[UCI]I804:24011-200000359998

February 2021
Master's Degree Thesis

Use of monitoring system for Weld defect characterization in GMAW by PCA-SVM Classifiers

Graduate School of Chosun University

Department of Naval Architecture and Ocean Engineering

Aswin Krishnaswamy

Use of monitoring system for Weld defect characterization in GMAW by PCA-SVM Classifiers

PCA-SVM 분류기에 의한 GMAW 용접시 결함 특성화를 위한 모니터링 시스템에 관한 연구

February 25, 2021

Graduate School of Chosun University

Department of Naval Architecture and Ocean Engineering

Aswin Krishnaswamy

Use of monitoring system for Weld defect characterization in GMAW by PCA-SVM Classifiers

Advisor: Prof. Sung-Min Joo

A thesis submitted in partial fulfillment of the requirements for a master's degree in engineering

October 2020

Graduate School of Chosun University

Department of Naval Architecture and Ocean Engineering

Aswin Krishnaswamy

아스원의 석사학위 논문을 인준함

위원장 조선대학교 교수 안규백(인)

위 원 조선대학교 교수 권영섭(인)

위 원 조선대학교 교수 주성민 (인)

2020년 11월

조선대학교 대 학 원

Table of Contents

List of Figures iii

List of Tables v

ABSTRACT vi

1. Theory and background 1

1.1 Overview 1

1.1.1 Gas Metal Arc Welding (GMAW) 3

1.1.2 Support Vector Machine (SVM). 4

1.2 Objectives 9

1.3 Thesis layout 10

2. Experimental details 12

2.1 Materials and welding setup 12

2.1.1 Materials used 12

2.1.2 Welding setup 15

2.1.3 Welding parameters 19

2.1.4 Experimental Flow 21

2.2 Weld feature extraction 23

2.2.1 Feature selection by PCA 23

2.2.2 Weld defect definition 25

2.2.3 Flaw diagnosis and detection by SVM 29

2.2.4 Algorithm development 31

3. Results and Discussion 32

3.1 Results of GMAW welding and Data extraction. 32

3.2 Parameter selection. 38

3.2.1 PCA with Two class classifier 38

3.2.2 PCA with Multi-class classifier 39

3.3 . Process and results of development of machine learning algorithms 40

4. Conclusion 42

References 43

List of Figures

Figure 1. Schematic representation of proposed model. 3

Figure 2. Representation of margin distance in classification......................... 4

Figure 3. Representation of Two class classifier model................................. 6

Figure 4. Representation of Multi class classifier model. 7

Figure 5. Flow chart of thesis layout 11

Figure 6. Schematic representation of SS400 steel plate dimensions. 13

Figure 7. Schematic representation of SS400 plates, a, b and c represents the metal plates with 6 mm thickness and d, e and f represents with 9 mm thickness. 14

Figure 8. Welding setup a) is the welding transformer and b) is the wire feeder (Longrun Model No. : 500L-C3)... 15

Figure 9. Welding speed controller Hiweld set on gear driven rail mechanism... 16

Figure 10. Hiweld weld monitoring system for Voltage-current monitoring. 16

Figure 11. Weld specimen setup in Aluminum block jig with copper strip backing. 17

Figure 12. SS400 metal plates used for the welding experiments. 17

Figure 13. Flow diagram of the process. 22

Figure 14. Representation of kernel function 24

Figure 15. Good weld achieved from the experiments a) being the welded sample and b) is the closer view... 26

Figure 16. Burn through specimen resulted after the experiment a) being the welding sample and b) is closer view of the weld back bead. .. 27

Figure 17. Lack of fusion and incomplete penetration sample obtained a) is the welding sample and b) being the closer view of the back bead 28

Figure 18. Graphical representation of RBF kernel categorizing the data points in 2D and 3D space. 30

Figure 19. Raw data representation and labelling in Excel. 31

Figure 20. Welded samples used for data mining. 32

Figure 21. Weld results classified as Good weld (a) is the front bead, (b) is the back bead, (c) current and (d) voltage characteristics of the weld specimen. 35

Figure 22. Weld results classified as Lack of fusion (a) is the front bead, (b) is the back bead, (c) current and (d) voltage characteristics of the weld specimen. 36

Figure 23. Weld results classified as Burn through (a) is the front bead, (b) is the back bead, (c) current and (d) voltage characteristics of the weld specimen. 37

Figure 24. Algorithm for Two-class classifier. 41

Figure 25. Algorithm for Multi-class classifier. 41

List of Tables

Table 1. Chemical composition of low carbon SS400 steel plate (%). 12

Table 2. Welding parameters for GMAW (Plate thickness 6 mm and 9 mm) 19

Table 3. Parameters for welding wire and shielding gas (All cases) 20

Table 4. Test results of classifiers 40

ABSTRACT

PCA-SVM 분류기에 의한 GMAW 용접시 결함 특성화를 위한 모니터링 시스템에 관한 연구

Use of monitoring system for Weld defect characterization in GMAW by PCA-SVM Classifiers

아스윈 크리슈나스와미 Aswin Krishnaswamy
지도교수: 주성민
조선대학교 선박해양공학과

주요한 소재이용기술로서의 용접공정은 지난 수십 년간 다양한 산업군에서 적용되어 왔다. 본 논문에서는 조선, 플랜트, 보일러 및 자동차 산업과 같은 제조업 분야에서 완성도 높은 핵심 생산 공정으로 자리잡은 용접기술에 최근 폭발적 잠재력을 보여주고 있는 기계학습 기법을 적용하여 용접결함의 검출기법을 고도화 하고자 하였다. 최근 기계학습 기법을 포함한 인공지능 분야의 각 산업계로의 확산적 전파가 가속화 되고 있는 상황이며, 과학기술분야뿐 아니라 생체 의학, 기후, 문화, 사회의 전반에 큰 변혁을 이끌고 있다. 본 연구에서는 GMAW 용접 시에 발생하는 용접결함을 실시간 모니터링 하여 결함의 존재유무와 종류를 선별해 낼 수 있는 혁신적 방법론을 제시하고자 하였다. 이를 위해서 본 연구에서는 용접 시 발생되는 전류와 전압파형을 모니터링 하여 획득한 파형신호를 인식하여 결함의 유무와 종별을 진단하도록 SVM 알고리듬을 작성하여 기계학습을 진행하였다. 기계학습을 위한 SVM 알고리듬은 이진 분류와 다중분류 기법을 적용하였으며 학습데이터에 따른 용접 결함의 인식 및 분류 성공률을 검토하였다. 본 연구에서 개발된 모델을 활용함으로써 GMAW 외의 다른 용접 프로세스에도 충분히 적용이 가능하며 자동화 등 용접분야에서의 다양한 시스템적 응용방식으로 활용이 가능한 접근방식을 제안하고 있다.

1. Theory and background

1.1 Overview

Welding is one of the key technologies which serves an important role of joining process in manufacturing industries such as ship building, automobile and aircraft industry [1]. Gas Metal Arc Welding (GMAW) being one of the most popular welding techniques all around the world. The complex nature of welding makes it challenging to correlate the process variables and parameters in developing a mathematical model. GMAW is an arc welding process which joins the metal by the arc created by the welding transformer with varying voltage and current. The feeding wire which serves as the medium of filler, melts into molten pool by the arc generated in the process resulting in a joint. However, defects tends to form in the welds and makes the joints vulnerable to failure. It is imperative to study the main cause of weld defects by approaching the welding parameters such as voltage, current, welding speed, Contact Tip Work Distance (CTWD), current-voltage arc characteristics etc [2]. Moreover, it is essential to monitor the geometrical parameters simultaneously reason being effecting the weld quality [3].

In recent years there are many research studies on welding with regard to machine learning [4]. Visual inspection of weld defects being the most commonly used technique for detecting the flaw. The use of Non-destructive alternatives for defect recognition have several limitations for the setup in large scale industrial production due to the complexity of the device setup. According to Sergey et al there are not many reliable techniques in real-time defect recognition and the quality control is performed post process and the weld defects such as porosity, undercut, spatter and undercut is not possible to recognize in real-time [5]. The use of X-ray for defect recognition is highly limited in large scale production due to high coat and risk for humans due to the radiation [6-8]. However, there are disadvantages in the visual inspection since imperfection inside the weld cannot be detected with naked eye and there is an elevated risk for human error.

In this study voltage-current monitoring system is tremendously applicable for flaw detection seeing advantage for less human exposure to harmful radiation from X-ray and moderate cost and easy setup of the above system [9]. Several researchers have made attempts in this field of voltage monitoring, Wu et al found that seven statistical parameters involved in monitoring of the wave shows the variations in step disturbance and it is also found that after statistical process control based identifying system, rate for identification of normal and abnormal welding condition was significantly higher [10]. The integration of data-driven monitoring system with machine learning methods has been widely applied manufacturing industries, and most effective being support vector machine (SVM) [11-14].

1.1.1 Gas Metal Arc Welding (GMAW)

Gas Metal Arc Welding is a welding process in which an electric arc forms between a consumable wire electrode and the workpiece, which melts the wire to melt and join. With the wire electrode, a shielding gas is fed through the welding gun to shield the welding process from external conditions. GMAW can be of automatic and semi-automatic. A welding transformer is used to supply constant voltage and direct current power to the welding electrode. There are four different types of metal transfer namely short-circuiting, spray, globular and pulsed-spray. However, in this study short-circuiting type was used. In the case of wire used, there are mainly two type's solid wire and flux-cored wire and solid wire was used in all experiments. The welding gun can be of manual and automatic and due to the process done is automated the automatic welding gun was attached to the automated weld rail system to initiate and end the power supply for the welding process [15, 16]. Schematic representation of proposed model is shown in Figure 1.

