IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)
e-ISSN: 2278-1684,p-ISSN: 2320-334X, Volume 20, Issue 6 Ser. II (Nov. – Dec. 2023), PP 53-62
www.iosrjournals.org

Corrosion Detection By Using MachineLearning

Pranav Nair
Department Of Mechanical EngineeringVishwakarma Institute Of Technology Pune, India

Abstract
Corrosion is a prevalent issue in the mechanical industries. Corrosion gets developed because of environmental variables such as temperature, humidity and acidic nature of the liquids. There are different techniques for detecting and monitoring corrosion development, both destructive and non-destructive. Visual inspection is a common technique of surface corrosion analysis, but manual inspection is extremely dependent on the inspecting person's abilities and expertise. The findings of the manual inspection are qualitative and may be biased, may result into the accidents because of incorrect analysis. Corrosion must be accurately detected in early phases to prevent unwanted accidents.

Date of Submission: 18-12-2023 Date of Acceptance: 28-12-2023

I. Introduction

Corrosion of metals and metallic structures is a worldwide problem. Because of corrosion there is huge loss of property, infrastructure. Corrosion of metal takes place because of purity of metal. The energy of metal in purest form is very high so it has a natural tendency to convert in to metallic oxide, or salt of many type depending upon the environmental condition or process environment. Because of conversion of metal into various salt and oxide, metal acquires a stable position in the environment. Failures of metal component or a structure occurs if it is does not able to manage stresses imposed during operation or use of.

One of the example of accident due to corrosion is the Bhopal gas leakage .It is probably the biggest industrial disaster in Indian and world history. Union Carbide India Limited (UCIL) pesticide plant was manufacturing phosgene, monomethyl amine (MMA) methyl isocyanate (MIC) and the pesticide carbaryl, known as Sevin. This plant was located at a distance of 5-6 kilometers from the center of Bhopal, the state of Madhya Pradesh. The incident took place in the night between 2nd and 3 rd December 1984. Many technical experts and specialist in the field of Corrosion claimed that this incident took place because of corrosion in various machine parts of MIC plant3 .It was found that most of the plant's MIC related safety systems were malfunctioning and many valves and lines were in poor condition.

Therefore, this feasibility study has focused on automatic corrosion detection. This project created an autonomous classifier that enabled detection of rust present in pictures. The challenge associated with this approach was the fact that the rust has no defined shape and colour. Also, the changing landscape and the presence of misleading object may lead to miss-classification of the images. Furthermore, the classification process should still be relatively fast in order to be able to process large amount of images in a reasonable time.

II. Literature review

We have gone through different projects and papers related to this topic. The first one was Image Classification via Support Vector Machine 1 Information and Engineering College, Capital Normal University, Beijing 100048, P. R. China 2 Agresearch Ltd, New Zealand, the objective of the above paper was some discussions of the support vector machine's theory are introduced firstly, then the solving process of the support vector machine objective function is shown. Secondly, in order to accomplish the task of classification, the images in the experiences are dealt with. At last getting help from the support vector machines, the images are classified successfully

Research Article; Image Processing-Based Pitting Corrosion Detection Using Metaheuristic Optimized Multilevel Image Thresholding and Machine-Learning Approaches was our on second priority to study. For metal infrastructure elements, corrosion negatively affects their durability and operability. It is reported that corrosion is a dominant form of defects with 42% of frequency of failure mechanisms in engineering structures . therefore, recognition as well as diagnostics of corroded areas is an important task in periodic structural heath surveys. e surveys' outcome significantly helps owners or maintenance agencies to judge the effectiveness of the currently employed protection methods and to prioritize rectifying measures .

