CAAP Quarterly Report

Date of Report: *April 6, 2019*

Prepared for: *Thomas Finch (Project Manager) and Joshua Arnold (CAAP Program Manager), U.S. DOT Pipeline and Hazardous Materials Safety Administration*

Contract Number: *693JK31850005CAAP*

Project Title: *Low-variance Deep Graph Learning for Predictive Pipeline Assessment with Interacting Threats*

Prepared by: *Hao Zhang (Colorado School of Mines) and Yiming Deng (Michigan State University)*

Contact Information: *Dr. Hao Zhang, Department of Computer Science, Colorado School of Mines; 1500 Illinois St., Golden, CO 80401; Phone: 303-273-3581; Email: hzhang@mines.edu*

For quarterly period ending: *April 7, 2019*

Business and Activity Section

(a) Contract Activity

No contract modification was made or proposed in this quarterly period. No materials were purchased during this quarterly period.

(b) Status Update of Past Quarter Activities

Pipeline networks are essential to gather, transport and distribute gas and hazardous liquid. Oil and gas revolution have been held back by aging pipeline infrastructure. Cause loss of life and safety issues Cost millions of dollars in environmental cleanup, infrastructure repair, property damage, and production loss. More over to truly achieve full energy independence by using both conventional and unconventional oil & natural gas to power our economy, the development and revival of pipeline infrastructure with seamless integration of computational algorithms and improved pipeline assessment methods, and the advancement of nondestructive evaluation (NDE) inspection technology and data analytics to assure next-generation pipeline system's usability, integrity, safety and reliability is imperative. A recent report has identified significant interacting threats and shown that, among many failures reported in the past few decades, loss of containment (LoC) may appear to have resulted from a single cause of threat but, in reality, has resulted because of circumstances that aggravated the threats[1]. However, characterization of coincident and overlapping defects, due to interacting threats, which can lead to pipeline integrity failure has not been well addressed. For example, the interacting nature of corrosion and mechanical damage consisting of a scrape or gouge is still debatable. Thus, a clear definition of coincident damage features and identification of interacting threats is a major existing technical gap.

Another critical gap is that existing pipeline inspection techniques are designed for specific threat conditions and does not provide a holistic assessment of interactive threats' impact, which results in the large sensor data variance and associated uncertainties leading to significant difficulty in resolving overlapping damage thereby affecting integrity management decision adversely. A general framework for identification of coincident damage features with better quantification of material physical and geometric characteristics will be of critical importance for enhancing pipeline assessment methods and models that reduces variance[2]. Joint learning and artificial intelligence algorithms focusing on semantic reasoning and multisensory fusion for anomaly recognition and prediction will be the critical computational component of MSU contribution. Moreover, there are different vendors and operators who conduct various kind of testing on the general operations and anomalies in pipelines across the country.

Figure 1 General Diagram of the Pipeline System

It will be useful to combine these separate databases and pool information across them for a unified analysis of the fault rates seen across different pipelines. Such data driven analysis will improve the prediction of future anomalies, which in turn can be used for better allotment of work force, inspection personnel and inventory management of pipeline repairing resources. Pipelines are considered the most efficient way of transferring fluids (oil and gas) over long distances. However, despite this high-efficiency, there has been reasons for concern over the last years, principally because a large part of the existing pipeline networks is coming to the end of their useful life. Consequently, it is necessary to be able to monitor, evaluate and to guarantee their structure as a whole, taking precautions against leaks and consequently protecting the environment and the population. Nondestructive tests (NDT) coupled with signal processing techniques and artificial intelligence such as convoluted neural network (CNN) and other non-linear models like Generalized Linear models (GAM) have contributed significantly to evaluate the structural integrity of the pipelines.

Historical data has validated that multiple threats often collectively lead to failures. corrosion coincident with manufacturing defects often cause much earlier penetration of pipe walls. Multiple threats acting on pipe walls that increase the probability of failure than each of the individual threats. Technical Objectives focus on developing and evaluating new methods for low-variance interacting threats assessment.

In this reporting period, the research team performed comprehensive literature review, and made great progress toward achieving the technical objectives including:

(1) Investigating different threats and developing risk models for the threats.
(2) Designing an approach for image data segmentation: We implemented three segmentation approaches, one is based on unsupervised learning, one is based on graph theory and the other one is based on deep learning to segment objects from image data. The approach is currently evaluated using the magnetic tile defect dataset which consists of defect tiles greyscale images and pixel level mask annotation of the defect.

The PIs believe that education is a critical component of the CAAP project, and we will continue integrating research with educational activities to prepare the next generation scientists and engineers for the gas and pipeline industry. Similar education and outreach activities were conducted during this period of performance.

(c) Cost share activity

PI Zhang used his 11.29% yearly effort as the in-kind cost share to work on the project at the Colorado School of Mines. Co-PI Yiming Deng used his 6.07% yearly effort as the in-kind cost share to work on the project at the Michigan State University. The cost share was used following the approved proposal and no modification was made.

(d) Performed Research in This Period of Performance

1. Background

Pipeline integrity management regulations require identification and evaluation of preventive measures to reduce the likelihood of failure. These knowledges heavily based on identification of single or interacting threats. These “interacting threats” are more fatal as they result in a likelihood of failure greater than that due to either threat individually or merely the superposition of the threats. Interactive threats are the merge of two or more defects in a pipe segment, the result of which is more damaging than either of the individual threat themselves. For threats to be interacting they must occur concurrently and coincidentally under a given circumstance. The definition of pipeline threat used in this project is given as the initiating condition or event leading to loss of containment (LoC) in the pipeline system[1]. The framework of assessment is presented below:

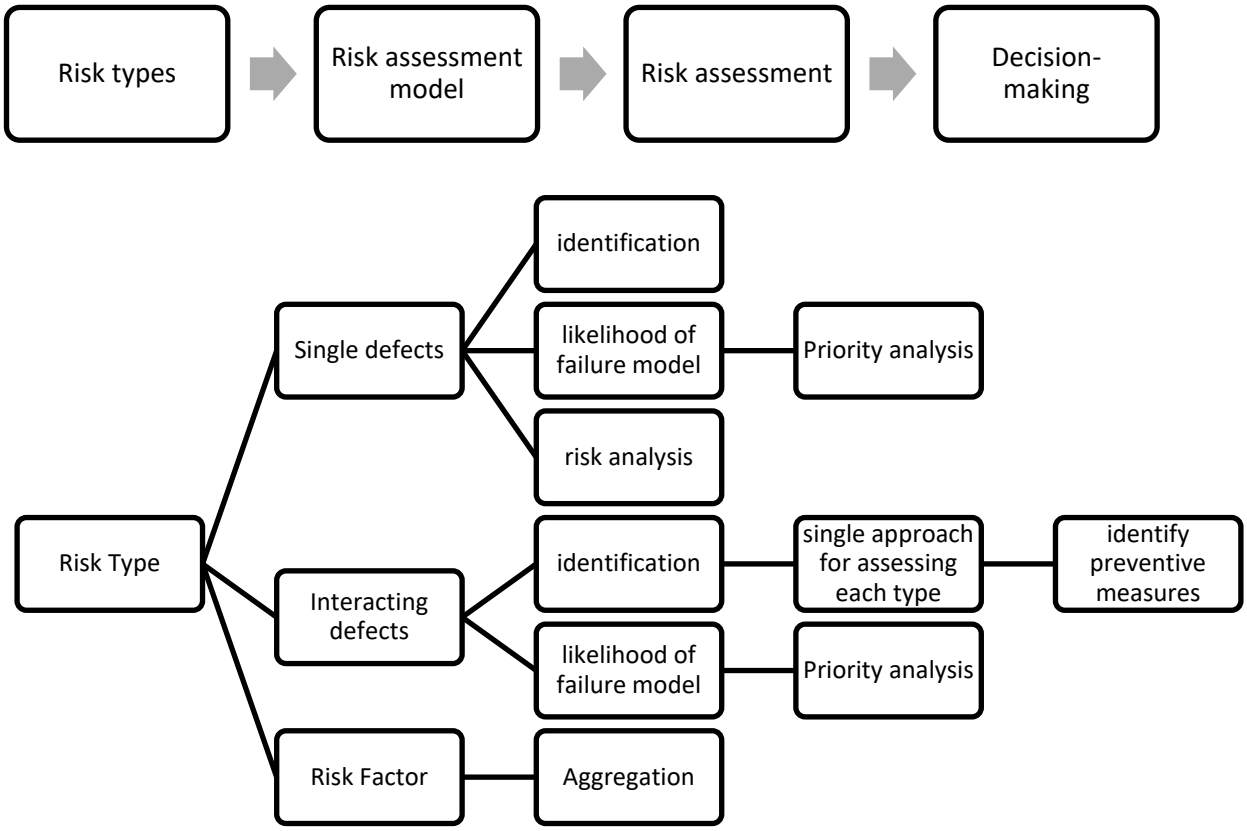


Figure 2 Flow Chart showing steps of Threat Assessment

By consensus pipeline operators in the U.S. accept that there are 22 "root causes" of pipeline accidents[3]. These 22 basic causes were identified from studies of "reportable incident" data sponsored by the Pipeline Research Council International (PRCI). Leaving out "unknown" and "miscellaneous" which are of little help in integrity assessment, we arrive at the following 20 basic causes (in effect, threats) that are usually considered in an integrity management plan:

Table 1 Different existing threats

1	External Corrosion (EC)
2	Internal Corrosion (IC)
3	Stress Corrosion Cracking (SCC)
4	Defective Pipe (DP)
5	Defective Pipe Seam (DPS)
6	Defective Fabrication Weld (DFW)
7	Defective Girth Weld (DGW)
8	Construction Damage (CD)
9	Malfunction of Control or Relief Equipment (MCRE)
10	Stripped Threads, Broken Pipe, or Coupling Failure (TSBPC)

11	Gasket Failure (GF)
12	Seal or Pump Packing Failure (SPPF)
13	Incorrect Operations (IO)
14	Third Party Damage (TP)
15	Previously Damaged Pipe (PDP)
16	Vandalism (V)
17	Earth Movement (EM)
18	Heavy Rains and Floods (HRF)
19	Lightning (LIGHT)
20	Cold Weather (CW)

However according to the ASME B31.8S standard out of the above twenty, nine primary threat conditions are identified which comes under three basic categories[4].

