Discover Artificial Intelligence

Discover

Research

Image-based and risk-informed detection of Subsea Pipeline damage

Rialda Spahić[1] · Kameshwar Poolla[2] · Vidar Hepsø[3] · Mary Ann Lundteigen[1]

Received: 2 May 2023 / Accepted: 5 June 2023
Published online: 12 June 2023

Abstract
As one of the most important assets in the transportation of oil and gas products, subsea pipelines are susceptible to various environmental hazards, such as mechanical damage and corrosion, that can compromise their structural integrity and cause catastrophic environmental and financial damage. Autonomous underwater systems (AUS) are expected to assist offshore operations personnel and contribute to subsea pipeline inspection, maintenance, and damage detection tasks. Despite the promise of increased safety, AUS technology needs to mature, especially for image-based inspections with computer vision methods that analyze incoming images and detect potential pipeline damage through anomaly detection. Recent research addresses some of the most significant computer vision challenges for subsea environments, including visibility, color, and shape reconstruction. However, despite the high quality of subsea images, the lack of training data for reliable image analysis and the difficulty of incorporating risk-based knowledge into existing approaches continue to be significant obstacles. In this paper, we analyze industry-provided images of subsea pipelines and propose a methodology to address the challenges faced by popular computer vision methods. We focus on the difficulty posed by a lack of training data and the opportunities of creating synthetic data using risk analysis insights. We gather information on subsea pipeline anomalies, evaluate the general computer vision approaches, and generate synthetic data to compensate for the challenges that result from lacking training data, and evidence of pipeline damage in data, thereby increasing the likelihood of a more reliable AUS subsea pipeline inspection for damage detection.

Keywords Pipeline inspection · Anomaly detection · Damage detection · Risk analysis · Pattern recognition

1 Introduction

Monitoring and inspection are essential for operational subsea oil and gas pipelines. However, subsea oil and gas operations are complex, with a range of structures and systems, in complex and harsh subsea environment. As a critical asset for transporting oil and gas products over vast distances, subsea pipelines are exposed to a variety of environmental hazards. Hazard is defined as the source of harm [1]. Exposure to environmental hazards can damage the pipelines and cause severe personnel, environmental, and financial damage [2]. Therefore, proper inspection and maintenance of subsea pipelines are essential tasks for their safe and reliable functioning and operations. In case of an unexpected event, continuous monitoring (i.e., pressure drop monitoring for leak detection) notifies the pipeline shutdown system with the supervisory role of an operator [3]. Despite the worldwide safety record of subsea pipelines, comprehending and

✉ Rialda Spahić, rialda.spahic@ntnu.no; Kameshwar Poolla, poolla@berkeley.edu; Vidar Hepsø, vidar.hepso@ntnu.no; Mary Ann Lundteigen, mary.a.lundteigen@ntnu.no | [1]Department of Engineering Cybernetics, Norwegian University of Science and Technology, Høgskoleringen 1, 7034 Trondheim, Norway. [2]Department of Mechanical Engineering, University of California, Berkeley, 110 Sproul Hall Berkeley, Berkeley, CA 94720, USA. [3]Department of Geoscience and Petroleum, Norwegian University of Science and Technology, Høgskoleringen 1, 7034 Trondheim, Norway.

responding appropriately to complex situations as well as anticipating their consequences are crucial for the safety of offshore operations [4]. Since sending human operators offshore can be dangerous and expensive, autonomous underwater systems (AUS) are intended to assist human operators in inspecting offshore structures, especially long and vast subsea pipelines. With the development of subsea docking stations that allow AUS to reside on the seabed for months, trained operators have the flexibility and opportunity to use AUS to inspect pipelines when the situation calls for it [5].

Autonomy, as described by [6], is the capacity to act and make decisions without external assistance. For AUS, autonomy is typically achieved through artificial intelligence (AI) systems, the computer systems designed to mimic intelligent human behavior [6], by analyzing large amounts of incoming data collected in near-real-time or real-time by sensors and cameras attached to the AUS. For damage detection scenarios, the dominant AI approaches include [7, 8]:

- *Computer vision* methods for analyzing image data,
- *Machine learning* methods that learn from large amounts of data to find patterns, and
- *Anomaly detection* methods that identify and report irregularities, or anomalies, in data patterns.

