3.1 Introduction

Image segmentation is the process of dividing digital images into multiple meaningful regions or sets of pixels. Classically, image segmentation is defined as partitioning an image into separate regions which are homogeneous with respect to some characteristics such as gray level or texture. It is one of the most difficult and important task in image processing, particularly in the case of noisy or low contrast images, such as radiographic images of welds.

Often weld radiographs are checked and interpreted by human experts. However, interpretation of weld radiographs by humans is very subjective, inconsistent, labor intensive and sometimes biased. Therefore, various automated inspection techniques for weld radiographs were attempted worldwide over the past years. Computer vision is a key factor for the implementation of total quality within the different processes in industrial automation (Goumeidane et al., 2009).

The progress in computer science and artificial intelligence techniques have allowed the defect detection and classification to be carried out by using digital image processing and pattern recognition tools (Mekhalfa et al., 2007).

Some parts of the material in this chapter have appeared / been communicated as the following research papers:
Computer vision applications often require segmentation of digital imagery into semantically meaningful regions. The segmented regions can provide a basis for subsequent tasks such as object detection and recognition (Qin and Clausi, 2010).

The segmentation of weld defects has been addressed by several researchers. In most of the papers, an automated vision system has been introduced to detect the welding defects from the radiographic images, using various image processing techniques. After successful segmentation, necessary information such as length, width, area and perimeter of the defects are calculated. Research in this area continues largely because no satisfactory results that allow the detection of the totality of defects without false alarms are available. Moreover, it is not possible to determine which lines of investigation will improve the overall result.

In the present thesis we have proposed a novel method to segment defects in the digital radiographic images of weld. The method is devised for detecting and discriminating discontinuities in the digital radiographic weld images. The method is based on the fact that the defective areas absorb more energy and thus the defects appear darker in the image (Valavanis and Kosmopoulos, 2010). However, Tungsten inclusion defect is the only type of defect which appears whitish.

### 3.2 Proposed Technique for Image Segmentation

From the observation of the horizontal line profiles of weld images without defects, we found one common feature of almost all the line profiles: each profile has a bell shape like a Gaussian curve, as shown in Figure 3.1. If there are any defects, the profile curves shape changes. Also the curve features are different in defects that appear darker and in defects which appear lighter. The shapes of profiles for defects appearing dark and light are shown in figures 3.2 and 3.3.
Figure 3.1: Sample line profile without defect, a. Original profile, b. After 2DLMF

Figure 3.2: Sample line profile of image with defects appearing darker

a. Original profile, b. After 2DLMF
Segmentation of Weld Defects

Figure 3.3: Sample line profile of image with defects appearing brighter

a. Original profile, b. After 2DLMF

In the process of segmentation, we checked images line by line for finding the defects. We separated the bell shape to two sections, increasing gray values (IG) and decreasing gray values (DG) sections (refer Figures 3.4 and 3.5). In IG section, any decreasing gray value is an anomaly, and in DG section, any increasing gray value is an anomaly.

Figure 3.4: Increasing gray values section
As a first step, we found Region of Interest by using mean of elements of each line. Then we detect these anomalies by using threshold in the segmentation section.

### 3.2.1 Region of Interest

The region of interest (ROI) is a selected subset of pixels in the digital radiographic images. It is the boundaries of a weldment and weld bead (Figure 3.6). As a first step we calculate $Th$ by using autothreshold algorithm.

**Algorithm:** Autothreshold

**Input:** A line profile (LP) of Radiographic images after 2DLMF, $T_i = 0$

**Output:** Threshold $Th$

Find minimum ($min$) and maximum ($max$) of gray value of LP.

Calculate $Th = \lfloor (min + max) / 2 \rfloor$.

If ($Th = T_i$) then return $Th$

Else $T_i = Th$

Find $Z_1$ as $mean$ of LP gray values $< T_i$

Find $Z_2$ as $mean$ of LP gray values $> T_i$

Return $Th = \lfloor (Z_1 + Z_2) / 2 \rfloor$.

End
The three algorithms given here identify the region of interest.

Algorithm: ROI
Input: Radiographic Image after 2DLMF of size \((m \times n)\)
Output: Region of Interest \((esta, este)\)

For \(i=1\) to \(n\)

Consider Line Profile (LP) size of \((1 \times m)\).

Find the position and maximum gray scale of LP as \((sx, max)\).

Find the \textit{mean} of LP.

Find \(j\) in \(1\) to \(sx\) as that pixel \textit{sta} when \(|\text{LP}(j)\text{-mean}|\) is minimum \((ysta)\).

\[esta = \text{Editsta (sta)}\]

Find \(j\) in \(sx\) to \(m\) as that pixel \textit{ste} when \(|\text{LP}(j)\text{-mean}|\) is minimum \((yste)\).

\[este = \text{Editste (ste)}\]

Return \(esta, este\)

End for
End ROI

Note that algorithm ROI has steps to edit \textit{sta} and \textit{ste}. We have to edit start point \textit{sta} and end point \textit{ste} of the algorithm above, because parts of defective weld area may be omitted as shown in Figure 3.6.

