

CAAP Quarterly Report

Date of Report: *April 7, 2019*

Contract Number: 693JK31850007CAAP

Prepared for: *Robert Smith, Project Manager, PHMSA/DOT*

Project Title: *A novel structured light based sensing and probabilistic diagnostic technique for pipe internal corrosion detection and localization*

Prepared by: *Mohand Alzuhiri, Yuhao Wang, Dr. Yiming Deng, Dr. Yongming Liu*

Contact Information: *Dr. Yiming Deng (MSU) and Dr. Yongming Liu (ASU)*

For quarterly period ending: *April 7, 2018*

Business and Activity Section

(a) Generated Commitments

Project abstract: Internal corrosion in pipes is dangerous due to multiple factors contributing to its development. Degradation of the pipeline health is susceptible to hazard due to failure. To prevent such failures, a major challenge for the maintenance crew to detect and repair corrosion still prevails due to difficult and expensive accessibility during scheduled maintenance. The proposed method will focus on the development of novel structural light-based imaging for internal corrosion detection, which simplifies the detection process while achieving superior spatial resolution. The proposed approach will develop an endoscopic structured light scanning tool that is based on phase measurement profilometry (PMP). The developed system will be simple to fabricate and easy to be used by maintenance personnel with minimal skillset due to its intuitive scans. The structured light system will be developed to generate high-resolution reconstructed images representing surface texture with high accuracy. Based on the images, additional processing capabilities developed using Bayesian updating technique will give the capability of automatic classification and identification of different types of precursors. A convolutional neural network-based corrosion detection method will provide automated detection, which further minimizes the operator involvement. The uncertainty quantification technique will be integrated to enhance the probability of detection and to quantitatively determine the damage size and location.

Based on the identified challenges and state-of-the-art techniques, a complete solution is needed to detect, localize and evaluate internal corrosion in metal pipelines. Our proposed solution is to develop a phase measurement profilometry (PMP) structured light-based tool to detect and locate any surface defects and damages on the pipe wall, which includes specifically corrosion. Internal corrosion will be effectively detected and evaluated by checking the amount of material loss, color change, spread and pattern in the pipe wall simultaneously with a capability of integrating with ILI platforms. This optical inspection is also integrated with a set of numerical tools to evaluate structured light sensor information in order to automate the analysis of inspection data. The specific technical objectives/goals of the proposed research are:

- Design and develop a PMP structured light-based in-line inspection endoscopic scanner. The deliverables include:
 - design a new SL module to produce patterns with high resolution and contrast.
 - develop a new scheme to calibrate the new PMP based projector and camera(s).

- Develop a new reconstruction algorithm called moving phase measurement profilometry (MPMP) to exploit the system movement of the scanner along the pipe to enhance the quality of the reconstruction and detection.
- Evaluate the suitability of different optical methods like stereo cameras to enhance the performance of corrosion detection.
- Develop a convolutional neural network-based model for the automatic detection and classification of corrosion damages from the provided structured light sensor data to mitigate the need for manual analysis of 3D sensor massive data.
- Develop an uncertainty quantification technique to enhance the probability of detection and to quantitatively determine the damage size and location.

Educational Objectives: Another major objective of the proposed effort is to inspire, educate and train Ph.D. and MS students to address pipeline safety challenges, potentially as a career after their graduation. If funded, two Ph.D. students from both universities and several MS/undergraduate students will be included in this CAAP program. They will be trained and educated in science and engineering to address pipeline safety and integrity challenges. The PIs believe education is a critical component of the CAAP project, and we will integrate research with educational activities to prepare the next generation scientists and engineers for the gas and pipeline industry. Specific educational objectives include:

- Inspiring, educating and training the graduate students at MSU and ASU as research assistants for pipe integrity assessment and management. Our previous successful CAAP projects have produced several engineers, researchers and summer intern in gas and pipeline industry,
- Integrating research topics from this effort with the existing undergraduate research programs at MSU, e.g. ENSURE program at the College of Engineering and ASU to involve undergraduate students in pipe safety research.
- Improving the curriculum at MSU (e.g., Nondestructive Evaluation) and ASU (e.g., Machine Learning and Artificial Intelligence) using the scientific findings and achievement from the proposed research,
- Adapt research topics from this project to student projects in seminar, senior design, and project courses, in order to make educational impacts on broader groups of students,
- Encourage the graduate research assistants involved in this project and students in the courses to apply for internships at USDOT/PHMSA and industry to practice their learned skills and gain practical experiences in areas related to pipe safety and integrity.

The above-mentioned goals and objectives of this CAAP project will be well addressed and supported by the proposed research tasks. Development, demonstrations and potential standardization to ensure the integrity of pipeline facilities will be carried out with the collaborative effort among two different universities and our industry partner, Gas Technology Institution. This MSU-ASU-GTI team has successfully completed several PHMSA projects including “Slow Crack Growth” study, which was ranked No. 2 overall in all core PHMSA projects in 2017. The quality of the research results will be overseen by the PIs and DOT program manager and submitted to high-profile and peer-reviewed journals and leading conferences. The proposed collaborative work provides an excellent environment for the integration of research and education as well as tremendous opportunities for two universities supported by this DOT CAAP funding mechanism. The graduate students supported by this CAAP research will be heavily exposed to ILI, NDE, reliability and engineering design topics for emerging pipeline R&D technologies. The PIs have been actively encouraging students to participate in past and ongoing DOT

projects and presented papers at national and international conferences. Students who are not directly participating in the CAAP project will also benefit from the research findings through the undergraduate and graduate courses taught by the PIs and attending university-wide research symposium and workshop.

(b) Status Update of Past Quarter Activities

Task 1 – 3D Acquisition sensor design

This task is specified to develop structured light-based scanning sensor for internal corrosion detection, which simplifies the detection process and decreases the scan time.

- Explore the suitability of stereo vision techniques for damage shape reconstruction.
- Develop an endoscopic structured light scanner.
- Fabricate an easy to use interface for intuitive scans.

