Research Journal of Applied Sciences, Engineering and Technology 2(3): 295-301, 2010
ISSN: 2040-7467

Submitted Date: March 26, 2010 Accepted Date: April 16, 2010 Published Date: May 10, 2010

Recognition of Welding Defects in Radiographic Images by Using Support Vector Machine Classifier

[1]Xin Wang, [2]Brian Stephen Wong and [1]ChingSeong Tan
[1]Department of Mechatronics and Biomedical Engineering, Faculty of Engineering and Science, University of Tunku Abdul Rahman (UTAR), Kuala Lumpur, 53300, Malaysia
[2]School of Mechanical and Aerospace Engineering, Nanyang Technological University (NTU), 50 Nanyang Avenue, 639798, Singapore

Abstract: Radiographic testing method is often used for detecting defects as a non-destructive testing method. In this paper, an automatic computer-aided detection system based on Support Vector Machine (SVM) was implemented to detect welding defects in radiographic images. After extracting potential defects, two group features: texture features and morphological features are extracted. Afterwards SVM criteria and receiver operating characteristic curves are used to select features. Then Top 16 best features are used as inputs to a designed SVM classifier. The behavior of the proposed classification method is compared with various other classification techniques: k-means, linear discriminant, k-nearest neighbor classifiers and feed forward neural network. The results show the efficiency proposed method based on the support vector machine.

Key words: Image processing, radiographic testing, support vector machine, welding defects,

INTRODUCTION

Radiographic Testing (RT) is one of the most important nondestructive testing techniques for welding inspection. It is based on the ability of X-rays or gamma rays to pass through metal and other materials opaque to ordinary light, and produce photographic records by the transmitted radiant energy (Hayes, 1997). Because different materials absorb either X-ray or gamma rays to different extents, penetrated rays show variations in intensity on the receiving films. RT can examine the internal structure of a weld. Traditionally, experienced interpreters evaluate the weld quality based on radiography. It is time and manpower consuming work. In addition, human interpretation of weld quality based on film radiography is very subjective, inconsistent and sometimes biased. Therefore, it is desirable to develop a computer-aided system to assist interpretation of radiographic images to increase the objectivity, accuracy and efficiency of radiographic inspection.

Currently there is a great deal of work and research on the development of automated systems for inspection and analysis of radiographs. In our view, the major steps of an automatic detection system are the film digitization stage, pre-processing of images, and identification of defects. These developments rely mostly on techniques such as image processing, feature extraction, and pattern recognition. The pattern classification stage is one of the most important steps in the implementation of an automated radiographic inspection system. Silva *et al.* (2001, 2002) applied a linear classifier to recognize defects. Later Romeu *et al.* (2004) used nonlinear classifiers by implemented neural networks to detect defects. Mery and Berti (2003) proposed to use a statistical classifier to recognize defects. Liao (2004) proposed another approach using fuzzy clustering methods.

SVMs are a relatively new generation of techniques for classification problems. The process of learning involves identification of a model based on the training data that are used to make predictions about unknown data sets. SVM was first introduced by Boster *et al.* (1992) and discussed in more detail by Vapnik (1995, 1998). There are various reasons for preferring SVM to other classification methods. First of all, the most distinguishing property of SVM is that it minimizes the structural risk, given as the probability of misclassifying previously unseen data. Typical pattern classification methods tend to minimize the empirical risk, which is given as the probability of misclassification errors on the training set. More specifically, Vapnik and Chervonenkis (VC) proposed a well studied theory that places reliable bounds on the generalization of linear classifiers and therefore indicate the control parameters on the complexity of linear functions in kernel spaces (Vapnik, 1995). The VC dimension of a function is defined as the maximum number of points that can be shattered by that function. It has been shown that once the VC dimension

Corresponding Author: Xin Wang, Department of Mechatronics and Biomedical Engineering, Faculty of Engineering and Science, University of Tunku Abdul Rahman (UTAR), Kuala Lumpur, 53300 Malaysia

of the family of decision surfaces is known, so is the upper bound for the probability of misclassification for test data of any possible probability distribution (Vapnik, 1995). Second, SVMs pack all the relevant information in the training set into a small number of support vectors. Since the hypersurface is computed with the information from the supports vectors only, the computational efficiency of the classification of a test case is increased by the ratio of the number of data set points over the number of support vectors.

