Applying a Neural Network Based Classification Scheme to Inspect Manufactured Assemblies

Patrick J. Sincebaugh

U.S. Army Research Laboratory
Materials Directorate
Aberdeen Proving Ground, MD 21005
psinceb@arl.mil

Abstract

This paper describes how an image classification system integrates a comprehensive image analysis library with a neural network classifier to automate the inspection of manufactured assemblies and solve image pattern recognition problems. The capabilities of popular image processing algorithms and neural network paradigms are discussed. A case study provides details on how an image classification system was implemented to automate the inspection of artillery fuzes.

Introduction

Manual inspection techniques are often time consuming, costly, and prone to errors caused by human fatigue and subjectivity. The manufacturing community has attempted to overcome the inherent problems associated with manual inspection techniques via the development of automated inspection systems. The objective of developing an Automated Visual Inspection Software(AVIS) system was to create a robust, user-friendly tool which can be used to automate the visual inspection of manufactured assemblies. The goal of using the AVIS system is to analyze and process an image scene in order to recognize the scenes content. This is accomplished by integrating a comprehensive image processing library with a neural network classification scheme.

Implementing the AVIS system can be described by 5 phases: problem constraint, image acquisition, image preprocessing, feature extraction, and classification. Constraining the problem is often the most critical step in utilizing the AVIS system. It is very difficult and often impossible to analyze low quality images, therefore it is important to control background scenes, object parameters, and image characteristics. Image acquisition is the process of converting an image into a digitized format which can be further analyzed by a computer. Image preprocessing involves image conditioning and segmenting the image into meaningful regions of interest which can be separately analyzed by the classification scheme. Feature extraction and enhancement involves processing the image such that relevant features are accentuated while redundant and non-pertinent features are suppressed. The classification phase involves assigning a class to an object or image based on the analysis of its extracted features. The classifier learns to assign certain class representations during its training phase. Figure 1 summarizes and provides examples for each phase.

Five Phase Approach to Image Classification

Problem Constraint

The first and most important step in utilizing an image classification system is to properly constrain the problem. It can be very difficult and often impossible to analyze low quality or uncontrolled images. The easiest approach to avoiding poor quality images is to ensure that quality equipment is being used. If this is not economically feasible, then filtering and correlation techniques must be applied to reduce noise in the image. Another helpful step in constraining the problem is to use artificial scenes and lighting when possible. Natural scenes can be difficult to analyze due to changing weather, lighting effects, etc. Controlling the background scene will reduce the complexity of analyzing the objects of interest. Natural lighting typically requires illumination compensation using artificial lighting. Another method which can greatly reduce the complexity of the problem is to fix or limit the perspective of the image source in relation to the scene. A non-fixed perspective leads to object translation (shift in x or y position), rotation (angular displacement), and size variations which can turn a simple monomorph image into a complex polymorph or xenomorph problem.

Image Acquisition

Image acquisition is the process of converting an image into a digitized format which can be further analyzed by a computer. The AVIS system can read images from a file (Tagged Image File Format or Flat Image File Format) or acquire them from external sources such as a video

