An Efficient Detection and Classification Method for Landmine Types Based on IR Images Using Neural Network

A. Mahmoud and H. Farouk

Abstract:---Infrared Image characteristics have some interesting capabilities that may assist in the detection of shallowly buried objects, in particular to help in the identification of landmine contaminated areas. This paper presents some preliminary results for the detection of buried Anti-Personnel Landmines (APLs) using an infrared imaging system. We describe an algorithm for the detection of landmine candidates by exploiting features in the images after extracting the object from background. Different threshold levels are applied to select groups of pixels that correspond to the object, and are the ones that could indicate a target position to the produced binary images. Unsupervised Self Organizing Map neural network was employed to differentiate among the land mines for better choice of the suitable removal method. Our test results approved more than 98% detection accuracy.

Keywords:--APL, Feature extraction, Neural network, IR images

I. Introduction

MORE than 26,000 people are killed or maimed by mines every year, which is equivalent to one victim, every 20 min. For example, in Cambodia one out of every 236 people is a landmine amputee. The casualty ratio rises to one out of every 140 people in Angola, which has more mines than people in addition to fatal casualties and enormous financial losses. In Cambodia, approximately 40% of the rice fields have been mined and abandoned. Most tragic is that many victims are children and most mine-afflicted countries are poverty stricken, as well.

Because of the potentially catastrophic results of unintentional mine encounters, the process of detecting and removing mines, called demining, is particularly important. Manual demining is extremely dangerous; around one deminer has been killed for every 2,000 mines removed, with even more civilian victims. The cost to purchase and lay a typical antipersonnel mine ranges from $3 to $30, while the cost to remove a single mine ranges from $300 to $1000. Because mines can be made of both metallic and nonmetallic materials, detection using only conventional metal detectors cannot give a promising result. Report also indicate metal detectors are subject to many false Image Processing-Based. High labor cost and the slow pace involved are encouraging development of other techniques. Although some military demining equipment has been developed and used during the Gulf War by the US Army, civilian related demining, called humanitarian demining, is quite different from the military work.

The object of humanitarian demining is to find and remove abandoned landmines without any hazard to the environment. These landmines were intended for military use when they were planted, but their duty has expired. Furthermore, humanitarian demining equipment is required to be more accurate than military-purpose equipment because the military can afford a certain degree of casualty risk.

Various techniques in the area of sensor physics, signal processing, and robotics have been studied during the last decade. Most mine detection techniques depend on sensors, signal processing, and decision processes. For the sensor part, ground penetration radar (GPR), infrared (IR), and ultrasound (US) sensors are reviewed. For the signal processing and decision parts, a commonly used set of image processing techniques including filtering, enhancement, feature extraction, and segmentation are surveyed. Segmentation is used to extract a mine signal from various competing signals, the mine detection algorithms work with inhomogeneous background.

In section 2 a brief history for mines detection is introduced, in section 3, a complete description for the proposed classification algorithm among some predefined mine types is demonstrated. In section 4 we present and discuss the results. In section 5 the work conclusion will be given.

II. Mine Detection Overview

Signal and Image Processing Techniques

2D information for mine location is the most important feature used for mine detection independent of the sensor used.

The sensor output can be represented in the form of 2D data, which can be considered an image. In the next table will give a comparison between common mines detection methods [1-7].

Many recent algorithm were developed as the geometrical feature-based sensor fusion framework [21], combining GPR and IR, as an effective technique for detection and classification of APM which reduces the false alarm rate significantly. In this algorithm they consider the basic geometrical shape descriptor features of an object and construct a feature vector for each of the objects. These feature vectors are used to train a probabilistic neural network (PNN) for the classification of APMs.

