy. Signal Processing Conference, 2006 14th European. IEEE, 2006.

[16] P.M. Mather and Magaly Koch. Computer processing of remotely-sensed images: an introduction. John Wiley & Sons, 2011.

[17] P.-H. Hsu, Feature extraction of hyperspectral images using wavelet and matching pursuit. ISPRS Journal of Photogrammetry and Remote Sensing 62, no. 2 (2007): 78-92.
[18] X. He, D.Cai, S. Yan and H.J. Zhang, Neighborhood preserving embedding. Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. Vol. 2. IEEE, 2005.
[19] J. W. Sammon, A nonlinear mapping for data structure analysis. IEEE Transactions on computers 5 (1969): 401-409.
[20] B.-C. Kuo, and D. A. Landgrebe. Nonparametric weighted feature extraction for classification. Geoscience and Remote Sensing, IEEE Transactions on 42, no. 5 (2004): 1096-1105.
[21] R. Ablin, and C. H. Sulochana.A survey of hyperspectral image classification in remote sensing. International Journal of Advanced Research in Computer and Communication Engineering 2.8 (2013): 2986-3000.
[22] Tbarki, K., Said, S. B., Ksantini, R., & Lachiri, Z . RBF kernel based SVM classification for landmine detection and discrimination. In 2016 International Image Processing, Applications and Systems (IPAS) (2016) (pp. 1-6). IEEE.
[23] Ahmed, M., Seraj, R., & Islam, S. M. S. The k-means algorithm: A comprehensive survey and performance evaluation. Electronics, 2020, 9(8), 1295.
[24] Chang, Chein-I. Orthogonal subspace projection (OSP) revisited: A comprehensive study and analysis. IEEE transactions on geoscience and remote sensing 43.3 (2005): 502-518.
[25] Lundberg, Magnus, and Irene YH Gu. 3D matched filter for detection of land mines using spatio-temporal thermal modeling. Detection and Remediation Technologies for Mines and Minelike Targets V. Vol. 4038. SPIE, 2000.
[26] Giannakis, Iraklis, et al. Numerical modelling and neural networks for landmine detection using ground penetrating radar. 2015 8th International workshop on advanced ground penetrating radar (IWAGPR). IEEE, 2015.
[27] Huang, Xin, and Liangpei Zhang. A comparative study of spatial approaches for urban mapping using hyperspectral ROSIS images over Pavia City, northern Italy. International Journal of Remote Sensing 30.12 (2009): 3205-3221.
[28] Zhang, Zijun. "Improved adam optimizer for deep neural networks." 2018 IEEE/ACM 26th international symposium on quality of service (IWQoS). Ieee, 2018.
[29] Sineglazov , V., & Kot, A. (2021). Design of hybrid neural networks of the ensemble structure. Eastern-European Journal of Enterprise Technologies, 1(4 (109), 31–45. https://doi.org/10.15587/1729-4061.2021.225301
[30] Zgurovsky, M., Sineglazov, V., Chumachenko, E. (2021). Formation of Hybrid Artificial Neural Networks Topologies. In: Artificial Intelligence Systems Based on Hybrid Neural Networks. Studies in Computational Intelligence, vol 904. Springer, Cham. https://doi.org/10.1007/978-3-030-48453-8_3
[31] Zgurovsky, M., Sineglazov, V., Chumachenko, E. (2021). Classification and Analysis Topologies Known Artificial Neurons and Neural Networks. In: Artificial Intelligence Systems Based on Hybrid Neural Networks. Studies in Computational Intelligence, vol 904. Springer, Cham. https://doi.org/10.1007/978-3-030-48453-8_1
[32] Zgurovsky, M., Sineglazov, V., Chumachenko, E. (2021). Classification and Analysis of Multicriteria Optimization Methods. In: Artificial Intelligence Systems Based on Hybrid Neural Networks. Studies in Computational Intelligence, vol 904. Springer, Cham. https://doi.org/10.1007/978-3-030-48453-8_2
[33] V.M. Sineglazov, K.D. Riazanovskiy, O.I. Chumachenko, (2020). Multicriteria conditional optimization based on genetic algorithms. In: System research and information technologies, No. 3 (2020) . DOI: https://doi.org/10.20535/SRIT.2308-8893.2020.3.07