Corrosion Detection By Using Support Vector Machines

Prathamesh Jugulkar[1], Sumit Patil[1] , Mita Dhaygude[1] , Pradnya Kesarkar[1] , Disha Kurkure[1]

Dr. Ganesh Dongre[2]

[1]Research Scholar, Department of Mechanical Engineering VIT Pune

[2]Dean R&D ,VIT Pune

[1]prathamesh.jugulkar21@vit.edu , [1]sumit.patil21@vit.edu , [1]mita.dhaygude21@vit.edu , [1]pradnya.kesarkar21@vit.edu, [1]disha.kurkure21@vit.edu

[2]ganesh.dongre@vit.edu

Abstract - Corrosion is a prevalent issue in the mechanical industry. Due to environmental factors like temperature, humidity, and the acidic composition of the liquids, corrosion might develop. Both destructive and non-destructive methods can be used to identify and keep track of the progression of corrosion. Although physical inspection is highly reliant on the skills and knowledge of the inspecting person, visual inspection is a typical technique for surface corrosion analysis. The results of the manual inspection are qualitative and subject to bias, which could lead to mishaps due to poor analysis. For the purpose of preventing unintended incidents, corrosion must be accurately recognized in the early stages.

Keywords – Predictive Maintenance , Machine Learning, SVM, Dataset .

I. Introduction

Corrosion of metals and metal structures is a worldwide problem. Property and infrastructure suffer severe losses as a result of corrosion. Because metals are so pure, corrosion can happen. Because the energy of the metal in its purest form is so great, it naturally has a tendency to change depending on the surrounding environment or process environment into a metal oxide or salt of numerous varieties. The metal gains a stable location in the environment as a result of being transformed into different salts and oxides. When a metal structure or component is unable to withstand the stresses placed on it during operation or use, failure results.

One example of an accident caused by corrosion is the Bhopal gas leak. Most likely, it is the largest industrial catastrophe in Indian and global history. Phosgene, monomethyl amine (MMA), methyl isocyanate (MIC), and the insecticide carbaryl, often known as Savin, are all made at the pesticide plant owned by Union Carbide India Limited (UciL). The location of this factory is 5–6 kilometer’s outside of Bhopal, Madhya Pradesh's city Centre. The incident happened on the eve of December 2 and 3, 1984. Many technical experts and corrosion specialists asserted that corrosion in various MIC3 plant mechanical sections was to blame for this disaster. It was discovered that many of the valves and pipes were in bad condition, and the majority of the safety systems connected to the MIC were not functioning properly.

Therefore, automatic corrosion detection was the main focus of this feasibility study. This project developed an independent classifier that made it possible to identify the presence of rust in photos. The difficulty with this strategy was that rust does not have a well defined shape or colour. Additionally, shifting scenery and the presence of deceptive objects can result in incorrect image classification. In order to handle a huge number of photos in a timely manner, the categorization procedure should still be quite quick.

II. Literature Survey

We have read a number of initiatives and essays on this subject. The first one involved image classification using Support Vector Machines. Capital Normal University's Information and Engineering College, Beijing 100048, P.R. New Zealand's China Ag Research Ltd .The purpose of the aforementioned article was to first discuss the theory of Support Vector Machines before outlining the process for resolving their objective function. Second, there are visuals in experiences that help with classification. The photos are correctly classified with the aid of support vector machines.

The detection of pitting corrosion utilizing image processing, meta-drastically optimized multilevel image thresholding, and machine learning was our second study goal. Corrosion reduces the service life and operability of metal infrastructure components. According to reports, corrosion accounts for 42% of failure mechanisms in engineering structures, making it the most common type of defect. Therefore, it is crucial for maintaining normal structural health to identify and diagnose corroded areas. The results of the survey are extremely helpful to owners or maintenance organizations when determining the efficacy of the current protective strategies and setting priorities for remedial activities.