In addition, *risk assessment and analysis* are common and well-established approaches for identifying *what can go wrong* in operations and offering a list of hazards, as potential sources of harm, the likelihood, sequence of events and consequences of hazards [9].

In recent years, due to the success of remotely operated vehicles (ROVs) that are manually controlled, pipeline inspection research has considered the potentials of AI technologies employed by AUS, such as underwater drones. Therefore, there is an increase in interest for the potential of image-based inspection by computer vision techniques through cameras attached to AUS, such as image classification, object detection, and image segmentation [2, 10–16]. However, the existing research for image-based inspections with AUS is particularly oriented toward image color and shape reconstruction and unsupervised methods due to the complexity of underwater conditions, poor visibility, and a significant lack of training data.

Despite the abundance of available research, the remaining obstacles to reliable operations with AUS stem from the underrepresentation of evidence of pipeline damage in data, which contributes to data imbalances that can lead to inaccurate data analysis results and misleading data pattern findings. In addition, there is a significant lack of training data for computer vision and data-driven methods to learn the patterns of potential dangers in order to detect them efficiently and reliably. Unfortunately, a significant number of the detected anomalies represent insignificant data, also known as noise, which further mislead the data analysis conclusions and disrupt the AUS operations decision-making system.

In this paper, we focus on analyzing industry-provided subsea pipeline images captured by underwater drones for external damage detection, introducing risk-informed training processes for the anomaly detection methods and evaluating the detected anomalies by isolating potential the anomalies that represent pipeline damage. The focus of this research paper is on utilizing risk analysis knowledge and semi-supervising computer vision methods for subsea pipeline images for early identification of pipeline damage while separating them from insignificant anomalies (noise and false alarms). The objective is to provide the missing training data while limiting the amount of manual labor to annotate the training images, and therefore to limit the frequency of false alarms generated by autonomous systems and to identify pipeline damage as early as feasible while increasing the scope of anomaly detection capabilities during visual monitoring and inspection. Therefore, the contributions of this paper can be summarized as:

- Analysis of external damage on subsea pipelines on raw, industry-provided data.
- Generation of synthetic data through a seamless blending of known anomalies, as defined by risk assessment and analysis methods, for a more reliable computer vision and anomaly detection.
- Review of computer vision challenges, such as monochromatic images and large images that necessitate extensive computational power to analyze.
- Proposal of a methodology to address the lack of training data, imbalanced data, and data quality for image-based subsea pipeline damage detection.

This paper consists of eight sections. Section 1 reviews related work on image-based subsea pipeline analysis. Section 2 describes the key problems with the general computer vision methods for subsea pipeline analysis. Section 3 describes the challenges of underwater computer vision for offshore inspections. Section 4 describes the anomalies identified by risk assessment and analysis experts from the oil and gas industry, describes expectations on what types of anomalies can occur, their damage potential, and the likelihood of occurrence. Section 5 presents the

data, methods, evaluation metrics, and the case study. The resulting methodology is summarized and illustrated in Sect. 6. Section 7 discusses the results and observations from the case study. Finally, Sect. 8 concludes the work and presents future research directions.

2 Related work

The efficiency and reliability of damage detection are vastly enhanced by computer vision. During visual inspection, environmental conditions and appropriate image collection are essential for obtaining high-quality images for image analysis [17]. Computer vision is a type of real-time, in-line detection that requires the analysis of vast quantities of data, often including redundant information, and a high-dimensional feature space. The primary obstacles of general computer vision applications are the computation speed required for real-time operations and the detection intelligence required to differentiate between significant and redundant information [17]. Recent efforts in computer vision have centered on general algorithms for the efficacy and precision of visual inspections [17, 18], the necessity of integrating multiple detection technologies [19], and the improvement of real-time performance with less computational power [18, 20]. The restrictions of computational power are particularly critical in applications with autonomous systems, such as underwater drones and other mobile vehicles [21]. However, underwater computer vision for subsea structures inspection is facing additional challenges, such as poor visibility, and lack of training data [21, 22]. Subsea pipelines are exposed to various environmental factors that can compromise their integrity and contribute to various types of damage. Due to this, substantial research has been conducted on inspecting subsea pipelines to look for damage.