A. Review of the corrosion and its inspection methods:

Pipelines are one of the important methods to transport natural gas and hazardous liquid products throughout the United States. They are attractive because of their high efficiency when compared to railroad and trucks. While metal pipelines provide are rigid and can withstand high pressure levels, their deterioration due to environmental condition is a real concern for safety of the operation. Different type of corrosions can attack the pipe walls which leads to reduction in the thickness of the pipe wall and then pipe failure. PHMSA database indicates that 12% of pipeline incidents in the United States were caused by internal corrosion from 1998 – 2017. The significant factors caused those incidents are the degrading nature of iron alloys, the age of pipelines and pipe joints design. The project will mainly focus on detecting and monitoring the corrosion condition of those types of pipeline that both newly installed and existing pipelines. The early detection and inspection are necessary to minimize the impact of internal corrosion on the pipe integrity. The internal corrosion can cause localized metal loss on the interior surface of pipeline system resulting reduction of the wall thickness. In some worst-case scenarios, the loss of pipeline material can result the leakage of natural gas and hazardous liquid product from pinhole or cracks. The main concern of this project is to detect internal corrosion inside gas transmission pipelines. Department of transportation data indicate that 60% of all transmission and gathering pipelines incidents are caused by the internal corrosion among all the corrosion types [1]. In the past years, in order to clean the pipeline internal corrosion, the pipeline industry has used scrubbing and scraping devices. However, after the more accurate and smarter in-line inspection (ILI) tools had been developed, the detection of corrosion, cracks, laminations, deformation, etc. became easier and more reliable. With the cutting-edge data processing techniques, the data from pipeline corrosion incidents can provide variety of information in order to help pipeline repairment and replacement. There are several ILI tools described below. The magnetic flux tools including magnetic flux leakage (MFL) tool and transverse MFL/ Transverse Flux Inspection tool (TFI)[4] can identify the metal loss by applying temporarily magnetic field to the internal wall of pipeline to detect the change in distribution of the magnetic flux resulted by defects on the internal pipeline. The TFI tool uses different magnetic field from the MFL tool in order to be able to detect more types of corrosion. The Ultrasonic Tools (UT) includes compression wave ultrasonic testing tools and shear wave ultrasonic testing tools . The first tool equipped transducers in order to emit ultrasonic signal perpendicular to the surface of the pipe. Through processing the timing of return signals, the thickness can be determined. The second tool can generate shear wave to the pipeline wall through angular transmission of UT pulses through a kind of liquid coupling medium to detect the different physical damages on the wall of the pipe. There are tools that can detect a certain type of defect also, such as geometry tool. However, all of the tools including above ILI tools cannot determine most of defects including physical damages and chemical corrosions in the same time. To make the corrosion detection method more universal, the project will focus on the development of novel structural light-based imaging for internal corrosion detection, which simplifies the detection process while achieving superior spatial resolution. The proposed approach will develop an endoscopic structured light scanning tool that is based on phase measurement profilometry (PMP). The structured light system will be

developed to generate high-resolution reconstructed images representing surface texture with high accuracy.

B. Structured light sensor design:

Basic structured light sensor consists of two main components which are the light projector and the camera. The light projector is used to project spatially or temporally coded light patterns on the scanned object surface. The projected image is deformed by the shape of the scanned object according to its height changes. The deformations in the light pattern are then captured by the camera and the height is then extracted by measuring the amount of shift in the captured pattern when it is compared to the original projected pattern. Our proposed sensor takes it to the next level by exploiting the sensor movement to improve the quality of the reconstructed images. Our original sensor uses two camera and projector that are placed in front each other and measure the pipe defects by using a stereo orientation in cylindrical geometry. This design enabled us to scan pipes with diameters below 2 inches. The large pipe diameter (8 inches) of gas transmission lines gives us the flexibility to improve the hardware used in the sensing module to improve its performance. Two improvement has been added:

- The scan will be performed in rectangular geometry where the camera is facing the scanned wall
- An additional camera has been added to give the capability of stereo vision.

Scanning in rectangular geometry will simplify the reconstruction process where we don't need to transform the data to polar coordinate. In this process we are also eliminating the error introduced during the polar transformation. Another advantage is the ability to use multiple cameras to cover the 360 degree around the scanner which will increase the scanning spatial resolution. The addition of the stereo capability to the system will be used to track the system movement inside the pipe to act as initial seed to the moving PMP system. A schematic of the proposed sensor is shown in Figure1. The letter C refers to the digital camera and the letter P refers to the light projector.

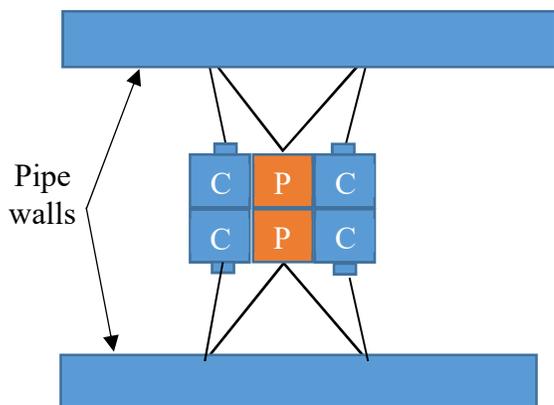


Figure 1: Structured light sensors schematic

An experimental setup has been fabricated to serve as a test bed to evaluate the scanning system outside the pipe. A schematic of the experimental setup is shown in Figure 2.a. A fabricated experimental test bed is shown Figure 2.b. The experimental setup consists of two stereo cameras that have the capability to record a full HD images with 30 frames per second. The cameras are attached to low distortion 100-degree field of view lenses to be able to reconstruct scenes at close range. The cameras are attached to a digital light projector that provide the required patterns to be able to reconstruct the scene in front of the camera. Figure 3 shows the acquired images from the system with the use of random dots patterns. Figure

4 shows the acquired images from the system with the use of sinusoidal patterns. The current system is also attached to moving platform to enable performing moving PMP scans.

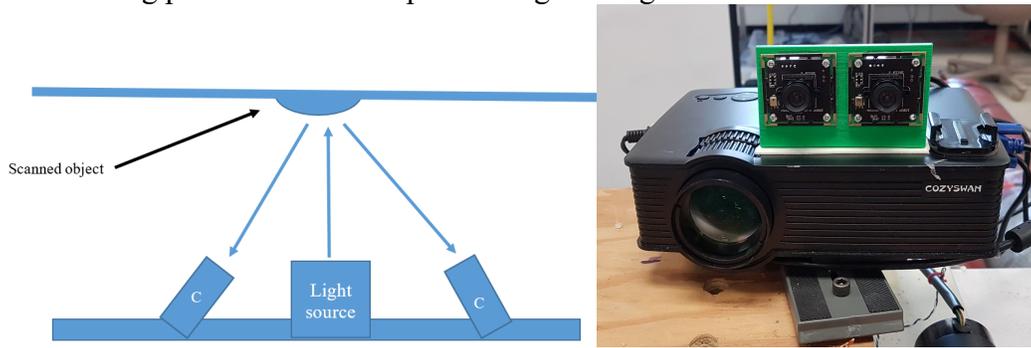


Figure 2:a) schematic of the scanning system, b) implemented experimental testbed



Figure 3: Stereo images with random patterns



Figure 4: Stereo images with sinusoidal pattern

C. Sensor calibration:

Sensor calibration is performed to extract the system parameters that are required to perform the scanning process. The calibration process also includes correction of the distortion caused by the projection lenses. The camera calibrations including single camera, stereo camera and projector camera measure the parameters of image sensor of the imaging sensor. In the camera system, three sets of parameters are needed for accurate reconstruction. These parameters are the internal matrices of the camera and projectors, and the stereo parameters between the projector and camera (rotation and translation vectors).