A classification of surface corrosion degree using a convolutional neural network is another research related to this subject that we looked at. Sanjay Kumar Ahuja, Manoj Kumar Shukla, and Kiran Kumar Ravula Kollu, among others, thus some crucial concepts we looked at are common issue in the oil and gas sector is corrosion. Iron pipes are typically used to transfer gas and oil. Large fields of pipelines are dispersed above ground, below ground, and even underwater. Environmental factors including temperature, humidity, and the acidic composition of fluids all contribute to the development of corrosion.

There are numerous methods for both destructive and non-destructive corrosion detection and monitoring. Visual inspection is frequently used to analyses surface corrosion, however manual inspection is very reliant on the skills and knowledge of the inspector. Qualitative results from manual inspections can be skewed and cause mishaps as a result of improper analysis. To prevent unintended mishaps, corrosion must be precisely identified in its early phases. This study will describe a deep learning-based solution based on computer vision for classifying corrosion in accordance with ISO-8501 standards. The assessment results are fair and reasonably comparable to those of a visual check.

Artificial intelligence-based corrosion detection: a comparison of traditional computer vision methods and a deep learning model. For the automatic identification of metal corrosion, we compare traditional computer vision approaches with a deep learning approach in this work (rust). The traditional method classified pixels depending on how many of them had particular red components. The OpenCV libraries were utilized by the Python code to compute and classify the photos. We used Caffe, a potent framework created at the "Berkeley Vision and Learning Center," for the Deep Learning technique (BvlC). The experiment involved classifying the photos and comparing the overall accuracy of the two methods. The findings reveal some intriguing information about both strategies. The accuracy of the OpenCV-based model was 69% across all metrics. It displayed low accuracy for the "stainless" classification (57%) and excellent accuracy for the "rust" classification (nearly 90%), as we had anticipated. The explanation is simple: since all "rusty" photos contained red components, the computer had no trouble identifying them. The "rustles" class, however, does not require the presence of red pixels to signify the existence of rust.. Therefore, the model has merely identified the red component and incorrectly classed it as "rust," decreasing the accuracy of "non-rust," when we provide an image of a red apple. For instance, the OpenCV system labelled all four of the photos in Figure 1 as "cut" whereas only two of them were accurate.

We also gave a few out-of-focus test photographs to help with the few false negatives, which appeared to be mostly caused by poor image illumination, color-related difficulties, or rust patches that were too small (less than 0.3% of the image). The categorization of surface corrosion degree using a convolutional neural network is different research we looked at for this subject Manoj Kumar Shukla, Kiran Kumar Ravula, and Sanjay Kumar Ahuja Kollu, among the key issues we covered was the fact that corrosion is a common issue in the oil and gas sector. Iron pipes are typically used to transfer gas and oil. Large fields of pipelines are dispersed above ground, below ground, and even underwater. Environmental factors including temperature, humidity, and the acidic composition of fluids all contribute to the development of corrosion. There are a number of destructive and non-destructive corrosion detection and monitoring methods available.

Visual examination is a typical method for analyzing surface corrosion, although manual inspection is very reliant on the skills and knowledge of the inspector. Qualitative results from manual inspections can be skewed and cause mishaps as a result of improper analysis. To prevent unintended mishaps, corrosion must be precisely identified in its early phases. This study will describe a deep learning-based solution based on computer vision for classifying corrosion in accordance with ISO-8501 standards. The assessment results are fair and reasonably comparable to those of a visual check.

III. BACKGROUND

Image classification can be done in a variety of ways. The majority of classifiers, including maximum likelihood, minimum distance, neural networks, decision trees, and support vector machines, reach a conclusion regarding the land cover class and need a training sample. A clustering-based technique, such as K-mean, K-NN, or ISODATA, is an unsupervised classifier, but fuzzy-set classifiers are soft classifiers that may produce more accurate results. This study's goal is to apply a Support Vector Machine (SVM) for image classification.