Zhou et al. [7] described the challenge of locating anomalies during subsea exploration. Using a context-enhanced autoregressive network that learns semantic dependence based on conditional probability to identify the anomaly in low-visibility underwater images weighted by both image reconstruction loss and feature similarity loss, they proposed a deep-learning-based anomaly detection framework to identify unknowns in a complex underwater environments for autonomous robots. With sufficient training data with images of marine animals, they successfully demonstrated their method for detecting marine animals as anomalies on a large, imbalanced dataset.

Samnejad et al. [23] explored ways to reduce the time-to-value and overall cost of the subsea pipeline inspection by replacing the laborious task of manually searching for anomalies through unorganized data with an efficient workflow through a set of neural network methods and substantial computational power from cloud-based services. The authors [23] presented a digital solution that integrates the value of visual data collected and aggregated over decades of inspection campaigns with computer vision technologies to detect and classify structure and equipment anomalies autonomously. However, the 20,000 images for the training dataset were annotated manually, requiring intense labor.

Bastian et al. [24] visually inspected and characterized external corrosion in pipelines located on land using a convolutional neural network (CNN). They proposed a CNN for detecting and classifying corrosion on four levels: no corrosion, low, medium, and high corrosion. Despite high accuracy and promising results, the authors [24] encountered several issues that made CNN misclassify corrosion, such as leaves, deposits on the pipeline, and the corrosion-like landscape surrounding the pipelines. They highlighted the need for pipeline images to contain background information, or context, for training. Among the classified corroded pipelines, there were images with background clutter that the CNN model could not distinguish. They emphasized the importance of pipeline images containing context or background information for training purposes a recommended a more localized pipeline inspection approach for more reliable results in differentiating corrosion levels. On land pipelines, however, image-based damage detection encounters fewer challenges with hazy, monochromatic images than on subsea pipelines, making the subsea pipeline inspection task more challenging.

Khan et al. [25] investigated methods for estimating subsea pipeline corrosion based on the color of the corroded pipeline. The authors [25] encouraged incorporating the color correction methods into a robotic system for subsea pipeline corrosion inspection, even in real-time to address the visibility challenges for underwater images. They proposed an algorithm for image restoration and enhancement to reduce blur and improve the color and contrast of underwater images that were tested on experimentally collected and publicly available hazy underwater images.

3 Problem description

Underwater computer vision for offshore inspections with autonomous systems is receiving greater attention and the methods need to mature for reliable and safe anomaly detection operations. The primary challenges that pique the interest of both the research community and the industry are:

- *Imbalanced data* is a frequent obstacle in data-driven analysis, such as with most machine learning and anomaly detection techniques. The difficulty is most apparent in anomaly detection applications where anomalies may reflect important information, such as potential pipeline damage. Due to the scarcity of damage evidence in everyday operations, the collected data consists of the vast majority of non-anomalous situations, making it difficult for algorithms to learn patterns about anomalies, report them, and not eliminate them as noise, which is the information that misleads data analysis [26].
- *Training data* is generally sparse in AI-based data-driven approaches. There is a saturation of applications tested with accessible training data; nevertheless, unsupervised algorithms that do not require annotated data are becoming increasingly popular as more data becomes available [27]. Yet, due to the complexity and inexplicability of these techniques, there is a growing interest in discovering automated methods to annotate massive amounts of data and save laborious manual effort. Creating training data is being explored from different perspectives, among others, generating data from simulations, using AI tools for automatic annotation, or through transfer learning where data is learned from one application and tested on a different one.
- *Image quality and visibility* are computer vision applications' most persistent and obvious obstacles. Due to the nature of water as a medium, underwater photos frequently need to be corrected to avoid incorrect lighting and color, causing them to appear predominantly blue or green. In addition, seawater may include a high concentration of plankton and other marine organisms that can obscure photographs. For subsea pipelines, layers of material such as sand and biological deposits referred to as fouling and biofouling, limit the view of the pipeline surface, and inhibit inspection. Hence, many underwater computer vision applications concentrate on reconstructing the image's color, shape, and overall item visibility.
- *Computing power* is another challenge for computer vision applications, because images are often very large and need substantial computing power and processing time. A weakness of prominent neural network algorithms is the necessity to resize or downscale images to improve processing speed, which may result in a substantial loss of information from the resized images. Sliding-window approaches are used in applications where the larger regions of image need analysis without substantial resizing or in case of substantial information loss due to resizing [28].