The next paper that we have studied for this topic is Surface Corrosion Grade Classification using Convolution Neural Network, Sanjay Kumar Ahuja, Manoj Kumar Shukla, Kiran Kumar Ravulakollu so some important point we have studied is Corrosion is a prevalent issue in the oil and gas industry. Usually, pipelines made of Iron are used for oil and gas transportation. The pipelines are large and distributed over big fields above the ground, underground and even underwater. Corrosion gets developed because of environmental variables such as temperature, humidity and acidic nature of the liquids. There are different techniques for detecting and monitoring corrosion development, both destructive and non-destructive. Visual inspection is a common technique of surface corrosion analysis, but manual inspection is extremely dependent on the inspecting person's abilities and expertise. The findings of the manual inspection are qualitative and may be biased, may result into the accidents because of incorrect analysis. Corrosion must be accurately detected in early phases to prevent unwanted accidents. This paper will present a computer vision-based approach in combination with deep learning for corrosion classification as perISO-8501 standard. The findings of the assessment are unbiased and in a fair acceptable range similar to the outcomes of the visual inspection.

corrosion detection using A.I. : a comparison of standard computer vision techniques and deep learning model, In this paper we present a comparison between standard computer vision techniques and Deep Learning approach for automatic metal corrosion (rust) detection. For the classic approach, a classification based on the number of pixels containing specific red components has been utilized. The code written in Python used OpenCV libraries to compute and categorize the images. For the Deep Learning approach, we chose Caffe, a powerful framework developed at "Berkeley Vision and Learning Center" (BVLC). The test has been performed by classifying images and calculating the total accuracy for the two different approaches. The results show a few interesting facts about the two approaches. The OpenCV based model showed a total accuracy (in all the images) of 69%. According to our expectations, it presented a reduced accuracy (57%) for the "non-rust" classification, while it had great accuracy for the "rust" classification (almost 90%). The reason for this is pretty clear: all the "rusty" images had red components, so it was easy for the algorithm to detect it. However, for the "non-rust" class, the presence of red pixels does not necessary imply the presence of rust. So when we pass a red apple picture, the model just detected red component and misclassified it as "rust", reducing the "non-rust" accuracy. All the four pictures in Figure 1 for example, have been classified by the OpenCV algorithm as "rust", while only two of them are actually correct. The few false negatives involved (where there was rust but it was not correctly detected), seemed were due mainly to the bad illumination of the image, problems associated with colour (we also provided few out of focus test images), or the rust spot was too small (less than 0.3% of the image).

The next paper that we have studied for this topic is Surface Corrosion Grade Classification using Convolution Neural Network, Sanjay Kumar Ahuja, Manoj Kumar Shukla, Kiran Kumar Ravulakollu so some important point we have studied is Corrosion is a prevalent issue in the oil and gas industry. Usually, pipelines made of Iron are used for oil and gas transportation. The pipelines are large and distributed over big fields above the ground, underground and even underwater. Corrosion gets developed because of environmental variables such as temperature, humidity and acidic nature of the liquids. There are different techniques for detecting and monitoring corrosion development, both destructive and non-destructive. Visual inspection is a common technique of surface corrosion analysis, but manual inspection is extremely dependent on the inspecting person's abilities and expertise. The findings of the manual inspection are qualitative and may be biased, may result into the accidents because of incorrect analysis. Corrosion must be accurately detected in early phases to prevent unwanted accidents. This paper will present a computer vision-based approach in combination with deep learning for corrosion classification as perISO-8501 standard. The findings of the assessment are unbiased and in a fair acceptable range similar to the outcomes of the visual inspection

III. Background

Image classification

There are various approaches for image classification. Most of classifiers, such as maximum likelihood, minimum distance, neural network, decision tree, and support vector machine, are making a definitive decision about the land cover class and require a training sample. On the contrary, clustering based algorithm, e.g. K-mean, K-NN or ISODATA, are unsupervised classifier, and fuzzy-set classifier are soft classification providing more information and potentially a more accurate result. The aim of this study is to applying Support Vector Machine (SVM) for image classification.