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

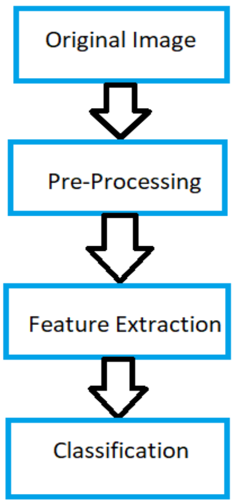


Fig. 1 Block Diagram of Model

IV. THEORY OF SUPPORT VECTOR MACHINE

One of the most well-liked supervised learning methods for classification and regression issues is the Support Vector Machine, or SVM. However, in machine learning, classification issues are where it is most frequently applied.

The SVM algorithm's objective is to establish the best decision boundary or line that may divide the n-dimensional space into classes so that it will be simple to later assign a new data point to the appropriate category. This decision boundary is the best one for categorizing the data points. The SVM hyperplane is referred to as this best bound.

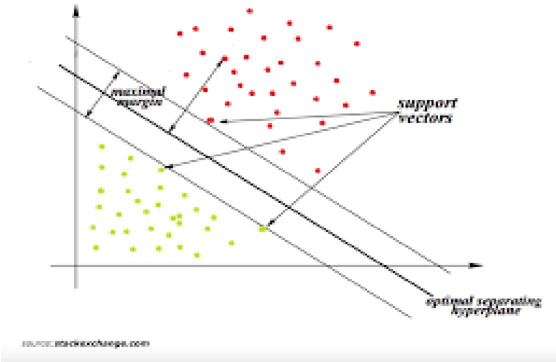


Image 2 Support Vectors

Svm selects extreme points/vectors that help in creating the hyperplane. These extreme cases are called support vectors, and thus the algorithm is called a support vector machine. Consider the diagram below, in which two different categories 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 using a single straight line, then such data is called linearly separable data and the classifier is used as a linear SVM classifier.

Non-linear svm is used for non-linearly separated data, which means that if the data set cannot be classified using a straight line, then such data is referred to as non-linear data and the classifier used is called non-linear svm classifier.

The svm's hyperplane and support vectors algorithm:

Hyperplane: we need to find the best decision boundary to help classify the data points, but there can be multiple lines/decision boundaries to separate classes in n-dimensional space. The svm hyperplane is a name for this best bound.

The dimensions of the hyperplane depend on the elements present in the dataset; as shown in the figure, the hyperplane will be a straight line if there are 2 elements. And the hyperplane will be a two-dimensional plane if there are three elements.

We always build hyperplanes with a maximum edge, or the maximum distance between data points.The closest data points or vectors to the hyperplane that have an impact on the position of the hyperplane are referred to as the support vectors. These vectors are referred to as support vectors since they support the hyperplane.

V. METHODOLOGY

a. Data Collection and Preprocessing

The SVM model's training data set for corrosion photos is downloaded from the Internet. These photographs were initially turned into an array, which we then reduced to a 156x156 image size for the model training dataset. The images were then given the labels 0 and 1. In which 1 denotes uncorroded and 0 indicates corrosion. After the dataset has been labelled, we divide it into a training set and a test set in an 80:20 ratio for training and validation.

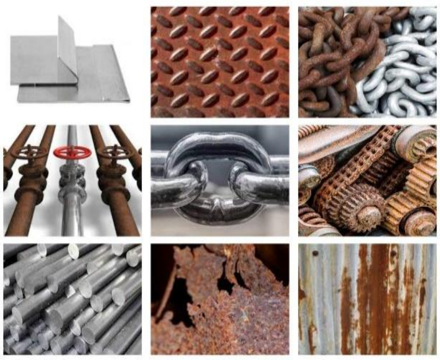


Image 3 Data Set Images

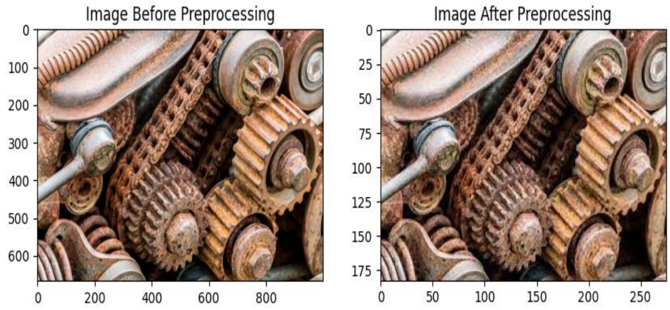


Image 4 Image Before And After Pre-processing

b. Model Setup and Construction

The basic need for a machine learning model is a dataset. After creating a dataset, we can easily build a model and train it using the training data. We created an SVM model and trained it. We performed hyperparameter tuning of the model. For tuning purposes, we have given C values as 0.1, 1, 10, and 100 and gamma values as 0.0001, 0.001, 0.1, 1. Kernels used for tuning are linear, poly, RBF, and sigmoid.