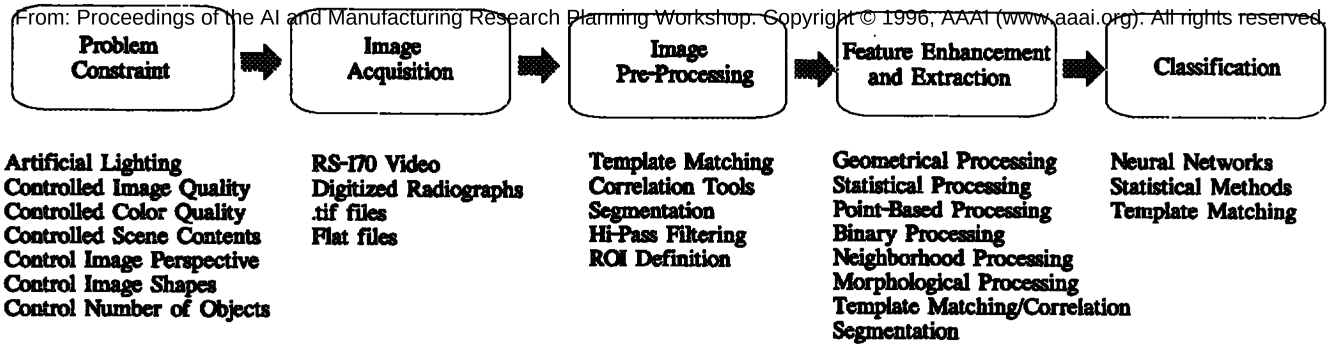


Figure 1. Five Phase Approach to Image Classification

camera or real-time x-ray machine. The AVIS system is configured with a frame grabber board which can be used to acquire monochrome or color images. Irrespective of the source, the digitized images are composed of binary digits representing small picture elements, or pixels. The frame grabber board stores the image in a 512 by 512 pixel array. Monochrome image pixels are composed of 8 bits which represent intensity levels ranging from 0 (dark or black) to 255 (bright or white). Color images are composed of 24 bits which represent the amount of reflectance (intensity), the amount of color (saturation), and spectral color value (hue) of each pixel. Once the image is acquired and stored in the frame grabber buffer, it can be further analyzed by computer invoked algorithms.

Image Pre-Processing

The purpose of image pre-processing is twofold. First, in order to analyze the objects of interest in an image, it is often necessary to perform filtering operations to remove noise and enhance the quality of the image. The amount and type of filtering required is dependent upon how well the problem was constrained. The second pre-processing step involves segmenting the image into meaningful regions of interest (ROI) or chips, which can be separately analyzed by the classification scheme(s). This procedure can greatly reduce the complexity of the problem since the typical visual inspection problem does not require the analysis of the complete image. Two of the most common types of image pre-processing techniques are template matching/correlation and segmentation.

Many image processing problems necessitate the identification of objects within an image. In an ideal setting, the problem would be constrained so that the object was always in the same location within the image. Unfortunately this is not the case in many applications. Template matching algorithms in conjunction with correlation tools can be utilized to locate the object(s) of interest within an image. The template matching algorithm first creates a mask by replicating the object of interest. The mask is then compared to all unknown objects in the image. This is accomplished by scanning the mask across the image, looking for a match which satisfies the correlation threshold that has been specified. This technique is very rarely exact due to inherent image characteristics such as noise. Therefore, a certain tolerance level should be specified.

Image segmentation involves partitioning an image into unique regions based on texture, color, etc. This technique can be applied to locate objects within an image by identifying changes in spatial parameters, energy, and boundary constraints. Image segmentation can be a powerful tool for developing automated systems.

Feature Extraction and Enhancement

The successful application of neural networks to solve pattern recognition problems is often dependent upon the effects of data preprocessing. This dependence can become even more critical when utilizing neural networks to solve image analysis problems.

Preprocessing not only reduces the complexity of a problem, but can also be utilized to enhance and extract relevant features from the ROI. Experience has shown that many image analysis problems can not be solved by analyzing raw pixel values.

The AVIS system provides a comprehensive library of algorithms to support geometrical, statistical, point by point, binary, neighborhood, morphological, template matching/correlation, and segmentation processing.

Geometrical processing techniques manipulate the geometrical properties of an image (i.e. angle, size, viewpoint). Geometrical processing techniques are often used to create additional inputs for Higher Order Neural Networks (HONN). This enables the classifier to be trained more accurately and efficiently by providing a better description of the input vector.

Statistical processing methods are based on the analysis of the image histogram. For example, histogram analysis techniques are based on counting the number of pixels representing each intensity level (i.e. 0 to 255 for gray scale). Many processing and analysis techniques perform better if the gray levels are equally distributed. Histogram equalization is an algorithm which can be used to accomplish this.