The main steps in the image classification process are shown in the following diagram:

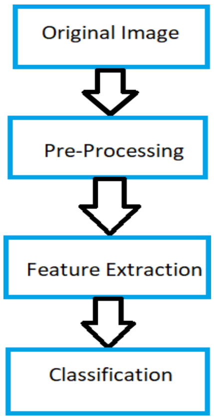
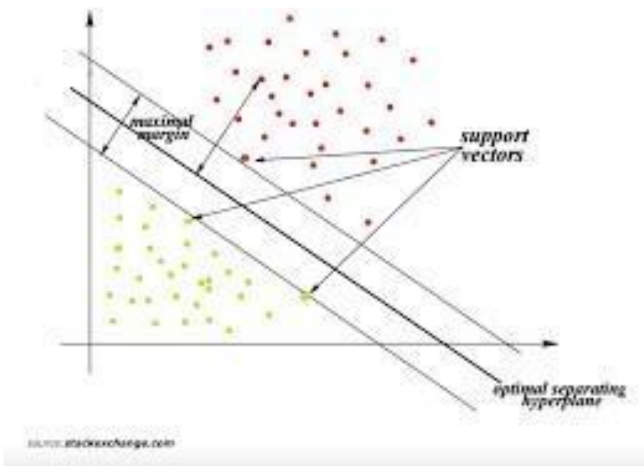


Image 1

IV. THEORIES OF SUPPORT VECTOR MACHINE

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.



SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

SVM can be of two types:

Linear SVM: Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.

Non-linear SVM: Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

Hyperplane and Support Vectors in the SVM algorithm:

Hyperplane: There can be multiple lines/decision boundaries to segregate the classes in n- dimensional space, but we need to find out the best decision boundary that helps to classify the data points. This best boundary is known as the hyperplane of SVM.

The dimensions of the hyperplane depend on the features present in the dataset, which means if there are 2 features (as shown in image), then hyperplane will be a straight line. And if there are 3 features, then hyperplane will be a 2-dimension plane.

We always create a hyperplane that has a maximum margin, which means the maximum distance between the data points.

Support Vectors:

The data points or vectors that are the closest to the hyperplane and which affect the position of the hyperplane are termed as Support Vector. Since these vectors support the hyperplane, hence called a Support vector.

V. METHODOLOGY

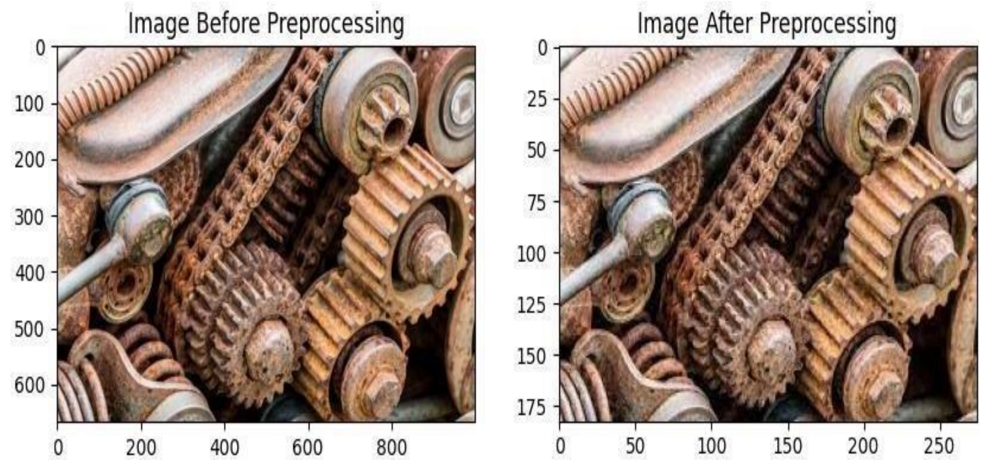
Data Collection and Pre-processing

Image collection and Image Pre-processing:

The dataset of corrosion images for training the SVM model is obtained from internet. These images we first converted into array and then the array is resized into 156x156 size of image for the dataset to train the model, the images were labelled as 0 and 1. In which 0 stands for corroded and 1 stands for non corroded. Once the dataset is labelled, we divided it into the training set and testing set for training and validation in the ratio of 80:20.