To date, there are no satisfactory results that allow the detection of the weld discontinuities without false alarms. The objective of this study is to develop a system based on SVM to improve the process of an automated welding detection system.

MATERIALS AND METHODS

This study was conducted in Nanyang Technological University from 2004 to 2006. The algorithm was developed using MATLAB.

The automated welding defect detection system is realized using three major techniques: preprocessing the radiographic images, feature extraction and pattern classification. After preprocessing, the 'hypothetical defects' are segmented. However, only some of them are defects and the others are false alarms. Subsequently, the feature extraction is centered principally on the measurement of properties of the regions. Finally, classification divides segmented regions into specific regions according to extracted features, assigning each region to one of a number of pre-established groups, which represent all possible types of regions expected in the image. In the detection problem, the classes that exist are only two: 'defects' or 'non-defects', whereas the recognition of the type of the defects (for example, porosity, slag, crack, lack of penetration, etc.) is known as classification of flaw types. In this paper, a classification system based on statistical learning theory, called the Support Vector Machine (SVM) is applied to the defect detection.

Digitalization of radiographic film: In this study, films are digitized using a laser digitizer with photomultipliers, manufactured by Computerised Information Technology Ltd, UK. The scanner provides an optical density range from 0 to 4.1. The films are digitized with a spatial resolution of 100 and a gray level resolution of 8 bit per pixel.

Preprocessing of the images: The radiographic image has low contrast and high noise and the image background is not uniform. Noise on digitized radiographic images usually appears as randomly dispersed pixels having different values of intensity in relation to their neighbors (Kehoe and Parker, 1990). Adaptive wavelet thresholding and adaptive histogram equalization techniques are used to remove the noise and improve the contrast of the radiographic image (Wang and Wong, 2005). Then the radiographic image is segmented using multi-level thresholding based on maximum fuzzy entropy (Wang and Wong, 2005) as shown in Fig. 1.

Feature extraction: Image features are the most basic characteristics used to distinguish an image from others. Most researchers extract the morphological features associated with the shape and size of the defects. Mery and Berti (2003) proposed to extract texture features. In this study, both morphological features and texture features are used to build a set of pattern classifier data inputs. The complete set of features is shown in Table 1. For each potential discontinuity, 87 features are extracted.

Texture-based features: For our problem, the description of texture of the defect in the radiographic image, two texture features: the co-occurrence matrix and Gabor filter are extracted. These features give a measurement of how often one grey value will appear in a specified spatial relationship to another grey value on the image.

Haralick *et al.* (1973) proposed 14 measures of textural features, which are derived from the co-occurrence matrices. Those used in this work for extracting features in the defect detection of radiographic images are:

Shannon Entropy: $f_E = -\sum_i \sum_j p(i,j) \log p(i,j)$

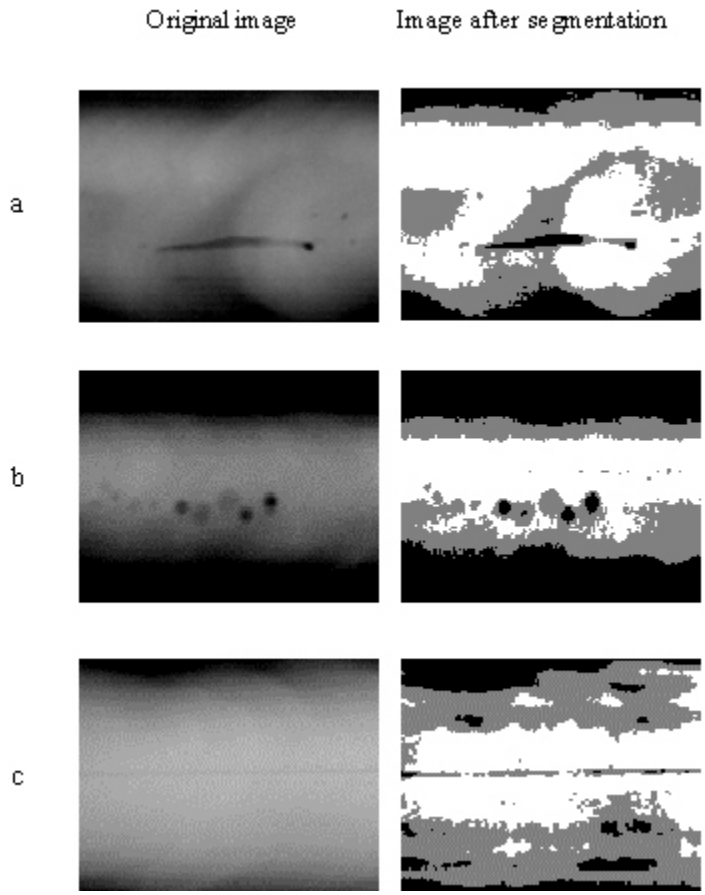
Contrast $f_C = -\sum_i \sum_j (i-j)^2 p(i,j)$