Figure 3.6: Interest region of Sample line profiles before edit operation
Algorithm: Editsta

Input: sta, ysta, Th

Output: Edited Start point (esta)

For j=1 to sta
    If (LP (j) >= ysta), then return j and stop
Endif
Endfor

For j=1 to sta
    If (LP (j) >= LP (j+1)) //find when LP is decreasing
        k = j
        While (LP (k) >= LP (k+1)) //detect local minimum of LP
            k = k + 1
        Endwhile
        If ((LP (j) – LP (k) ) >= Th ), then return j and stop //correct start point
    Endif
Endfor

Return sta

End

Figure 3.7: Interest region of Sample line profiles after Editsta
Algorithm: Editste

Input: \( ste, yste, Th \)
Output: Edited end point (\( este \))

Algorithm:

For \( j = m \) to \( ste \), step (-1)
   If \( (LP(j) >= yste) \), then return \( j \) and stop
Endif
Endfor

For \( j = m \) to \( ste \), step (-1)
   If \( (LP(j) >= LP(j-1)) \) //find when LP is decreasing
      \( k = j \)
      While \( (LP(k) <= LP(k-1)) \) //detect local minimum of LP
         \( k = k - 1 \)
      Endwhile
   If \( ((LP(j) - LP(k)) >= Th) \), then return \( j \) and stop //correct start point
   Else break
   Endif
Endif
Endfor

Return \( ste \)

End

Figure 3.8: Final Interest region of Sample line profiles after Editste
3.2.2 Proposed Segmentation Technique

At the first step of our segmentation technique, we defined a difference vector ($\text{diff}$) of size $\text{stea} = \text{este} - \text{esta} + 1$. The elements of this vector are difference between successive LP values in the region of interest ($\text{esta}, \text{este}$) (Figure 3.10). We then calculate the standard deviation ($\text{SD}$) of this vector. The two horizontal lines above and below 0 in the $\text{diff}$ graph at a distance of $\text{tol} = \lfloor \text{SD} \rfloor$ is tolerance region (figure 3.11). If $\text{diff}$ gets beyond this region then it is considered to be anomaly.
After calculation **diff**, **stea** and **tol**, we found size of anomaly in the interest region of digital radiographic images, line by line by using two techniques. Algorithms ADD (Anomaly Detection in defects that are seen Dark) and ADL (Anomaly Detection in defects that are seen Light) complete the process of segmentation of defective portions in weld images. Figure 3.12 shows multiple anomalies in a weld image where defects appear dark. Figure 3.13 shows segmentation of defective portion in the radiographic weld image.
Algorithm: ADD

Input: \( \text{diff, esta, este, stea, Th} \)
Output: Sections of anomalies in LP

For i= 1 to \( \text{stea} \)

If \( \text{diff(i)} \geq \text{tol} \) //beginning of anomaly

\[ \text{anbeg}=\text{i} \]

Endif

j=i+1

While \( \text{diff(j)} \geq \text{tol} \) or \( -\text{tol} < \text{diff(j)} < \text{tol} \) // j within tolerance region

\[ j=j+1 \]

Endwhile

While \( \text{diff(j)} < -\text{tol} \)

\[ j=j+1 \]

Endwhile

\[ \text{anend}=j \]

Locate \( \text{max} \) (maximum of diff within anbeg and anend)

Locate \( \text{min} \) (minimum of diff within anbeg and anend)

\[ x_1=\text{max}+\text{esta} \]

\[ x_2=\text{min}+\text{esta} \]

\[ y_1=\text{LP}(x_1) \]

\[ y_2=\text{LP}(x_2) \]

If \( |y_1- y_2| \geq \text{Th} \) then

Output \( x_1, x_2 \) //region of anomaly in LP

Endif

i=\( \text{min} \)

Endfor
Algorithm: ADL

Input: \textit{diff}, \textit{esta}, \textit{este}, \textit{stea}, \textit{Tr}

Output: Sections of anomalies in LP

For \( i = 1 \) to \( \text{stea} \)

If \( \text{diff}(i) \leq -\text{tol} \)  
   //beginning of anomaly

\hspace{1cm} \textit{anbeg} = i

Endif

\hspace{1cm} j = i + 1

While \( \text{diff}(j) < -\text{tol} \) or \( -\text{tol} < \text{diff}(j) < \text{tol} \)  
   // j within tolerance region

\hspace{1cm} j = j + 1

Endwhile

While \( \text{diff}(j) > \text{tol} \)

\hspace{1cm} j = j + 1

End while

\hspace{1cm} \textit{anend} = j

Locate \textit{max} (maximum of \textit{diff} within \textit{anbeg} and \textit{anend})

Locate \textit{min} (minimum of \textit{diff} within \textit{anbeg} and \textit{anend})

\hspace{1cm} x_1 = \textit{min} + \textit{esta}

\hspace{1cm} x_2 = \textit{max} + \textit{esta}

\hspace{1cm} y_1 = \text{LP}(x_1)

\hspace{1cm} y_2 = \text{LP}(x_2)

If \( |x_1 - x_2| \geq \text{Tr} \) then

\hspace{1cm} Output \( x_1, x_2 \)  
   //region of defect in LP

Endif

\hspace{1cm} i = j

Endfor

End
Shown in Figure 3.13 is the segmented defective portion in the radiographic weld image where defects appear dark.

Figure 3.13: Segmented region of defects appearing darker

Figure 3.14: Location of anomaly in \textit{diff}
Figures 3.14 and 3.15 show the location of anomaly in $\text{diff}$ and correspondingly in the LP for weld images where defects appear lighter. In this segmentation $\text{Tr}$ is based on API-1401 American Petroleum Institute Standard, (2001) for defects that are brighter.