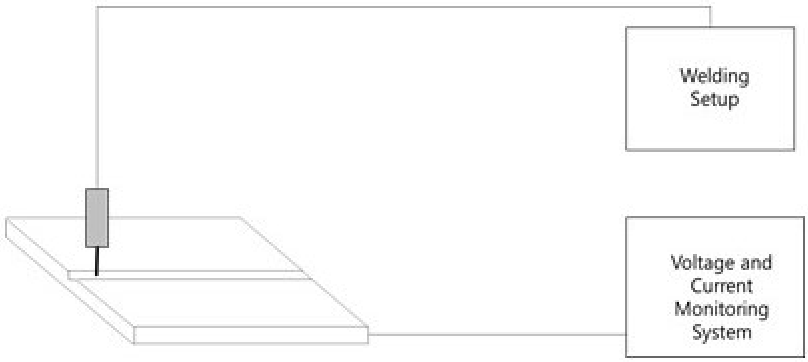


Figure 1. Schematic representation of proposed model.

1.1.2 Support Vector Machine (SVM).

Support Vector Machine is a supervised machine learning algorithm which helps in sorting the data into suitable categories. The algorithm is trained to categorize the datasets SVM is a useful technique for data classification. The main objective of the SVM is to compute and find a high-dimensional plane also called as hyperplane in N-dimensional space (where N is the number of features) and to classify data points distinctively. There are many possible ways to separate the data points in the hyperplane. However, hyperplane with maximum margin between the data points in both classes is selected, increase in margin distance reinforces for the future data points for classification as claimed by Tasadduq et al [17, 18] is shown in Figure 2. The selection of the hyperplane depends on the number of input features, for two input features only one hyperplane is needed if the input parameter is more than two the number of hyperplanes needed is more than one and hence complexity increases. This algorithm can be used for both classification and regression tasks.

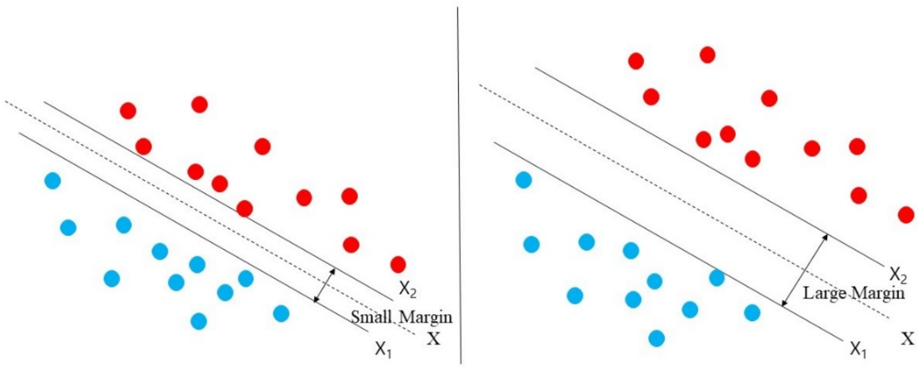


Figure 2. Representation of margin distance in classification.

Even though it’s considered that Neural Networks are easier to use than this, however, sometimes unsatisfactory results are obtained. A classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one target values and several attributes. The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. Classification in SVM is an example of supervised Learning. Known labels help indicate whether the system is performing in a right way or not. This information points to a desired response, validating the accuracy of the system, or be used to help the system learn to act correctly. A step in SVM classification involves identification as which are intimately connected to the known classes. This is called feature selection or feature extraction. Feature selection and SVM classification together have a use even when prediction of unknown samples is not necessary. They can be used to identify key sets which are involved in whatever processes distinguish the classes [12, 18-20].

1.1.2.1 Two class classifier

Classification is the important process in support vector machine, it is a process of differentiating the outcome by classes and in two classifier, classification has been done with two outcomes, which is good weld and weld with defects. This SVM model needs labelled data for the training process. Training process consists of algorithm to analyze the input data and identifies a pattern in multi-dimensional feature. This multi-dimensional is also called as hyperplane.

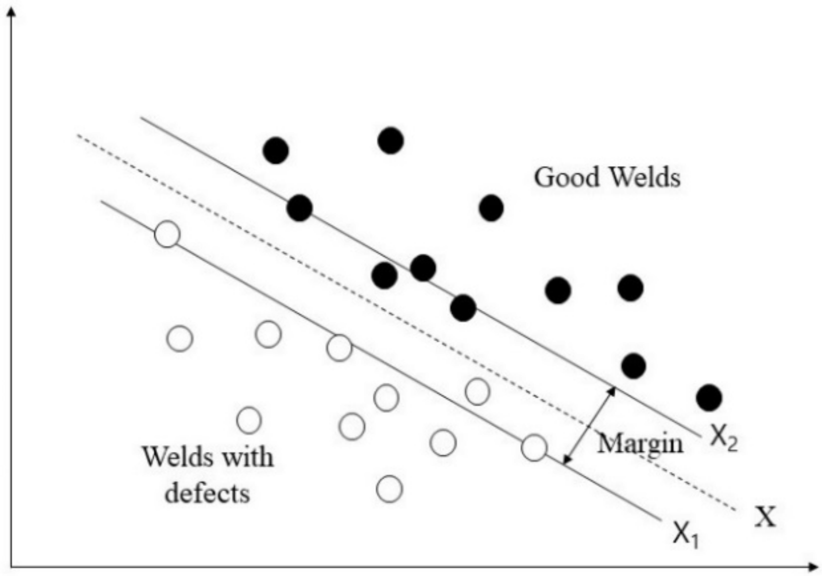


Figure 3. Representation of Two class classifier model.

All the datasets are represented in this plane, the datasets are grouped in certain categories in a way that there is wide margin from the X- hyperplane shown in Figure 2. Yi et al explored in the case of binary classification but had to conclude that the use multiple parameters in this classification doesn't have desired results [20]. The use of PCA in two class classification is effective since the position of the datasets are linear [18]. In this research primarily the model was created to make sure binary classification of weld defect will be recognized with the dataset which was obtained from the experiments. Figure 3 represents the two way classification of the weld data in graphical way for defining the dataset.

1.1.2.2 Multi class classifier

In this type of classifier, more defined weld flaws were classified into burn through, lack of fusion and good welds. The classification were done with the voltage factor which is used for classifying the dataset in this research. The need to use for multiple parameters arises at most cases in classification problem. Initially SVM problems were introduced to solve two-class classification problem. However, SVM only works when datasets in hyperplane is linear, this is when kernel function is used. The important task is section of the suitable kernel for arranging the datasets and computation of separation of the hyperplane in [18, 21].

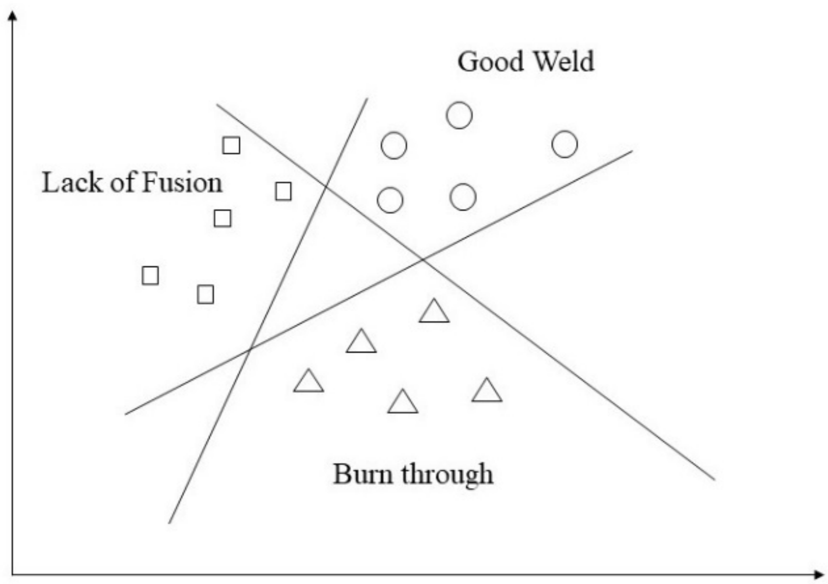


Figure 4. Representation of Multi class classifier model.

The separation of the hyperplane in three classes in this research is depicted in Figure 4. The datasets are in different hyperplane it makes the datasets spread across the high-dimensional space which is converted simple linear form for the algorithm to understand easily. In the case of welding there are multiple weld defects which occur, this multi-class classification is ideal for real world classification problems. Figure 4 shows how the dataset are classified into three types as represented in the graph [22].

1.2 Objectives

Automated welding is the future of welding technology, and data-driven monitoring system controlled by machine learning helps in predicting the defect and fault diagnosis. With all the aspects taken into account from above section, the goals and ideas for this research are set and are as follows:

- To develop a real-time weld defect monitoring system for weld defect recognition.
- To establish an Automated GMAW setup for welding of the work pieces and to acquire the weld results with abnormal or normal welds.
- To determine the relationship between voltage-current changes and weld quality through weld monitoring system.
- To investigate the number of the feature parameters used for cross validation.
- To develop an algorithm with PCA-SVM in two class and multi-class classifiers for defect recognition through feature selection.
- To achieve a successful weld defect recognition model through SVM classifiers in real-time process.

1.3 Thesis layout

The thesis has been organized in following steps, Chapter 2 briefly explains about the working of Gas Metal Arc Welding, parameters which has to be considered while performing joining operation, type of materials used for the experiments and dimensions of the base metals. This chapter also explains the weld defects which are burn through and lack of fusion and reason for selecting these weld defects. Furthermore this chapter describe the concept of Support vector machine (SVM) and the mathematical formula description which aid to understand the logical working of supervised machine learning model. In Chapter 3, the results of welding experiments have been analyzed and the data from the voltage-current monitor were collected. Moreover, an algorithm was designed under SVM and the data collected were used to develop the learning algorithm for weld defect prediction. Finally a new set of welding experiments were performed to test the algorithm. Eventually Chapter 4 concludes the thesis and extends some conclusions from the study.