Table 1. Comparison between mines detection methods

Method	Risk Profile
Greatcore GDM1 remote landmine detector	- Detection from up to 30 feet - no injury risk during detection. - Can detect landmines with low metal content, resulting in higher success.
Metal manual landmine detector	- Close proximity to landmines. - False positives of 1000 for every 1 landmine. - Low success with landmines of low metal content.
Use of animals for mine detection	- Time and investment taken to train these animals. - Indeterminate false positives. - Lacking compliance with U.N. standards for humanitarian demining.
Use of plants for landmine detection	- Still experimental with indeterminate false positives. - Issues of ecological control of a new genetically engineered specie. - Question of meeting U.N. standards for humanitarian demining.
Bacteria for landmine detection	- Certain explosive chemicals are yet undetectible. - Question of meeting U.N. standards for humanitarian demining.
Nuclear detection	- Still theoretical.
Unmanned landmine detection vehicles	- Relatively high cost operations. - Logistics of transporting and servicing these vehicles.

The main core of the detection algorithm is the library mine images we have. In this work we will depend on IR library of images for learning the neural network and testing it for detecting and classifying mine. The image ibrary source was Joint Research Centre (JRC) Landmine Signatures Database. It was one of the European Commission (EC) R&D Program on Humanitarian Demining [8]. This library has more than 45 IR images for 6 mine types in different image capture angle.

A. Filtering

Noise is unavoidable in the output of most sensors. Two filtering techniques were given as example for noise removal, the Wiener filter and the alternating sequential filter, were reviewed for the purpose of noise removal. The Wiener filter is also known as the minimum mean square error (MMSE) filter in the image processing area. The alternating sequential filter is based on gray-scale morphology [9-13].

B. Gray-Scale Morphology

Mathematical morphology has been applied primarily for binary image processing. Basic or the first level functions, such as dilation and erosion, are performed by structure elements with various shapes and sizes. Repetition of the basic functions forms the second level functions, such as opening and closing. By appropriately combining those operations, a regionbased processing, such as boundary extraction, region filling, and thinning, can be realized. Gray-scale morphology can provide more complicated processing, such as gradient extraction, contrast enhancement, and regionbased segmentation (watershed algorithm) as well as noise removal and smoothing which are typical applications of binary morphology.

In order to define morphological operators, the structuring element should be defined first. A structuring element can be considered as a simple matrix or a small window that represents a certain local property of the whole image. A structuring element defines the region of support around the origin, and it adds an offset value to each pixel on the defined region of support [24].

C. Feature Extraction

Both GPR and IR sensors produce huge amounts of data for Cscanned image sequences and dynamic thermography, respectively. Combination of two or more different types of sensors results in multiple, heterogeneous data. Extraction of the desired features from the large-scale, heterogeneous data is a daunting task. Orthogonal transformations can serve as a tool for removing redundant data and analyzing the desired property in large-scale data. The discrete cosine transform (DCT) used in image and video compression is one example of this type of transformation, and the discrete Fourier transform or singular value decomposition is another.

In this paper we used MATLAB R2008 including some toolbooks like signal processing and neural network to develop our classification algorithm in the simplest way. In the following subsections we'll briefly describe each function used [14-20].

III. The Proposed Classification Algorithm

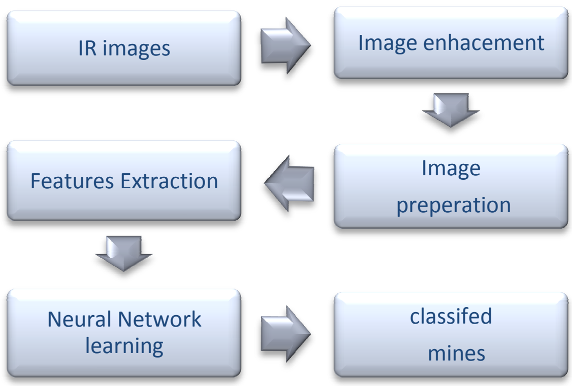


Fig 1. Proposed algorithm block diagram

In the proposed algorithm as shown in the block diagram in Figure 1, we followed the coming main steps:

1. Image Preparation
2. Image Enhancement
3. Feature Extraction
4. Building Neural network model
5. Mines classifications

A. Image preparation

In this phase we did:

i. Image scaling to set all images to 340x340 pixels
ii. Converting the image to Black and White image (BW)

B. Image Enhancements

We applied many function (using MATHLAB) as an image enhancement step before applying the features functions which are:

1. BWareaopen

Morphologically open binary image (remove small objects)

$$BW2 = bwareaopen(BW,P) \quad (1)$$

Where BW2 removes from a binary image all connected components (objects) that have fewer than P pixels, producing another binary image, BW2.

2. Bwtraceboundary

This function will be used to trace object in binary image

$$B = bwtraceboundary(BW,P,fstep,conn) \quad (2)$$

Where B specifies the connectivity to use when tracing the boundary. *P* is a two-element vector specifying the row and column coordinates of the point on the object boundary where you want the tracing to begin. *fstep* is a string specifying the initial search direction for the next object pixel connected to P.

3. Strel

Strel : Create morphological structuring element (STREL)

SE = strel('diamond', R) (3)

It creates a flat, diamond-shaped structuring element, where R specifies the distance from the structuring element origin to the points of the diamond. R must be a nonnegative integer scalar as shown in Figure 2

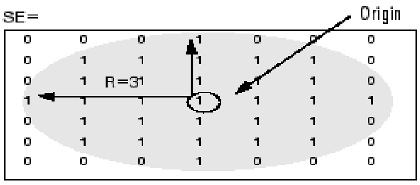


Fig 2. The diamond shape of strel function

4. Imclose

We use imclose to join the circles in the image together by filling in the gaps between them and by smoothing their outer edges.

$$IM2 = imclose(IM,SE) \quad (4)$$

This function performs morphological closing on the grayscale or binary image IM, returning the closed image, IM2 as shown in Figure 3

.

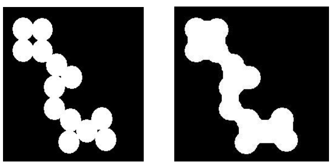


Fig 3. Joined and enjoined circles

5. imfill

Fill image regions and holes

$$BW2 = imfill(BW,'holes') \quad (5)$$

It fills holes in the binary image BW. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

C. Features extractions

As phase 3 we go to feature extraction step

We applied many function on each image and finally we got a vector for features for each image this vector length was 9 These 9 features were:

1. Mean
2. 6 coefficients of 2d DCT
3. Center and Radius
4. Difference

First: Mean

In mathematics and statistics, the arithmetic mean, often referred to as simply the mean or average when the context is clear, is a method to derive the central tendency of a sample space. The term "arithmetic mean" is preferred in mathematics and statistics because it helps distinguish it from other averages such as the geometric and harmonic mean.

Suppose we have sample space $\{x_1, \ldots, x_n\}$. Then the arithmetic mean *A* is defined via the equation

$$A := \frac{1}{n}\sum_{i=1}^{n} x_i \quad (6)$$

Because of the list is an image pixels values, and then the mean of that population is called an Image mean.

Second: 2D DCT

The Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT) helps separate the image into parts (or spectral sub-bands) of differing importance (with respect to the image's visual quality). The DCT is similar to the discrete Fourier transform: it transforms a signal or image from the spatial domain to the frequency domain (Figure 4).

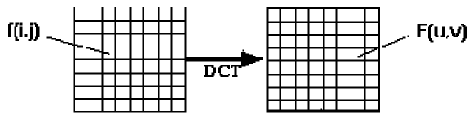


Fig 4. Discrete cosine Transform

DCT Encoding

The general equation for a 2D (*N* by *M* image) DCT is defined by the following equation:

$$F(u,v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1}\sum_{j=0}^{M-1} \Lambda(i).\Lambda(j).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j) \quad (7)$$

The basic operation of the DCT is as follows:

- The input image is N by M;
- f(i,j) is the intensity of the pixel in row i and column j;
- F(u,v) is the DCT coefficient in row k1 and column k2 of the DCT matrix.