Model Setup and Construction

The basic need for machine learning model is the dataset. After creating the dataset we can easily create model and train it with the training data. We created the SVM model and trained it. We have performed hyper parameter tuning of the model. We have given C value as 0.1, 1, 10, 100 and gamma values as 0.0001, 0.001, 0.1, 1 for the tuning purpose. The kernel used for tuning are linear, poly, rbf, sigmoid.

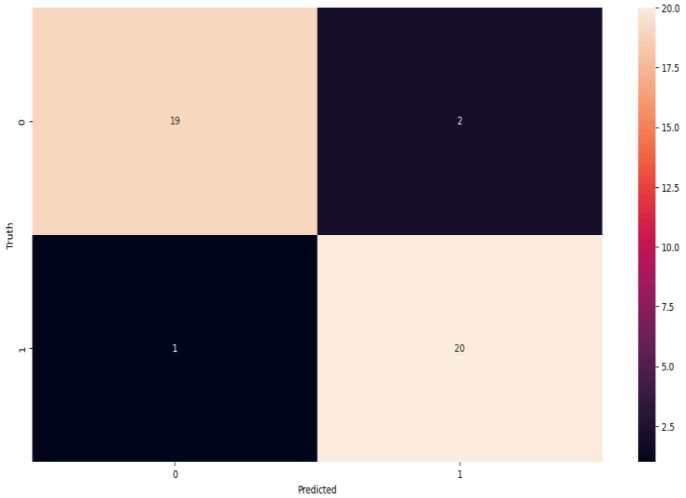


Confusion Matrix :

It is the simplest way to measure the performance of a classification problem where the output can be of multiple type of classes. A confusion matrix is just a table with two dimensions viz. Actual and Predicted and furthermore, both the dimensions have True Positives (TP), True Negatives (TN), False Positives (FP), False

Negatives (FN) as shown below:
Explanation of the terms associated with confusion matrix are as follows −

True Positives (TP) − In this case both actual class & predicted class of data point is 1. True Negatives (TN) − In this case both actual class & predicted class of data point is 0.
False Positives (FP) − In this case the actual class of data point is 0 & predicted classof data point is 1.
False Negatives (FN) − In this case the actual class of data point is 1 & predicted class of data point is 0.



Confusion Matrix of Linear Kernel
From the figure above we get the values :

True Positive (TP) : 20

It states that our prediction model has predicted that the machine will not fail 20 times which is predicted correctly and efficiently followed by the actual model.

False Positives (FP) : 1

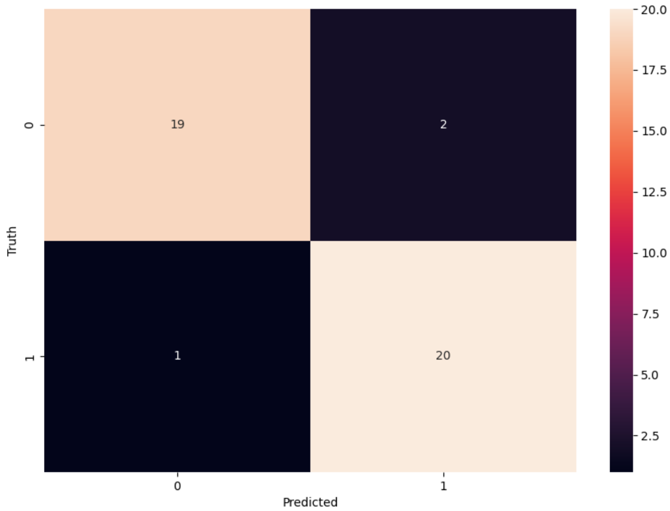
It states that our prediction model has predicted that the machine will not fail 1 times but it is incorrectly predicted as it does not fail in actuality.

False Negatives (FN) : 2

It states that our prediction model has predicted that the machine will not fail 2 times but it is incorrectly predicted as the machine fails in actuality.