Single camera calibration:

The pinhole imaging system can be described as a combination of the intrinsic matrix, rotation, and translation matrices. These matrices give us the ability to project Cartesian world points to camera coordinates. Therefore, a camera system can be described by the following equation

$$S = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_{11} & r_{12} & r_{31} & t_1 \\ r_{21} & r_{22} & r_{32} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$K = \begin{bmatrix} fx & 0 & 0 \\ s & fy & 0 \\ cx & cy & 1 \end{bmatrix}$$

where u, v are the coordinates of the corresponding image point, x, y, z are the world coordinate of the points, $r(ij)$ and $t(i)$ are the rotation and translation parameters and s is an arbitrary scaling factor. Here K is the intrinsic camera matrix that specifies the focal lengths and the coordinates of the principal points and represents a projective transformation from the 3-D camera's coordinates into the 2-D image coordinates.

In addition to that, the distortion coefficients are needed to correct effect of the radial and tangential distortion from the use of lenses. The radial distortion is caused by light rays bend more at the edge than the optical center of the lens. The tangential distortion will occur when the lens and image plane are not parallel. Radial distortion can be represented as follows:

$$x_{\text{distorted}} = x(1 + k_1*r^2 + k_2*r^4 + k_3*r^6)$$

$$y_{\text{distorted}} = y(1 + k_1*r^2 + k_2*r^4 + k_3*r^6)$$

x and y indicate the normalized image coordinates which is undistorted pixel locations. The k_1, k_2 , and k_3 indicate the radial distortion coefficients of the lens. Tangential distortion can be represented as follows:

$$x_{\text{distorted}} = x + [2 * p_1 * x * y + p_2 * (r^2 + 2 * x^2)]$$

$$y_{\text{distorted}} = y + [p_1 * (r^2 + 2 * y^2) + 2 * p_2 * x * y]$$

The x and y indicate the normalized image coordinates which is undistorted pixel locations. The p_1 and p_2 indicate the tangential distortion coefficients of the lens. In this process, since we are using a flat calibration plane, a simplified 2D camera calibration scheme that was inspired by Zhang et al. [3] is used to calibrate the camera with a 2D checkerboard. The z coordinates of the board points are set to zero, therefore the camera model is simplified to become as follows:

$$S = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_{11} & r_{12} & t_2 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = S \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = M \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

In the experiments, calibration is done by taking multiple images at different orientations for a predefined plane with known fixed points. In this step, we estimate the intrinsic parameters, distortion parameters, and camera pose for each calibration frame by using MATLAB Camera Calibration function.

Stereo camera calibration:

Stereo camera calibration is needed to map the pixels between the left and right camera. Those parameters describe the rotation and translation of one camera to the other. From the single camera setup we can extract each camera location with reference to the calibration board. Both of the images from stereo cameras are corrected to remove the effect of the lens distortion. Comparing with the single camera calibration, the system of stereo camera calibration can be described as the product of intrinsic matrix and extrinsic matrix including the combination of rotation, and translation matrices. The difference between single camera and stereo camera calibration is that the rotation and translation matrices are the common matrices shared by both of the cameras. The camera system can be described by the following equation.

$$S = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} R \\ t \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$K = \begin{bmatrix} fx & 0 & 0 \\ s & fy & 0 \\ cx & cy & 1 \end{bmatrix}$$

Because we use same type of cameras for system design, both of the cameras are assumed to have the same intrinsic matrix (described in the single camera calibration part). The common matrices for the rotation and translation matrices and the relationship between the first camera location matrix in world coordinate and the second camera location matrix in world coordinate can be described by the following equations.

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = R_{12} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} + T_{12} \quad (1)$$

From equation (1), we can derive the following equations:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = R_1 \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T_1 \quad \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = R_2 \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T_2 \quad (2)$$

From equation (2), we can derive the following equation:

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = R_2 R_1^{-1} \left\{ \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} - T_1 \right\} + T_2 \quad (3)$$

Compared with equation (1) and (3), we can derive the following equations:

$$R_{12} = R_2 R_1^{-1} \quad T_{12} = -R_2 R_1^{-1} T_1 + T_2$$

In the above equations, $x_1 y_1 z_1$ indicate the camera 1 and $x_2 y_2 z_2$ indicate the camera 2 locations in the world coordinate. R_{12} Indicates the common rotation matrix and T_{12} indicates the common translation matrix. In order to find R_{12} and T_{12} , we can find the rotation and translation matrices for each of camera $R_1 T_1$ and $R_2 T_2$. Because of calibration, both of the cameras should acquire the same matrix in world coordinate from a same source. We can establish the equation to find $R_{12} T_{12}$ by using $R_1 T_1$ and $R_2 T_2$. The procedure of stereo camera calibration in the experiments is similar to the single camera calibration. The difference is that the calibration is done by taking multiple images by the left and right cameras at different orientations for a predefined plane with known fixed points in the same time. The whole

calibration procedures are based on the MATLAB Stereo Camera Calibration function. The distortion coefficients for both cameras have been measured in the single camera calibration part.

Projector camera calibration:

In the camera projector calibration procedure, similar equations to the stereo cameras can be used because the projector can also be modeled as a pinhole camera. The only difference is that the projector cannot see the world by itself, therefore we need the camera assistance during the calibration process. A calibration procedure is under development to relate the pixels on the projector image plane to the associated stereo cameras. The procedure will depend on the use Fourier transform profilometry to extract the 3D shape from a single frame while calibrating the system.