- **Time-Dependent Threats** are issues with a pipeline that occur over time. Though these threats can be partially mitigated by adding protective measures to pipe infrastructure (i.e., protective coatings), they are typically addressed by ongoing and periodic assessments, in line inspections, and pressure testing. Time-dependent threats are the driving force behind the development of assessment tools recommended in regulations aimed at integrity management programs.
  - Internal Corrosion
  - External Corrosion
  - Stress Corrosion Cracking

- **Resident Threats** are threats that do not grow over time; instead they tend to act when influenced by another condition or failure mechanism. These threats are most effectively managed by close monitoring after post-installation/construction pressure tests, which helps determine the initial state of the pipeline. Studies have shown that most pipeline defects are unaffected by the pressure changes that naturally occur during transmission operations, and as a result, they are considered to be "stable"
  - Manufacturing
  - Fabrication
  - Construction

- **Time independent threats** are issues that occur as a result of an outside force. The most common of these threats is pipeline damage caused by erosion, operator error during excavation and/or operations. Unlike time-dependent threats, which require periodic assessments, the most effective way for owners and operators to avoid time-independent issues is through proactive prevention. Following proper operating procedures, adhering to best practice guidelines for maintenance and repair operations, and ensuring that all onsite personnel have been properly trained before engaging in excavation activities are all critical to managing such risks.
  - Human error
  - Excavation Damage
  - Earth movement, outside force or weather.

Among them the most predominant are the Internal Corrosion and External Corrosion. Internal corrosion caused by the materials are transported by the pipeline. Pipeline structures consisting of metals are degraded by the liquids and gases that they carry as these gases and liquids are corrosive whereas external corrosion of pipeline is caused by the external environment. The acidity and the moisture content of the soil are the most common factors for soil degradation. Fatigue though not considered in the above table has significant contribution in causing the threats. Threats listed above exist in a given pipeline segment with a probability to cause failure, thus it is essential for pipeline integrity management. However, these threats could interact and results in a higher probability of failure than that associated with each of the individual threats. Definition of interacting threats is given by Kiefner project, stated that Two or more threats acting on a pipe or pipeline segment that increase the probability of failure to a level greater than the effects of the individual threats acting alone. Two conditions must be satisfied, that is threats act concurrently and coincidentally. Only in this way, condition caused by interacting threats will be more severe than coexisting individual threats. It is important to note that threats are not necessarily interacting simply because they exist at the same location on a pipe or pipeline.

On the other hand, if LoC is resulted from multiple causes which are the circumstances that aggravate single threat. It is apparent that aggravated individual threat is different from the individual threat themselves. For example, a seam manufacturing threat (DPS) might be considered stable in a reasonably steady-state operating pressure regime after a hydrostatic test to a test-pressure-to-operating-pressure of a low ratio. However, if the operation involves large frequent swings in pressure, a seam manufacturing defect may grow large enough to cause a failure at the operating pressure after a period of time. It can be seen that from here that individual threat and circumstance are interacting with each other. Hence if the operating process involves frequent changes in pressure within the pipeline or by overlooking any of the factors like change in product quality, throughput rates can also exacerbate the situation leading to the rise of aggravate threats. In fact, it is necessary to consider aggravated threats in an integrity management plan. In general, several of the 20 recognized threats can be aggravated if the following circumstances exist.

- When a high level of axial stress is present in addition to the normally expected hoop stress arising from internal pressure, several of the above 20 recognized threats can be aggravated.
- When repeated or fluctuating stress, strain or stress intensities is applied on structural components at a particular location which is known as Cyclic loading. Under this circumstances it leads to degradation in that location often referred to as fatigue degradation, then also the threats can be aggravated. However the pressure-cycle induced fatigue is another significant aggravating factor which can cause anomalies like small corrosion pits, small stress corrosion cracks, dents to grow into large fissures with the passage of time.
- Existence of factors like external loading that may develop due to frost heave or due to human settlement or other time varying factors such as temperature change, change in soil moisture or change in nature of the transported fluid can make the situation unstable leading to aggravation of defects.

2. Research on Task 1 and Task 3

Converting Aggravated Threats to Interacting Threats

In the case of any defect that is not large enough to cause an immediate failure but could cause a failure at a later time if caused to grow by pressure-cycle-induced fatigue, the aggravated threat from the defect (whether it be CD, DP, DPS, EC, IC, PDP, or SCC) will be defined as the interaction of the threat with the threat of incorrect operations (IO). The rationale for this definition is that the aggravated threat can be mitigated by periodic hydrostatic testing, in-line-inspection, direct assessment or by pressure testing and that failure of an operator to recognize and monitor the potential for defect growth and apply the appropriate remedies in a timely manner constitutes incorrect operation. Hence, the following aggravated individual threats are transformed to the following interacting threats.

Table 2 Threats transformation

CD aggravated by pressure-cycle-induced fatigue becomes CD + IO.
DP aggravated by pressure-cycle-induced fatigue becomes DP + IO.
DPS aggravated by pressure-cycle-induced fatigue becomes DPS + IO.
EC aggravated by pressure-cycle-induced fatigue becomes EC + IO.
IC aggravated by pressure-cycle-induced fatigue becomes IC + IO
PDP aggravated by pressure-cycle-induced fatigue becomes PDP + IO.
SCC aggravated by pressure-cycle-induced fatigue becomes SCC + IO.
Wrinkle bend subjected to pressure-cycle-induced fatigue CD + IO.

Threats assessment

The definition of threats assessment is given as systematic process in which hazards from pipeline operation are identified and the probability and consequences of potential adverse events are analyzed and estimated. Federal gas and hazardous liquid pipeline safety integrity management (IM) regulations contain requirements for the uses of risk assessments by pipeline operators. Threat is defined as a measure of potential loss in terms of both the likelihood (or frequency of occurrence) of an event and the magnitude of the consequences from the event[5]. A standard conceptual definition of risk used to structure threat assessment is given by the equation:

$$Threat = Likelihood \times Consequence$$

Threat assessment considers the likelihood of failure from all potential and existing threats at each segment along the pipeline. In addition, each receptor category may experience different consequence levels from a pipeline failure, depending on the failure mode (e.g., leak vs. rupture event) and location of the failure (e.g., proximity to receptors such as population and environmentally sensitive areas). Integrity management (IM) involves the identification and evaluation of the preventive measures, continuous IM evaluation for close monitoring of the integrity threats as well as determining the assessment interval for proper detection of the threats.

Threat assessment model

Threat assessment models are adopted to support risk assessment.
A set of algorithms or rules that use available information and data relationships to perform risk assessment. A model is a simplified representation of a pipeline system and represents the relation of important risk factors. There are several types of model that are employed in pipeline risk analysis.

Table 3 pipeline risk model

Model category	Inputs	Outputs	Algorithm
Qualitative	Qualitative (descriptive and relational)	Qualitative	"Matrix" Mapping Inputs to Outputs
Relative assessment /index	Qualitative and Quantitative	Quantitative – unit-less	Risk Index Scoring
Quantitative System	Qualitative	Quantitative - with units (e.g., probability, frequency, expected loss	Quantitative System Model
Probabilistic	Quantitative, including probability distributions	Probability distributions	Quantitative System Model

Threat assessment approach

In this section different threat assessment approaches have been elucidated. For integrity assessment of the concerned pipeline segments, a threat priority is to be established. This threat value comprised of a range of numbers reflecting the overall likelihood of failure and the number of consequences leading to that failure. While constructing the threat assessments the following components are commonly used:

- Evaluating the likelihood of failure and consequences
- Identification of potential events hampering the integrity of pipeline.
- Performing threat ranking and analysis based on increasing likelihood of failure from the most common threats (most common is corrosion).
- Continuous monitoring and updating the threat assessment.
- If the threat estimate exceeds the pre-defined threshold of minor scratches to be considered as defects then raise flag and take impromptu actions.