Point based processes are operations that manipulate and analyze individual pixel values. Point based operations are often implemented to enhance the regions of an image which possess similar gray scale values, while suppressing the remainder of the image. This is an effective technique for analyzing an object that possesses a different intensity level than the image background.

Binary processing techniques manipulate the individual bits from which the image is composed. The bits are combined with themselves using a logical operator (OR, AND, XOR). This method can be used to create additional inputs for a Higher Order Neural Network. Applying binary processing techniques often provides a network classifier with a stronger model of the input patterns.

Neighborhood processing is based on the principle of adjusting the value of a pixel based upon the values of its neighbors. Neighborhoods are defined as either a 3x3 or 5x5 template formed by the surrounding pixels. The center pixel value is determined by the result of convoluting a filter template with the neighborhood template. This procedure does not allow for the processing of image edges because it is impossible to define a full neighborhood template. Neighborhood processing is often used for noise smoothing and edge enhancement. It is important to use caution when filtering noise, ensuring that pertinent data is not removed from the image.

Morphological processing pertains to the study of shapes. The objective of morphological processing is to extract features of an image by analyzing the image structure. Morphological processing alters the image, resulting in an image from which relevant features may be more easily extracted. The two most commonly used morphological processes are erosion and dilation. Erosion is a procedure where the image shrinks in a uniform manner. This procedure is often used to eliminate small regions of noise. Conversely, dilation is a procedure where the image expands in a uniform manner. Dilation is often used to bridge gaps in contours.

Classification

The AVIS system utilizes neural networks to provide image pattern recognition capabilities. Neural networks were chosen over traditional classification methods because of their inherent robustness for solving different types of pattern recognition problems (Gorman and Sejnowski [1988], Dietz, Kiech, and Ali [1989], and Sincebaugh and Green [1996]). Recent studies have shown that neural networks can often perform equally or better compared to traditional statistical classification methods. Khotanzad and Lu compared the performance of a Multi-Layer Perceptron (MLP) neural network, Bayes algorithm, the nearest neighbor rule, and the weighted minimum-mean-distance rule to solve image analysis problems (Khontanzad and Lu [1990]). In all of the experiments described, the MLP network outperformed the three traditional statistical classifiers.

The AVIS system allows the user to interactively select a neural network paradigm. AVIS currently supports the backpropagation, Delta-Bar-Delta, Extended-Delta-Bar-Delta, and High Order Neural Network (HONN) paradigms.

The backpropagation paradigm is a fully connected, feed forward type of network which implements a supervised learning algorithm (Rumelhart, Hinton, and Williams [1986]). The network architecture is composed of an input layer, one or two hidden layers, and an output layer. Each layer is composed of a number of processing elements, or nodes. Typically each node in a layer is connected to every node in the adjacent layers. Each connection is assigned a weight which represents the strength of the connection. The input vector represents each of the pixel values (or processed results) within the defined region of interest (see figure 2). The network learns how to correctly map the output vector based on an iterative training process. The algorithm uses the Delta rule, whereby the weights associated with each node are adjusted by an amount calculated as a function of the output error and the network learning rate. The network is considered successfully trained when the mean squared error of the output is less than a specified limit.

The Delta-Bar-Delta paradigm is similar to the backpropagation paradigm with the exception that each weight has an individual learning rate which is dynamically updated (Jacobs [1988]). The learning weights are varied based on information of the network error. If the error surface gradient has the same sign for many iterations, the corresponding learning rate is increased. If the error surface gradient changes sign frequently, then the learning rate is decreased. Dynamically changing the learning weights often results in a faster learning rate compared to the backpropagation algorithm.