D. Algorithms evaluation with simulation environment:

Simulation environment is used to evaluate the performance of different setups, algorithms and different scanning objects. The simulation environment was setup to have similar configuration to the experimental setup. The simulation environment has 2 stereo cameras and single light projector. In order to evaluate the algorithm for scanning a surface with random shape which the case for the corroded surfaces, we are simulating the reconstruction of a random surface. A picture of the scanned surface is shown in Figure 5. Direct stereo images of the surface produced by simulation is shown in Figure 6. Stereo images with the introduction of random pattern are shown in Figure 7. Another test was performed when phase shifting is introduced by moving the scanner beside the imaged target while projecting a sinusoidal pattern. Four images are acquired at angles 0,90,180,270 and then rectified as shown in Figure 8. A direct reconstruction o these images is shown in Figure 9. The results show that the surface has been reconstructed successfully but we also notice the periodic sinusoid that we expected with direct reconstruction.

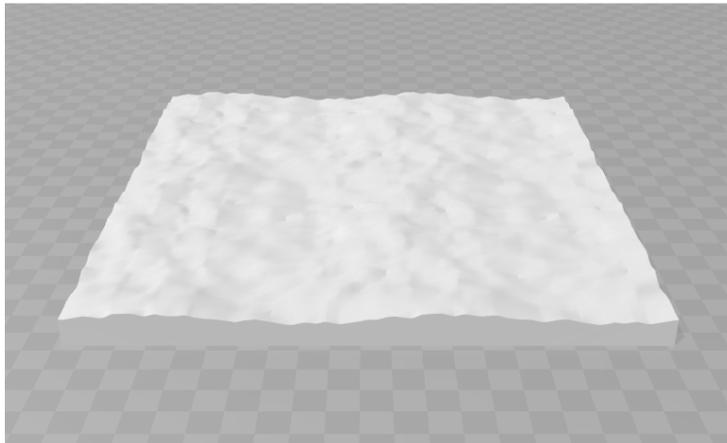


Figure 5: Simulated random surface

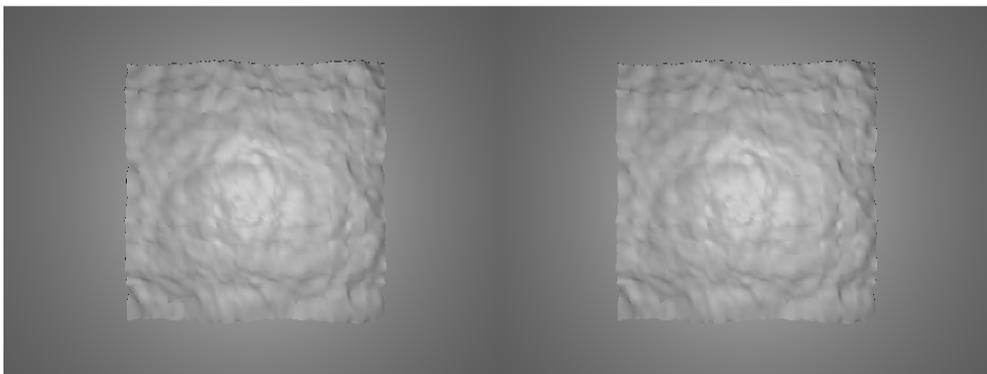


Figure 6: Direct stereo images

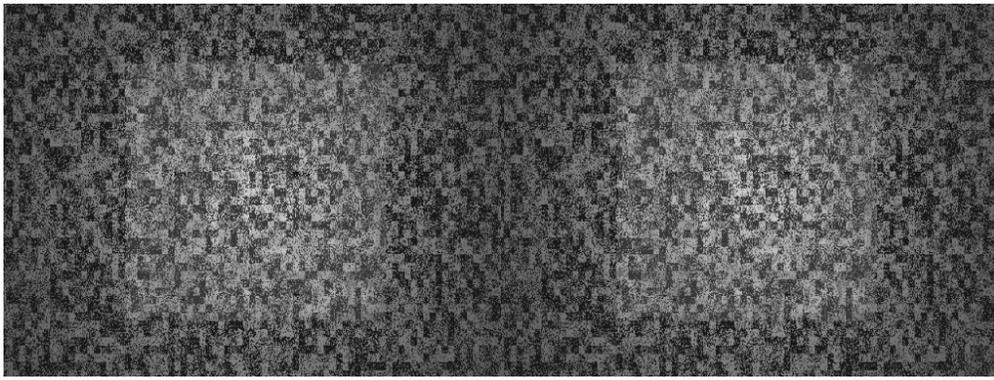


Figure 7: Stereo images with random pattern

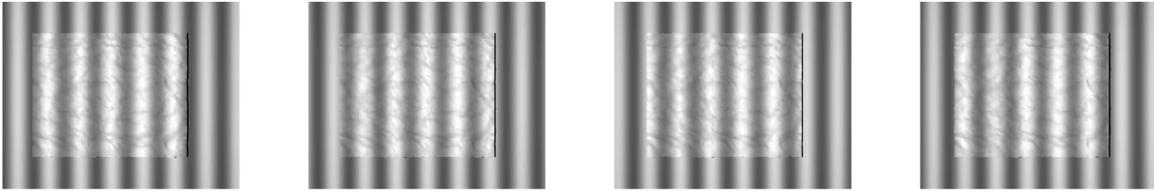


Figure 8: Four rectified images from moving sensor scanning a surface with random defects

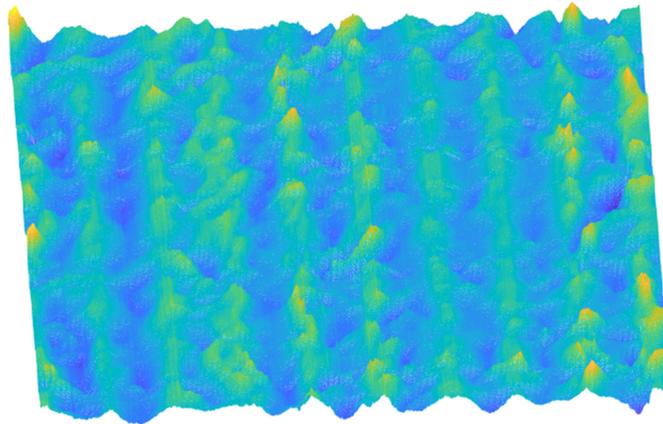


Figure 9: Reconstructed surface with moving phase shifting

Another simulation was performed by simulating the reconstruction of the wall of pipe section with moving PMP. A pipe section with cylindrical shape was simulated. A picture of the simulated pipe is shown in Figure 10. The scanning was moved along the pipe section to create 270 degrees movement. The sequence of the recorded images is shown in Figure 11. A reconstruction of the pipe wall section is shown in Figure 12. The middle section was reconstructed correctly because it was taken as the reference for our phase measurements. The amount of oscillation increases toward the sides due to the difference in height. The next section describes the mechanism to suppress the oscillation the reconstructed profile.