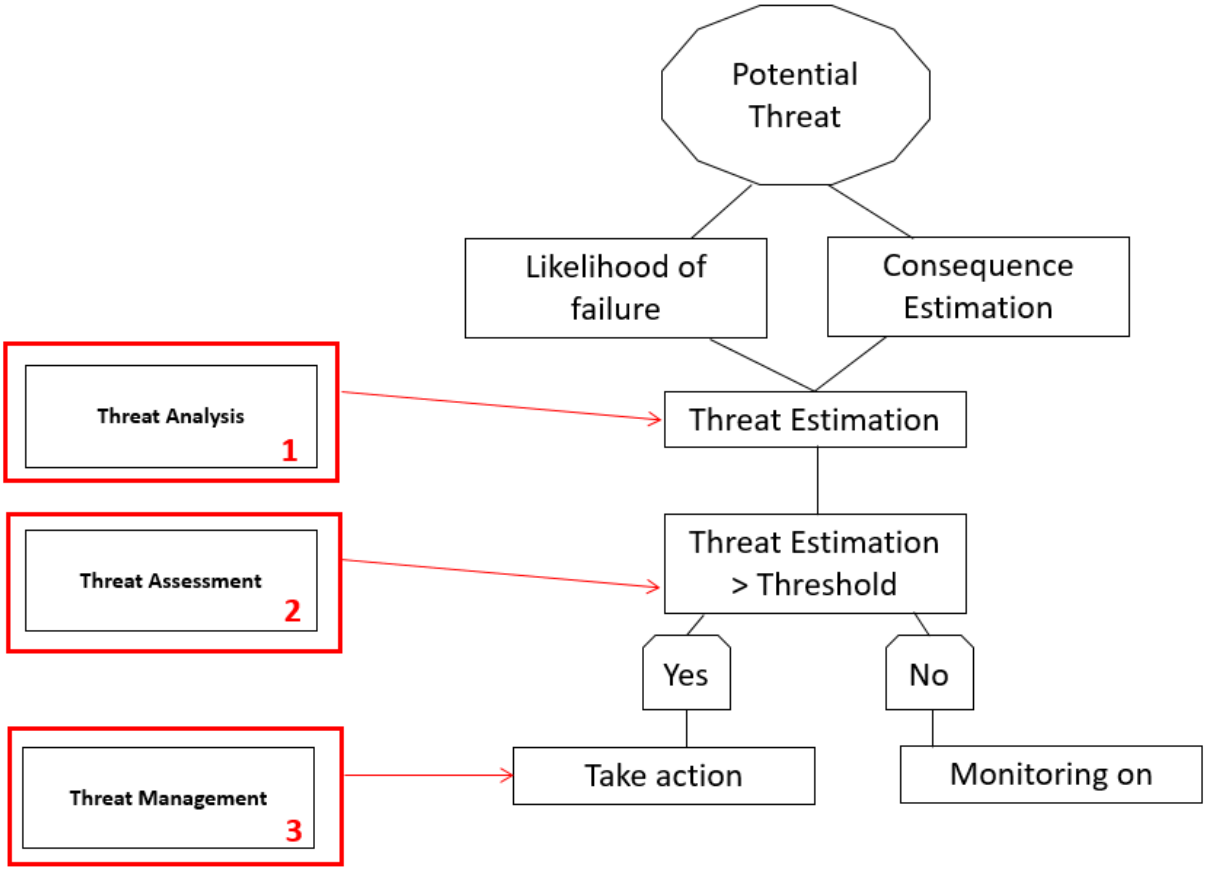


Figure 3 Basic flow chart of Threat management model

The four approaches for integrity management threat assessment are as follows[6]:

1) ***Subject Matter Experts*** *(SMEs).* SME from the operating company combined with information from mostly literature and surveys, is used to provide a range of numeric value describing the likelihood of failure and the consequence. After the analysis of pipe segments, relative likelihood and consequence are assigned thereby the risk is calculated. This process is time consuming and not cost effective as it involves manual intervention to large extent.
2) ***Relative Assessment models.*** Here the assessment is more rigorous and accurate as the assessment is done on more extensive historical datasets. The information from historical data is served as an auxiliary information here to construct new models. Based on the auxiliary information the threats which occurred in the past are given more weightage while constructing new model. This is basically a problem of Transfer learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.
3) ***Scenario-Based Models.*** This risk assessment approach creates models that generate a description of an event or series of events leading to a level of risk, and includes both the likelihood and consequences from such events. This method usually includes construction of event trees, decision trees, and fault trees. From these constructs, risk values are determined.

4) ***Probabilistic Models.*** This is the most demanding assessment approach as it has the added feature of representing the uncertainty in the model inputs by probability distribution and the resulting ability to produce model outputs. A Probabilistic model with input and output distributions is particularly effective for identifying integrity assessment intervals through its ability to support evaluations of the uncertainty in the predicted probability of failure given actual integrity assessment results.

Here we will study the probabilistic model in detail. In case of interactive threats, the interaction mechanism driving the threats increases the total likelihood of failure (LOF) than that of the individual threats. Multiple threats interact to result in premature failure at a location on the pipeline. To provide an appropriate likelihood of failure estimate, the risk model should account for the interaction of multiple threats. In addition, identification of effective preventive measures and evaluation of the effect of preventive measures on reducing the likelihood of failure should include consideration of interacting threats.
External corrosion of pipe damaged by denting is an example of interacting threat where, if the pipe damage result in the damage of external coating in an ineffective cathodic protection then the likelihood of failure will increase.

Say there are two defects which are interacting with each other. Then the total LOF is given by:

$$P_T = P_{D1} + P_{D2} + P_I$$

Where,

P_T: represent the total failure probability.
P_{D1}: represent the failure probability due to defect 1 only.
P_{D2} : represent the failure probability due to defect 2 only.
P_I: represent the increased failure probability which comprised of both $(P_{D1}|P_{D2})$ and $(P_{D2}|P_{D1})$.

Modelling of Consequence

For accurate threat assessment apart from LOF, number of consequences is of prime importance. Consequence represents the extent of loss associated with the leakage in pipeline system. The consequence portion of a pipeline risk model encompasses the scenarios following a pipeline failure. The threat model uses the factors driving those scenarios and the interrelationships among risk factors to estimate the overall consequence of failure to potential receptors. Hence for estimation of the consequences, the model has to be comprised of input variables as follows:

Table 4 Input variables

Damage Type	The sort of damage the product will occur to the receptors due to the leakage.
Rate of release and volume	The rate and volume of the gas or liquid or vapor being released is to be measured
Dispersion Characteristics	The duration, place and the reason behind the leakage has to be determined.

Receptors	To determine who are affected and to what extent
Loss	Loss of property or lives has to be estimated.

Hence to evaluate the consequence as a function of the input variables will result in the following function as follows:

$$Failure\ Consequence = f(D_T, V, D, R, L)$$

Where,

D_T: represents the type of damage due to leakage.
V: represents the rate and volume of the leaked product in a certain time interval
D: represents the Dispersion characteristics.
R : represents the receptors impacted by the leakage
L : Measure of the losses as a result of the leakage.

Probabilistic assessment model

Here Input distributions should be chosen by considering the range of possible values for the inputs and how the possible values are distributed over the range. Statistical methods, such as Bayesian analysis (where a set of existing information can be updated by additional pipeline data), can be used to choose distributions of the given data. Applying methods to estimate input variable uncertainty is especially important for risk models that apply only point estimates as input values rather than probability distributions. Applying a point estimate to represent a variable's underlying data is convenient, but if the data is spread over a wide range, this should be understood and handled in a deliberate manner.

There are two general type of failure measures, one in which the threats are concentrated at explicit locations, and we know in beforehand about those locations. Then we can use location specific probability P_{fi} to determine the likelihood of failure. However, there are some defects which are distributed along the pipeline and for them the locations are not known in prior. Then we can use distributed probability P_d to calculate LOF. Hence if the segment of the pipe i.e. where the interaction is happening is L (say), then the total failure probability (in per mi year) is given by:

$$P_f = P_d + \frac{1}{L}\sum_{i=1}^{n} P_{fi}$$

For time dependent damages (corrosion, cracks) it is easy to anticipate the behavior of the defect with time. Hence data from earlier databases can be used as auxiliary information to detect the growth of defects with time. Here the integrity management process comprises of the following steps:

- To assess existing damage severity.
- Assess the growth behavior over time by Bayesian update.
- Thereby managing integrity.

However for Time independent random damages (Third party damages or sudden ground movement) we have to assess the likelihood of event occurrence.
For time-dependent damage the failure rate (per mi-yr) is given by:

$$Fail_Rate_{per\ mi-yr} = No\ of\ defects_{per\ mi} * POF\ per\ defect_{per\ yr}$$

The probability of failure (POF) as a function of time will depend on the defect growth rate, variability in pipe properties like change in yield stress and toughness and different load characteristics.
For time-independent threats the probability estimation is given by:

$$Fail_Rate_{per\ mi-yr} = Event\ Freq_{per\ mi} * POF_{\ per\ event}$$

Here the 3rd party damage (event occurrence) has to be taken into account for finding the POF.

Simulation based on COMSOL

Steel pipelines are often subjected to complex stress/strain conditions in oil and gas industry. In our simulation[7], wall of the pipe is subjected to a certain amount of pressure when gas or liquid flow through within the pipe. The internal radial pressure has reshaped dimension of the pipe. For simplicity here, longitudinal strain due to surrounding soil movement is not considered here. Pipeline is loaded with external stress[8] whose value is predefined so that it is not too large to cause deformation. Since strain produced within wall of the pipe and external stress has reached an equilibrium where energy is not lost during the process of stressing and straining. The process is reversible. Therefore, we only need to define the stress lower than yield stress and keep deformation being elastic. Young's Modulus is a measure of stiffness of an elastic material. It is used to describe the elastic properties of objects like wires, rods or columns when they are stretched or compressed. The model geometry is comprised of a 3D high strength alloy steel pipeline with outer diameter (OD) and inner diameter (ID) defect located in the center. The elastic-plastic stress simulations are performed here using a small strain plasticity model and von-Mises yielding criterion. The pipeline has the length of 100 cm and wall thickness of 20 mm. The corrosion defect on the exterior side of the pipeline is cubic in shape with a size of 10 by 10 mm and a variable depth. High-strength alloy steel is selected as the material with strain-stress curve being carefully predefined.

Table 5

Density	7580 kg/m^3
Young's modulus	$200GPa$
Poisson' ratio	0.33
Initial yield stress	$250\ MPa$

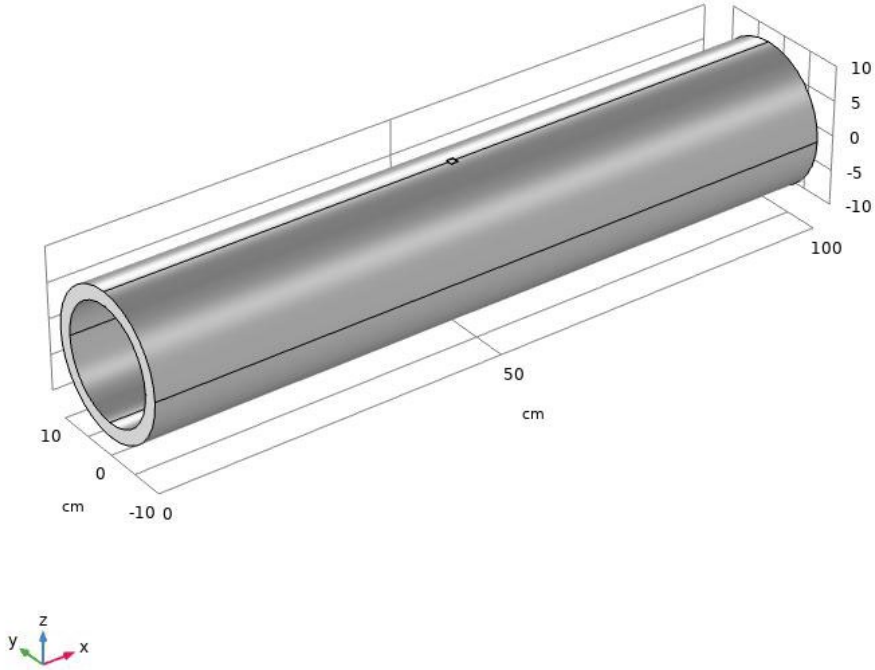


Figure 4 Pipeline simulation on COMSOL

Wall of the pipe is subjected to a certain amount of pressure when gas or liquid flow through within the pipe. In addition to stress from internal pressure, pipelines are subjected to significant longitudinal strain due to surrounding soil movement. The effect of elastic and plastic deformation on pipeline corrosion is demonstrated below.