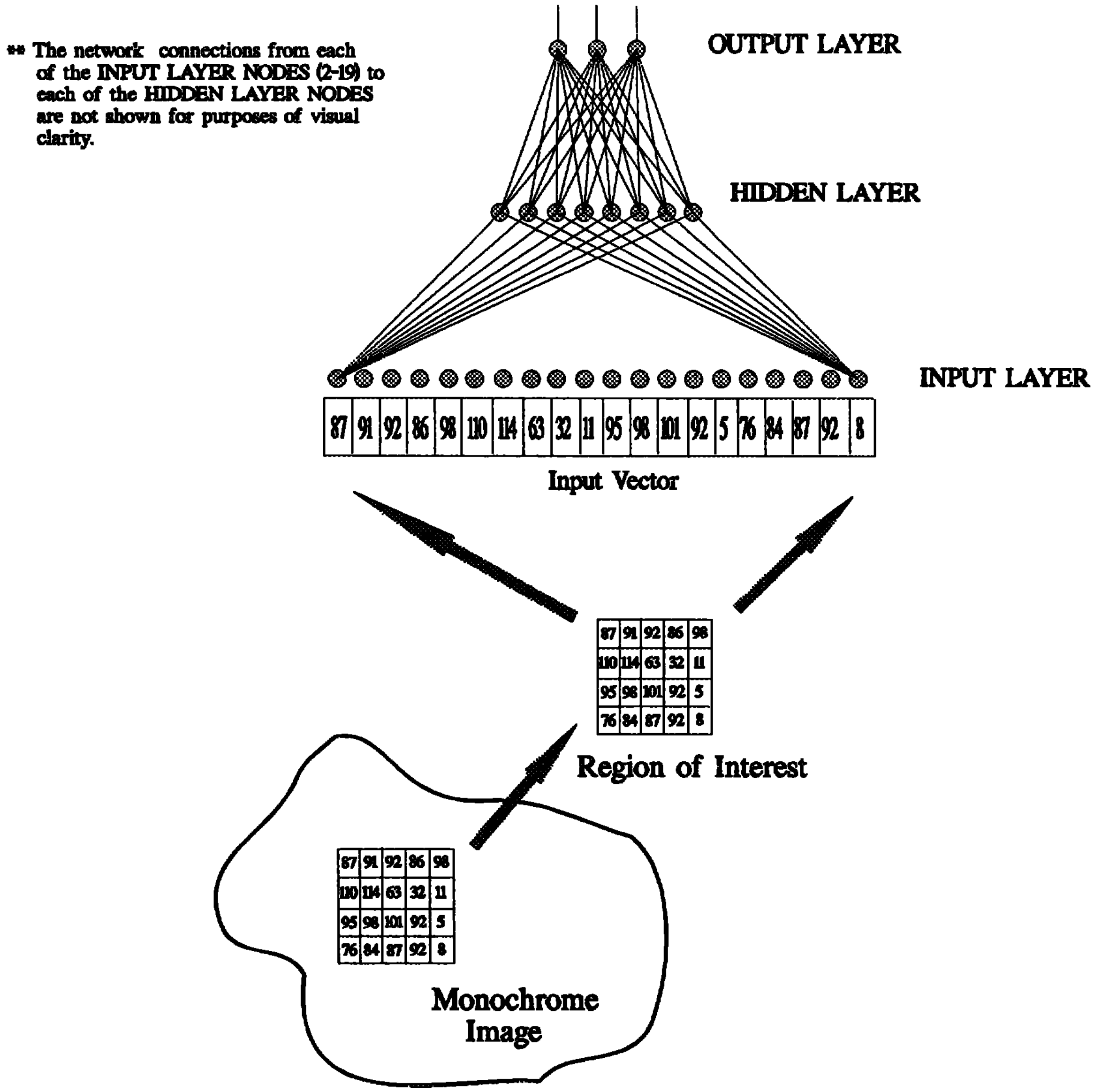


Figure 2. Image Region of Interest Representing Input Vector to Backpropagation Network

The Extended Delta-Bar-Delta algorithm (Minai, and Williams [1990]) is an enhancement to the Delta-Bar-Delta algorithm. The Extended Delta-Bar-Delta algorithm applies an exponential decay to the learning rate increase, adds in a momentum component, and limits the learning rate and momentum. This method often results in a faster learning rate compared to the backpropagation and Delta-Bar-Delta algorithms.