Figure 3.15: Location of $x_1$ and $x_2$ in Region of Interest

Figure 3.16: Radiographic image after 2DLMF

Figure 3.17: Region of Interest

Figure 3.18: Segmented region of defects appearing brighter
3.3 Comparative Study

We have compared our method with Region Growing and Mean shift segmentation by using ground truth.

3.3.1 Ground truth

In machine learning, the term "ground truth" refers to the accuracy of the training set classification for supervised learning techniques. This is used in statistical models to prove or disprove research hypotheses. Then "ground truth" refers to the process of gathering the proper objective data for this test.

Ground truth is a term used in remote sensing; it refers to information collected on location. Ground truth allows image data to be related to real features and materials. The collection of ground-truth data enables calibration of remote-sensing data, and aids in the interpretation and analysis of what is being sensed. Examples include cartography, meteorology, analysis of aerial photographs, satellite imagery, radiographic images and other techniques in which data are gathered at a distance.

More specifically, ground truth may refer to a process in which a pixel on a weld image is compared to what is there in reality in order to verify the contents of the pixel on the image. In the case of a segmented image, it allows supervised segmentation to help determine the accuracy of the segmentation performed by the measure calculating like: peak signal to noise ratio (PSNR), structure similarity index measure (SSIM) and Euclidian distance. Therefore, results accuracy increase by minimizing errors in the proposed segmentation methods.

3.3.2 Region Growing

Region growing algorithm is an algorithm that appends pixels or sub regions into larger regions based on predefined criteria. Hence the region of interest i.e. the weld defect is obtained directly through similarity detection. The algorithm involves selecting a seed pixel, which is the representative pixel of the region of interest. Selecting a set of one or more seed pixels often can be based on the nature of the problem. In this application, it is known that pixels of detective weld tend to have
the minimum allowable digital gray values in ROI in the images for defects which appear darker and maximum value related to defects which appear lighter. All pixels with nearly same intensities are grouped together. Grouping nearly similar pixels is achieved by selecting a threshold value.

### 3.3.3 Mean Shift

The mean shift algorithm is a powerful clustering technique, which is based on an iterative scheme to detect modes in a probability density function. It has been utilized for image segmentation by seeking the modes in a feature space composed of spatial and color information by Kaftan et al. (2008).

Mean shift is an unsupervised clustering algorithm by Fukunaga and Hostetler (1975), which estimates the gradient of a probability density function to detect modes in an iterative fashion. Hence, image segmentations that take color/intensity-similarity as well as local connectivity into account can be obtained by applying this algorithm to the combined spatial-range domain. Mean shift segmentation has been successfully applied to several applications by Comaniciu and Meer (1999, 2002) and Christoudias and et al. (2002).

However, for larger images or applications where processing time is very crucial, the mean shift segmentation algorithm might be still too times consuming. The mean shift is based on the EDISON framework (Christoudias et al. (2002); Comaniciu and Meer (2002)).

### 3.4 Experiment and results

The proposed segmentation methods are evaluated by comparing with region growing and Mean shift segmentation methods and ground truth. Accuracy of the proposed methods can be inferred from Figures 3.19 to 3.21, where the average of PSNR, SSIM and Euclidean distance between obtained segmentation and ground truth, are given for 196 defective images, covering various types of defects.
It can be seen that the proposed segmentation techniques have much higher PSNR and SSIM. We computed Euclidean distance between LP of the ground truth and proposed, mean shift, region growing methods. The proposed method has least Euclidean distance from ground truth.

![Figure 3.19: Average PSNR values](image1)

![Figure 3.20: Average SSIM values](image2)
In this chapter, two automatic segmentation techniques for the digital radiographic images of weld are proposed. The first method is devised to segment defects that appear darker in the digital radiographic images and the second method related to defects which appear brighter.

We compared proposed segmentation techniques with mean shift and region growing methods using ground truth. The accuracy of proposed techniques was evaluated by PSNR, SSIM measures and Euclidean distance. The result of the experiments demonstrated proposed segmentation techniques have higher accuracy compared to mean shift and region growing methods.

3.5 Conclusion

Figure 3.21: Average Euclidean distance
Chapter 4

Classification of Weld Defects

4.1 Introduction

The problem of image classification has drawn considerable attention in the Computer Vision community. The continued effort of the research community in the last few years resulted in many novel approaches for image classification.

Image classification, including object recognition and scene classification, remains to be a major challenge to the computer vision community. Perhaps one of the most significant developments in the last decade is the application of local features to image classification. Image classification methods can be roughly divided into two broad families of approaches:

1. Learning-based classifiers: Requires an intensive learning/training phase of the classifier parameters (e.g., parameters of SVM). These methods are also known as parametric methods.

2. Nonparametric classifiers: Classification decision based directly on the data and requires no learning/training of parameters. The most common non-parametric methods rely on Nearest Neighbor Distance Estimation (NNDE). A special case of these is the “Nearest-Neighbor-Image” classifier, which classifies an image to a class of its nearest (most similar) image in the database (Varma and Ray, 2007).