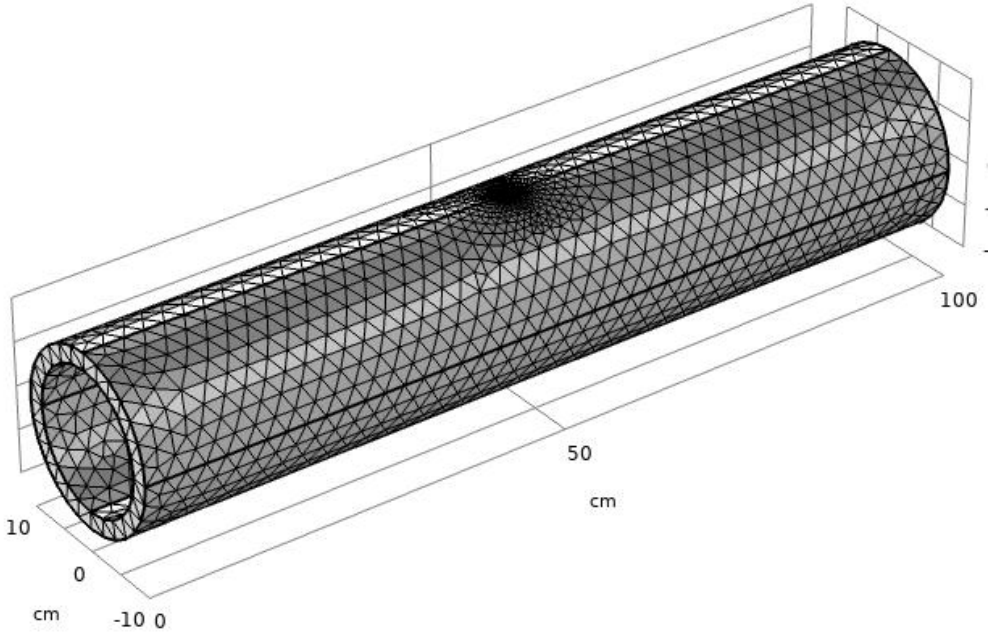


Figure 5 Mesh Plot of the entire pipeline simulation

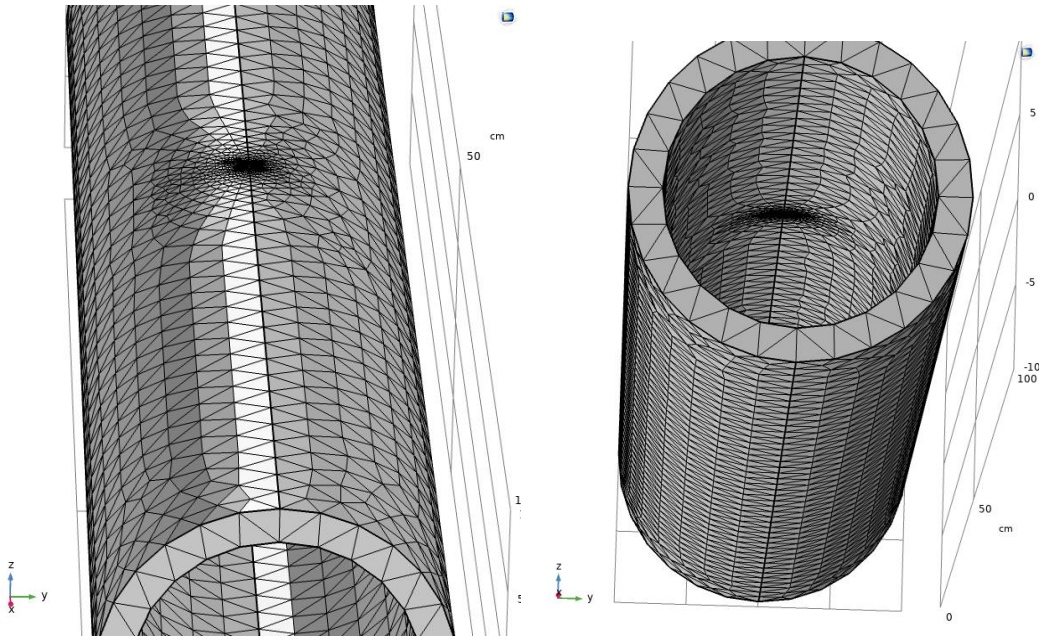


Figure 6 Left: OD defect, Right: ID defect

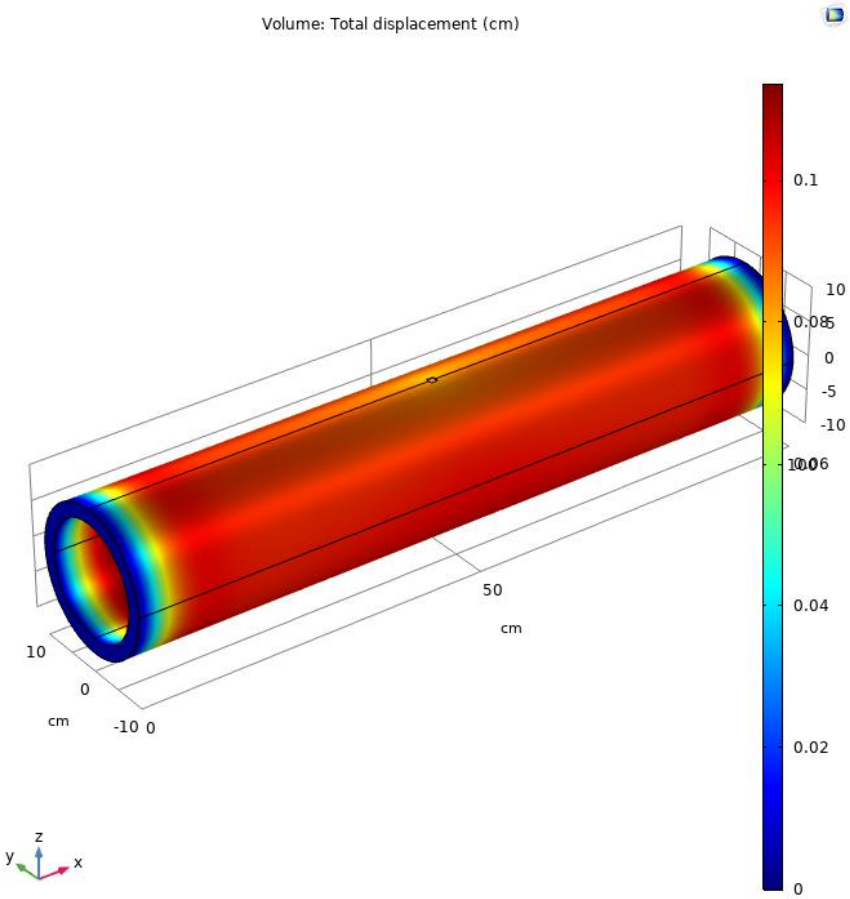


Figure 7 Total displacement distribution with pipeline under inner pressure

Von Mises stress distribution

Von Mises stress is a value used to determine if a given material will yield or fracture. It is mostly used for ductile materials, such as metals. The von Mises yield criterion states that if the von Mises stress of a material under load is equal or greater than the yield limit of the same material under simple tension — which is easy to determine experimentally —, then the material will yield. When a body in an initial state of equilibrium — or undeformed state — is subjected to a body force or a surface force, the body deforms correspondingly until it reaches a new state of mechanical equilibrium or deformed state. The inner body forces are the result of a force field, such as gravity, and the surface forces are forces applied on the body through contact with other bodies. The relations between external forces — which characterize what is called the stress — and the deformation of the body, which characterizes strain, are called Stress-Strain relations. The figures below illustrate Von Mises stress distribution when pipeline is loaded with pressure of 600 MPa on axial direction

