

CHAPTER 1

INTRODUCTION

The oil, gas, chemical and petro-chemical industries operate on hundreds of kilometres of pipelines, often transporting hydrocarbons and non-hydrocarbons fluids. Each year, there are hundreds of pipeline failures, resulting in pollution, loss of transportation capacity, loss of oil and gas availability, and repair expenses. These pipelines are prone to external corrosion, under support corrosion (crevices), corrosion under insulation (CUI), corrosion under painting and many other types of the same. Sometimes the corrosion may lead to cracking as well (known as Stress Corrosion Cracking). Therefore, inspection of pipelines at regular intervals and maintaining the integrity of pipeline system is a task of the highest priority.

The state-of-the-art inspection methods include magnetic flux leakage (MFL), ultrasonic testing, long range guided wave inspection, external corrosion direct assessment (ECDI) etc. All these methods have their own limitations, such as determination of stress corrosion cracking (SCC), limited range of inspection, finding small pitting defects, efficiency in identification of defects, and others. Inspection of a pipeline system ranging a few kilometres may not be efficient (more time consuming) using most of the existing methods. There is, therefore, an urgent need for the development of a quick, reliable method for the detection of corrosion in pipeline systems

1.1 Machine learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Machine learning algorithms are often categorized as supervised or unsupervised.

- Supervised machine learning algorithms can apply what has been learned in the past to new data using labelled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.
- In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labelled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabelled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabelled data.
- Semi-supervised machine learning algorithms fall somewhere in between supervised and unsupervised learning, since they use both labelled and unlabelled data for training – typically a small amount of labelled data and a large amount of unlabelled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi supervised learning is chosen when the acquired labelled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring unlabelled data generally doesn't require additional resources.
- Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behaviour within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal. Machine learning enables analysis of massive quantities of data. While it generally delivers faster, more accurate results in order to identify profitable opportunities or dangerous risks, it may also require additional time and resources to train it properly. Combining machine learning with AI and cognitive technologies can make it even more effective in processing large volumes of information.

1.2 Image processing

Signal processing is a discipline in electrical engineering and in mathematics that deals with analysis and processing of analog and digital signals, and deals with storing, filtering, and other operations on signals. These signals include transmission signals, sound or voice signals, image signals, and other signals etc.

Out of all these signals, the field that deals with the type of signals for which the input is an image and the output is also an image is done in image processing.

It can be further divided into analog image processing and digital image processing.

1.2.1 Analog image processing

Analog image processing is done on analog signals. It includes processing on two dimensional analog signals. In this type of processing, the images are manipulated by electrical means by varying the electrical signal. The common example include is the television image.

Digital image processing has dominated over analog image processing with the passage of time due its wider range of applications.

1.2.2 Digital image processing

The digital image processing deals with developing a digital system that performs operations on a digital image.

An image is nothing more than a two dimensional signal. It is defined by the mathematical function $f(x,y)$ where x and y are the two co-ordinates horizontally and vertically.

1.3 Motivation

The oil, gas, chemical and Petro-chemical industries operates on hundreds of kilometres of pipeline. Each year there are hundreds of pipeline failures resulting into pollution, loss of transportation capacity and repair expensive. Recognition of corrode pipelines in unhabitated area is quite an arduous process. Therefore, inspection of pipelines at regular intervals and maintain the integrity of pipeline systems is very important. Using the knowledge of modern machine learning and image processing techniques, we can develop a tool for detecting corrosion in pipelines without on ground inspection.