Some parts of the material in this chapter have appeared / communicated as the following research paper:

1. Classification of Welding Defects in Radiographic Images. In the process for publication in the International Journal of Pattern Recognition and Image Analysis (Springer)
Non-parametric classifiers have several very important advantages that are not shared by most learning-based approaches:

1. Can naturally handle a huge number of classes.
2. Avoid over fitting of parameters, which is a central issue in learning based approaches.
3. Require no learning/ training phase. Although training is often viewed as a one-time preprocessing step, retraining of parameters in large dynamic databases may take days, whereas changing classes/training-sets is instantaneous in non-parametric classifiers.

In this chapter we use a novel Non-parametric classifier and compare the proposed procedure with KNN and SVM classifier methods.

### 4.2 Proposed Technique for Image Classification

The present problem of pattern classification involves discrimination between two classes: profiles without defect (ND) and profiles with defects (D). Defective class (D) include five classes: Under-Cut (UC), Lack-Of-Penetration (LOP), Incomplete Fusion side wall (IF), Gas Porosity (GP) and Tungsten Inclusion (TI).

The type of defect detection step in weld radiographs in the process of assigning a set of objects the respective class labels: Non-Defect (ND), Tungsten Inclusion (TI), Gas Porosity (GP), Lack-Of-Penetration (LOP), Incomplete Fusion side wall (IF) and Under-Cut (UC) as shown in Figure 4.1. In order to classify each of the obtained objects, a set of geometrical features is extracted which is then used as input in the proposed classification method.

In the first level of tree, after segmentation digital radiographic images of weld, we classify images into two classes, Defect and Non-defect image. In Non-defect image, there are no anomalies and the images are segmented to have only one part, as shown in Figure 4.2.
Figure 4.1: Proposed classification method
Segmented defective images are classified in two classes: defects appearing brighter and appearing darker. Then brighter image have been detected as Tungsten Inclusion and darker image can have two classes of defects: round and linear types. Gas Porosity detected by round defect and we identify three linear type defects: Lack-Of-Penetration, Incomplete Fusion side wall and Under-Cut.

4.2.1 Feature definition and extraction

The next stage after profiling lines is the feature extraction in terms of individual and overall characteristics of the defects. This represents a great reduction in image information from the segmented input image and ensures that the subsequent classification of defect type is efficient.

In the segmentation stage esta and este are computed for each LP. The boundaries of anomalies in each LP are shown in 4.3 and 4.4. l and w be length and width of the defect are shown in Figure 4.5. To discriminate round and linear defects we simply use the feature called cir which is calculated as the ratio of length of the defect in the direction of weld seam to the width of the defect (l/w). If this is close to 1 then the defect is classified as round else it is linear. The images where defects appear as bright the defects are TI (Figure 4.7 and 4.12). The round defect in images where
defects appear dark, the defect is Gas porosity (GP) (Figure 4.8 and 4.13). For classification of linear defects we extract features that are describes in subsequent paragraphs.

Figure 4.3: Sample digital radiographic image with defects
a. Image after 2DLF.  b. Image showing Start, Maximum and End point arrays
c. Image with the line of msta, msx, mste and two means.
In our experiments we have used geometrical attributes to evaluate a defective linear segment in digital radiographic images of weld. A system that tries to identify defects should exploit information sources well. During segmentation \textit{esta}, \textit{este} and \textit{sx} are located for each LP. These are nothing but start, end points of the region of interest (ROI) in LP and position of maximum gray value in LP. The arrays of \textit{n} values obtained from each LP are shown in Figure 4.3 b. We now find the mean of these values derived from each LP and denote these as \textit{msta}, \textit{mste} and \textit{msx}. The horizontal lines at \textit{msta}, \textit{mste}, \textit{msx}, at mean of \textit{msta} and \textit{msx} and at mean of \textit{mste} and \textit{msx} are shown in Figure 4.3 c. Using the positions of these five horizontal lines we define our features for classification of linear dark defects. Also shown in Figure 4.4 and 4.6 are these five lines in a segmented image.

A set of five features are computed from the positions of the five horizontal lines. All these parameters are calculated automatically in segmented images. These features and their calculation formulas are given below.
Position feature $P_1 (= \frac{h_1}{H})$: $h_1$ is the displacement of the $msx$ line from the axis of defect and $H$ is the total span of the defective area ($mste - msta$). With this feature it is possible to separate Lack-Of-Penetration, frequently located in the seam centre (Figures 4.9 and 4.14).

Position feature $P_2 (= \frac{h_2}{H})$ and $P_3 (= \frac{h_3}{H})$: Here $h_2$ is distance between the lines $msta$ and axis of defect. Similarly $h_3$ is distance between the lines $mste$ and axis of defect. With these parameters it is possible to separate Under-Cut, usually present at the border of region interest (Figures 4.10 and 4.15).

Position features $P_4 (= \frac{h_4}{H})$ and $P_5 (= \frac{h_5}{H})$: Here $h_4$ is distance between the line at mean of $msta$, $msx$ and axis of defect. Similarly $h_5$ is distance between the line at mean of $mste$, $msx$ and axis of defect. With these parameters it is possible to separate Incomplete Fusion side wall, usually present between the border and center of region interest (Figure 4.11 and 4.16).

Figure 4.6: Illustration of the parameters used in five features extracted.
4.2.2. Proposed classification method

In this section, we propose two novel automated processes of classification into five types of defects: Tungsten Inclusion (TI), Gas Porosity (GP), Lack-Of-Penetration (LOP), Incomplete Fusion (IF) and Under-Cut (UC).