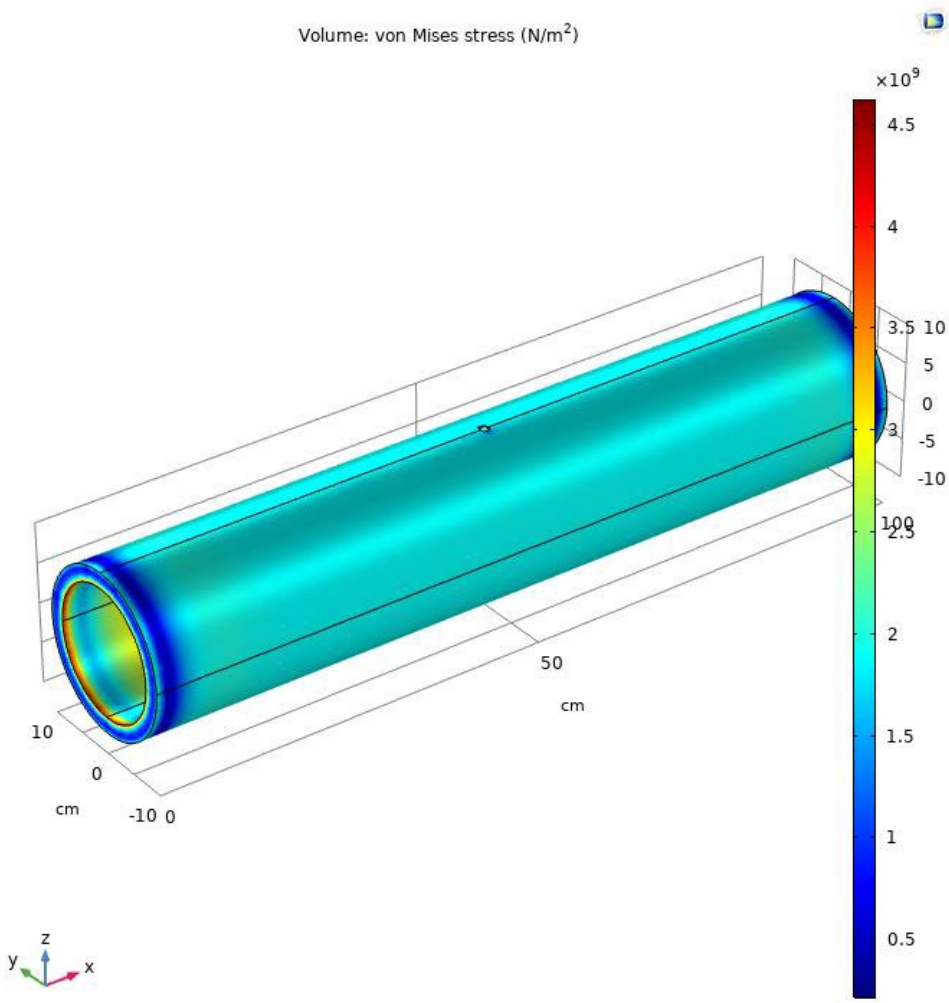


Figure 8

Interacting threats observation

We make some observation on interacting relations between OD and ID defect. The size of OD and ID regarding depth is classified as small (1mm), middle (5mm) and large (9mm), 9 combinations are illustrated in table below.

Table 6: Combination of OD and ID

OD / ID	S	M	L
S	SS	SM	SL
M	MS	MM	ML
L	LS	LM	LL

Total displacement distribution in terms of 9 combinations are shown in the graph below:

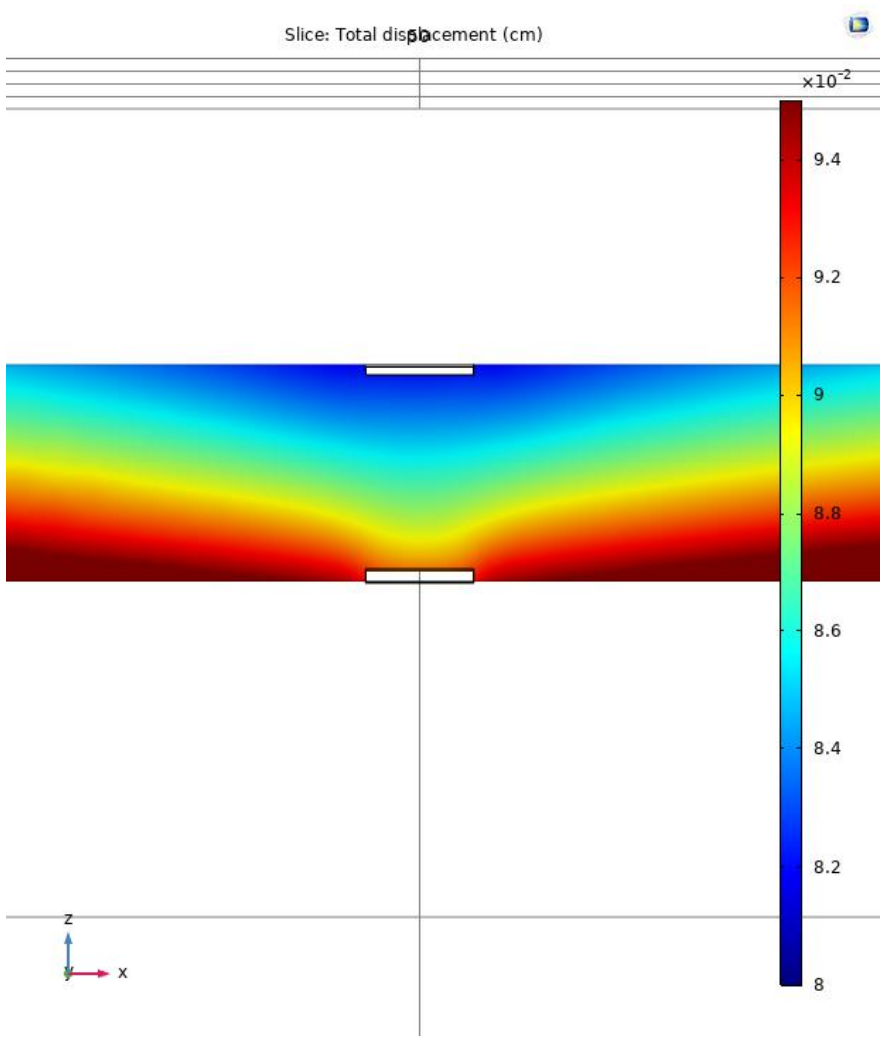


Figure 9: Total displacement distribution for SS

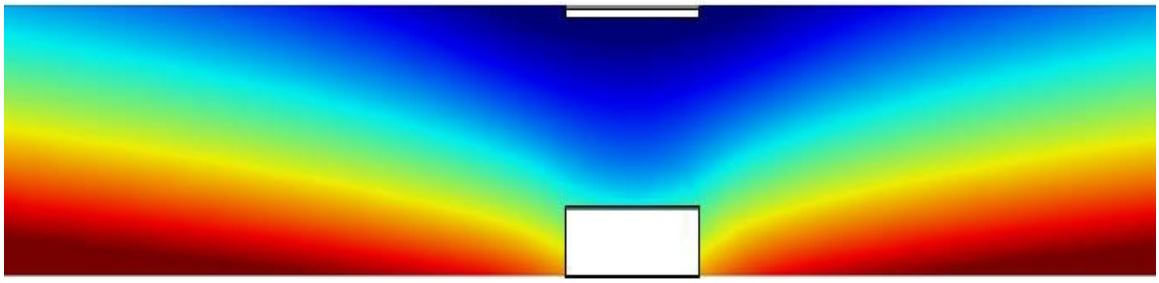


Figure 10: Total displacement distribution for SM

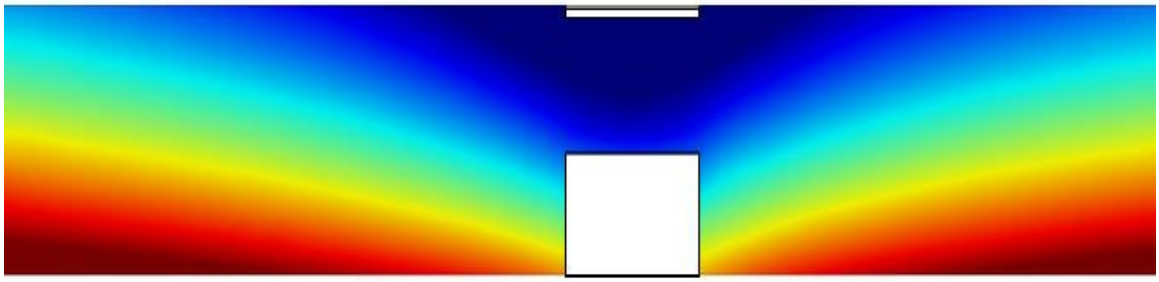


Figure 11: Total displacement distribution for SL

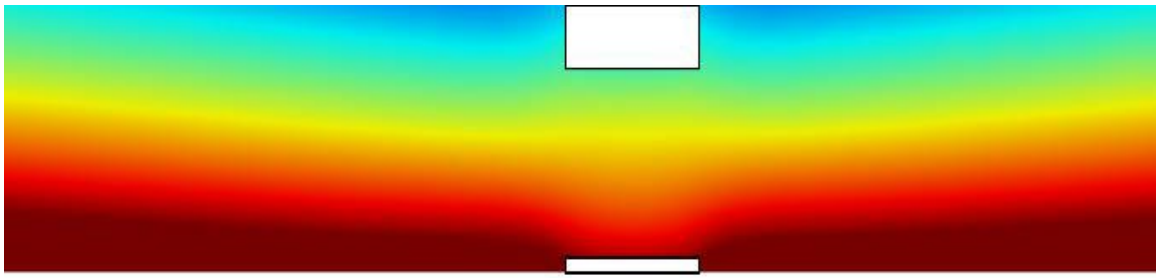


Figure 12:Total displacement distribution for MS

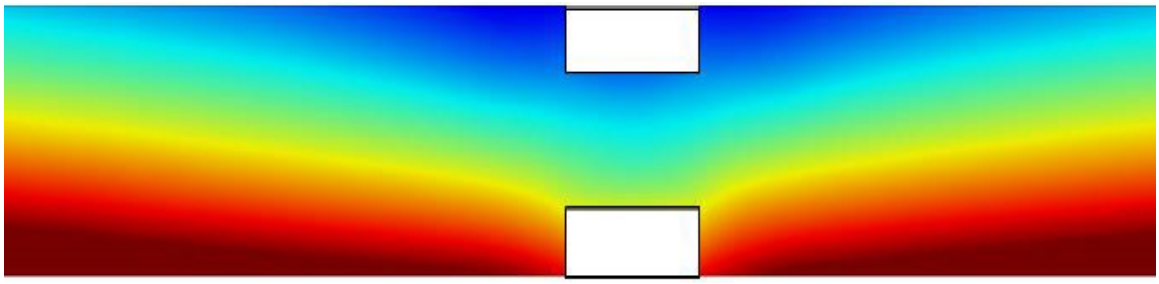


Figure 13:Total displacement distribution for MM

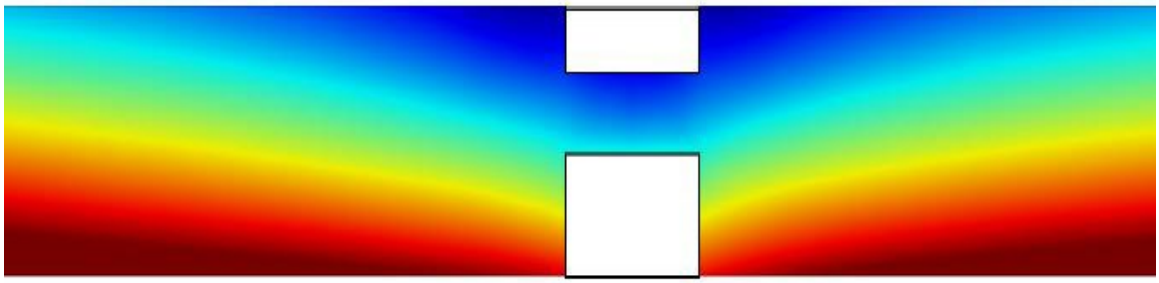


Figure 14:Total displacement distribution for ML

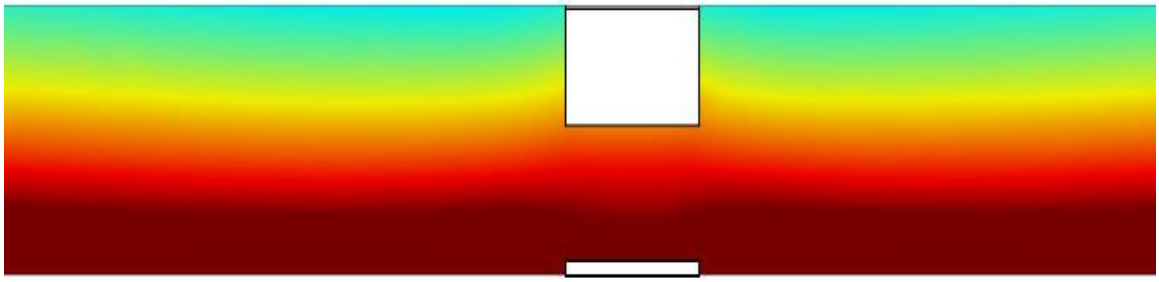


Figure 15:Total displacement distribution for LS

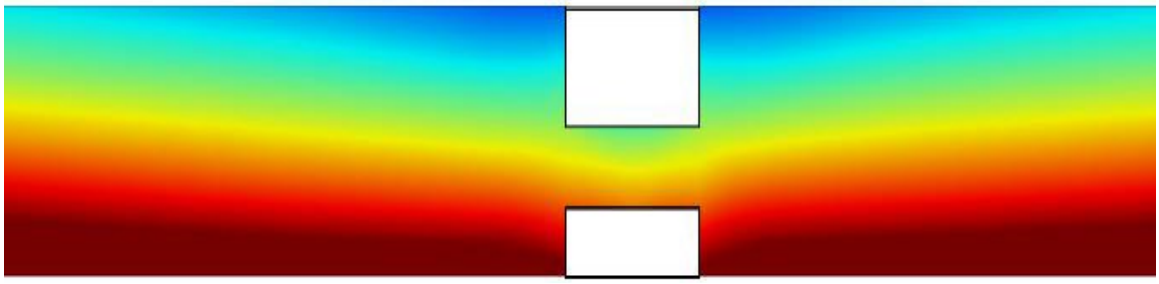


Figure 16: Total displacement distribution for LM

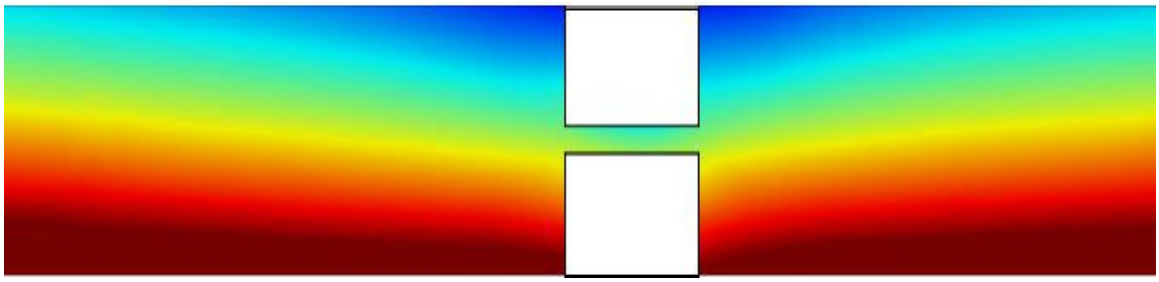


Figure 17: Total displacement distribution for LL

Conclusion

In this report, types of threats, threat assessment and modelling, decision making are covered which illustrate work flow of interacting threat analysis. The different sorts of interactive threats

and the importance to detect them properly has been studied here. In the later stages we will build our own mathematical models using the NDE data. Here we have used finite element modelling to simulate the defect models and thereby feeding them to different classification algorithms like CNN, LSTM so that we can predict the shape and size of unknown interactive defects with better accuracy. We focus on threat combinations that are likely to occur by frequency analysis in terms of historic dataset. Thus, threats repeatedly observed under inspection will be studied and simulated in our future research. In COMSOL simulation study, a pipeline under interior pressure is simulated. Interaction between OD and ID defects can be observed in terms of displacement distribution. More detailed and quantitative description will be given in our further studies.

3. Progress on Task 2 – Image Segmentation for Extracting Interacting Threats

Image segmentation refers to the decomposition of an image into a collection of disjoint regions, which is essential to identify Regions of Interest (ROIs) for threat characterization and also a key stage for further feature extraction and threat classification. We implemented three methods for image segmentation, which are K-means clustering, region proposal generation and deep learning based methods respectively.

3.1 Approach

Unsupervised learning

We use k-means clustering algorithm which is an unsupervised learning method to do image segmentation [15], its essence can be regarded as a kind of pixel clustering process. Based on the calculation of the similarity between pixels, we can classify the class of each pixel, the fundamental process is as followings:

- Initialize the number of cluster, ***k***
- Repeat the following steps until converge
  - Calculate the similarity between each pixel and the center of each cluster
  - Assign the current pixel to the cluster with largest similarity
- Recalculate the center of each cluster
- Output the classification of each pixel and the center of each cluster

In our implementation, we use graph value to calculate the similarity. But traditional k-means algorithm is sensitive to the initial center of each cluster, so we use the k-means++ algorithm [16] which can make the distance between ***k*** clustering centers to be initialized be as large as possible. The value of ***k*** is a hyper-parameter which can be obtained experimentally, in our implementation, we set ***k*** to 2.

The traditional segmentation methods aim to assign each pixel in an image to one of k cluster, like k-means. Ideally, we would like to generate a complete segment for each object in the image. But actually it is very hard to realize. If the foreground object merged with the background, like they have similar color, then the classic segmentation methods will failure.

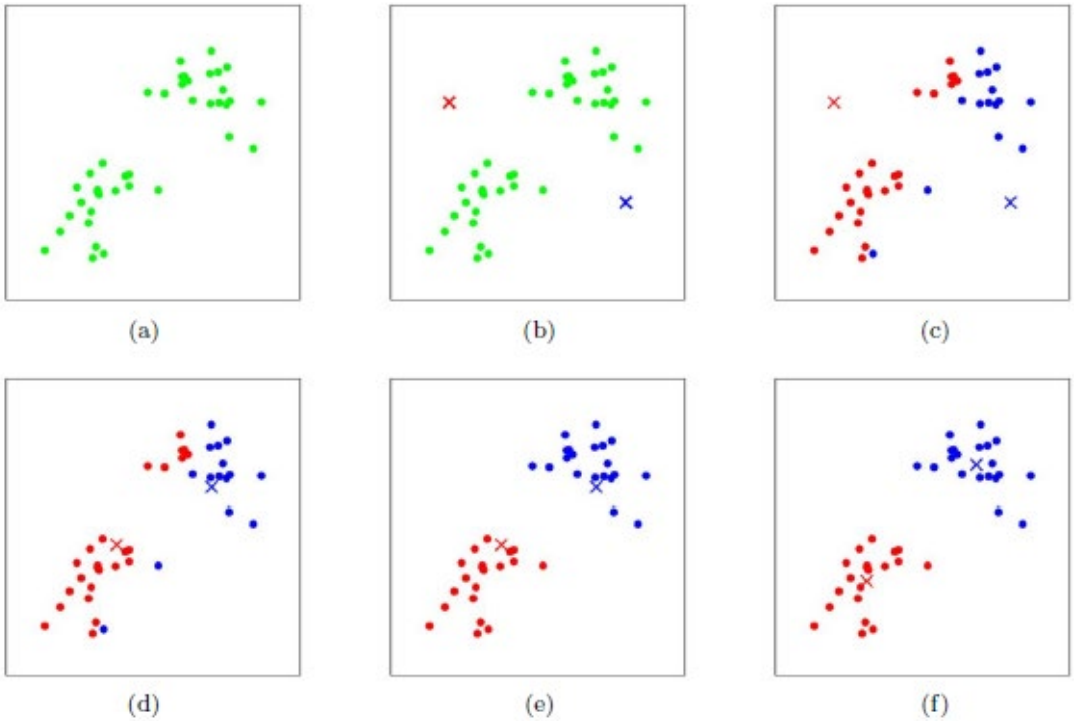


Figure 18 k-means clustering

Graph theory

Nowadays, lots of state-of-the-art image segmentation methods use region proposal [17][18] as pre-processing step to generate candidates' regions which include our target defects. In this report, we implemented a similarity grouping based method to generate proposal regions instead of sliding-window based method because the later one is very time consuming.