VI. Results

Confusion Matrix is the easiest method for gauging how well a classification task performs when the result can take many different forms. The confusion matrix is just a two-dimensional table with True and Predicted as well as True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) as shown below.

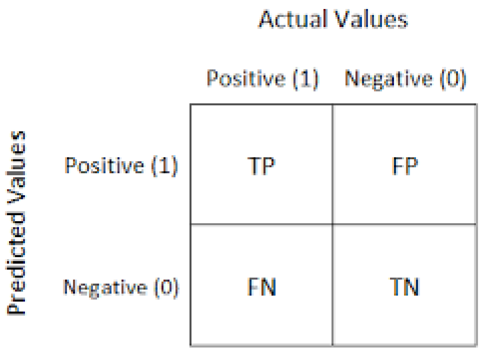


Image 5 Confusion Matrix

The terms connected to the confusion matrix are defined as follows:

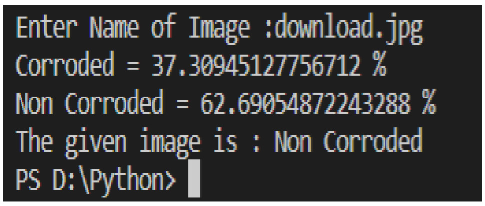
True Positive (TP) In this instance, the data point's true and anticipated classes are both 1.

True Negatives (TN) - In this situation, the data point's true and projected classes are both 0.

False Positive (FP) - In this scenario, the data point's projected class is 1 while its real class is 0.

False Negative (FN) - In this scenario, the data point's real class is 1 while its projected class is 0.

Image 6 Test Image Image



7 Test Image Console Output

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 1 Comparison of Various Kernels

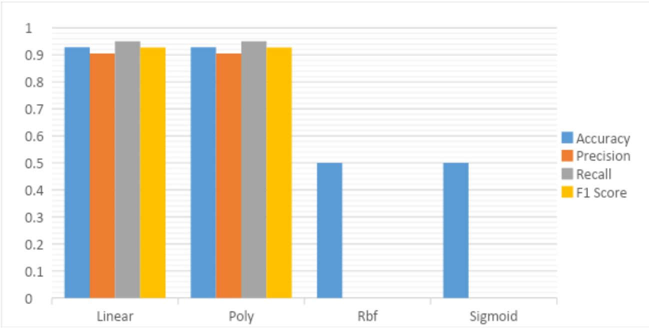


Image 8 Comparison of Various Kernels

VII. CONCLUSION

In this study, we presented a comparison between four support vector machine cores for corrosion detection. The kernels used are linear, poly, RBF, and sigmoid. We trained the model with more than 165 images and tested it with 42 images. Accuracy, precision, recall, and f1 scores of linear and poly cores are the same. On the other hand, the RBF and sigmoid kernels have an accuracy of about 0.5. The two kernels have the same results. Looking at the graph, we can say that linear and poly kernels would work effectively if we compare these two kernels with the other two kernels.

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

VIII. REFERENCES

[1] Image Processing-Based Pitting Corrosion Detection Using Metaheuristic Optimized Multilevel Image Thresholding and Machine-Learning Approaches, Nat-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 Ravula Kollu 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

[4] MODEL, Luca Petricca, Tomas Moss, Gonzalo Figueroa, and Stian Brown

[5] 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

[6] Image Classification using Support Vector Machine and Artificial Neural Network Article in International, International Journal of Information Technology and Computer Science · May 2012