This study is primarily focused on establishing a real-time weld defect recognition system with the help of voltage-current weld monitoring system and supervised machine learning for immediate defect recognition for industrial applications. Even though weld defects such as porosity, undercut and internal crack cannot be visually recognized, visually recognizable defects were used as the parameters in the study. Voltage and current while welding were used as the data for the algorithms.

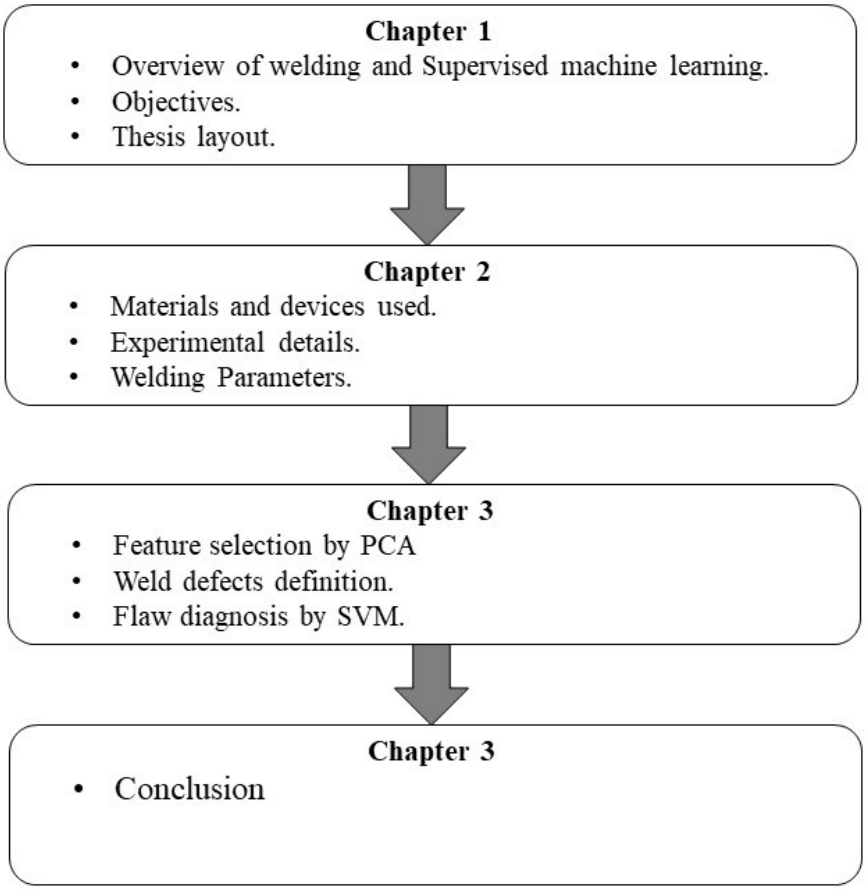


Figure 5. Flow chart of thesis layout

2. Experimental details

2.1 Materials and welding setup

2.1.1 Materials used

Low carbon SS400 steel has been selected for the welding experiment, three types steel plates of dimension were used for the welding experiments. The welding groove were machined to 0°, 30° and 35°, such that the voltage-current weld characteristics of the welds can be determined with the difference in the flow of the consumable welding wire in the welding groove. The length and width of the steel plates are of 500mm x 95 mm and with thickness of 6 mm and 9 mm. The chemical composition of the SS400 steel plates are listed on Table 1.

Table 1. Chemical composition of low carbon SS400 steel plate (%).

C	Si	Mn	P	S
0.12-0.20	0.30	0.30-0.70	0.045	0.045

The schematic representation of the metal plates is shown in Figure 5, the length of the plate is selected to 500 mm to have a reasonable weld time for data extraction.

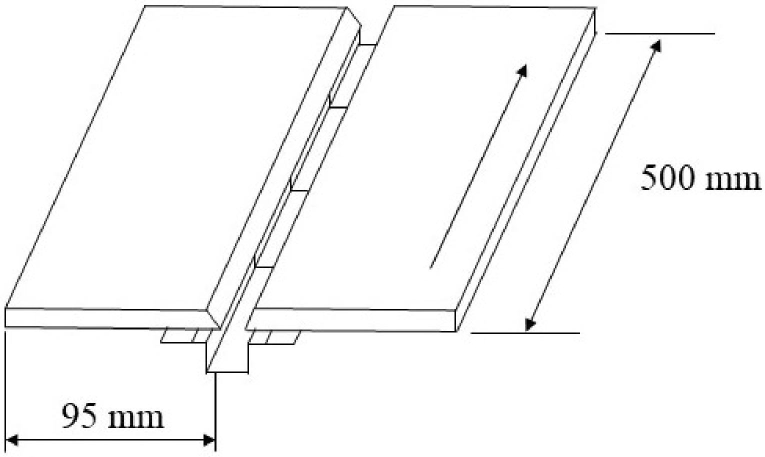
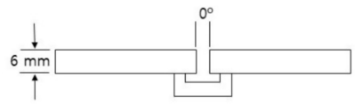
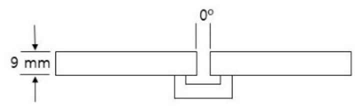


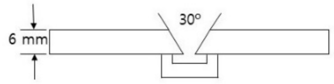
Figure 6. Schematic representation of SS400 steel plate dimensions.



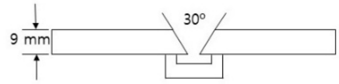
(a)



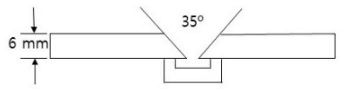
(d)



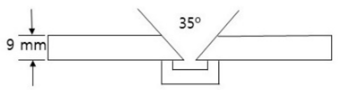
(b)



(e)



(c)



(f)

Figure 7. Schematic representation of SS400 plates, a, b and c represents the metal plates with 6 mm thickness and d, e and f represents with 9 mm thickness.

The machined welding grooves in the respected groove angle and the thickness of the metal plates is schematically represented in Figure 6.

2.1.2 Welding setup

The welding transformer is from Longrun inverter Model. No: 500L-C3 shown in Fig 4.a, the wire feeder is from the Longrun, Fig 4.b shows the carriage setup from Hiweld weld speed controller. The welding jig is made from aluminum block machined to fit a copper backing along the weld line shown in Figure 5. The voltage-current monitoring system is Hiweld weld monitoring system, this system is connected to the welding transformer and the wire feeder to monitor the current, voltage and wire feeding rate. The parameters are monitoring in screen with voltage, current and weeding rate graphs as shown in Fig 6.

(a) (b)

Figure 8. Welding setup a) is the welding transformer and b) is the wire feeder (Longrun Model No. : 500L-C3)

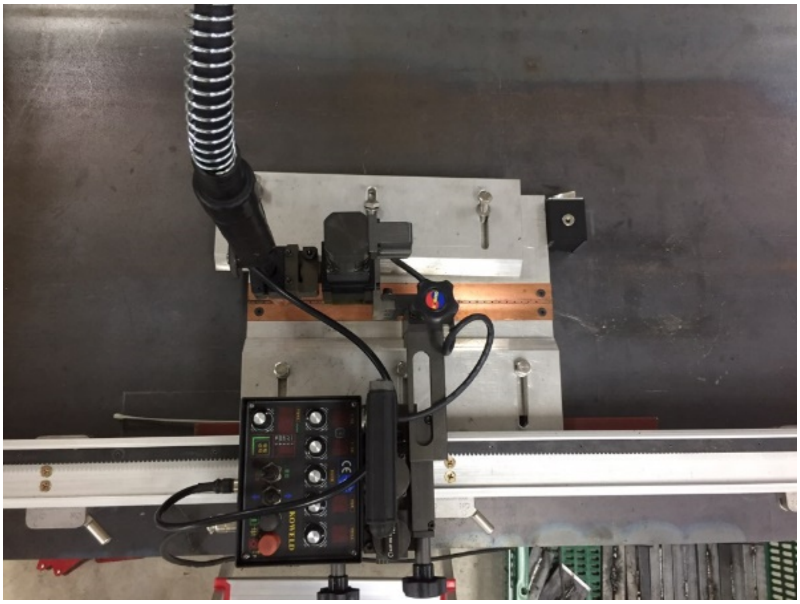


Figure 9. Welding speed controller Hiweld set on gear driven rail mechanism

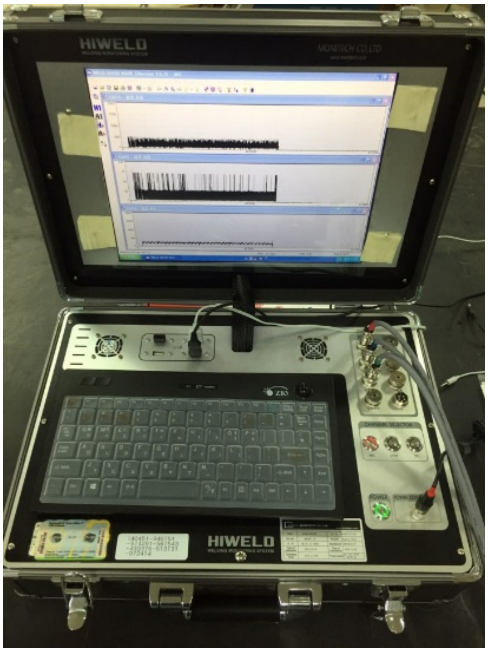


Figure 10. Hiweld weld monitoring system for Voltage-current monitoring.