Autonomous systems powered by computer vision have great potential to detect subsea pipeline damage. However, as offshore operations prioritize the reliability and maturity of emerging technologies, it is necessary to investigate options for generating more training data and reducing the need for Black-box algorithms to be closer to permanently employing autonomous underwater systems for remote operations. It is also important to determine if the image resizing, which is often required to reduce needed computational power during image analysis poses a considerable information loss and reduces the chances of reliable anomaly detection.

4 Anomalies as risk factors

General visual inspection of subsea pipelines, traditionally performed by ROVs is one of the most common inspection methods for determining the pipeline's integrity and identifying areas of increased risk [29]. The operators who manually control the ROV during the pipeline inspection are trained and experienced in detecting anomalies on and around pipelines. The following is a set of the common anomaly criteria for general visual inspection of subsea pipelines established by the best practices in industry [29]:

1. Any evidence of **fluid leakage**.
2. Any external **corrosion** on the exposed metal or outer sheath.

3. Any external **damage, deformation, and bending** on the pipe surface, anodes or other components.
4. Any **debris** blocking the visibility of the pipeline, including litter and other seabed debris, and sediments, is known as fouling. The visibility is also impeded by an abundance of **marine growth**, known as marine fouling or biofouling. The anomaly is considered if more than 50% of the surface is covered within 10 ms. Additionally, debris considered an anomaly are **objects in the nearby vicinity**, up to 1 m, of a pipeline that can cause damage or obstruct visibility, such as large boulders.
5. **Ineffective pipeline support**, including ineffective seabed support.

Accordingly, Table 1 shows a summary of anomalies as risk factors that can contribute to pipeline failure. Table 1 illustrates each risk factor's potential damage analyzed, from extensive to minor damage, and compared to its expected occurrence probability, from most probable to least probable occurrence of damage [30].

Table 1 shows the general representation of anomalies and the expectation of their occurrence probability, However, the exact probability and anomalies that are identified as damage are typically calculated within a specific operation context. It is crucial for the UAS that detects anomalies to have information or knowledge of the major risk contributing factors associated with the subsea pipelines to adjust expectations and reporting in regions where the likelihood of the most extensive damage potential is higher.

5 Case study

5.1 Data description

The dataset for this case study consists of an imbalanced set of 166 subsea pipeline images captured with an autonomous underwater drone, provided by domain experts from the oil and gas industry. There are 126 images without anomalies and 38 images with anomalies or mechanical damage on the surface of the pipeline. We used 35 additional images without anomalies to generate synthetic mechanical damage images. This was done to balance out the anomalous and non-anomalous images and test if the synthetic data is sufficiently realistic to improve the network learning process. The images are in high resolution and do not require shape or color recovery. However, the nature of the mechanical damage makes it difficult to distinguish the damage from marine growth on the pipeline surface, as both share irregular patterns and similar colors, posing a challenge to distinguish between small-scale damage and marine growth. The original size of

Table 1 Risk factors contributing to pipeline failure, adapted from [30]

Potential Hazard	Damage Potential			Probability of Occurence		
	Extensive	Moderate	Minor	Most Probable	Expected Occurance	Least Probable
Leakage, explosion						
External Corrosion						
Material Deficiency						
Debris						
Marine Fouling						
Object Dragging (Anchor, boulder)						
Erosion, soil transport and bottom phenomena						

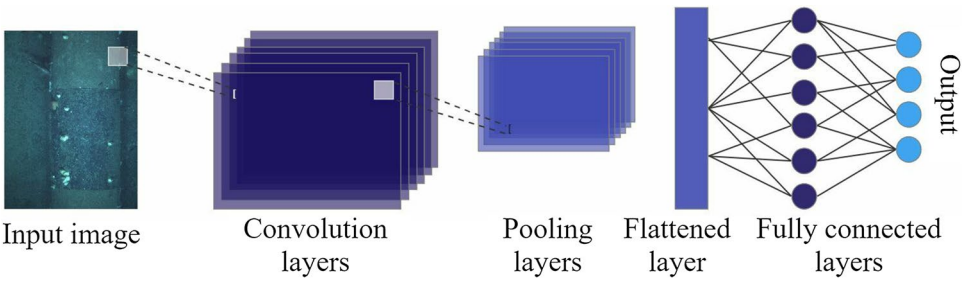

Fig. 1 Building blocks of CNN, adapted from [31]

Table 2 Confusion Matrix for Binary Classification, adapted from [34]

Confusion matrix	Actual positive class	Actual negative class
Predicted positive class	True Positive (TP)	False Negative (FP)
Predicted negative class	False Positive (FP)	True Negative (TN)

each image was 4096 x 2304 pixels, however, due to computational resources required during CNN training, the images were reduced to 224 x 224 pixels where mechanical damage is still visible on the pipeline.