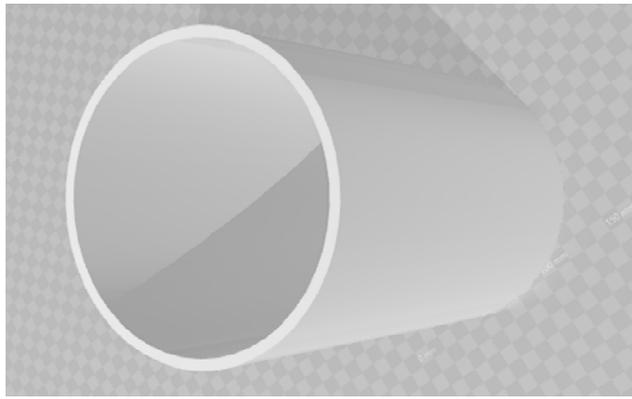


Figure 10: Scanned pipe section

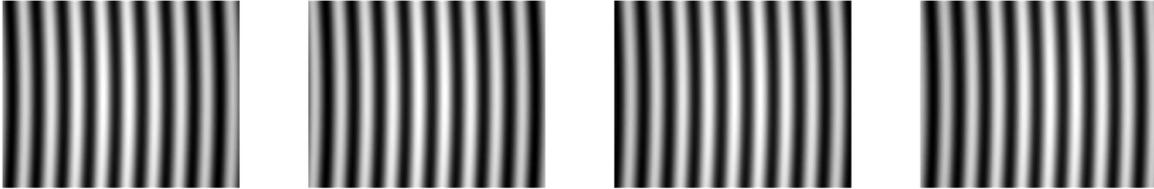


Figure 11: Four rectified images from moving sensor inside the pipe

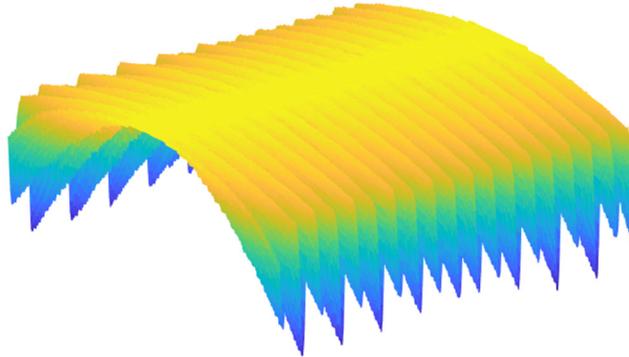


Figure 12: Reconstruction of pipe wall section with moving PMP

Discussion and future work:

We have designed a new structured light probe to increase the system spatial resolution and increase its robustness. The new system also uses an extra camera to add stereo 3D reconstruction capability to the system to facilitate the tracking of the system movement. An experimental testbed was fabricated to simplify the testing of the system outside the pipe. The stereo setup calibration was completed and currently working developing a calibration procedure for the projector. An algorithm to compensate for the oscillation in reconstruction resulted has been developed and tested in 1 D. The future work will include miniaturizing the sensor to be fit inside the pipe diameter and continue with the development of the reconstruction algorithm.

Task 3. Automated corrosion detection and uncertainty quantification

Overview:

Pipeline inspection is a critical component of preventive and condition-based maintenance in the oil & gas, chemical and sewage treatment industries. The infrastructure is designed to last decades, and improper maintenance has significant safety and economic costs. The interactions between the fluids flowing through the pipe and the inner wall cause the degradation of the mechanical properties of the pipe over time, and early prognosis of damage can prevent catastrophic events that can even be fatal. The primary driver for Artificial Intelligence-based solutions for in-line inspection is the sheer size of piping infrastructure in refineries and cities. Full automation of the inspection procedure can reduce costs and drive higher reliability because automated inspection techniques are not susceptible to human error.

Previous efforts focused on using the structured light technique to reconstruct the pipe geometry and extract features from the reconstructed geometry for classification using the Naïve Bayes algorithm. Another approach was to use Deep Neural Nets (DNN) to classify pipes into one of two categories; damaged, or not damaged. The basis of this approach is a powerful technique called transfer learning, where a deep neural network is pre-trained with a huge volume of image data, which do not pertain to the problem at hand. With a large enough training set, the features extracted by this pre-trained neural net can generalize to other, possibly very different problems, with a far lower number of training examples. The network used for this purpose was the Inception v3 network, which was pre-trained on the ImageNet database, which consists of more than 1 million labeled images. The Convolutional Neural Network (CNN) layers within Inception automatically extract features, with the level of abstraction of features increasing with each successive layer (Fig 1). The CNN-based approach performed better than the Naïve Bayes approach, and this report summarizes the efforts to extend the previous capabilities, identifies bottlenecks in the solution and proposes solutions to these bottlenecks.

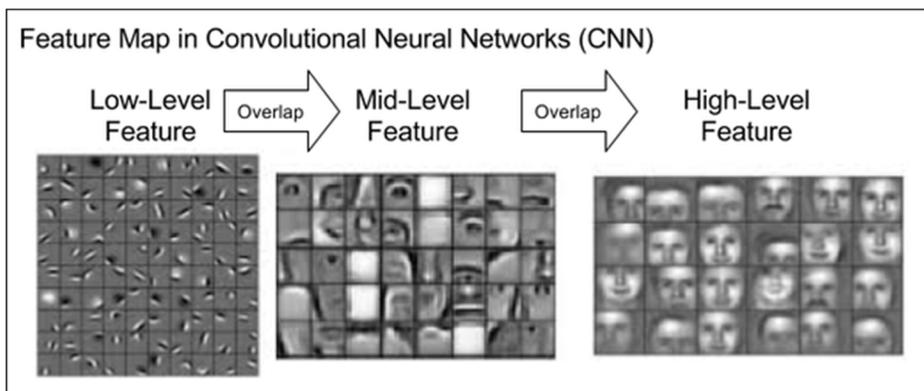


Figure 13: Feature extraction by successive layers of a Convolutional Neural Net

Given that the end objective is to assess the risk to failure of a pipeline section, there are two ways to formulate a solution. One way would be to assign a risk score for each image of a pipeline section based upon the detected damage, and the other would be to directly specify the locations of the areas of damage, be it corrosion, water accumulation, structural deformation or cracks on the wall. The first method relies on the detection of the presence of each of these defects, and a weighted sum can be

calculated based upon the detection result, which directly translates to a risk score. Therefore, the first problem is arguably a classification problem, while the second is an object detection problem.