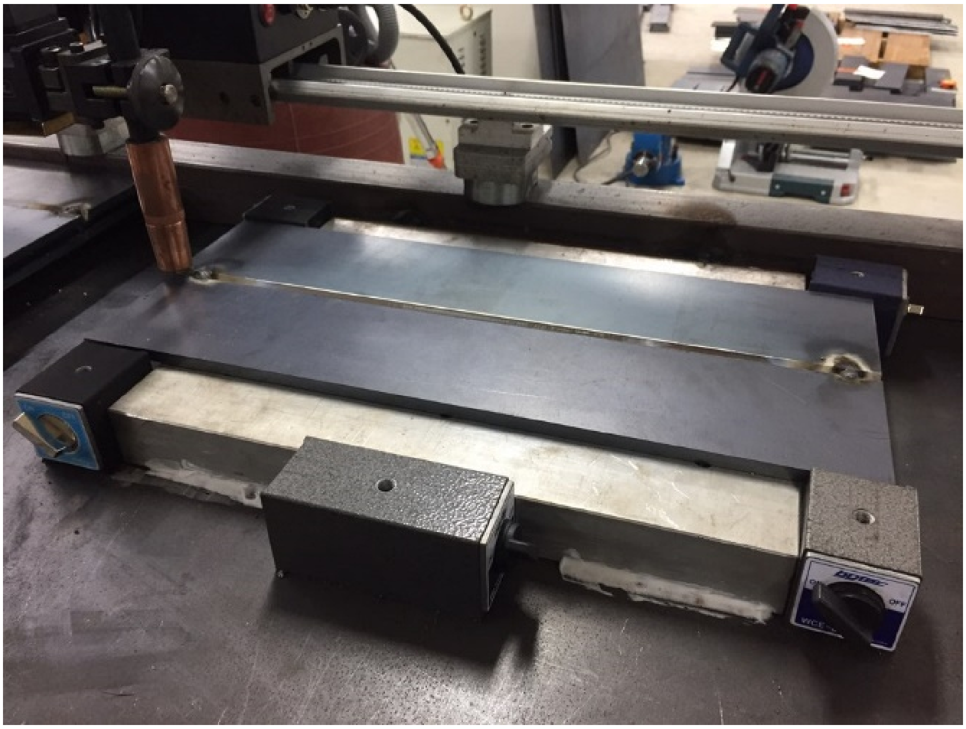


Figure 11. Weld specimen setup in Aluminum block jig with copper strip backing.

Figure 12. SS400 metal plates used for the welding experiments.

The welding specimens are set on aluminum jig with copper backing and the welding specimen are fixed with strong magnets at all sides to reduce distortion while welding process is on progress. The jig setup is shown in Figure 11. In Figure 12 the SS400 metal plates used for the welding are shown in the picture.

2.1.3 Welding parameters

The welding parameters for the weld were selected through hit and trial experiments. Experiments were carried by regulating voltage, current, wire feed rate and weld speed. The parameters are recorded in the Table 2.

Table 2. Welding parameters for GMAW (Plate thickness 6 mm and 9 mm)

Welding Parameters	Values
Position	1G
Voltage (V)	23-26
Current (A)	150-200
Shielding gas (ℓ /min)	15-20
Torch angle (deg)	90
No. of welds	1
Welding speed (cm/min)	45-50
Root thickness (mm)	2
Gap size (mm)	1.0-1.5
Groove angle (deg)	0, 30, 35
CTWD (mm)	15-20
Backing material	Copper

The experiments where conducted in 1G position with a uniform welding speed of 40 cm/min and constant wire feed rate of 350 cm/min. The fused weld were successfully achieved with weld with good welds, burn through and Lack of fusion. All the data were carefully recorded and used for analysis

Table 3. Parameters for welding wire and shielding gas (All cases)

Case	Wire	Shielding gas (ℓ /min)
1G	1.2 Ø metal wire	15-20

The technical specification of wire is AWS A5.18/ ASME SFA-5.18 and the thickness of wire and shielding gas rate is recorded on Table 3.

2.1.4 Experimental Flow

The detailed process of the whole has been simplified into a flow diagram, for easy study in Figure 13. The first stage of the experiment starts with the GMA Welding process, the input parameters current, voltage and weld travel speed are regulated through adjusting former parameters for welding. The results of the welding named as good weld, burn through and lack of fusion, the voltage-current characteristics of the welding is collected from the weld monitor system.

The raw data from monitoring system is recorded and saved in excel file. The results of the welding are labelled to the respective datasets. The second stage of the process involves voltage wave analysis through classification models. However, there are two different classification models used in this research, the same method of data extraction and labeling is used before running the data into respective classification model.

The final stage of this research involves the use of feature extraction where the datasets in higher dimensionality reduced to lower dimension for the computer to easily understand and execute in the algorithm. The final result of the respective model in weld defect recognition is executed at the last.

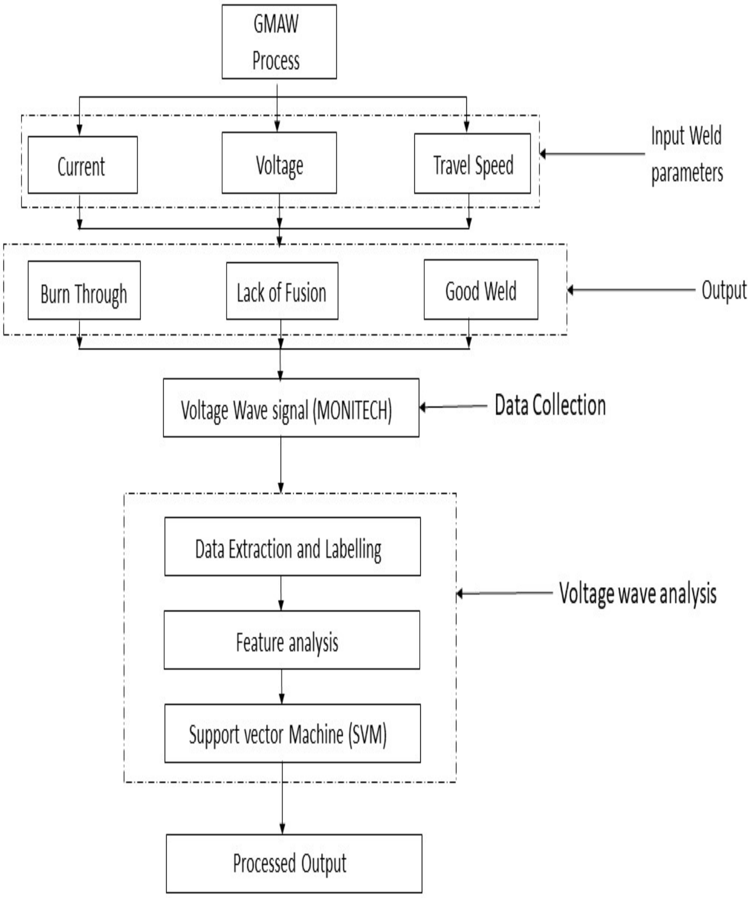


Figure 13. Flow diagram of the process.

2.2 Weld feature extraction

2.2.1 Feature selection by PCA

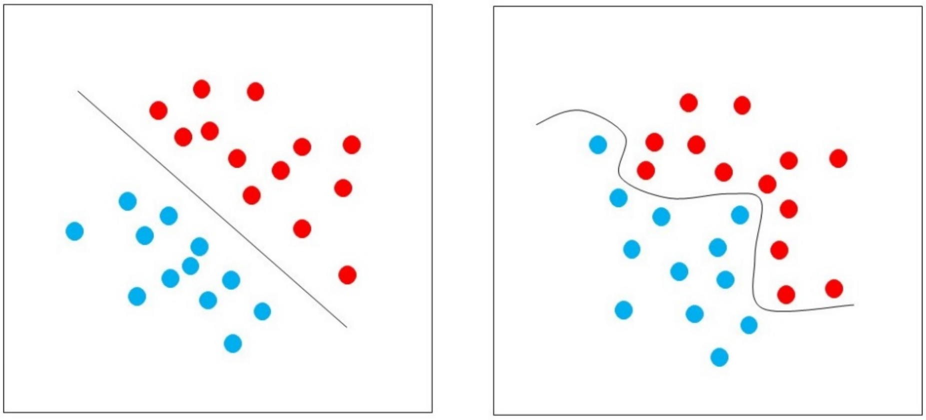
The Principal Component Analysis (PCA) is basically a tool for feature selection and to select and rearrange the features according to the magnitude i.e., from the highest value to the lowest value. The PCA seeks to replace N variable by K < N in uncorrelated linear projections of the original values. Linear combinations are particularly attractive because they are simpler to compute and analytically traceable [23, 24].

Given, $x \in R^N$ find an N x K matrix U such that:

$$y = U^T x \in R^K \text{ where } K<N$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ . \\ . \\ x_N \end{bmatrix} \xrightarrow{f(\mathbf{x})} \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_K \end{bmatrix}$$

The above matrix represents the projection from the N^{th} dimensional space to the K^{th} dimensional space. In mathematical point of view, finding an optimum mapping $\mathbf{y}=f(\mathbf{x})$ is equivalent to optimize an objective function. The data points are randomly spread around the higher dimensional plane, since the feature parameters is more than two the selection of optimum hyperplane is difficult. The use of kernel parameter helps in selecting a hyper plane as shown in Figure 14. When the kernel value is k=1 the hyperplane created by the algorithm is a straight line in linear classification this is not a problem. While k value is increased (k>1) the hyperplane optimizes the feature in a hyperbolic function. When the datasets are non-linear the kernel classification fails. However Radial Basis Kernel (RBF) is effective in this study since RBF only requires two parameters this was achieved by You et al [25].



k=1 k>1

Figure 14. Representation of kernel function

The k principal component are selected by the importance of their explained variances, and each values attributes with the varying degree of every single components. The feature extraction is based on largest variance criteria, where principal component will be used as a new feature instead of the original ones.

The most commonly appeared problem with using PCA is that measurements from all of the original variables are projected in lower dimensional space such that only linear relationships are considered in feature selection based on PCA- based methods [26].