The Higher Order Neural Network (HONN) paradigm, also known as a Functional Link Network (FLN), is another variation of the backpropagation algorithm (Pao [1989]). The patented FLN model utilizes the same supervised learning algorithm as the backpropagation paradigm, but uses an expanded input set to facilitate the solution of non-linear problems. The expanded input set is obtained by processing the base set of inputs and appending the results to the input vector. Two common processing techniques include functional processing and cross product processing. Functional processing can either manipulate the individual elements of the input vector or perform a functional analysis on the vector as a whole. For example, input vector (A,B,C) can be processed as follows:

> *Example1:*
> Taking the sine of each individual element, results in the new input vector (A, B, C, sin(A), sin(B), sin(C)).
>
> *Example 2:*
> Calculating the minimum of the input vector results in the new input vector (A, B, C, minimum(A,B,C)).

Cross product processing consists of calculating the cross product of each element of the input vector. For input vector (A, B, C) this results in the new input vector (A, B, C, AA, AB, AC, BB, BC, CC).

The expansion of the input vector often results in a stronger model of the input patterns, allowing the HONN to be taught more complicated separating surfaces than equivalent backpropagation networks. Second-order and third-order HONN's are equivalent to quadratic and cubic classifiers respectively. While expanding the input vector can increase a networks ability to learn, it can also lead to very computationally intensive problems. Careful thought must be given to implementing this approach because it can often result in huge numbers of input nodes. For example, if a 9 pixel by 9 pixel region of interest is being processed, a traditional first-order neural network will require 81 input nodes. A second-order HONN approach will require 6,561 input nodes, while a third-order approach requires 531,441 input nodes. Since the number of input nodes can expand rapidly based on the HONN approach, it is critical not to expand the input vector more than necessary. Using the smallest possible ROI will also result in smaller input vectors. The AVIS system does not support the hardware implementation of HONN's, therefore only small problems have been solved using this approach. More complex problems have utilized a multilayer approach.

Case Study

Neural networks have been successfully applied to automate the process of inspecting artillery fuzes. Previous artillery fuze inspection techniques included the manual inspection of radiographic images. Each fuze is represented by two radiographs, one depicting an axial view (figure 3), the other a transverse view (figure 4). The inspector was responsible for inspecting 17 critical regions for each fuze. This inspection method was a tedious, time consuming procedure which often resulted in errors that could be attributed to factors inherent to human inspection, such as fatigue and inspector subjectivity. These errors were overcome by the development of an automated test system. The remainder of this section will describe how the procedures for implementing an image classification system described earlier were followed to inspect the actuator of an artillery fuze. The actuator is the black curved band which is located at about 3/4 of the height and transverses from the left side to the right side of the fuze (see figure 4). A common problem with the artillery fuze is that more than one actuator is installed.

The first step in developing the automated fuze inspection system was to properly constrain the problem in order to reduce the complexity of classifying the image. A robotically controlled system was integrated with a x-ray unit to produce high quality digitized radiographs. A jig was developed to ensure that each fuze was aligned in the same position. This eliminated problems associated with object translation. The fuze to x-ray source distance was kept constant, ensuring that the fuze image size remains constant. Limiting the perspective of the image source (x-ray unit) in relationship to the scene (fuze + background) eliminated problems such as object rotation. Also, the fuze background was kept constant, making it easier to detect the fuze object in the overall image.