1. Methodologies and Results:

1.1. Previous work: Binary Classification with Transfer Learning

The simplest approach to the problem is to classify each incoming frame as either damaged, or not damaged, based on features extracted using a CNN. Previous work focused on using the Inception v3 neural network to classify the images based on transfer learning. The optimal model produced an accuracy of 87.3%

1.2. Multi-class classification with transfer learning

The pre-trained neural nets were used to classify the types of damage in each image. Here, the dataset consisted of 1620 simulated images, and was trained on the same network as the binary classifier. The difference in this case was the classification categories. Each image was classified into one of five categories:

- a. Good
- b. Impingement damage
- c. Dent damage
- d. Slit damage
- e. Squeeze damage

The pre-trained model performed very well in this categorization, giving a test accuracy of 93.5% for a test dataset consisting of 420 images (Fig 3). The false predictions were mainly around the misclassification of impingement damage to slit damage and vice versa.

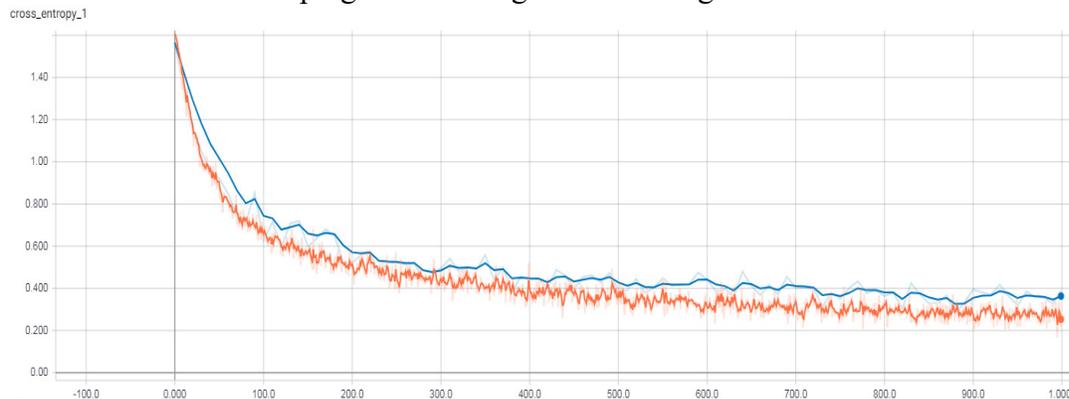
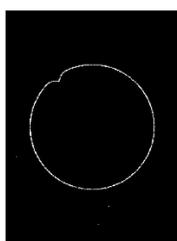
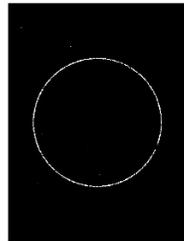


Figure 2: Cross entropy loss for the 5 class classifier for deformation damage

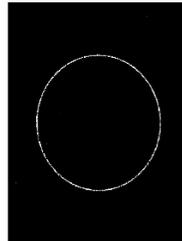
Real images obtained from YouTube video sources were used to train the neural network on classifying images based on three categories: a. corrosion, b. crack or c. good. While the program converged very quickly and produced a test accuracy of 99%, it was most likely due to test images being sourced from the same video as the training set, reducing the potential variability, and the corrosion examples were of images with large regions of corrosion.



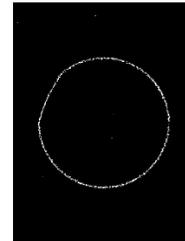
(a) Impingement
(Score = .97)



(b) Slit
(Score = .84)



(c) Squeeze
(Score = .60)



(d) Dent
(Score = .98)

Figure 3: Test images with their highest prediction scores

Small regions of corrosion will not be resolved, as the classifier works on predicting the class with information from the entire image, and not the corrosion spot. To demonstrate the shortcomings of this naïve approach to classifying crack and corrosion damage, we use images from videos that are not part of the training set to classify the type of damage. This is demonstrated in Fig. 5 where the classifier incorrectly classifies a corroded image to being good with a confidence of 88.7%, and a good image with the crack class with a confidence score of 79.4%. The problem of not being able to capture the specific features of cracks or corrosion using a traditional classifier requires us to investigate more sophisticated techniques of localizing the spots of damage in the pipe.



(a) Crack (Score = .91) (b) Good (Score = .90)

Figure 4: Highest predictions for out of sample test images by the Inception network



(a) Good (Score = .88) (b) Crack (Score = .79)

Figure 5: Misclassified images by the Inception classifier

1.3. Damage Localization and Detection: Fully-Supervised Learning

The previous sections used simulated images to classify the types of damage in each image. Real world pipeline images are many orders of magnitude more complicated than the simulated images. While the simulated images abstract away all the noise and redundant features in the real world, the real-world image is also affected by occlusion, noise, and other artifacts. Real-world images are also much more informative in that they contain far more features in comparison to simulated images. The simulated images are effectively the result one would get upon applying laser profiling to the real video feed. While this method is well suited to capture the structural deformations on the profile of the pipeline, such as dents and impingements, it does not capture other risk factors such as cracks, corroded patches and water accumulations. Moreover, the probability of finding only one type of damage in each frame of the video feed is minimal. Most of the training data consisted of at least two types of defects and had up to eight regions of potentially dangerous defects.

The detection task was attempted using YOLO, a supervised object detection algorithm which first determines whether an object exists in the image, classifies the object according to its category, and then localizes the object using bounding boxes. The YOLO algorithm is used here due to its fast performance compared to architectures such as R-CNN. YOLO can also detect cases where there are two or more types of overlapping defects in the same region. The YOLO algorithm divides the image into a 7x7 grid and performs the detection task on each. Each of the cells predict whether there is an object in it and calculates a confidence score for it. The confidence score is calculated as:

$$\text{Confidence Score} = P(\text{object}) * IOU_{Pred}^{Truth}$$

The resulting bounding box is characterized by its center relative to the grid cell, length, width and confidence score. Each grid cell predicts a conditional class probability, and this is multiplied with the confidence score on a test image. This gives the confidence score for the class in a grid cell

$$\begin{aligned} \text{Confidence Score}(\text{Class}) &= P(\text{Class}_i|\text{object}) * P(\text{Object}) * IOU_{Pred}^{\text{Truth}} \\ &= P(\text{Class}_i) * IOU_{Pred}^{\text{Truth}} \end{aligned}$$