Angular Second Moment: $f_A = -\sum_i \sum_j (p(i,j))^2$

Inverse Difference Moment:

$$f_I = -\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$$

In Eq. (1-4), p(i, j) refers to the normalized entry of the cooccurrence matrices. That is $p(i, j) = P_d(i, j)/R$, where R is the total number of pixel pairs (i, j). For a displacement vector $d = (dx, dy)$ and image of size $M \times N$, R is given by $(M - dx)(N - dy)$. In this work, the texture features are extracted for four directions (0, 45, 90 and 135°) for distance $d = 1$.



(a) Slag inclusions (b) Porosity (c) Incomplete penetration
Fig. 1: Image after preprocessing

Table 1: Extracted features for defect detection

Feature No.	Feature description
From *f*1 to *f*64	Gabor features
From *f*65 to *f*80	Co-occurrence matrix features
From *f*81 to *f*87	Morphological features

The Gabor functions are Gaussian-shaped bandpass filters, with dyadic treatment of the radial spatial frequency range and multiple orientations, which represent an appropriate choice for tasks requiring simultaneous measurement in both space and frequency domains. The Gabor functions are a complete (but a nonorthogonal) basis set given by:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x2}{\sigma_x^2}+\frac{y2}{\sigma_y^2}\right)\right]\exp(2\pi jU_0x) \quad (5)$$

Where σ_x and σ_y denote the Gaussian envelope along the *x* and *y* axes, and u_0 defines the radial frequency of the Gabor function.

The Gabor features, denoted by g_{mn}, are defined as the average output of G_{mn}, ie, it yields $K \times S$ features for each segmented window:

$$g_{mn} = \frac{1}{P_W q_W}\sum_{i=1}^{P_W}\sum_{j=1}^{q_W} G_{mn}(i,j) \quad (6)$$

Where the filtered windows,

$$G_{mn} = \left\{\left[f_{mn}(x,y)_r * W(x,y)\right]^2 + \left[f_{mn}(x,y)_i * W(x,y)^2\right]\right\}^{1/2}$$

The size of G_{mn} is $p_w \times q_w$.

Morphological features: This category features descriptors comprised of the area, length and shape factors of potential defects.

Area: Area is defined as the number of pixels interior to or on the potential defect boundary.

Length: Length is the projection of the potential defect along the major axis of the ellipse that has the same normalized second central moments as the potential region.

Width: Width is the projection of the potential defect along the minor axis of the ellipse that has the same normalized second central moments as the potential region.

Elongation: Elongation computes the ratio between the length and width of the potential defect, which takes a valued between 0 and 1.

Orientation: The orientation of the potential defect stretching in the weld is calculated as the angle (in degrees) between the horizontal line (x-axis) and the major axis.

Ratio of width to area (RWA): RWA computes the ratio between the width and the area of the potential defect.

Compactness: This feature measures the object shape that is calculated by

$$compactness \frac{perimeter}{area}$$

Where perimeter is the number of boundary points around the defect area. A circular object has a smaller compactness valued than a non-circular object.