5.2 Image classification with neural networks

One of the elements of data analysis through machine learning is the discovery of discriminant data features. Discovering discriminant data features in images can be particularly challenging and requires complex methods inspired by visual cortex processing in the brain that are capable of learning a substantial number of features and extracting patterns [31]. We will focus on a deep learning method CNN or convolutional neural networks. The CNN model consists of convolutional layers whose primary function is to learn and extract the features required for efficient image comprehension [31]. The objective of the convolutional layer, modeled over neuronal cells, is to extract features such as edges, colors, texture, and gradient orientation. Convolutional layers, see Fig. 1, are composed of convolutional filters or kernels. The kernels are convolved across the width and height of the input image. CNN intuitively learns filters that are activated upon encountering edges, colors, textures, and other image properties. The pooling layer performs nonlinear downsampling of convolved features and reduces the computational power necessary to process the data by reducing dimensionality [31]. The output of pooling is the subdivision of its input into a collection of rectangle patches. Depending on the pooling method selected, each patch is replaced with a single value [31]. There are two main types of pooling, maximum and global average pooling. *Global average pooling* is the more interpretable of the two types because it enforces correspondence between feature maps and categories through the creation of micro-networks [32]. Global average pooling is a structural regularizer that prevents overfitting, a phenomenon in which the CNN model provides accurate predictions for training data but not test data. *Maximum pooling*, or Max pooling performs linear separation, and provides a maximum network that is more potent and achieves higher performance with less computational power by assuming that instances of latent concept lie within a convex set [32].

Although CNN is considered a less explainable approach in image analysis applications, numerous efforts have been made to enhance its explainability. Particularly for image classification and object detection tasks, *localized anomaly detection* is one of the most effective methods for explaining which local regions of an image have been selected for classification. Typically, local regions are depicted using attention maps, which highlight feature regions deemed (by the trained model) crucial for satisfying the training criteria [33]. An example of an attention mask is a highlighted class region on the image, such as mechanical damage, which helps to explain why this image has been classified by CNN as mechanical damage or anomaly. Localized anomaly detection is crucial not only for determining if the classification occurred for the correct reason but also for understanding CNN's learning patterns and identifying noise during classification (i.e., analyzing highlighted regions that do not represent the accurate class).

5.2.1 Evaluation metrics

The evaluation metrics are used to assess the general performance of a trained method, such as a classifier that classifies two or more classes from a given set of data [34]. Various metrics can be evaluated based on the application's requirements.

Accuracy is one of the most common metrics that counts the total amount of correct classifications on the unseen data. The correct and incorrect classification results can also be illustrated with a confusion matrix, such as in Table 2.

The confusion matrix consists of the total numbers of correctly and incorrectly predicted classes, and the numbers of actual classes, to determine true and false positive and negative predictions [34]. *True positive (TP)* and *true negative (NP)* represent the total number of accurately predicted classes, where the predictive method (i.e., classifier) accurately predicted the instances of a positive class and the instances of a negative class. Alternatively, *a false positive (FP)* and *false negative (FN)* represent the total numbers of incorrectly predicted positive and negative classes. Typical evaluation metrics that are calculated through a confusion matrix are accuracy, error rate, sensitivity, specificity, precision, recall, F-measure, and averaged measures of each of these metrics [34].