True Negatives (TN) : 19

It states that our prediction model has predicted that the machine will fail 19 times which is predicted correctly and efficiently followed by the actual model.



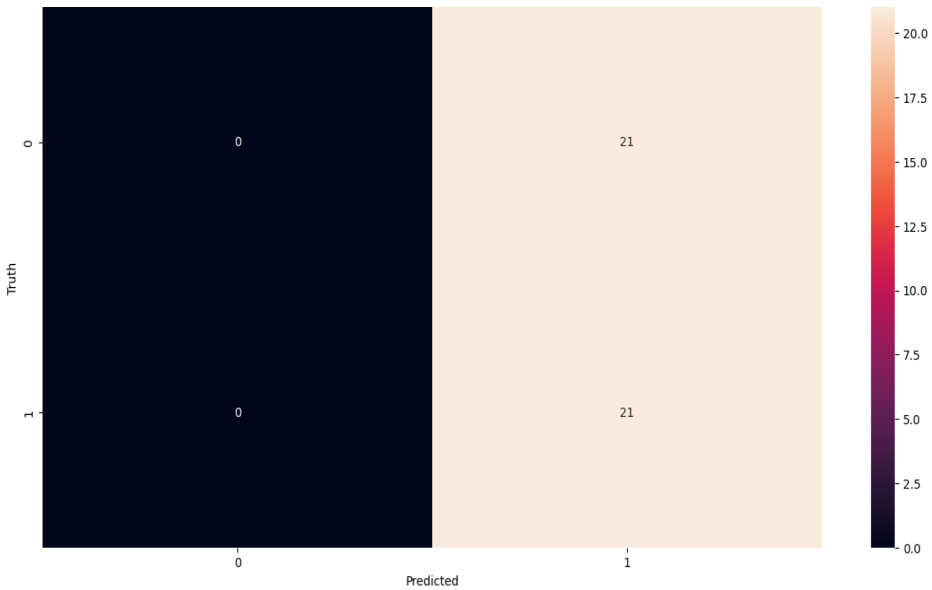
Confusion Matrix of Poly Kernel
From the figure above we get the values :

True Positive (TP) : 20
It states that our prediction model has predicted that the machine will not fail 20 times which is predicted correctly and efficiently followed by the actual model.

False Positives (FP) : 1
It states that our prediction model has predicted that the machine will not fail 1 times but it is incorrectly predicted as it does not fail in actuality.

False Negatives (FN) : 2
It states that our prediction model has predicted that the machine will not fail 2 times but it is incorrectly predicted as the machine fails in actuality.

True Negatives (TN) : 19
It states that our prediction model has predicted that the machine will fail 19 times which is predicted correctly and efficiently followed by the actual model.



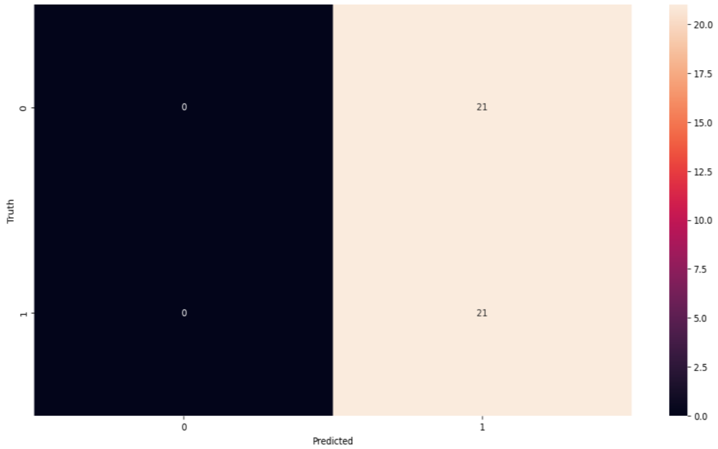
Confusion Matrix of rbf Kernel
From the figure above we get the values :

True Positive (TP) : 21
It states that our prediction model has predicted that the machine will not fail 21 times which is predicted correctly and efficiently followed by the actual model.