2.2.2 Weld defect definition

The weld defects are defined according to the international standard EN ISO 13919-1 [27]. The flaws were chosen to classify the defects happens in welds. The defects were classified in three type's good weld, burn through and lack of fusion. Even though there are defects such as porosity, undercut and cracks these defects cannot be visually inspected. All those defects which cannot be recognized through visual inspection need more complicated techniques such as X-ray tomography, ultrasonic detector and spectrometer, these setups are expensive and it cannot be used in large scale industries. However, in this study the defect were Figure 15 shows the three class of weld defects happened in the back bead.

2.2.2.1 Good weld

Good weld can be defined as weld with uniform width, which shows no burn marks caused by overheating. The weld should have a good penetration such that back bead is clearly visible and lacks porosity or inclusions. The good weld achieved in this research is shown in Figure 15.

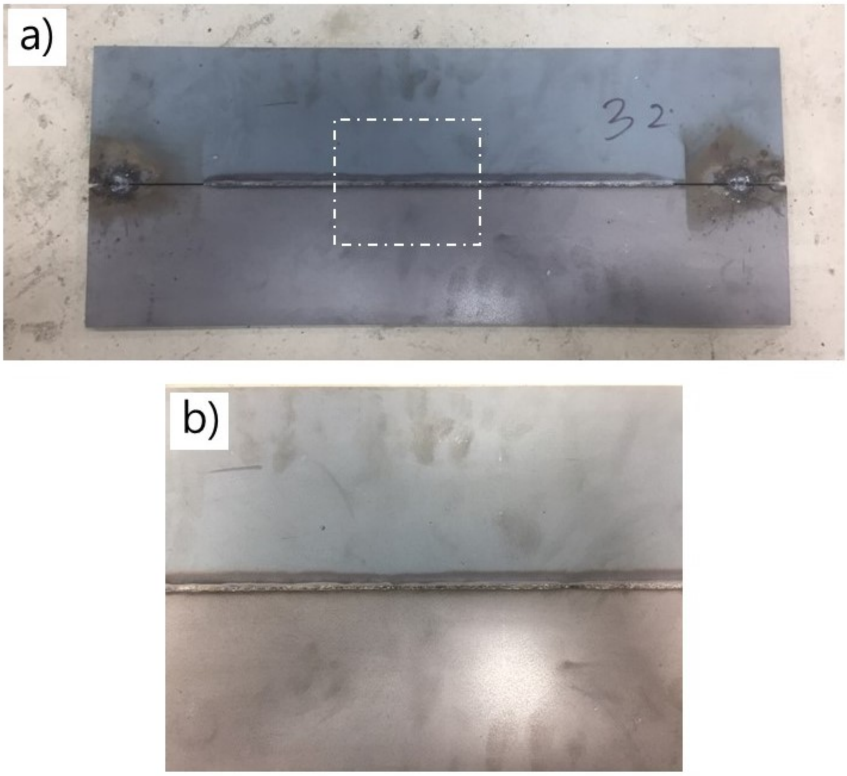


Figure 15. Good weld achieved from the experiments a) being the welded sample and b) is the closer view.

2.2.2.2 Burn Through

Burn through can be defined as collapse of the molten pool due to loss of control or excessive penetration which results in cavity in the root of the weld.

This defect is also caused by excessive amperage during welding which cause a cavity to appear. The burn through sample is shown in Figure 16.

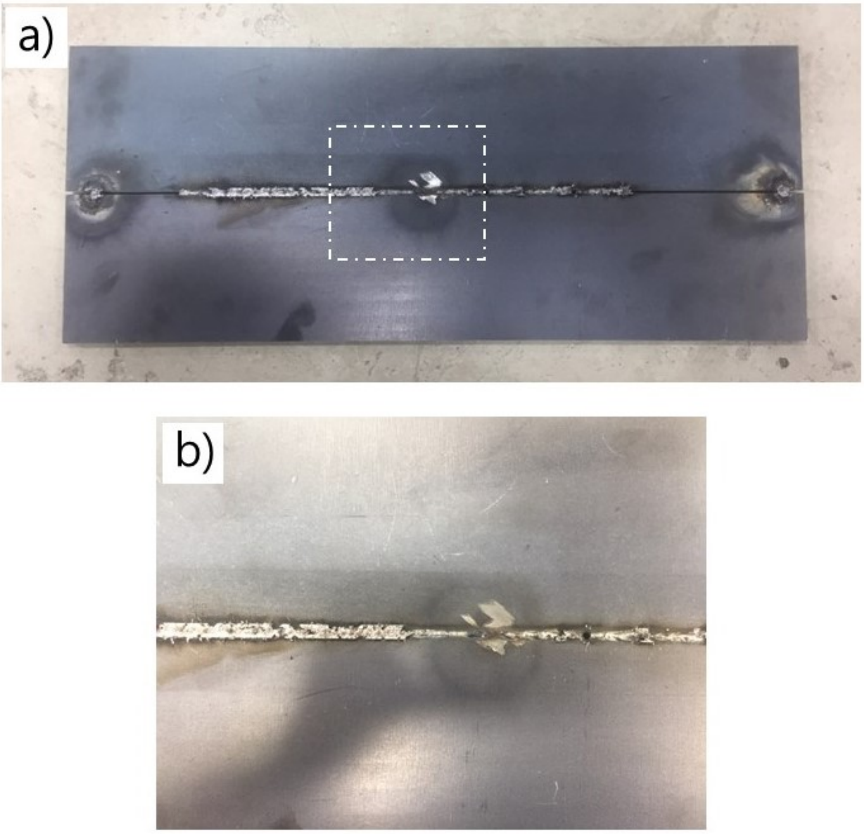


Figure 16. Burn through specimen resulted after the experiment a) being the welding sample and b) is closer view of the weld back bead.

2.2.2.3 Lack of fusion

Lack of fusion or incomplete penetration can be defined as there is poor adhesion in the root of the base metal, the molten pool will not flow till the root of the bead resulting no fusion in the weld root. This is mainly happens when there is less gap clearance between the weld specimens, arc length, current and the angle of the electrode. Figure 17 shows the lack of fusion weld sample.

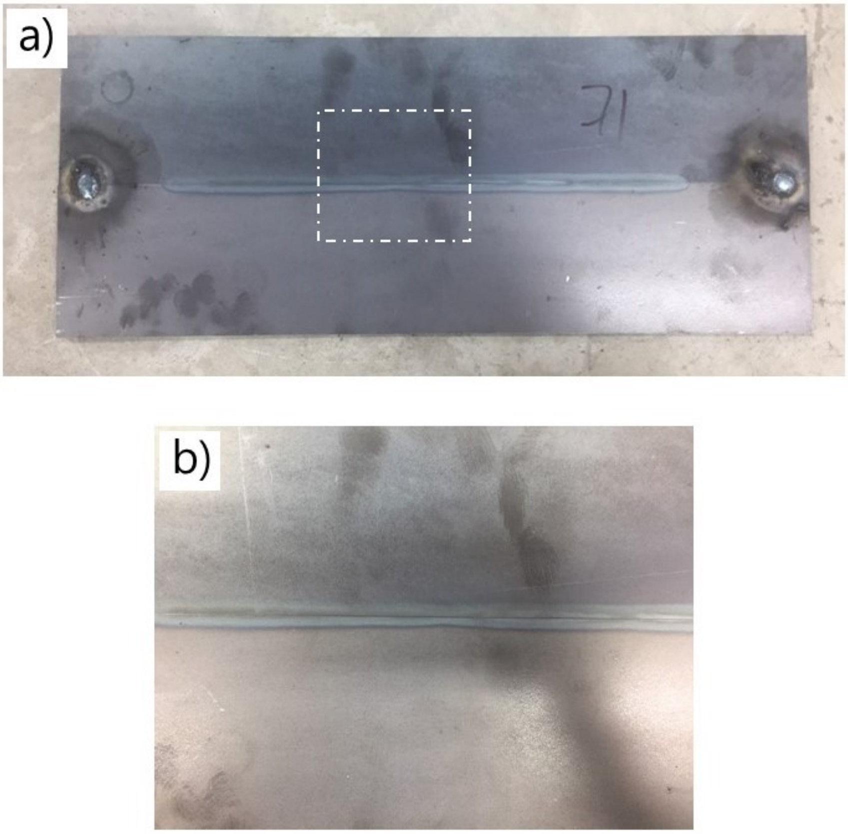


Figure 17. Lack of fusion and incomplete penetration sample obtained a) is the welding sample and b) being the closer view of the back bead .

2.2.3 Flaw diagnosis and detection by SVM

The Support Vector Machine is one of the commonly used for machine learning method for classification, regression and other machine learning techniques. However, in this research use of support vector machine is based on the classification problem [28].

$$\not{f}(s) = sign\left(\sum_{t=1}^{D} \lambda(t) y_{wr}(t) \kappa(S(t)S) + \bar{b}\right)$$

Where, $\not{f}(s)$ is the training sample function, $y_{wr}(t)$ is a class label of the t^{th} training sample $S(t)$, D is the number of the training samples, $\lambda(t)$ is a Lagrangian multiplier, and $\bar{b}$ is a bias. The Kernel function κ used in this paper is a radial basis function (RBF) Kernel [25, 29, 30].

$$\kappa(S(t)S) = exp(-\gamma \|S(t) - S\|^2)$$

Where, γ is a kernel parameter. The RBF Kernel nonlinearly maps samples into higher dimensional space, thus the linear classification can be performed in this space. Gaussian RBF (Radial Basis Kernel) is used in high-dimensional matrix and it is popularly used in SVM due to effectiveness while dealing with large number of features. This Kernel function value depends on distance from the data points to the origin. In this Kernel support vectors has less effect in the hyperplane while in linear Kernel it is exactly the opposite.

However, when the regularization parameter values increase the model get and overfits and if regularization parameter decrease it underfits. As well in kernel when the value increases the model will get overfit and vice versa when value is decreased.