The current implementation attempted to detect both cracks and corrosion spots without using pre-trained model weights, but did not achieve detection, as the loss function gradient across iterations flattened out and the loss oscillated within the range 3-3.5. Typical ranges for loss functions to achieve detection is at the order of 0.1. This oscillation suspected to be because of a combination of factors, such as low quality training images, and the choice to use tiny-YOLO, a network that is lighter and optimized for greater than real time performance at the expense of detection performance. Reducing the learning rate to up to 1E-6 resulted in a gradual reduction of the loss function at the rate of approximately 0.1 units per thousand iterations, which is exceedingly slow. Currently, the images are being re-annotated with only one type of defect (corrosion), to study whether the algorithm successfully detects corrosion in pipes, before moving on to detecting two types of damage simultaneously. Improvements in the quality and quantity of the training dataset, as well as training the data on YOLO v3, a larger network with a mean Average Precision (mAP) of 51.5 at 45 fps performance, is expected to improve the outcomes of the study. A reason why YOLO could be a feasible detector is the relatively low variability in features of cracks or corrosion spots, compared to those seen in generic objects such as chairs and airplanes.

1.4. Weak-Supervision and Unsupervised Learning

The learning pipeline consists of data preparation and feature engineering, followed by the application of a learning algorithm to predict results given new data. The problem of feature engineering was difficult before the advent of deep CNNs, as features had to be manually engineered using techniques such as Discrete Cosine Transforms, Histogram pooling, DAISY, SIFT, etc. Each of these feature extractors were able to function well within their assigned applications but failed upon generalization to new applications. Deep CNNs, on the other hand, became an automated method to extract features given a large set of images. Given good features, learning algorithms do a great job predicting unseen data. The only remaining bottleneck is data preparation. Data preparation in supervised learning usually consists of annotations and labeling. Weakly-supervised and unsupervised techniques are currently being explored to find new ways to solve this problem.

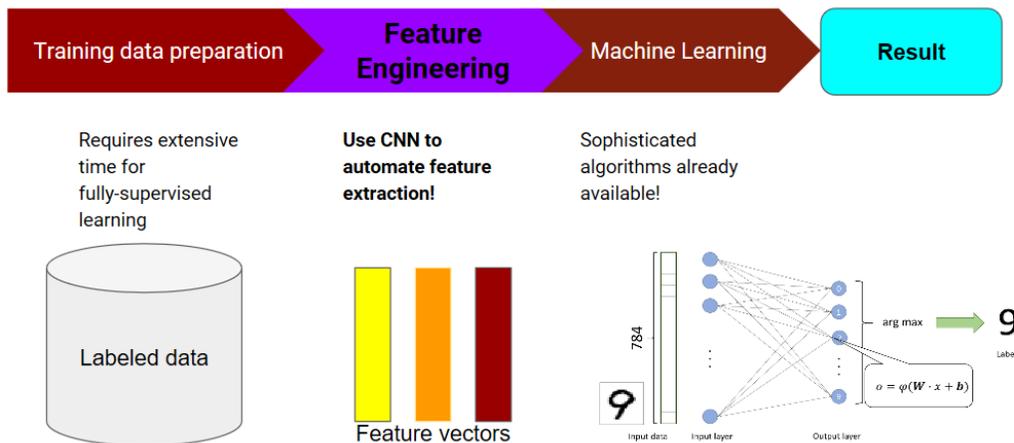


Figure 6: Current workflow for pipeline damage detection has a bottleneck in data engineering

A solution to the problem of data-preparation is data programming- a method that uses heuristics and labeling functions to generate probabilistic labels for the training data using generative models. While this method has been proven to work well on text mining, information extraction from images using data programming has not been explored extensively. This is because images are interpreted by CNNs

as pixels and writing labeling functions or heuristics over raw pixels is impractical. Therefore, unsupervised localization of damage becomes challenging. A possible approach to this problem can be to initially segment the image in an unsupervised manner, and then write heuristics over the segmented image. An auto-encoder architecture that utilizes Convolutional Neural Networks can segment images in an unsupervised fashion. The W-Net architecture used in this study consists of an encoder CNN that produces a segmentation layer of the same spatial size as the image, and this is decoded using another CNN that reconstructs the image from the segmentation. The two CNNs are based on the U-Net architecture that is used for semantic segmentation. The downward portion of the encoder U-Net captures the features of the input image through a mixture of convolution layers, ReLU non-linear activation, and max pooling. The max pooling operation selects the dominant activations from an image to produce the next layer of the convolutional net. The downward portion of the U-Net, therefore, can be thought of as a mechanism to extract features, regardless of their relative positions in the image. The next part of the encoder U-Net consists of a series of up-convolution layers with learned kernels and ReLU activations. This step produces a segmentation of the image. For an image with K classes, the output of the encoder is a $224 \times 224 \times K$ prediction for each pixel in the image. The loss function associated with the encoder is the normalized cut loss, which finds the best graph cut to produce the segmentation. The second U-Net takes as input the segmented image, and outputs the reconstructed image from the features of the segmented image. The loss associated with the entire net is the reconstruction loss, and the minimization of the reconstruction loss updates the parameters in both the encoder and decoder.