We define these five defects and discuss the procedure of two classification algorithms in the two different subsections 4.2.2.1 and 4.2.2.2.

4.2.2.1. Defects definition

Tungsten Inclusion is unique to the Gas tungsten arc welding (GTAW), also known as tungsten inert gas (TIG) welding process. This discontinuity occurs in most metals welded by the process, including aluminum and stainless steel. The TIG method of welding produces a clean homogeneous weld that radiographer can interpret it easily.

Tungsten is a brittle and inherently dense material used in the electrode in tungsten inert gas welding. If improper welding procedures are used, tungsten may be entrapped in the weld. In the digital radiographic images (tungsten is more density than aluminum or steel). Hence it shows up as a brighter round area with a distinct outline on the radiography (Figures 4.7).

Figure 4.7: Tungsten inclusion in radiographic image
Gas Porosity is the result of gas entrapment in the solidifying metal. Porosity can take many shapes on a radiograph but often appears as dark round or irregular spots or specks appearing singularly, in clusters, or in rows. Sometimes, porosity is elongated and may appear to have a tail. This is the result of gas attempting to escape while the metal is still in a liquid state and is called wormhole porosity. All porosity is a void in the material and it will have a higher radiographic density than the surrounding area as shown in Figure 4.8.

![Figure 4.8: Gas porosity in radiographic image](image)

Lack-Of-Penetration (LOP) or Incomplete penetration (IP) occurs when the weld metal fails to penetrate the joint (Figure 4.9). It is one of the most objectionable weld discontinuities. Lack-Of-Penetration increases natural stress from which a crack may propagate. The appearance on a radiograph is a dark area with well-defined straight edges followed by the land or root facing down the center of the weldment. That is located near maximum points array of the weld seam msx.

![Figure 4.9: Lack of penetration in radiographic image](image)
Under-Cut (UC) is an erosion of the base metal next to the crown of the weld Figure 4.10. In the radiograph, it appears as a dark irregular line along the outside edge of the weld area. That is located near start point *msta* or end point arrays *mste*.

![Figure 4.10: Undercut in radiographic image](image)

Incomplete Fusion side wall (IF) or incomplete fusion is a condition where the weld filler metal does not properly fuse with the base metal Figures 4.11. Appearance on radiograph: usually appears as a dark line or lines oriented in the direction of the weld seam along the weld joining area. That is located between *msta* and *msx* or *mste* and *msx*.

![Figure 4.11: Incomplete fusion in radiographic image](image)
4.2.2.2. Classification process

After feature extraction in the third level of classification tree (Figure 4.1) we have segmented images to brighter and darker defects. Tungsten Inclusion defects are detected from brighter section using TI algorithm. In this algorithm, we used length \( l \) and \( cir \) features.

---

**Algorithm:** TI

**Input:** Segmented defective bright weld images, \( Ths, Tha, l, w \)

**Output:** Tungsten Inclusion defect

For each defect

Calculate \( cir \, (l/w) \) feature

If \((cir-1) \leq Ths \) \&\& \((l \leq Tha)\)

Then Output “defect is TI”

Else Output “defect is not TI”

Endif

Endfor

End TI

---

Approved size of defects based on American Petroleum Institute Standard (API 1104) has been defined, in our data set \( Ths \) equal 0.5 and \( Tha \) is 9.
We found other four defects from segmentation of darker defects by algorithm DDC. In this level, we used length \((l)\) and \(cir\) \((= l/w)\) features for separating the round and linear defects as shown in classification flow chart (Figure 4.1) for detecting Gas Porosity and three linear defects: Lack-Of-Penetration, Incomplete Fusion and Under-Cut by using five position features \(P_1, P_2, \ldots, P_5\).

**Algorithm: DDC**

**Input:** Segmented defective image (defects seen dark), \(Ths, Tha, l, w\)

**Output:** GP, LOP, IF and UC

1. Calculate features \(cir, P_1, P_2, P_3, P_4, \) and \(P_5\).
2. Calculate \(dec = \min(P_1, P_2, P_3, P_4, \) and \(P_5)\).
3. If \(( (cir-1) \leq Ths) \&\& (l \leq Tha)\)
   - Then Output is “GP” and stop
4. Else “Defect is linear shape”
5. Endif
6. If \((P_1 = dec)\)
   - Then Output is “LOP” and stop
7. If \((P_2 = dec)\) or \((P_3 = dec)\)
   - Then Output is “UC” and stop
8. Output is “IF” and stop
9. Endfor
10. End DDC
4.3 Comparative Study

We compared our classification techniques with two standard classification methods; Support Vector Machine (SVM) and K-Nearest Neighbor (KNN).

4.3.1 SVM Classification

SVM is one of the best known methods in pattern and image classification. It is designed to separate a set of training images two different classes, \((x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\) where \(x_i\) is in \(\mathbb{R}^d\), d-dimensional feature space, and \(y_i\) in (-1,+1), the class label, for \(i=1, ..., n\). When the two classes are linearly separable in \(\mathbb{R}^d\), we wish to find a separating hyperplane which gives the smallest generalization error among the infinite number of possible hyper planes. Such an optimal hyperplane is the one with the maximum margin of separation between the two classes, where the margin is the sum of the distances from the hyperplane to the closest data points of each of the two classes. These closest data points are called Support Vectors (SVs).