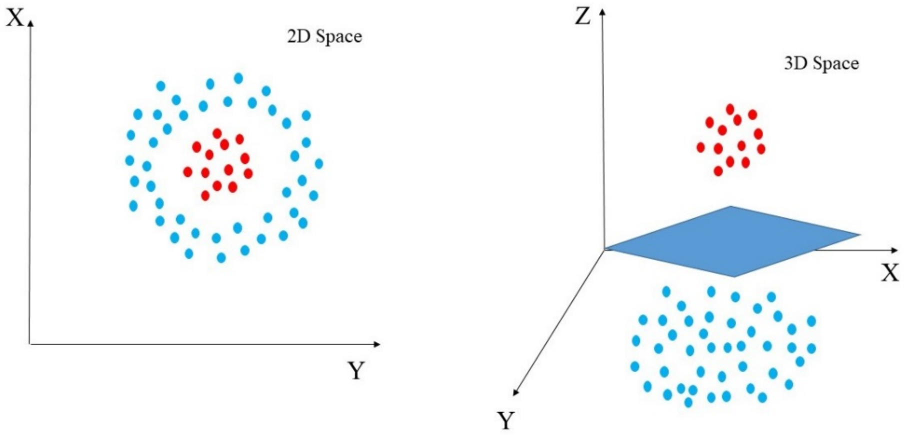


Figure 18. Graphical representation of RBF Kernel categorizing the data points in 2D and 3D space.

After solving by dual optimization problem, the decision function was obtained.

$$\max_{\lambda}\left(\sum_{t=1}^{D}\lambda(t)-\frac{1}{2}\sum_{t,l=1}^{D}\lambda(t)\lambda(l)y_{wr}(t)y_{wr}(l)K(S(t),S)\right)$$

Subject to $\sum_{t=1}^{D}\lambda(t)y_{wr}(t)=0,\qquad 0\leq\lambda(t)\leq C$

Where $S(t)$is the support vector corresponding to non-zero $\lambda(t)$ and C is the regularization constant, these parameters can be obtained by a grid search, because only two parameters, namely, the Kernel parameter γ and the regularization parameter C, have to be optimized.

2.2.4 Algorithm development

The data from the weld monitoring system is converted machine readable format for data input. The data's were arranged in excel format which is shown in Figure 19.

Flaw																	
BT	31.15234	25.92773	26.46484	27.49023	26.31836	24.07227	21.48438	19.82422	19.43359	19.48242	19.62891	20.11719	20.89844	21.72852	22.02148	22.70508	22.31445
BT	25.73242	26.12305	27.73438	28.32031	26.26953	23.38867	21.67969	22.02148	21.43555	20.50781	21.09375	22.55859	24.07227	25.09766	25.68359	25.68359	25.58594
LOF	28.80859	27.44141	28.22266	27.7832	24.85352	21.92383	20.16602	19.72656	19.67773	20.01953	20.80078	22.07031	23.48633	24.6582	25.43945	25.58594	25.04883
GW	24.12109	24.70703	26.07422	25.24414	22.50977	20.01953	18.70117	18.60352	19.28711	19.43359	19.67773	20.21484	20.94727	21.5332	23.48633	24.16992	22.26563
GW	21.67969	21.97266	22.85156	21.82617	19.43359	16.89453	16.21094	16.74805	16.35742	16.79688	17.13867	17.62695	17.82227	18.31055	18.79883	19.62891	19.58008
GW	28.85742	23.04688	23.29102	24.51172	23.53516	21.875	18.79883	18.21289	18.11523	18.4082	18.55469	18.79883	19.53125	20.50781	22.36328	21.77734	21.5332
GW	23.14453	22.55859	23.92578	23.29102	21.04492	19.04297	17.23633	17.67578	17.77344	17.62695	17.82227	18.45703	19.14063	19.72656	20.06836	20.26367	20.75195
GW	22.16797	23.53516	25.58594	25.78125	24.16992	21.09375	19.04297	18.4082	18.11523	18.4082	18.60352	19.38477	20.16602	21.14258	21.14258	21.92383	22.70508
BT	27.97852	26.2207	26.51367	27.05078	25.3418	22.60742	19.82422	19.0918	19.04297	19.38477	19.87305	20.70313	21.77734	22.31445	22.80273	23.29102	23.68164
BT	19.58008	16.16211	15.91797	15.77148	15.625	15.52734	5.66406	3.17383	3.02734	2.92969	2.88086	2.83203	2.92969	2.97852	2.88086	3.02734	3.125
BT	22.99805	20.50781	18.16406	17.87109	18.55469	17.43164	17.13867	17.23633	17.38281	18.06641	18.45703	18.75	20.55664	10.15625	5.95703	5.46875	4.88281
BT	31.64063	35.44922	35.69336	36.03516	36.47461	37.59766	37.93945	38.33008	39.0625	38.7207	38.67188	37.64648	38.18359	37.93945	37.35352	37.64648	37.45117
GW	28.85742	26.66016	26.80664	27.05078	25.58594	22.75391	20.11719	18.99414	18.99414	18.84766	19.04297	20.21484	20.50781	20.99609	21.67969	21.97266	22.75391
BT	26.31836	26.41602	27.83203	28.125	25.53711	22.46094	20.11719	19.38477	19.53125	19.33594	19.82422	20.16602	20.89844	21.48438	22.02148	23.29102	23.19336
LOF	26.26953	23.68164	24.60938	25.19531	23.48633	21.04492	19.23828	18.26172	18.16406	18.65234	18.55469	18.79883	19.33594	20.16602	20.60547	21.43555	21.63086
LOF	20.01953	16.89453	16.55273	16.35742	15.86914	16.06445	15.9668	16.06445	9.27734	3.51563	3.36914	3.41797	3.41797	3.56445	3.56445	3.41797	3.32031
BT	25.3418	23.7793	25.58594	26.36719	25.43945	21.97266	19.28711	18.01758	18.11523	17.96875	18.16406	18.99414	19.0918	19.53125	20.26367	20.6543	21.48438
GW	26.9043	25.29297	26.2207	26.26953	23.68164	20.99609	19.38477	18.70117	18.60352	18.75	20.11719	20.45898	21.43555	22.65625	23.53516	24.21875	25.04883
BT	26.75781	27.49023	28.32031	26.66016	23.87695	21.09375	19.87305	19.72656	19.87305	20.16602	20.60547	21.09375	21.77734	22.41211	23.24219	23.82813	24.26758
BT	25.87891	23.19336	23.87695	24.6582	23.24219	21.04492	18.99414	18.65234	18.60352	18.70117	19.28711	20.45898	21.92383	23.68164	23.33984	23.82813	24.26758
BT	23.4375	24.51172	26.02539	25.92773	23.92578	21.77734	19.72656	18.79883	18.89648	19.43359	20.11719	19.53125	19.33594	20.41016	21.33789	22.02148	22.55859
BT	26.66016	26.85547	27.63672	26.46484	23.7793	21.28906	19.87305	19.72656	19.58008	19.77539	20.36133	21.14258	21.82617	22.41211	22.90039	23.63281	23.73047
BT	17.57813	16.21094	16.30859	16.65039	17.1875	17.28516	17.23633	17.38281	17.72461	16.99219	4.73633	3.85742	4.00391	3.90625	3.75977	3.71094	3.66211

Figure 19. Raw data representation and labelling in Excel.

The algorithm were developed using Python, the coding were done with the help of pandas as module. The classifier were selected as linear kernel classifier for the two class classifier model and RBF for the multi-class classifier model. Principal component analysis were performed in the algorithm for the datasets in both models.

The development of k-Nearest Neighbor (kNN) is on the process, since the model was successful with two-class and multi-class classifier, the predicting accuracy is believed to be higher and the prediction time can also be reduced.

3. Results and Discussion.

3.1 Results of GMAW welding and Data extraction.

A series of weld experiments was carried out with different parameters to demonstrate the ability of this thesis, the weld flaws were identified through weld monitoring system. The physical identification of the flaw was done through visual inspection. A total of 100 sampling data were obtained for this case. 80% of the sample were used for learning and rest 20% were used for testing the developed algorithm. The welded samples is shown in Figure 20. The data's collected were labelled according to the classes.

Figure 20. Welded samples used for data mining.

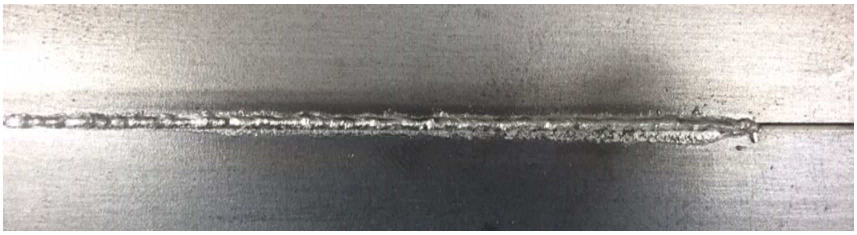
The welding line were perfectly calibrated and weld line were kept inline through all experiments. The results of the welding were categorized in good weld, burn through and lack of fusion. All welded specimens were tediously inspected in back bead and front bead through visual inspection. Figure 21 (a) shows the closer view of the front bead, the weld line is straight and the molten wire is uniformly spread in the weld groove, Figure 21 (b) represents the uniform back bead achieved. The voltage and current graph in Figure 21 (c) & (d) respectively, the current-voltage relation in good weld is that current were higher than 500 A at the start and end of the experiment, after 0.5 sec of welding the current were maintained below 500 A. Even though the current were higher it is believed that heat generated in the welding caused distortion slightly changing the weld line in the range of few millimeter. In the case of voltage graph the same trends of voltage spikes were recorded at start and end of the experiments, there were less spikes in the middle. Shashin et al came to conclusion that the spikes in the monitoring is caused by the thermal distortion [31, 32]. The author successfully found a relation between current and voltage in good welds if the current is below 500 A and voltage is between 24-28 V.