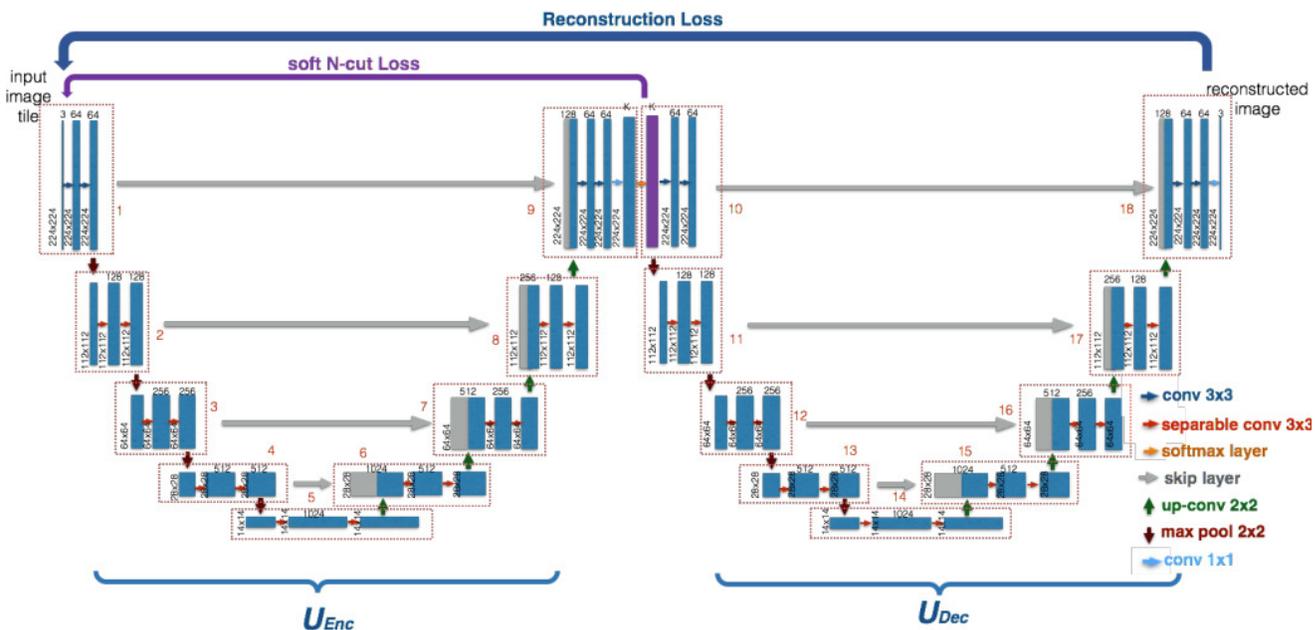
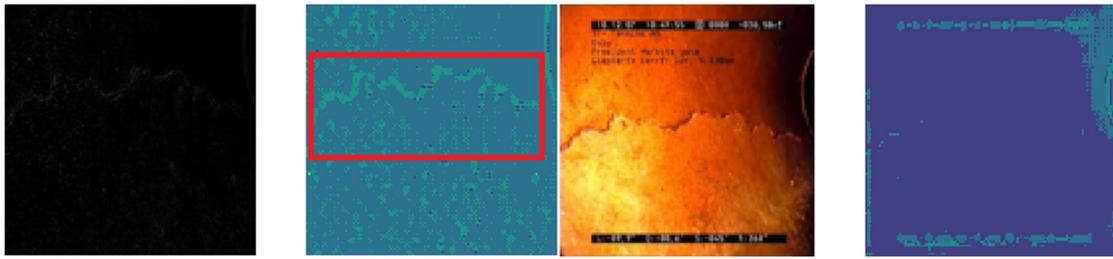
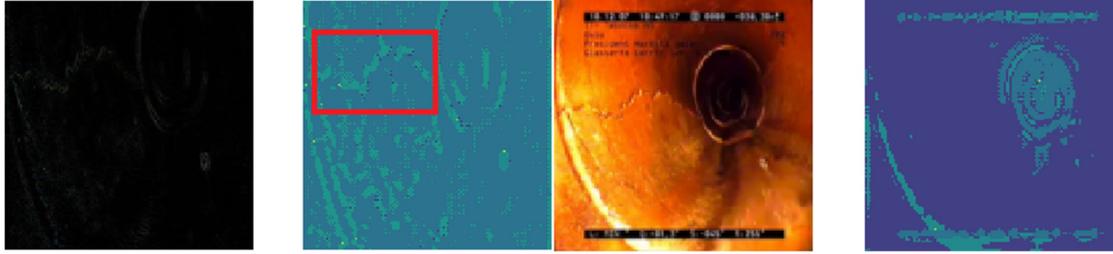


Figure 7: W-Net architecture

The W-Net architecture was trained for 100,000 iterations on the PASCAL VOC dataset. Raw images from the pipeline dataset were initially passed to the trained network but it did not produce usable segmentation results. To remedy this, an edge detector was applied to the test images, and this resulted in far better segmentations, especially for images with cracks. This segmentation is expected to improve with a model that has been fine tuned to produce a lower reconstruction loss.



(a) Image 1: Cracked pipe wall (isolated)



(b) Image 2: Cracked pipe wall with the pipe section in view



(c) Image 3: Corroded section

Figure 8: (a), (b), (c) show 3 samples where the segmentation is performed with (left) and without (right) edge detection and foreground text removal. The requisite segmented features are highlighted with a red box upon which heuristics must be written.

As described in the overview, the problem can be recast as a risk assessment problem, given the segmented image. Noisy heuristic labeling functions can be written over these segments to classify each image with a risk label. A pressing challenge, however, is to get the unsupervised segmentation to work across varieties of images, without too many incorrect predictions. These segments themselves can potentially work as localization algorithms for corrosion pits and cracks.

2. Conclusion

We have extended the previous work of binary classification by performing the classification task for multiple classes of deformation of the pipe section using simulated images. The test accuracy was 93.5% and most of the errors were due to the misclassification of impingement and slit damage. We also trained the YOLO algorithm to detect cracks and corrosion damage, so that it can work in real-time alongside Inception v3 to classify the deformation damage. The primary challenges in this project are the lack of availability of quality training data for YOLO and traditional deep neural net approaches, effectively utilizing the high information content of real images, and the bottleneck of manual annotation of training data for damage localization. The problem of annotating large amounts of training data can be alleviated using unsupervised and weakly supervised methods and work is ongoing to find potential methods that can cluster different types of corrosion and cracks automatically without any training data. The W-Net auto-encoder architecture consists of two U-Nets that segment an image followed by reconstruction from the segmentation. Applying the edge detection filter to images and removing the foreground text resulted in much better segmentation

results as compared to the raw images. The reconstruction loss of the W-Net was not minimized to the levels published in the original paper, and future work will focus on fine-tuning the network learning parameters to improve the segmentation results for reliable heuristics.

References:

- [1] P. S. Huang and S. Zhang, “Fast three-step phase-shifting algorithm,” *Appl. Opt.*, vol. 45, no. 21, p. 5086, 2006.
- [2] J. Geng, “Structured-light 3D surface imaging : a tutorial,” *IEEE Intell. Transp. Syst. Soc.*, vol. 160, pp. 128–160, 2011.
- [3] B. Li, Y. Wang, J. Dai, W. Lohry, and S. Zhang, “Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques,” *Opt. Lasers Eng.*, vol. 54, pp. 236–246, 2014.
- [4] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado, “A state of the art in structured light patterns for surface profilometry,” *Pattern Recognit.*, vol. 43, no. 8, pp. 2666–2680, 2010.
- [5] H. Guo, “3-D Shape Measurement Based on Fourier Transform and Phase Shifting Method,” p. 114, 2011.
- [6] S. Van der Jeught and J. J. J. Dirckx, “Real-time structured light profilometry: A review,” *Opt. Lasers Eng.*, vol. 87, pp. 18–31, 2015.