Let us suppose they are completely separated by a d-dimensional hyperplane described by

\[
wx + b = 0 \quad (4.1)
\]

The separation problem is to determine the hyperplane such that \(w.x + b \geq 1\) for positive examples and \(w.x + b \leq -1\) for negative examples. Since the SVM finds the hyperplane which has the largest margin, it can be found by maximizing \(1/\|w\|^2\). The training of the optimal parameters \((w^*, b^*)\) is transformed to the following quadratic programming task, Eq. (4.2), under the constraint Eq. (4.3) to correctly separate the training data.

\[
(w^*, b^*) = \arg\min_{w} \frac{1}{2}\|w\|^2 \quad (4.2)
\]

\[
y_i (w.x_i + b) - 1 \geq 0 \quad \forall i \quad (4.3)
\]
While using SVM for the problem of classification of defects in weld images, first level classification is round or linear. Round class is then split into defect types GP and TI. Images where defects are linear, first classification is (i) LOP, (ii) UC and IF. Finally UC and IF images are separated.

### 4.3.1 K-NN Classification

The K-nearest neighbor (KNN) classifier, a conventional non-parametric, has been both a workhorse and benchmark classifier (Xiaowei et al., 2009). Given a query vector \( x_0 \) and a set of \( N \) labeled instances the task of the classifier is to predict the class label of \( x_0 \) on the predefined \( P \) classes. The K-nearest neighbor (KNN) classification algorithm tries to find the \( K \) nearest neighbors of \( x_0 \) and uses a majority vote to determine the class label of \( x_0 \). Without prior knowledge, the KNN classifier usually applies Euclidean distances as the distance metric. However, this simple and easy-to-implement method can still yield competitive results even when compared to the most advanced machine learning methods.

### 4.4 Experiment and results

In the third chapter we segmented digital radiographic images of weld and classified images in two classes: Defective (\( D \)) and Non-defective (\( ND \)). In the present chapter with using proposed classification algorithms, we classified digital radiographic images of weld with defects (\( D \)) and compared our results with K Nearest Neighbor and Support Vector Machine classification methods. Data bases comprise of five different types of weld defects: Tungsten Inclusion, Gas Porosity, Lack-Of-Penetration, Incomplete Fusion and Under-Cut.

We classified Tungsten Inclusion (TI) defect which appear brighter in digital radiographic images of weld by using TI Algorithm. We have 959 line profiles from 314 round defects. Our technique in this stage has detected round defects with the accuracy of 95.54%. As we already mentioned that there is only one type of defect that appears bright in the digital radiographic images of weld defects. Our segmentation technique when compared with region growing method was more accurate, for defects that appear bright.
We classified four defects which appear darker by using DDC Algorithm. The first stage of this algorithm detects Gas Porosity (GP) by classifying defects in images as round and linear. We have 204879 line profiles from 1761 linear defects and 549 line profiles from 99 round defects. Our technique in this stage has detected round defects with an accuracy of 97.98%.
KNN and SVM are chosen for comparison for classification of round type weld defects. We tested these methods (KNN, SVM) nine times. The number of training levels varied from 10% to 90% with increments of 10% as shown in Table 4.1. The mean accuracy of results of classification methods are 96.67 and 96.10%, respectively. It is evident from the results that our proposed technique is better than KNN and SVM.
Table 4.1: The results of 413 Round defects classification by SVM

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>96.10</td>
<td>96.10</td>
<td>95.67</td>
<td>95.88</td>
<td>96.10</td>
<td>95.67</td>
<td>96.10</td>
<td>95.67</td>
</tr>
<tr>
<td>20</td>
<td>95.69</td>
<td>95.91</td>
<td>95.91</td>
<td>95.91</td>
<td>95.91</td>
<td>95.91</td>
<td>95.91</td>
<td>95.91</td>
</tr>
<tr>
<td>30</td>
<td>95.72</td>
<td>95.84</td>
<td>95.84</td>
<td>95.84</td>
<td>95.84</td>
<td>95.70</td>
<td>95.84</td>
<td>95.84</td>
</tr>
<tr>
<td>40</td>
<td>95.69</td>
<td>95.80</td>
<td>95.69</td>
<td>95.59</td>
<td>95.59</td>
<td>95.80</td>
<td>95.80</td>
<td>95.69</td>
</tr>
<tr>
<td>50</td>
<td>95.78</td>
<td>95.52</td>
<td>95.78</td>
<td>95.78</td>
<td>95.78</td>
<td>95.70</td>
<td>95.52</td>
<td>95.78</td>
</tr>
<tr>
<td>60</td>
<td>95.77</td>
<td>95.63</td>
<td>95.77</td>
<td>95.77</td>
<td>95.77</td>
<td>95.77</td>
<td>95.63</td>
<td>95.77</td>
</tr>
<tr>
<td>70</td>
<td>95.76</td>
<td>95.79</td>
<td>95.76</td>
<td>95.76</td>
<td>95.76</td>
<td>95.76</td>
<td>95.76</td>
<td>95.76</td>
</tr>
<tr>
<td>80</td>
<td>95.75</td>
<td>95.70</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
<td>95.70</td>
<td>95.75</td>
</tr>
<tr>
<td>90</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
<td>95.75</td>
</tr>
</tbody>
</table>