The weld results of lack of fusion for back bead, front bead, current and voltage were depicted in Figure 22 (a), (b), (c) and (d) respectively. The back bead in this weld were not clearly visible and quality of front bead was good. Moreover, the current were managed under 500 A at most times apart from start and end of the welding. Evidently the spikes in voltage wave signal were continuous through the experiment. Visible burn mark can be seem on the front bead. When the current is under 500A and voltage is higher than 24-28 V this relation can be found for this defect [31, 32].

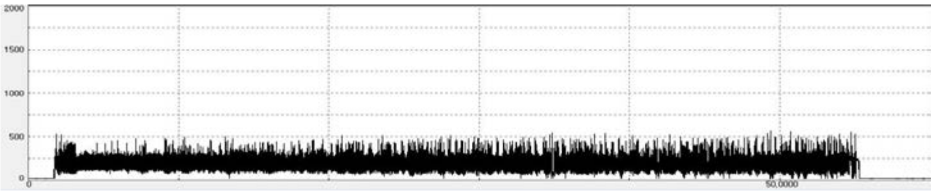
Furthermore burn through weld results also have a significant relation between current and voltage, where voltage signal fluctuations are most than in good weld in Figure 23 (c). However, the current were above 500 A throughout the welding and hence voltage and current for burn through relation can be brought when voltage spikes are higher than good weld voltage value and current is higher than 500 A, this relation can be brought for the burn through defect.

The voltage-current wave characteristics relation for weld different weld defects is used for labeling the defects according to their character.

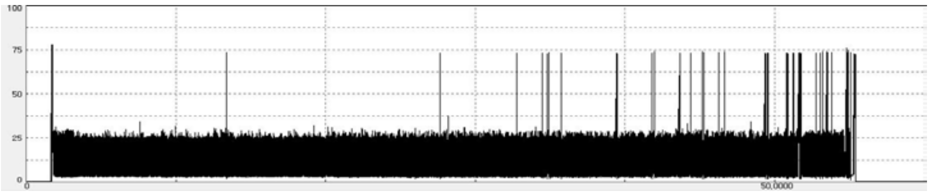
(a)



(b)

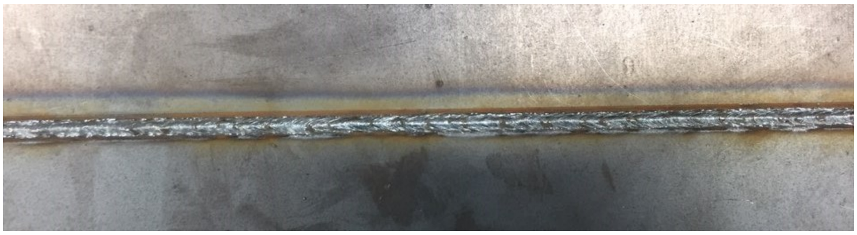


(c)



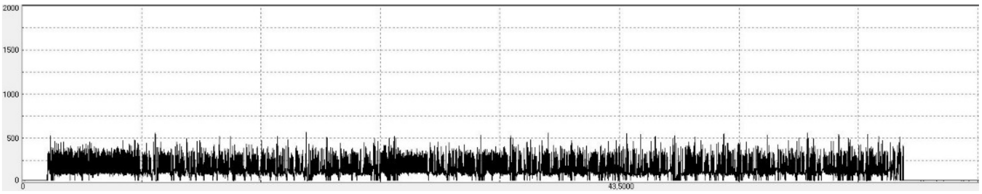
(d)

Figure 21. Weld results classified as Good weld (a) is the front bead, (b) is the back bead, (c) current and (d) voltage characteristics of the weld specimen.

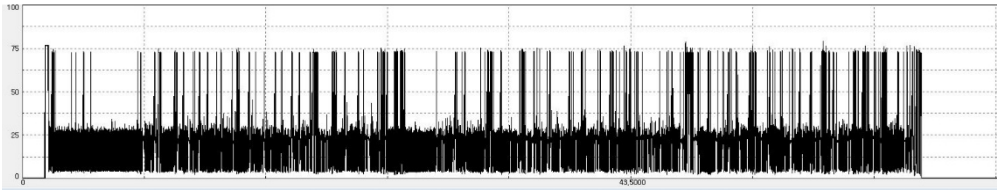


(a)

(b)



(c)

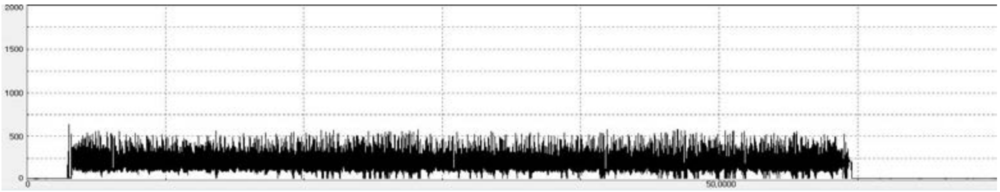


(d)

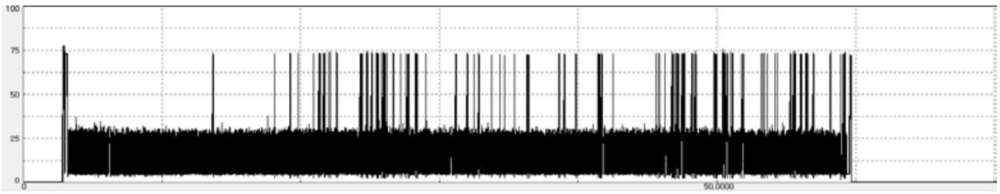
Figure 22. Weld results classified as Lack of fusion (a) is the front bead, (b) is the back bead, (c) current and (d) voltage characteristics of the weld specimen.

(a)

(b)



(c)



(d)

Figure 23. Weld results classified as Burn through (a) is the front bead, (b) is the back bead, (c) current and (d) voltage characteristics of the weld specimen.

3.2 Parameter selection.

The feature selection were done through the monitored data and the two class SVM classifier were used for the prediction of Good and bad classes. Later the same procedure were done with multi class SVM classifier for detailed model to recognize the defects as Good weld, Burn through and Lack of fusion.. The raw data of the voltage and current were grouped and a set of 80 sampling from the whole welded samples with class 1, class 2 and class 3 flaws were used to learn the machine in different subbands. The results of the accuracy were tested through PCA analysis [13-15].

3.2.1 PCA with Two class classifier

The used parameters in the two class classifier is set. The kernel parameter of the SVM was set at 0.125 and the regularization parameter was set at 512. The optimal parameter, subband and principal component (PC) numbers were determined by three fold cross validation. The mean classification accuracy of the SVM model were measured [18-20]. Multiple Correspondence Analysis (MCA) helps in dimensionality reduction with the uncorrelated variables where PCA technique is not possible [11, 12, 16].

$$\mathrm{MCA} = \frac{1}{5}\sum_{j=1}^{5}\left(\frac{a_j}{n_t} \times 100\right)$$

Where, e_{ij} and a_j are the prediction error and successful classification number of j fold cross validation. Table 4 shows the accuracy of the SVM model. For SVM higher accuracy were achieved when training data is increased which is shown in Table 4. As the number of sampling data is increased, the accuracy of the model can maintained at higher levels.

3.2.2 PCA with Multi-class classifier

In the case of this classifier, the data sets are non-liner and are complex in nature, this causes use of kernel as Radial Basis Kernel (RBF) [29]. The training and testing of the data was set at 80% for training purpose and 20% for testing.

$$\mathrm{k}(x_i, x_j) = exp\left(-\gamma\|x_i - x_j\|^2\right)$$

Where γ is the Kernel parameter, the use of RBF Kernel makes the datasets which is non-linear to simple linear phase and the linear classification can be performed. The Kernel is defined as RBF with value of γ as 0.001 and C as 1000 where C is the regularization parameter. Table 4 shows the precision of the model [11, 25, 26].

Three welding experiments were presented in order to prove the accuracy and effectiveness of the proposed model. The experiment included three geometrical parameters (good weld, blow through and lack of fusion). The three class classifier SVM model gave accurate predictions on most of the welding.

3.3 . Process and results of development of machine learning algorithms

The algorithm were developed through pandas data frame for both two-class and multi-class classifier algorithm. The Eigen vectors in the proposed model is from the three signals of voltage and current signals from the weld monitoring system. The wavelength features in voltage were used with fivefold cross validation with PCA. The domain features of eigenvector were extracted with time factor. Two sensors were used in the weld monitoring system.

To enhance classification accuracy datasets for training were increased in the case of two-class classifier it increased from 13.53% to 21.30%. As in this classifier since weld class is defined as good or bad welds, the model failed to identify some welds in good weld class because of the voltage characteristics is similar to weld with defects even though welds contained defects. Figure 24 shows the algorithm if two-class classifier. In the case of multi-class classifier the training accuracy of three weld classes were 73.8 % and testing accuracy is 68.5%. The reason of decrease in the testing accuracy is caused by data which have noise which confuses the algorithm predicting good as bad welds and vice versa. This problem arises when the data is insufficient. The algorithm for this classifier is shown in Figure 25. The test result in accuracy is tabled in Table 4.