As to similarity grouping based region proposal, one way to segment an image is based on graph cut theory, in this report, we treat image segmentation as a pixel classification problem, if a pixel belongs foreground (defect), then its class is 1, otherwise (background) it is 0. Let $\mathrm{L} = \{l_1, l_2, l_3, \cdots, l_i, \cdots, l_\mathrm{p}\}$, which p is the number of pixels and $\mathrm{l}_i \in \{0,1\}$. Obviously, an ideal segmentation has all the boundaries of defects located at foreground and background. We define the energy function as followings,

$$\mathrm{E}(L) = \partial R(L) + B(L)$$

where $\mathrm{R}(L)$ is called as region term which describes the penalty of all pixels belonging to foreground or background, $\mathrm{B}(L)$ is named as boundary term which describes the penalty of pixels on both side of a boundary, the larger pixel difference, the smaller the value of $\mathrm{B}(L)$ will be. If all the pixels are classified to the correct label, then the energy function is minimized. ∂ is used to control the weight between region term and boundary term, we set ∂ to 0.5.

We solve the energy function by using max-flow/min-cut algorithm. According to Boykov and Jolly's theory[19], we can convert the minimization of the energy function to a problem of finding graph min-cut. Let us define each pixel as an vertex, and all the pixels are fully connected, then we get a graph G=<V,E>, and then we add two terminal vertices ***s*** and ***t***, which describe foreground and background respectfully, each edge in the graph has an weight/cost denoting as w_e, we want to cut the graph into two disjoint subgraphs with the least cost, in another word, we want to find an edge subset C of E and minimize the following function,

$$\min \sum_{e \in C} w_e$$

This cut named min-cut which is equivalent to minimizing the original energy function due to Boykov and Jolly's theory. Since min-cut and max-flow are dual, we can easily use Ford-Fulkerson algorithm (find max-flow in a graph) to find the optimal solution. The whole process is as the following figure.

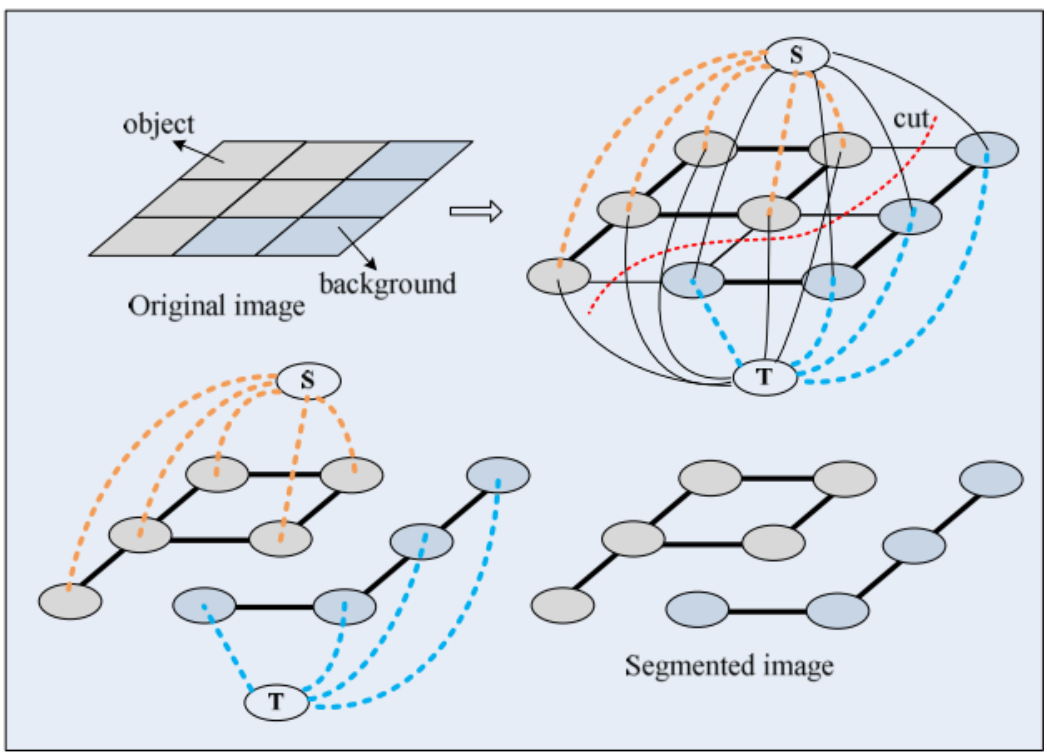


Figure 19 graph cut in image segmentation

Region proposal based methods focus on improving recall, which means that they aim to find as much candidates as possible but has poor performance on preserving the boundary of defect. Thus, we finally use deep learning based methods to do image segmentation, which have better performance of defect boundary according to our experiments.

Deep learning

In this report, we use DeepLab v3+ to segment defects from an image. The TensorFlow DeepLab v3+ model is the state-of-the-art semantic segmentation method that achieved 89.0% mean IoU in PASCAL VOC 2012 dataset and 82.1% mean IoU in Cityscapes dataset [10]. This network combined the advantages of spatial pyramid pooling and an encoder-decoder structure. The former method excels at encoding multiscale contextual information, effectively handling objects with different scales, and the latter method produces sharper object boundary by a coarse-to-fine recovery of spatial information.

Training a deep neural network from scratch for semantic segmentation typically require a large-scale dataset with more than 10,000 annotated images. For example, the leading semantic segmentation models trained on open benchmarking datasets PASCAL VOC, Cityscapes and ADE20K achieved more than 80% mean Intersection over Union (mIoU)[9] [10]. These benchmarking datasets consist of 16408, 3475, 22210 annotated images respectively.

But in the NDE field, there are few annotated open datasets for semantic segmentation. Such datasets are usually domain specific, proprietary, or only have a few images (500 images). With few annotated data available, it is difficult to train a deep neural network directly due to the following reasons: (1) the network might overfit to the few instances in the dataset, (2) insufficient data variance might slow down training and trap the model in local minimum.

A common approach to address insufficient data in deep neural network model development is transfer learning. In transfer learning, researchers take a starting point from a network pre-trained on large-scale open dataset like ImageNet, remove the last few layers, the output layer, the softmax layer and the logit layer and retrain these layers on a new domain with smaller annotated dataset. The intuition is, after training on a generic large-scale dataset, the deep neural network can learn to effectively extract salient features, these low-level features extraction pipeline can then be adapted to analyze images from a new domain. In recent years, successes have been reported in deep neural network transfer learning on medical imaging, where annotated data must be labeled by experts and is limited in size[10].

In the later experiment, we performed transfer learning on the state-of-the-art semantic segmentation deep neural network (TensorFlow DeepLab v3+) [9] [10] and tested the efficacy of the model on an open magnetic tile defect dataset [12] [13].

3.2 Experimental Program

Magnetic tile defect dataset

The *magnetic tile surface defect dataset* consists of defect tiles greyscale images and pixel level mask annotation of the defect. [12] [13] There are five defect classes in this dataset, blowhole, break, crack, fray and uneven. Each class has 50~120 instances. The selected samples in the dataset are shown as following.

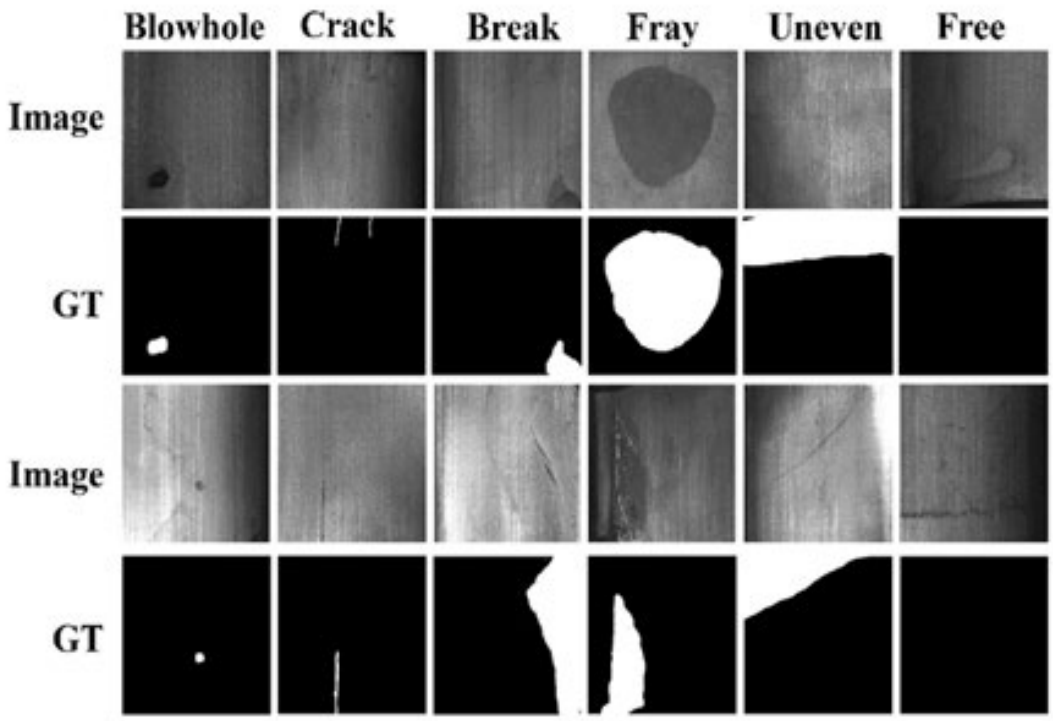


Figure 20 dataset illustration

Experimental design

Our k-means and region proposal methods were implemented in MATLAB. All the experiments were performed using a PC with 3.4GHz Intel Core i5 CPU and 32 GB RAM. DeepLab v3+ was implemented with TensorFlow and performed using Nvidia GPU GTX1080.

We use mIoU to evaluate our methods, IoU is defined as following, which describe the accuracy of segmentation result.

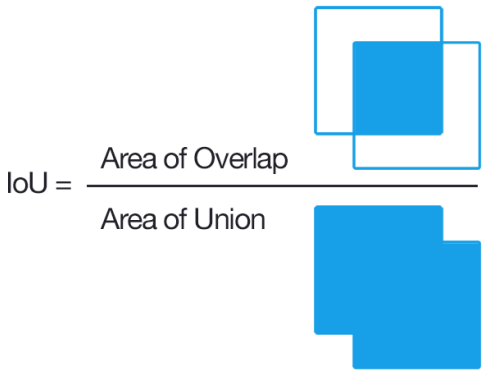


Figure 21 IoU concept explanation

For k-means, we set k to 2 and for region proposal, we set hyper parameter ∂ to 0.5. We randomly pick 100 images from our testing dataset and calculate the mIoU. For DeepLab v3+, we split the magnetic tiles surface defect dataset into training set, training validation set and validation set as table 1.