Table 4.2: The results of 413 Round defects classification by KNN

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>81.11</td>
<td>77.78</td>
<td>78.89</td>
<td>81.11</td>
<td>72.22</td>
<td>70.00</td>
<td>86.67</td>
<td>64.44</td>
</tr>
<tr>
<td>20</td>
<td>80.00</td>
<td>90.00</td>
<td>76.25</td>
<td>82.50</td>
<td>86.25</td>
<td>82.50</td>
<td>86.25</td>
<td>68.75</td>
</tr>
<tr>
<td>30</td>
<td>78.57</td>
<td>85.71</td>
<td>90.00</td>
<td>81.43</td>
<td>87.14</td>
<td>94.29</td>
<td>85.71</td>
<td>91.43</td>
</tr>
<tr>
<td>40</td>
<td>85.00</td>
<td>73.33</td>
<td>86.67</td>
<td>91.67</td>
<td>90.00</td>
<td>85.00</td>
<td>85.00</td>
<td>86.67</td>
</tr>
<tr>
<td>50</td>
<td>88.04</td>
<td>70.00</td>
<td>82.00</td>
<td>86.00</td>
<td>88.00</td>
<td>86.00</td>
<td>76.00</td>
<td>80.00</td>
</tr>
<tr>
<td>60</td>
<td>90.00</td>
<td>92.50</td>
<td>92.50</td>
<td>85.00</td>
<td>85.00</td>
<td>95.70</td>
<td>95.00</td>
<td>92.50</td>
</tr>
<tr>
<td>70</td>
<td>86.67</td>
<td>83.33</td>
<td>86.67</td>
<td>93.33</td>
<td>83.33</td>
<td>96.67</td>
<td>96.67</td>
<td>93.33</td>
</tr>
<tr>
<td>80</td>
<td>90.00</td>
<td>85.00</td>
<td>90.00</td>
<td>90.00</td>
<td>80.00</td>
<td>88.57</td>
<td>95.00</td>
<td>85.00</td>
</tr>
<tr>
<td>90</td>
<td>90.00</td>
<td>93.54</td>
<td>90.00</td>
<td>92.65</td>
<td>90.00</td>
<td>80.00</td>
<td>90.00</td>
<td>91.00</td>
</tr>
</tbody>
</table>
In the second stage of our classification technique in DDC algorithm, we detected three linear defects: Lack-Of-Penetration (LOP), Under-Cut (UC) and Incomplete Fusion side wall (IF) from 204879 line profiles from 1761 linear defects. The experimental results are shown in Tables 4.3. It can be seen from the table that Lack-Of-Penetration is the best recognized welding defect detection. Recognition ratio is 98.07%. However, approximately 2% of the total welding defects are not recognized by our technique.

Table 4.3: The results of defects classification proposed technique

<table>
<thead>
<tr>
<th>Type of Defects</th>
<th>Actual defects</th>
<th>No. of detected defective profiles</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOP</td>
<td>816</td>
<td>50740</td>
<td>48759</td>
</tr>
<tr>
<td>UC</td>
<td>379</td>
<td>67560</td>
<td>65048</td>
</tr>
<tr>
<td>IF</td>
<td>566</td>
<td>86579</td>
<td>81883</td>
</tr>
</tbody>
</table>

The same dataset is used again for comparison with performance of KNN and SVM methods for classification linear defects.

Tables 4.4, 4.6 and 4.8, demonstrate the best accuracy of LOP, UC and IF by SVM as 82.50, 86.61 and 76.57%, respectively. The accuracy of these three linear defects by KNN is 96.34, 86.32 and 87.72%, respectively. It can be seen from the report that our proposed classification technique is better than KNN and SVM classification methods.
Figure 4.14: Sample image with Lack-Of-Penetration

a. Digital Radiographic image    b. Segmented image    c. Detected defects
Table 4.4: The results LOP defect classification by SVM

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>78.29 76 81.71 77.14 79.29 81.14 76.00 70.29 70.29</td>
</tr>
</tbody>
</table>

Table 4.5: The results LOP defect classification by KNN

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>79.32 83.67 76.05 80.14 81.50 78.78 76.87 85.31 82.86</td>
</tr>
</tbody>
</table>
Figure 4.15: Sample image of Under-Cut
a. Digital Radiographic image   b. Segmented image   c. Detected defects
Table 4.6: The results UC defect classification by SVM

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>82.29 83.43 82.86 83.43 85.71 84.00 84.27 85.24 83.84</td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>85.19 84.62 86.32 84.62 <strong>86.61</strong> 84.90 84.05 82.85 85.20</td>
</tr>
<tr>
<td>30</td>
<td>83.30 83.68 83.11 84.44 83.11 84.63 83.11 81.78 85.01</td>
</tr>
<tr>
<td>40</td>
<td>84.50 83.78 83.07 84.21 84.07 82.36 83.50 83.93 84.64</td>
</tr>
<tr>
<td>50</td>
<td>85.91 85.23 83.07 83.98 82.95 83.30 84.55 84.55 83.75</td>
</tr>
<tr>
<td>60</td>
<td>84.19 82.39 83.81 83.71 83.05 83.90 83.90 83.90 84.38</td>
</tr>
<tr>
<td>70</td>
<td>83.60 83.12 84.33 84.74 83.77 84.01 82.87 81.74 82.95</td>
</tr>
<tr>
<td>80</td>
<td>82.74 83.38 83.52 83.45 83.81 84.59 83.45 83.74 84.02</td>
</tr>
<tr>
<td>90</td>
<td>82.70 84.41 83.71 82.45 82.89 80.43 82.32 83.46 83.84</td>
</tr>
</tbody>
</table>