Third: The Area, Center and Radius

In this step we calculate the object area based on calculating its center and radius using the enxt MATHLAB function

Regionprops

regionprops can be used to estimate the area enclosed by each of the boundaries

$$\text{STATS} = \text{regionprops}(L, I, \text{centroid}) \quad (8)$$

Where, STATS measures a set of properties (centroid) for each labeled region in the 2-D grayscale image I. *L* is a label matrix that identifies the regions in I and must have the same size as I and L is a two-dimensional array of nonnegative integers that represent contiguous regions. *Centroid* is 1-by-ndims(L) vector that specifies the center of mass of the region

Forth : Difference

diff(X) returns the differences calculated along the first non-singleton. We use this function. We apply this function on the image library we have to get difference between the input image with respect to the image in the test library

C. Building Neural network model

There is a weight plane for each element of the input vector. They are visualizations of the weights that connect each input to each of the neurons. (Darker colors represent larger weights.) If the connection patterns of two inputs are very similar, you can assume that the inputs were highly correlated. The best case when we achieve very different patterns for different classes of the inputs. In our case we classify between 6 categories of mines and tried to use different numbers of neurons to achieve the best results and finally came to the used parameters "input range, training accuracy, biasing, epochs, and the maximum number of iterations.

Training data used (please put specs of the used data "number of samples/each class, number of classes, condition 'orientation, env. Condition, repetition, ',".

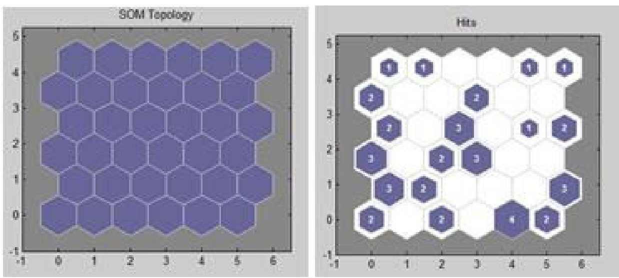


Fig 5. The network presentation as hexagonal network topology formed of 6*6 neurons (a) and resembles how many data points are associated with each neuron (b).

Self-organizing Map networks learn to detect regularities and correlations in their input and adapt their future responses to that input accordingly. Such networks learn to recognize groups of similar input vectors in such a way that neurons physically near each other in the competitive layer respond to similar input vectors and automatically learns to classify them. However the classes that the competitive layer finds are dependent only on the distance between input vectors. If two input vectors are very similar the competitive layer probably puts them in the same class. There is no mechanism in a strictly competitive layer design to say whether or not any two input vectors are in the same class or different classes.

It is best if the data are fairly evenly distributed across the neurons. In this figure, the data are concentrated a little more in the outer neurons, but overall the distribution is fairly even.

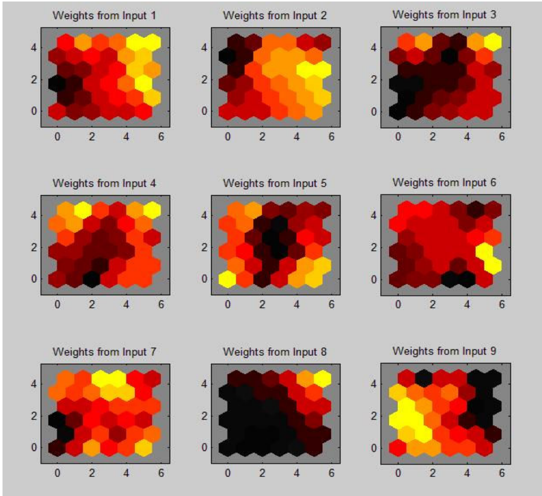


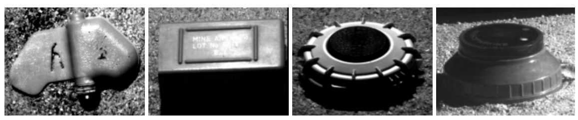
Fig 6. The weight plane figure of the 9 elements of one input vector "as an example" to the network.