Table 4. Test results of classifiers

		Classifier	Accuracy
SVM	PCA	Two Class	21.30%
	PCA	Multi Class	68.5 %

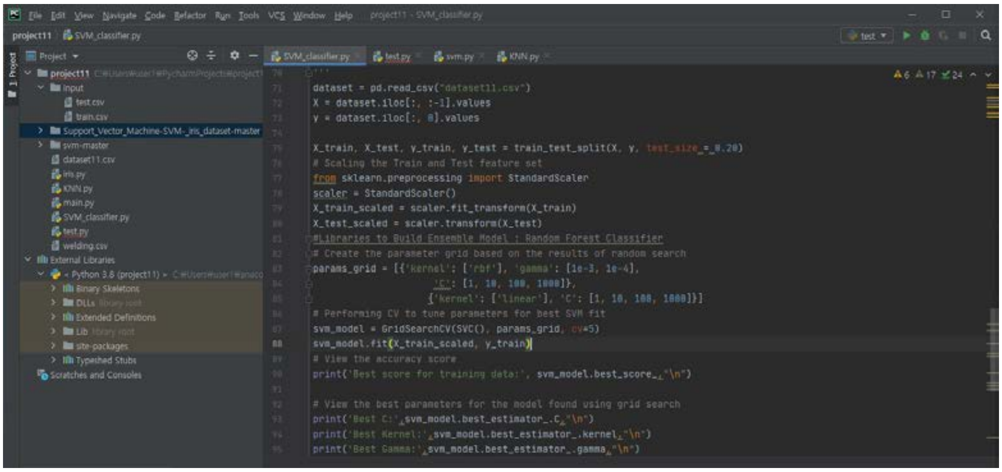


Figure 24. Algorithm for Two-class classifier.

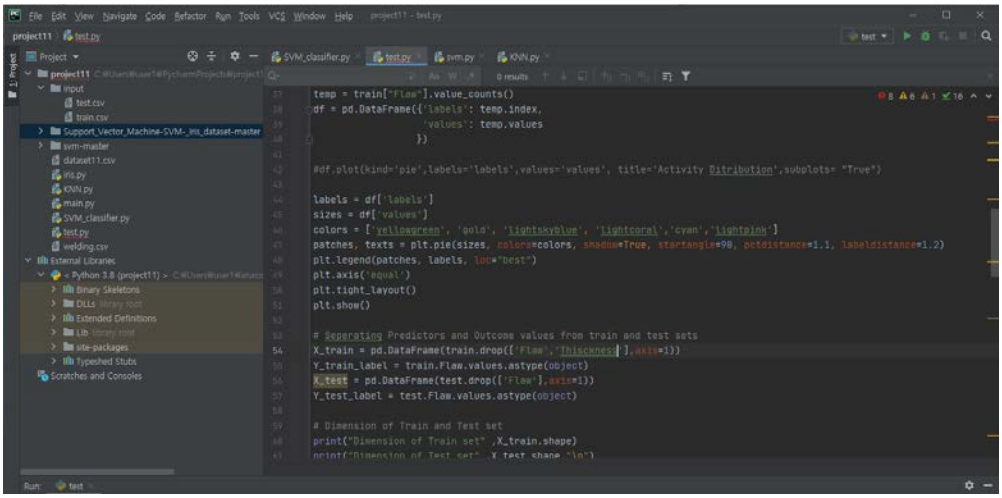


Figure 25. Algorithm for Multi-class classifier.

4. Conclusion

This study has proposed an innovative way to approach and realize the effect of GMAW process monitoring and weld defect diagnosis through PCA-SVM classifiers. The proposed method uses voltage-current monitoring system instead of X-ray visual sensor which is harmful for the human environment and not feasible setup in large scale production. The monitoring system helped to extract the voltage and current values of the welding status. The welded samples were thoroughly subjected to visual inspection. In addition, a two-class and multi-class SVM classifier model was successfully established with signal feature. The values from the monitoring system was successfully extracted by PCA. In addition the defect recognition accuracy had a significant increase from 21.3% to 68.5% when more defined multi-class algorithm was used in recognition with the three feature parameters which enables the algorithm to categorize the data efficiently. Experimental results shows that with the low-cost monitoring system, effective estimation and classification on the welding status and defects can be realized. The detailed training accuracies for good weld, burn through and lack of fusion were 97%, 70% and 94% respectively. It should be noted that training accuracy in burn through is lesser which depends on the presence of noise factor occurred while welding experiment. However, the precision of the model can be increased by increasing the dataset and the feature parameters. The current research has successfully created a model of defect recognition by PCA-SVM classifiers through weld monitoring system

References

1. Xiong, J. and G. Zhang, *Online measurement of bead geometry in GMAW-based additive manufacturing using passive vision.* Measurement Science and Technology, 2013. **24**(11): p. 115103.
2. Madigan, R., *Arc sensing for defects in constant-voltage gas metal arc welding.* Weld J, 1999. **78**: p. 322S-328S.
3. Juang, S. and Y. Tarng, *Process parameter selection for optimizing the weld pool geometry in the tungsten inert gas welding of stainless steel.* Journal of materials processing technology, 2002. **122**(1): p. 33-37.
4. Sumesh, A., et al., *Use of Machine Learning Algorithms for Weld Quality Monitoring using Acoustic Signature.* Procedia Computer Science, 2015. **50**: p. 316-322.
5. Shevchik, S.A., et al., *Laser welding quality monitoring via graph support vector machine with data adaptive kernel.* IEEE Access, 2019. **7**: p. 93108-93122.
6. Martin, A.A., et al., *Ultrafast dynamics of laser-metal interactions in additive manufacturing alloys captured by in situ X-ray imaging.* Materials Today Advances, 2019. **1**: p. 100002.
7. Thompson, M.K., et al., *Design for Additive Manufacturing: Trends, opportunities, considerations, and constraints.* CIRP annals, 2016. **65**(2): p. 737-760.
8. Roy, R.B., et al., *Weld defect identification in friction stir welding through optimized wavelet transformation of signals and validation through X-ray micro-CT scan.* The International Journal of Advanced Manufacturing Technology, 2018. **99**(1-4): p. 623-633.
9. Hu, S.J., et al., *Online monitoring system and method for a short-circuiting gas metal arc welding process*. 2004, Google Patents.
10. Wu, C., J. Gao, and J. Hu, *Real-time sensing and monitoring in robotic gas metal arc welding.* Measurement science and technology, 2006. **18**(1): p. 303.
11. You, D., X. Gao, and S. Katayama, *WPD-PCA-Based Laser Welding Process Monitoring and Defects Diagnosis by Using FNN and SVM.* IEEE Transactions on Industrial Electronics, 2015. **62**(1): p. 628-636.
12. Scholkopf, B. and A.J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. 2018: Adaptive Computation and Machine Learning series.
13. Burges, C.J., *A tutorial on support vector machines for pattern recognition.* Data mining and knowledge discovery, 1998. **2**(2): p. 121-167.

14. Wang, T., et al., *Real-time monitoring for disk laser welding based on feature selection and SVM.* Applied Sciences, 2017. **7**(9): p. 884.
15. Ibrahim, I.A., et al., *The Effect of Gas Metal Arc Welding (GMAW) processes on different welding parameters.* Procedia Engineering, 2012. **41**: p. 1502-1506.
16. Fratiello, P.J., D.K. Hartman, and L.E. Seufer, *Method and apparatus for GMAW welding*. 1992, Google Patents.
17. Imam, T., K.M. Ting, and J. Kamruzzaman. *z-SVM: An SVM for improved classification of imbalanced data*. in *Australasian Joint Conference on Artificial Intelligence*. 2006. Springer.
18. Kallas, M., et al. *Multi-class SVM classification combined with kernel PCA feature extraction of ECG signals*. in *2012 19th International Conference on Telecommunications (ICT)*. 2012. IEEE.
19. Mathur, A. and G.M. Foody, *Multiclass and binary SVM classification: Implications for training and classification users.* IEEE Geoscience and remote sensing letters, 2008. **5**(2): p. 241-245.
20. Liu, Y. and Y.F. Zheng. *One-against-all multi-class SVM classification using reliability measures*. in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005*. 2005. IEEE.
21. Chamasemani, F.F. and Y.P. Singh. *Multi-class support vector machine (SVM) classifiers--an application in hypothyroid detection and classification*. in *2011 Sixth International Conference on Bio-Inspired Computing: Theories and Applications*. 2011. IEEE.
22. Valavanis, I. and D. Kosmopoulos, *Multiclass defect detection and classification in weld radiographic images using geometric and texture features.* Expert Systems with Applications, 2010. **37**(12): p. 7606-7614.
23. Shih, J.-S., Y.-F. Tzeng, and J.-B. Yang, *Principal component analysis for multiple quality characteristics optimization of metal inert gas welding aluminum foam plate.* Materials & Design, 2011. **32**(3): p. 1253-1261.
24. Datta, S., et al., *Application of PCA-based hybrid Taguchi method for correlated multicriteria optimization of submerged arc weld: a case study.* The International Journal of Advanced Manufacturing Technology, 2009. **45**(3-4): p. 276.
25. You, D., X. Gao, and S. Katayama, *Multisensor fusion system for monitoring high-power disk laser welding using support vector machine.* IEEE Transactions on Industrial Informatics, 2014. **10**(2): p. 1285-1295.

26. Wang, J., et al., *Feature extraction in welding penetration monitoring with arc sound signals.* Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2011. **225**(9): p. 1683-1691.
27. Hobbacher, A. and M. Kassner, *On relation between fatigue properties of welded joints, quality criteria and groups in ISO 5817.* Welding in the World, 2012. **56**(11-12): p. 153-169.
28. Chen, B., et al. *A study of welding process modeling based on Support Vector Machines*. in *Proceedings of 2011 International Conference on Computer Science and Network Technology*. 2011. IEEE.
29. Chen, Y., H.-W. Ma, and M. Dong, *Automatic classification of welding defects from ultrasonic signals using an SVM-based RBF neural network approach.* Insight-Non-Destructive Testing and Condition Monitoring, 2018. **60**(4): p. 194-199.
30. Fan, J., et al., *Automatic recognition system of welding seam type based on SVM method.* The International Journal of Advanced Manufacturing Technology, 2017. **92**(1-4): p. 989-999.
31. Vyas, S.M., S.I. Shah, and G. Acharya, *Minimization of Distortion During Gas Metal Arc Welding Process: A Review*. Journal of Experimental & Applied Mechanics, 2017. **8**(3): p. 48-53.
32. Ribeiro, R.A., et al., *Application of cold wire gas metal arc welding for narrow gap welding (NGW) of high strength low alloy steel.* Materials, 2019. **12**(3): p. 335.