Pattern classification:

Theory of SVM: SVM is based on the structural risk minimization principle from computational learning theory, or better, and on minimization of the misclassification probability of unseen patterns with an unknown probability distribution of data (Vapnik, 1998). In its simplest, linear form, an SVM is a hyperplane that separates the data maximizing the distance between the hyperplane and the closest points of the training set (Fig. 2). Using a Lagrange function, the optimal hyperplane can be represented by a classifier function (Gunn, 1997):

$$f(x) = \mathrm{sgn}\left(\sum_{x_i \in SVs} a_i y_i K(x_i, x) + b \right) \quad (7)$$

$$b = -\frac{1}{2} \sum_{x_i \in SVs} a_i y_i [K(x_r, x_i) + K(x_s, x_i)] \quad (8)$$

where x_i is the ith training example and y_i is the correct output of the SVM for the ith training example. x_r

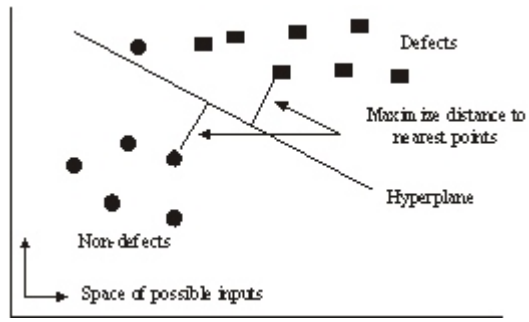


Fig. 2: A linear support vector machine

Table 2: Some kernel functions

Kernel Function	Definition	Parameters
Linear	$x_i . x$	
Gaussian RBF	$\exp(-\frac{\lVert x - x_i \rVert}{2\sigma^2})$	σ is a user defined value
Polynomial of degree d	$(x . x_i+1)^d$	d is a positive integer

is any support vector from positive examples and x_s is any support vector from negative examples. $K(x_i, x_j)$ is a kernel function to construct a mapping into a highdimensional feature space and to enable operations to be performed in the input space rather than the potentially high dimensional feature space. Some of the commonly used kernels include Linear, Gaussian RBF (Radial Basis Functions), polynomial functions, and sigmoid polynomials as shown in Table 2. Usually an RBF kernel is favored, because they are not sensitive to outliers and do not require inputs to have equal variances. So, we choose the Gaussian RBF kernel function. Platt (1998) developed a fast algorithm for training a SVM called Sequential Minimal Optimization (SMO), which made it possible for PC users to practice complex applications. In this paper, we implement this algorithm. The details of the training procedure are described in (Platt, 1998).

SVM based feature selection: In order to improve the prediction performance of the predictors and providing faster and more cost-effective predictors, an SVM-based feature selection algorithm is applied. However, this algorithm relies on a backward feature selection, which is computationally tractable but not necessarily optimal. We improve the performance of the feature selection result by combining SVM-based feature selection with a Receiver Operation Characteristic (ROC) feature selection process.

The SVM is provided with many statistics that allow one to estimate their generalization performance from bounds on the leave-one-out error, L. The leave-one-out error is the number of classification errors produced by the leave-one-out procedure which consists in learning a decision function from m-1 examples, testing the remaining one and repeating until all elements served as