The next step was to determine the critical regions of the image which needed to be inspected. The regions used during the manual inspection method served as a starting point for the automated inspection procedure. However, pre-processing techniques were utilized to greatly reduce the complexity of the problem. The entire digitized image of the artillery fuze is represented by 512 x 512 pixels. Therefore analysis of the entire fuze would result in a minimum of 262,144 inputs to a neural network. The complexity of this problem can be greatly reduced by selecting a smaller region of interest. One approach would be to analyze only the pixels which represent the actuator. This can be accomplished by defining a ROI consisting of 272 x 78 pixels, resulting in 21,216 network inputs. Although this is a significant reduction in problem complexity, further thought can result in a much simpler approach. A better solution to this problem is to define a ROI which represents a vertical slice through the actuator. The number of black pixels in

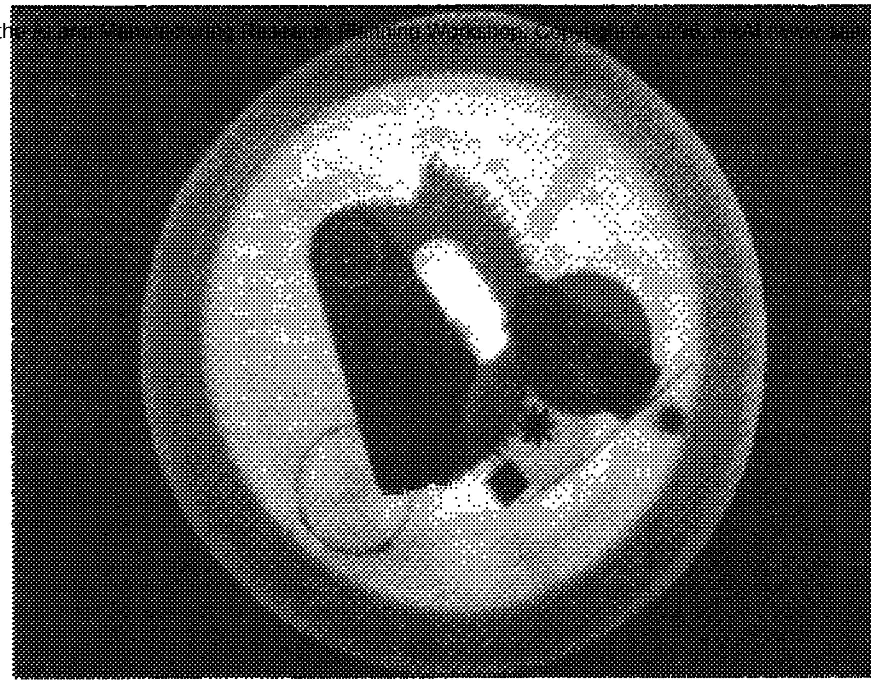


Figure 3. Axial View Radiograph of Artillery Fuze

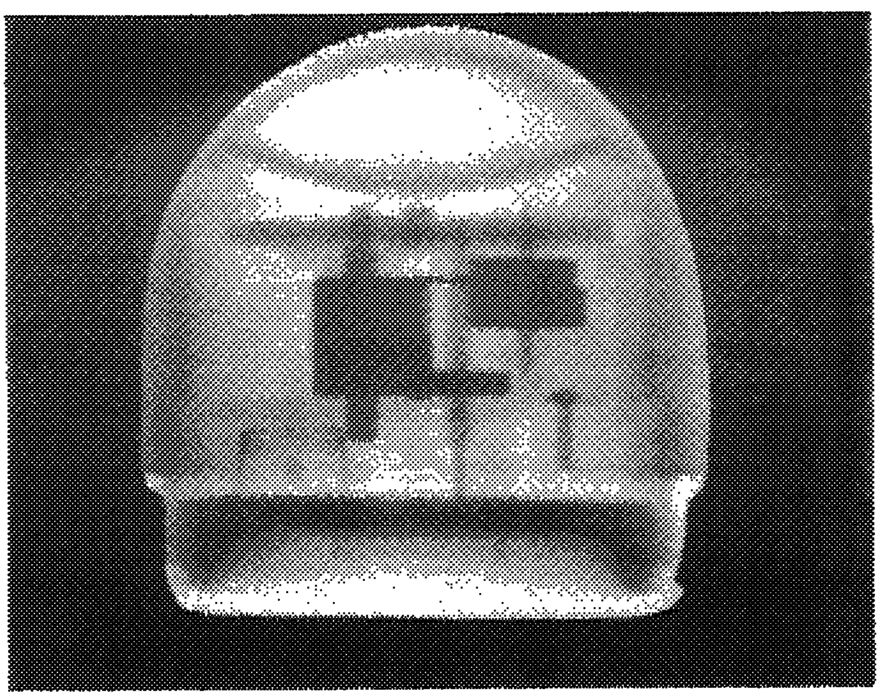


Figure 4. Transverse View Radiograph of Artillery Fuze

the middle of the ROI can be analyzed to determine the number of actuators present and whether the actuator is properly installed. Figure 5 shows a ROI which was defined to solve the actuator inspection problem. The ROI is composed of 6 x 40 pixels, resulting in 240 network inputs. Experience has shown that a well thought out approach to image pre-processing can result in a significantly simpler image inspection problem.

Once the region of interest was identified, further processing was performed to enhance the image and extract pertinent information. The effects of applying different image processing algorithms such as high pass filtering, median filtering, and edge detection were studied. It was concluded that histogram equalization algorithm resulted in the best image representation for inspecting the actuator. Applying the histogram equalization algorithm enhanced the contrast between the light and dark pixel intensities making it easier to distinguish the actuator. This method of representing the image data significantly reduced the training time for the neural network classifier. Figure 6 shows the difference between the unprocessed actuator ROI's and the ROI's which have been processed using