The Neural network used is Self Organizing Map composed of two dimensional 6*6 grid one , figure 5, with 9 inputs resembling the number of features of each image, as shown in figure 5. The network was initialized with values ranging from 0 to 1 and 1000 epochs with 0.001 training accuracy and one tuning phase neighborhood distance.

While training the neurons in a competitive layer are distributed to recognize frequently presented input vectors. The architecture for a competitive network is shown in figure 6 where the input vector P and weight matrix IW1,1 are accepted to produce a vector having S1 elements which present the negative of the distances between the input vector and iIW1,1 vectors formed from the rows of the input weight matrix. The net input n1 of a competitive layer is computed by finding the negative distance between input vector P and the weight vectors and adding the biases b. if all biases are zero the maximum net input a neuron can have is 0. This occurs when the input vector P equals that neuron's weight vector.

The competitive transfer function accepts a network input vector for a layer and returns neuron outputs of 0 for all neurons except for the winner which returns output 1, the neuron associated with the most positive element of net input n1. If all biases are 0 then the neuron whose weight vector is closest to the input vector has the least negative net input and therefore wins the competition to output a 1.

D. Mines classifications



(a) (b) (c) (d)

Fig 7. Mines types

We have infrared image library for number of mine types (4 types) listed from figure 7 with different positions to each type according to angle of making photo and different number of positions

IV. Results

After applying image processing and enhancement we described earlier features extracted were extracted and applied as input to the neural network engine. Each image is described by 6 vectors resembling the 6 features forming the input matrix for each image. After training the 6x6 neural network it classified the images into groups according to mines types. In figure 7 we group the neural network output neurons according to their response to each mine type. We achieved 100% matching the mines to images set we have. For the first type we have 16 images (we grouped with red color frame) and we can see the neuron number in side the set framed with red not repeated in any other set and so for the second type – 8 images and their corresponding neurons marked with green.

```
result =

  Columns 1 through 10

    [21]    [16]    [19]    [8]    [15]    [21]    [3]    [1]    [8]    [3]

  Columns 11 through 19

    [16]    [15]    [1]    [19]    [21]    [16]    [24]    [24]    [6]

  Columns 20 through 29

    [28]    [25]    [6]    [28]    [25]    [5]    [13]    [5]    [5]    [7]

  Columns 30 through 38

    [12]    [13]    [7]    [5]    [12]    [12]    [13]    [7]    [31]

  Columns 39 through 42

    [32]    [36]    [35]    [23]

>>
```

Fig 8. grouping the results based on mine types

For the third type – 13 images and their corresponding neurons marked with blue. For the forth type – 5 images and their corresponding neurons marked with magenta.

We notice that no neuron weight repeated from group to another which means no conflict in image classification among different mines types.

V. Conclusion

In this paper we developed a system for mines calcification based on IR images using artificial neural network. We used a library of IR images resembling Four different mine categories to train and test the developed algorithm. the results proved excellent recognition accuracy regardless the image size, imaging angle, and the added noise ratio. 98% classification accuracy may be achieved by increasing the training pool size and help increasing the number of classes.