False Positives (FP) : 0
It states that our prediction model has predicted that the machine will not fail 0 times but it is incorrectly predicted as it does not fail in actuality.

False Negatives (FN) : 21
It states that our prediction model has predicted that the machine will not fail 21 times but it is incorrectly predicted as the machine fails in actuality.

True Negatives (TN) : 0
It states that our prediction model has predicted that the machine will fail 0 times which is predicted correctly and efficiently followed by the actual model.



Confusion Matrix of Sigmoid Kernel
From the figure above we get the values :

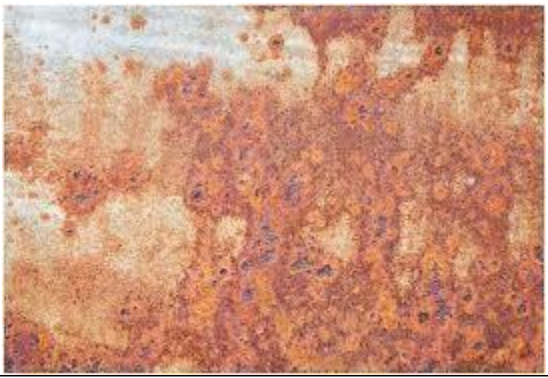
True Positive (TP) : 21
It states that our prediction model has predicted that the machine will not fail 21 times which is predicted correctly and efficiently followed by the actual model.

False Positives (FP) : 0
It states that our prediction model has predicted that the machine will not fail 0 times but it is incorrectly predicted as it does not fail in actuality.

False Negatives (FN) : 21
It states that our prediction model has predicted that the machine will not fail 21 times but it is incorrectly predicted as the machine fails in actuality.

True Negatives (TN) : 0
It states that our prediction model has predicted that the machine will fail 0 times which is predicted correctly and efficiently followed by the actual model.

Test Image Classification Results:

	
Enter Name of Image :download.jpg Corroded = 37.30945127756712 % Non Corroded = 62.69054872243288 % The given image is : Non Corroded PS D:\Python>	Enter Name of Image :images (1).jpg Corroded = 71.64964760640476 % Non Corroded = 28.35035239359526 % The given image is : Corroded PS D:\Python>

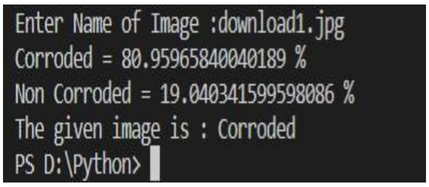
```
Enter Name of Image :images.jpg
Corroded = 68.73578394857115 %
Non Corroded = 31.264216051428857 %
The given image is : Corroded
PS D:\Python>
```

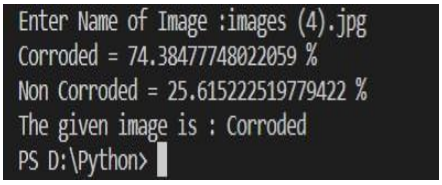
```
Enter Name of Image :images (5).jpg
Corroded = 58.88733737939171 %
Non Corroded = 41.112662620608276 %
The given image is : Corroded
PS D:\Python>
```

False Prediction

```
Enter Name of Image :images (2).jpg
Corroded = 55.12935742098853 %
Non Corroded = 44.87064257901147 %
The given image is : Corroded
PS D:\Python>
```

```
Enter Name of Image :images (6).jpg
Corroded = 53.50386578614306 %
Non Corroded = 46.49613421385692 %
The given image is : Non Corroded
PS D:\Python>
```