the histogram equalization process. The first 7 ROI's represent fuzes with two actuators installed. The remaining seven ROI's represent nominal fuzes. An Extended Delta-Bar-Delta algorithm was successfully implemented to classify the condition of the installed actuator for the artillery fuze. The network architecture consisted of one input layer, one hidden layer, and one output layer. Using the processed ROI's as input to the network resulted in 240 nodes, each representing a pixel intensity. The hidden layer was composed of 8 nodes and the output layer consisted of one node. An output value of zero corresponded to a nominal fuze, while and output value of one corresponded to a defective fuze. The network was trained until the Mean Squared Error between the desired output and the actual output was less than 0.10. The trained network correctly classified the condition of the artillery fuze with over 99% certainty.

Conclusions

Automation can alleviate many of the problems inherent to manual inspection techniques, such as human fatigue and subjectivity. This paper described techniques to automate visual inspection problems by integrating image processing algorithms with neural network classifiers. The capabilities of popular image processing algorithms and neural network paradigms were discussed. A 5 phase approach to solving image classification problems was presented. Details on the Problem Constraint, Image Acquisition, Pre-Processing, Feature Extraction and Enhancement, and Classification phases were provided. Details on how the 5 phase approach was implemented to automate the inspection of artillery fuzes were also provided. The image classification system was trained to inspect the artillery fuzes with almost 100% accuracy. The use of neural network classifiers has resulted in the development of a robust tool which can be utilized to solve various visual inspection problems. Future work will include applying the image classification techniques to inspect printed circuit boards. Research will also be conducted to improve the performance of neural networks which can be applied to image classification problems.

Acknowledgments

The author would like to thank the U.S. Army Artificial Intelligence (AI) Center for continuing to support our development of AI-based systems. The author would like to acknowledge the Johns Hopkins University Applied Physics Laboratory for their efforts inspecting M549 artillery fuzes.