References

[1] http://fragmentedlegs.blogspot.com/2008/06/comparison-of-landmine-detection.html
[2] http://www.humanitarian-demining.org.
[3] http://www.icbl.org/lm/
[4] http://www.mineaction.org/
[5] http://www.icrc.org.
[6] http://www.un.org/cyberschoolbus/banmines/minetypes.asp
[7] http://en.wikipedia.org/wiki/Land_mine
[8] http://apl-database.jrc.it/
[9] J. K. Paik, C. P. Lee, and M. A. Abidi, "Image Processing-Based Mine Detection Techniques Using Multiple Sensors: A Review", Subsurface Sensing Technologies and Applications: An International Journal, 3(3) (2002) 153-202
[10] J. MacDonald, J.R. Lockwood, J. McFee, T. Altshuler, T.Broach, L. Carin, R. Harmon, C. Rappaport, W. Scott and R. Weaver, "Alternatives for Landmine Detection", RAND Science and Technology Policy Institute, Library of Congress (2003).
[11] O. Lopera and N. Milisavljevic, "Prediction of the Effects of Soil and Target Properties on The Antipersonnel Landmine Detection Performance of Ground-Penetrating Radar", Journal of Applied Geophysics, 63 (1) (2007)13.
[12] Goldman, I. Cohen, "Anomaly Detection Based on an Iterative Local Statistics Approach", Signal Processing 84 (2004) 1225–1229.
[13] S. Reese, G. Sukthankar, R. Sukthankar, "An Efficient Recognition Technique for Mine-Like Objects Using Nearest-Neighbor Classification", Proceedings of Undersea Defense Technology Europe, June, (2003).
[14] N. T. Thanh, H. Sahli, and D. N. Hao, "Detection of Buried Landmines Using Infrared Thermography", Proceedings of the Fifth International Conference on Computational Heat and Mass Transfer, Canmore, Canada, June 18–22 (2007) 393.
[15] T. T. Nguyen, D. N. H. Mao, P. Lopez, F. Cremer and H. Sahli, "Thermal Infrared Identification of Buried Landmines", Proc. SPIE, Det. and Rem. Techn. for Mines and Minelike Targets X, Orlando FL, USA, Vol. 5794 (2005).
[16] S. Sjokvist, A. Linderhed, S. Nyberg and M. Uppsall, "Temporal Method for IR Minefield Feature Detection", SPIE Defence and Security Symposium, Detection and Remediation Technologies for Mines and Minelike Targets, Orlando, USA, 12-16 April (2004).
[17] F. Cremer, W. De Jong and K. Schutte, "Processing of Polarimetric Infrared Images for Landmine Detection", 2nd International Workshop on Advanced Ground Penetrating Radar, Delft, the Netherlands, May (2003).
[18] [40] F. Cremer, W. de Jong and K. Schutte, "Fusion of Polarimetric Infrared Features and GPR Features for Landmine Detection", 2nd International Workshop on Advanced Ground Penetrating Radar, Netherlands, May (2003).
[19] M. Nixon and A. Aguado, "Feature Extraction and Image Processing", Linacre House, Jordan Hill, Oxford OX2 8DP (2002).
[20] E. R. Davies, "Machine Vision: Theory, Algorithms, Practicalities", 3rd edition, by Elsevier Inc (2005).
[21] Baikunth Nath, Alauddin Bhuiyan, “A sensor fusion model for the detection and classification of anti-personnel mines”, International Journal of Innovative Computing and Applications (IJICA), Vol. 2, No. 1, 2009

Biographies

Dr. A. Tobal is an Associate Prof. at the Electronics Research Institute, Cairo, Egypt. He received his B. Sc., M. Sc. And Ph. D. from Faculty of Engineering, Cairo University in 1990, 1994 and 1999 respectively.His fields of research include embedded systems implementation, digital signal processing, Biological Neural Network, pattern recognition and Artificial Intelligence. tobal51000@yahoo.com

Dr. Farouk is an Assistant Prof. He joined the Electronics Research Institute, Egypt, in 1993. His fields of research are signal processing, image compression, video processing, video compression, video indexing and retrieval, video on demand, pattern recognition and machine vision.
Dr. Farouk received his Ph.D. at 2001 from Electronics & Communications Dept., Faculty of Engineering, Cairo Univ. and his M.Sc. at 1996 from Electronics & Communications Dept., Faculty of Engineering, Cairo Univ.
heshamali68@hotmail.com