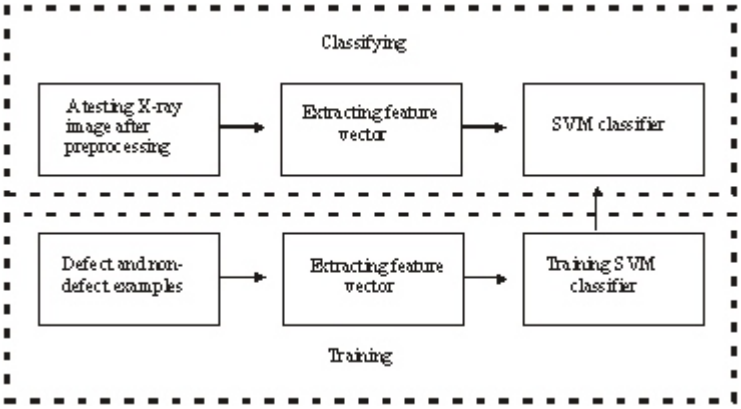


Fig. 3: The procedure of training and classifying

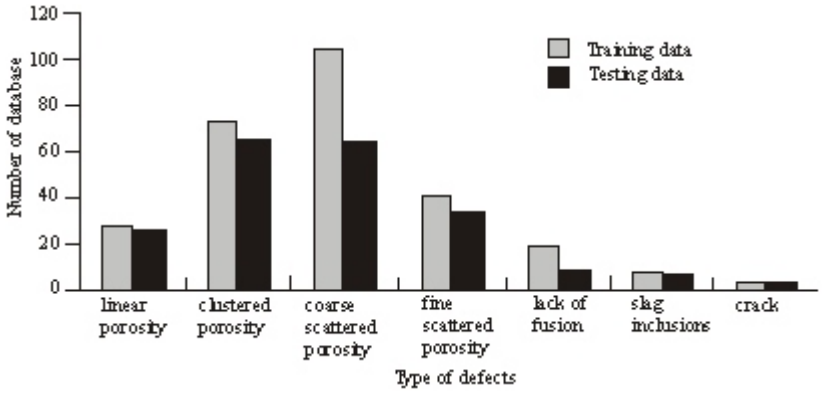


Fig. 4: Types and distribution of defects

a test example. The leave-one-out error is known to be an unbiased estimator of the generalization performance of a classifier trained on *m-1* examples. One of the most common *L* error bounds for SVM is the radius/margin bound (for a decision function with non-zero bias *b*) (Vapnik, 1995). The ROC analysis is commonly used to measure the performance of a two-class classification. In our case, each feature is analyzed independently using a threshold classifier. In this way, a hypothetical flaw is classified as a 'no-defect' (or 'defect') if the value of the feature is below (or above) a threshold value. At last, 16 features are selected by combining the top 12 features selected using SVM and the best 12 features selected using ROC. The complete set of selected features is shown in Table 3.

SVM training and defect detection: The SVM based classification technique consists of two steps: training examples and classifying unknown data. The procedure of training and classifying is shown in Fig. 3. There are two types of examples: positive examples and negative examples. In this case, the positive examples represent defects and the negative examples represent non-defects.

Table 3: Selected features

Feature type	Selected feature No.
Gabor features	$f5, f7, f8;, f12, f28, f12, f55$
Co-occurrence matrix features	$f65, f67, f69, f71, f72, f77, f79$
Morphological features	$f81, f82, f83, f84$

Firstly, we extract 16 features of the examples selected using SVM based feature selection method and train the SVM with these vectors. Then, we preprocess the testing image to extract 16 features, and then apply the trained SVM to decide between the defects and non-defects.

RESULTS AND DISCUSSIONS

In this study, 25 radiographic weld images are analyzed. In the preprocessing stage, 809 potential discontinuities are obtained, of which 473 are real defects. Types and distribution of welding defects is shown in Fig. 4. For each potential discontinuity, 16 features are extracted. 14 images with 475 potential defects are selected and used to train the SVM classifier. Special attention is paid to ensure that instances of all different defects types are included. The remaining 11 images with

Table 4: Performance of different classifier

Classifier	Error	TP	TN	FP	FN	Sn	1-*Sp*	AUC
Support Vector Machine	7.49%	193	116	19	6	96.98%	14.07%	0.94
Linear Discriminant	19.46%	154	112	20	45	77.39%	14.81%	0.88
K-nearest Neighbor	12.28%	170	123	12	29	85.43%	8.89%	0.91
Feed Forward network	14.67%	164	121	14	35	82.41%	10.37%	0.91
K-means	29.34%	113	123	12	86	56.78%	8.89%	0.70

where; TP is the number of true positives (correctly detected defects), TN is the number of true negatives (correctly detected non-defects), FP is the number of false positives (non-defects detected as defects), and FN is the number of false negatives (defects detected as non-defects). , Sn = $\frac{TP}{TP+FN}$, 1-S_p= $\frac{FP}{TN+FP}$ Ideally, Sn = 1 and 1-Sp = 0, this means that all defects are found without any false alarms. AUC is the area under the ROC curve. The AUC is a reasonable performance statistic for classifier systems.

Table 5: Feature selection influence for classification

Classifier	Error	TP	TN	FP	FN	Sn	1-*Sp*	AUC
Feature selection is applied before classification	7.49%	193	116	19	6	96.98%	14.07%	0.94
Feature selection isn applied before classification	19.96%	150	124	11	49	75.38%	8.15%	0.93