We selected the xception65_coco_voc_trainval model as backbone, which achieved 87.80% mean IoU on PASCAL dataset. We then removed the output layers and the logits layer and replaced them with an output layer representing the five defect classes. In this defect dataset, the blowhole, break and crack can be small, the contribution from a successful detection on defect pixels might be overwhelmed by the background pixels. If left unchecked, the network predicting a blank background could get a good value in the loss function. So, we set the relative weight of background pixel and foreground pixel as 1:10 in the loss function.

Table 7 Dataset organization

Classes	**Training Set**	**Training Validation Set**	**Validation Set**
Blowhole	80	19	16
Break	59	14	12
Crack	39	14	8
Fray	40	10	7
Uneven	71	18	14

Experimental result

We use mIoU to evaluate the performance of these three methods, the quantitative result is shown in the following table.

Table 8 IoU quantitative result

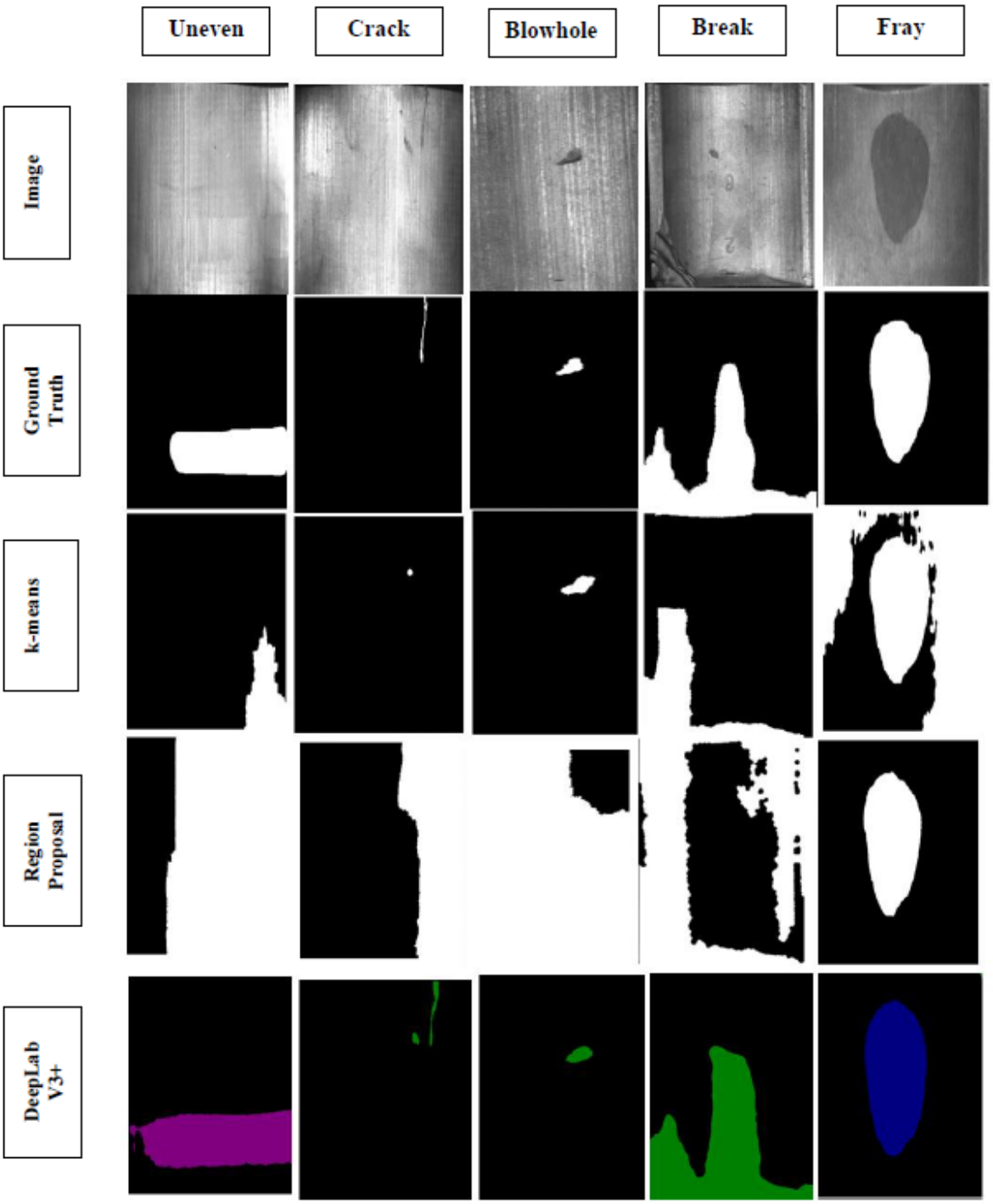
Method	K-means	Region Proposal	DeepLab v3+
mIoU(%)	29.43	20.14	53.60

We can see that deep learning based method outperforms than the others, the final model achieved 53.6% mean IoU on the magnetic tiles surface defect dataset. Region proposal based method is a little worse than k-means, as we discussed before, the previous method focus on recall rather than accuracy.

Given the qualitative results, we can see that DeepLab v3+ achieved the best performance on different kinds of defects. As to k-means and region proposal based method, they are influenced heavily by the color contrast between foreground and background and graph cut based region proposal method works well when there exist clear boundary of defect but fail on break defect which has fuzzy boundary.

Based on our experiment, it is also noticed that the DeepLab v3+ failed to distinguish between blowhole, break and crack defects. It is possible that these three classes of defect all look like small patches of darken pixels, so that the network was unable to distinguish. But for our purpose, proposing ROIs for further defect characterization and tracking, DeepLab v3+ is enough.

Figure 22 Qualitative results



3.3 Future work

Perform training on annotated defect data specific to our application.

In this period, we have proven that transfer learning is effective on a small dataset from a new domain. In the next period, if enough images of defect with mask annotations are obtained through FEM simulation or experiments, they can be used to train the semantic segmentation network, to provide ROI proposals specific to our application.

Perform Deep Neural Network fine tuning techniques.

There exist multiple fine-tuning techniques for improving the accuracy of Deep Neural Network models. Notable examples are (1) data augmentation, flipping the images to increase training data, so that the network can see more data, and detect orientation irrelevant features. (2) Longer training time, the network could reach higher accuracy with extended training. (3) Auto-DeepLab, perform hierarchical neural architecture parameter search automatically to improve network performance[14].

Perform hardware architecture level optimization.

Migrate the trained model into hardware accelerated framework such as Nvidia TensorRT to speedup inferencing. The hardware-accelerated TensorRT framework can speedup neural network inferencing by combining multiple computation instructions within a layer into one cycle, and handle tensor computation in 4x4x4 unit, perform tensor computation concurrently on each element of the 4x4x4 tensor, instead of iterating through 64 elements sequentially.

References

[1] Eduardo Munoz and Michael J. Rosenfeld, P.E., 0339-1403_PHMSA_Final_Report. 2016

[2] Chen, Y., et al., Failure assessment of X80 pipeline with interacting corrosion defects. Engineering Failure Analysis, 2015. 47: p. 67-76.

[3] America, I.N.G.A.o., Interacting Threats to Pipeline Integrity – Defined and Explained.

[4] Administration, P.a.H.M.S., Pipeline Risk Modeling Overview of Methods and Tools for Improved Implementation. 2018

[5] Technologies, M.S.C.-F., PHMSA Risk Model Work Group Meeting: Methods for Probability Estimation. 2016.

[6] Group, R.M.W., The “Risk Modeling Work Group” Discussion of Interactive Threats

[7] 5.3a, C.i.C.M., Modeling of mechanoelectrochemical coupling in pipeline corrsion.pdf

[8] 5.3a, C.i.C.M., Stress Corrosion.

[9] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff and A. Hartwig, "DeepLab: Deep Labelling for Semantic Image Segmentation," 18 January 2019. [Online]. Available: https://github.com/tensorflow/models/tree/master/research/deeplab.

[10] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," in Computer Vision -- ECCV 2018, 2018.

[11] C.-K. Shie, C.-H. Chuang, C.-N. Chou, M.-H. Wu and E. Y. Chang, "Transfer representation learning for medical image analysis," in Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, 2015.

[12] Y. Huang, C. Qiu, Y. Guo, X. Wang and K. Yuan, "Surface Defect Saliency of Magnetic Tile," in IEEE International Conference on Automation and Engineering, Munich, Germany, 2018.

[13] Y. Huang, "dataset of the upcoming paper "Saliency of magnetic tile surface defects"," [Online]. Available: https://github.com/abin24/Magnetic-tile-defect-datasets.. [Accessed 18 March 2019].

[14] L. Chenxi, L.-C. Chen, F. Schroff, H. Adam, H. Wei and F.-F. Li, "Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation," *arXiv:1901.02985,* 2019.

[15] Ray, Siddheswar, and Rose H. Turi. "Determination of number of clusters in k-means clustering and application in colour image segmentation." Proceedings of the 4th international conference on advances in pattern recognition and digital techniques. 1999.

[16] Arthur, David, and Sergei Vassilvitskii. "k-means++: The advantages of careful seeding." Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, 2007.

[17] Humayun, Ahmad, Fuxin Li, and James M. Rehg. "RIGOR: Reusing inference in graph cuts for generating object regions." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.

[18] Carreira, Joao, and Cristian Sminchisescu. "CPMC: Automatic object segmentation using constrained parametric min-cuts." IEEE Transactions on Pattern Analysis and Machine Intelligence 34.7 (2012): 1312-1328.

[19] Boykov, Yuri Y., and M-P. Jolly. "Interactive graph cuts for optimal boundary & region segmentation of objects in ND images." Proceedings eighth IEEE international conference on computer vision. ICCV 2001. Vol. 1. IEEE, 2001.