Table 4.7: The results UC defect classification by KNN

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>59.36 68.13 61.40 65.20 61.70 64.04 55.26 58.77 59.36</td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>63.82 74.01 77.30 69.08 70.72 67.11 70.72 66.12 74.67</td>
</tr>
<tr>
<td>30</td>
<td>73.68 71.43 73.68 75.19 82.33 73.68 76.69 77.07 75.19</td>
</tr>
<tr>
<td>40</td>
<td>82.89 79.39 75.44 69.74 74.56 71.49 76.32 78.95 74.56</td>
</tr>
<tr>
<td>50</td>
<td>80.53 75.26 75.79 78.95 75.79 77.89 <strong>86.32</strong> 76.84 74.74</td>
</tr>
<tr>
<td>60</td>
<td>82.89 74.34 78.95 73.03 78.95 78.29 81.58 83.55 76.32</td>
</tr>
<tr>
<td>70</td>
<td>78.95 81.58 75.44 81.58 78.95 82.46 85.96 82.46 85.09</td>
</tr>
<tr>
<td>80</td>
<td>81.58 84.21 84.21 85.53 81.58 84.21 82.89 81.58 76.32</td>
</tr>
<tr>
<td>90</td>
<td>92.11 92.11 89.47 81.58 89.47 81.58 86.84 86.84 81.58</td>
</tr>
</tbody>
</table>
Figure 4.16: Sample image of Incomplete Fusion

a. Digital Radiographic image  
b. Segmented image  
c. Detected defects
Table 4.8: The results IF defect classification by SNM

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>70.86 75.43 76.57 71.43 72.14 76.00 66.68 72.00 69.71</td>
</tr>
<tr>
<td>20</td>
<td>66.48 73.01 72.73 73.86 69.32 69.03 72.73 73.30 73.01</td>
</tr>
<tr>
<td>30</td>
<td>72.30 70.02 68.88 72.07 72.43 75.33 73.62 69.26 70.97</td>
</tr>
<tr>
<td>40</td>
<td>70.88 70.31 71.31 70.45 72.10 71.59 70.17 70.31 70.60</td>
</tr>
<tr>
<td>50</td>
<td>74.77 71.93 71.70 71.02 73.01 70.80 71.14 72.39 72.95</td>
</tr>
<tr>
<td>60</td>
<td>73.11 71.21 69.70 69.41 72.44 72.82 71.97 69.98 73.58</td>
</tr>
<tr>
<td>70</td>
<td>72.65 71.02 71.35 71.10 72.40 73.21 71.83 69.56 72.08</td>
</tr>
<tr>
<td>80</td>
<td>68.68 69.11 68.18 74.29 69.39 70.81 71.73 70.67 71.45</td>
</tr>
<tr>
<td>90</td>
<td>70.71 70.83 73.30 70.08 70.52 70.77 66.54 68.81 68.88</td>
</tr>
</tbody>
</table>

Table 4.9: The results IF defect classification by KNN

<table>
<thead>
<tr>
<th>Training %</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>60.00 49.61 47.65 52.55 53.53 59.22 57.06 53.92 49.80</td>
</tr>
<tr>
<td>20</td>
<td>59.16 59.60 65.56 60.71 66.67 62.91 61.59 64.24 61.15</td>
</tr>
<tr>
<td>30</td>
<td>72.04 68.77 66.75 71.79 66.75 62.97 63.73 73.55 65.49</td>
</tr>
<tr>
<td>40</td>
<td>70.00 72.06 75.29 71.76 72.35 71.47 70.59 69.41 72.35</td>
</tr>
<tr>
<td>50</td>
<td>79.51 72.79 74.20 73.85 72.79 69.61 76.68 76.33 74.91</td>
</tr>
<tr>
<td>60</td>
<td>69.60 75.77 75.77 76.65 77.09 78.85 73.13 75.77 76.21</td>
</tr>
<tr>
<td>70</td>
<td>80.00 82.35 78.24 75.29 82.94 75.88 70.59 76.47 71.76</td>
</tr>
<tr>
<td>80</td>
<td>83.33 81.58 83.33 86.84 75.44 84.21 72.81 82.46 79.82</td>
</tr>
<tr>
<td>90</td>
<td>77.19 84.21 84.21 87.72 73.68 80.70 82.46 80.70 82.46</td>
</tr>
</tbody>
</table>
4.5 Conclusion

Classification of welding defects is a useful research with strong application background. In this chapter we have developed a method to extract features and classify the defects from the digital radiographic images of welds and we compared the proposed method with K-Nearest Neighbor and Support Vector Machine. Experimental results have shown the classification method is useful for the linear defects: such as Lack-Of-Penetrations, Incomplete Fusion and Under-Cut, as well as round defects: Tungsten Inclusion and Gas Porosity. Generally, the method of classification of defects in welding is effective and it can reduce the effort of human being and increase the defect recognition efficiency. However, the method put forward is not 100% successful and has some error in classification. However this error is a very small percentage. Also the method developed is not designed to detect all welding defects. It is still very useful for classification and recognition of the weld defects that are common.