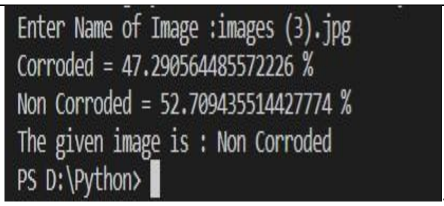
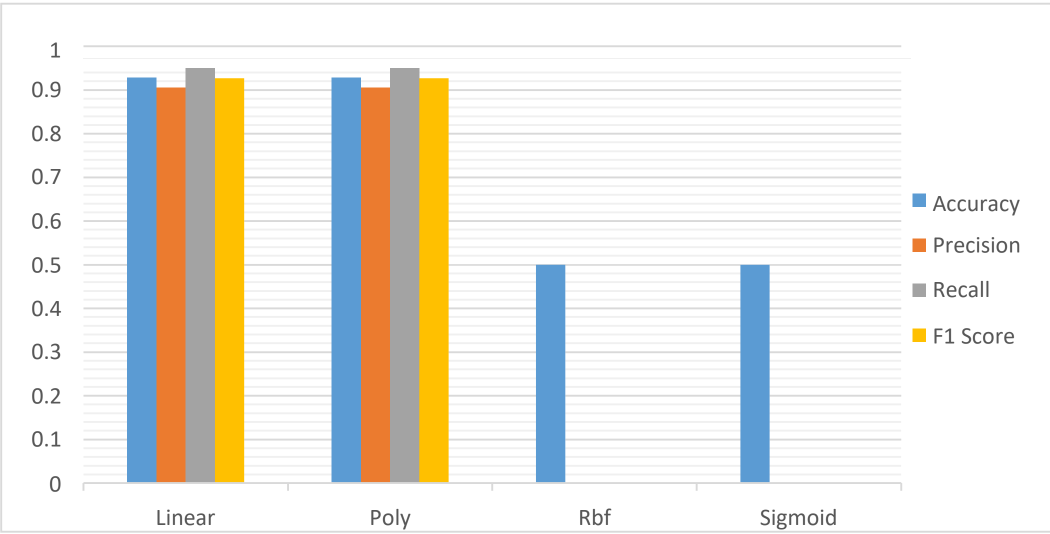


Table 1

Kernel	**Linear**	**Poly**	**Rbf**	**Sigmoid**
Accuracy	0.9285	0.9285	0.5	0.5
Precision	0.9047	0.9047	0	0
Recall	0.95	0.95	0	0
F1 Score	0.9268	0.9268	0	0

Table 2



VI. CONCLUSIONS

In this study we presented a comparison between four kernels of support vector machine for corrosion detection. The kernels used are linear, poly, rbf and sigmoid. We trained the model with more than 165 images and tested with 42 images. The accuracy, precision, recall and f1 score of linear and poly kernels are same. In other hand rbf and sigmoid kernels have accuracy about 0.5. These two kernels have same results. By looking the chart we can say that linear and poly kernels would perform if we compare these two kernels with other two kernels.

In future work we will seek to refine the model and train it with a new and larger dataset of images, which we believe would improve the accuracy of the SVM model.

References

[1]. Image Processing-Based Pitting Corrosion Detection Using Metaheuristic Optimized Multilevel Image Thresholding And Machine-Learning Approaches, Nhat-Duc Hoang, Mathematical Problems In Engineering Volume 2020, Article Id 6765274, 19 Pages

[2]. Surface Corrosion Grade Classification Using Convolution Neural Network, Sanjay Kumar Ahuja, Manoj Kumar Shukla, Kiran Kumar Ravulakollu International Journal Of Recent Technology And Engineering (Ijrte) Issn: 2277-3878, Volume-8 Issue-3, September 2019

[3]. Corrosion Detection Using A.I. : A Comparison Of Standard Computer Vision Techniques And Deep Learning Model, Luca Petricca, Tomas Moss , Gonzalo Figueroa And Stian Broen

[4]. Image Classification Via Support Vector Machine Xiaowu Sun , Lizhen Liu , Hanshi Wang , Wei Song , Jingli Lu, 2015 4th International Conference On Computer Science And Network Technology

[5]. Image Classification Using Support Vector Machine And Artificial Neural Network Article In International, International Journal Of Information Technology And Computer Science · May 2012