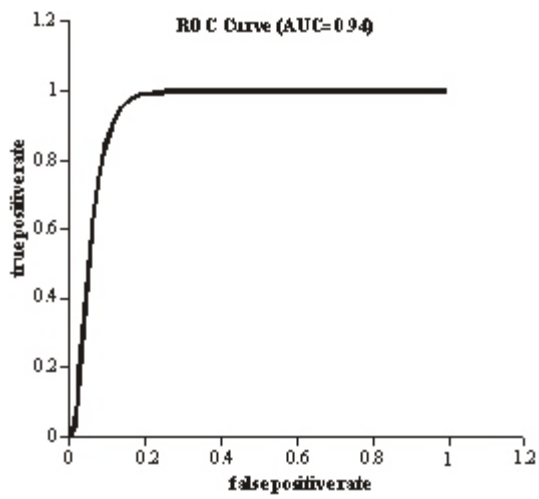


Fig. 5: ROC curve of SVM classifier

334 potential defects are used to test the classifier. Using this method, 96.98% of the existing flaws are detected with 14.07% of false alarms. The ROC curve obtained by the SVM classification with the sixteen selected features is shown in Fig. 5.

In the experiment, an SVM classifier and several types of popular pattern classifiers used for defect detection are implemented on the data set, and the results are compared with each other. The classifiers compared with SVM in this study are a k-means classifier, a linear discriminant classifier, a k-nearest neighbor classifier and a feed forward neural network. It is noteworthy that the k-means method is a typical unsupervised statistical classifier, the linear discriminant method and k-nearest neighbor classifier are typical supervised statistical classifiers, and the feed forward network is a typical artificial neural network. In other words, we are comparing the classification performances of the SVM with typical unsupervised/supervised statistical classifiers and an artificial neural network in this study.

The results are summarized in Table 4, where the error, true positives, false positives, false negatives, true negatives, sensibility, 1-specificity, and Area Under the Curve, AUC are tabulated for the mentioned classifiers. The best performance is obtained by the support vector machine.

In Table 5, the classification results with the selected best sixteen features and no feature selection are compared. We can see that feature selection can eliminate irrelevant variables to enhance the generalization performance of a given learning algorithm.

CONCLUSION

In this study, SVM is applied to recognize welding defects. The obtained results show the effectiveness of using SVM to detect defects. It has therefore succeeded to detect welding defects present in radiography, better than a k-means classifier, a linear discriminant classifier, a k-nearest neighbor classifier and a feed forward neural network. Indeed, the proposed based on SVM detection technique has provided satisfying results.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support from Nanyang Technological University.

REFERENCES

Boster, B., I. Guyon and V. Vapnik, 1992. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual ACM workshop on Computational Learning Theory, Pittsburgh, USA, pp: 144-152.

Gunn, S.R., 1997. Support vector machines for classification and regression. Technical report image speech and intelligent systems research group, Southampton University, Southampton, UK.

Haralick, R.M., K. Shanmugam and I. Dinstein, 1973. Textural features for image classification. IEEE T. Syst. Man. Cyb., 3: 610-621.

Hayes, C., 1997. ABC's of nondestructive weld examination. Weld J., 76(5): 46-51.

Kehoe, A. and G. Parker, 1990. Image processing for industrial radiographic inspection: image enhancement. Br. J. NDT, 32(4): 183-190.

Liao, T.W., 2004. Fuzzy reasoning based automatic inspection of radiographic welds: Weld recognition. J. Intell. Manuf., 15: 69-85.

Mery, D. and M. Berti, 2003. Automatic detection of welding defects using texture features. Insight, 45: 676-681.

Platt, J.C., 1998. Sequential minimal optimization: A fast algorithm for training support vector machines. Technical report.

Romeu, R., P. Luiz, L. Calo a and H. Marcio, 2004. Pattern recognition of weld defects detected by radiographic test. NDT & E International, 37: 461-470.

Silva, R., M. Siqueira, L. Calo a and J. Rebello, 2001. Radiographics pattern recognition of welding defects using linear classifiers. Insight, 43(10): 669-674.

Silva, R., M. Siqueira, L. Calo a, L. Sagrilo and J. Rebello, 2002. Evaluation of the relevant characteristic parameters of welding defects and probability of correct classification using linear classifiers. Insight, 44(10): 616-622.

Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer-Verlag, New York.

Vapnik, V., 1998. Statistical Learning Theory. Wiely, New York.

Vapnik, V., 1999. An overview of statistical learning theory. IEEE T. Neural Networ., 10: 989-999.

Wang, X. and B.S. Wong, 2005. Radiographic image segmentation for weld inspection using a robust algorithm. Res. Nondestruct. Eval., 16(3): 132-142.