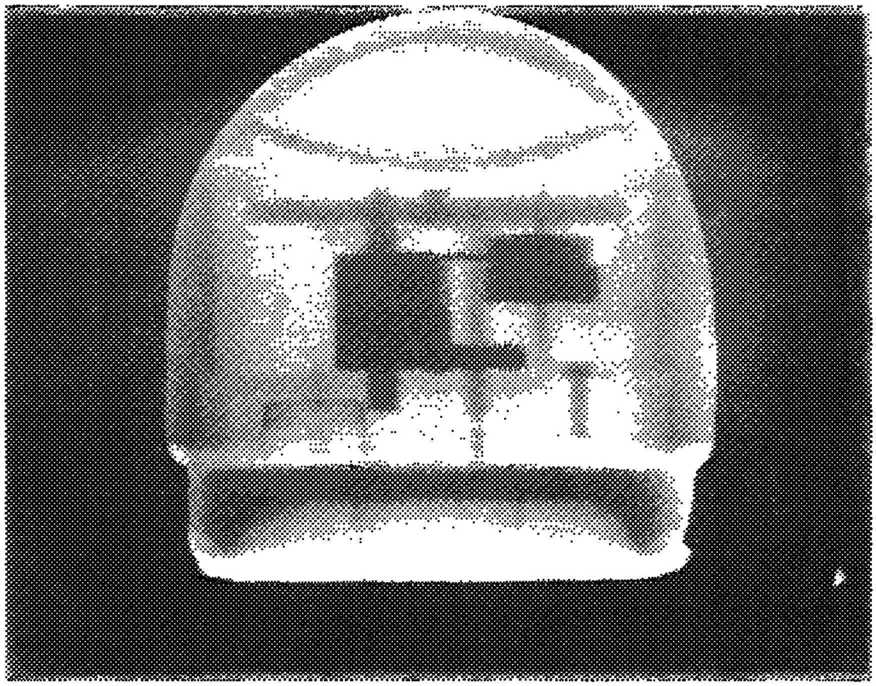
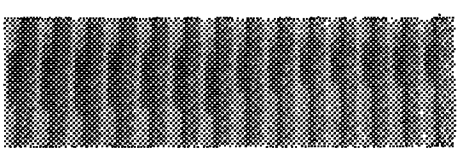
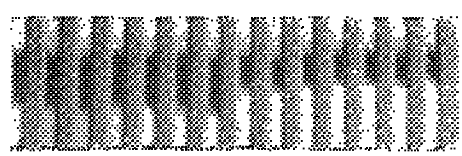


Figure 5. Defined Region of Interest for Inspecting Artillery Fuze Actuators



Unprocessed ROI's



Processed ROI's using Histogram Equalization Algorithm

Figure 6. Unprocessed and Processed Regions of Interest for Actuator Inspection

References

Dietz, W.; Kiech, E.; and Ali, M. (1989). Jet and Rocket Engine Fault Diagnosis in Real Time, *Journal of Neural Network Computing Technology, Design and Applications, Vol. I*, pp. 5-18.

Gorman, R. P.; Sejnowski, T.J. (1988). Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets, *Neural Networks, Vol. I*, pp. 75-89.

Jacobs, Robert A. (1988). Increased Rates of Convergence Through Learning Rate Adaptation, *Neural Networks Vol. I.*

Khotanzad, A.; Lu, J. (1990). Classification of Invariant Image Representations Using a Neural Network, *IEEE Transactions on Acoustics, Speech, and Signal Processing Vol. 38 No. 6*, pp. 1028 - 1038.

Minai, Ali, A.; Williams, Ronald, D. (1990). Acceleration of Back-Propagation Through Learning Rate and Momentum Adaptation, IJCNN-90-WASH DC Proceedings, International Neural Network Society and Institute of Electrical and Electronics Engineers, Inc. Lawrence Erlbaum Associates.

Pao, Yoh-Han (1989). *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley.

Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. (1986). Learning Internal Representations by Error Propagation, *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*. pp. 318-362. Cambridge: MIT Press.

Sincebaugh, P.; Green, W. (1996). A Neural Network Based Diagnostic Test System for Armored Vehicle Shock Absorbers,