$$Accuracy(acc) = \frac{tp + tn}{tp + fp + tn + fn} \tag{1}$$

$$ErrorRate(err) = \frac{fp + fn}{tp + fp + tn + fn} \tag{2}$$

$$Sensitivity(sn) or Recall(r) = \frac{tp}{tp + fn} \tag{3}$$

$$Specificity(sp) = \frac{tn}{tn + fp} \tag{4}$$

$$Precision(p) = \frac{tp}{tp + fp} \tag{5}$$

$$F - Measure(FM) = \frac{2 * p * r}{p + r} \tag{6}$$

Accuracy, calculated with Eq. 1, measures the ratio of correct predictions from the total number of predicted instances [34]. However, accuracy does not represent a reliable evaluation metric when the dataset is imbalanced. Due to the low representation of certain classes, many predictive models are unable to learn the patterns of poorly represented data and the inaccurate prediction becomes nearly invisible as compared to the prevalent number of highly represented classes. The accuracy of a predictive model can be high even when all of the underrepresented classes are predicted incorrectly. Depending on the needs of an application, other evaluation metrics are measured to determine the reliability of the model. Error rate measures the ratio of incorrect predictions from a total number of evaluated instances and it is calculated with Eq. 2. Sensitivity or Recall, calculated with Eq. 3, measures the proportion of correctly classified positive patterns, whereas Specificity (see Eq. 4) measures the proportion of correctly classified negative patterns [34]. With Eq. 5, Precision determines correctly classified positive patterns from the total predicted patterns of a positive class. Finally, F-Measure, calculated with Eq. 6, measures the harmonic mean between recall and precision [34].

5.3 Generating synthetic anomalies

Global image editing, such as resizing, shape reconstruction, and color correction, is a typical preprocessing step for image analysis tasks. However, achieving local changes that are restricted to a region of an image, such as object replacement, distortion, blending, cloning, and texture changes, can provide opportunities to manipulate images and create new, seamless, and realistic images. To balance the dataset and provide additional training data for image analysis, we generate synthetic anomalies, mechanical damage on pipeline surface, using the computationally efficient Poisson equation for local seamless blending. With the Poisson equation, we blend an extracted anomaly from anomalous images and seamlessly blend it into another image without anomalies.

$$\mathbf{v} = \nabla g \tag{7}$$

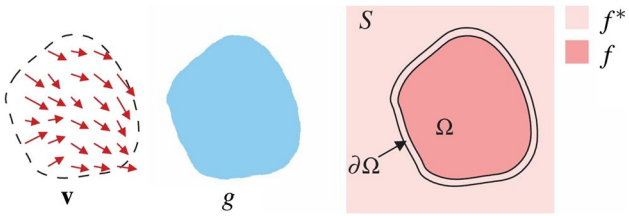

Fig. 2 Guided image interpolation, adapted from [35]

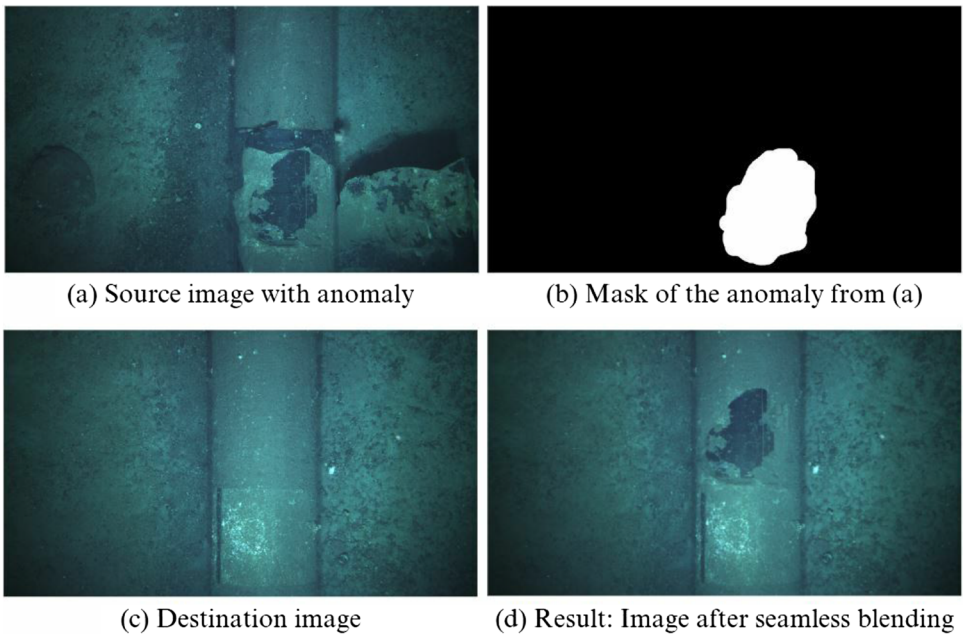

Fig. 3 Seamless blending of mechanical damage on a subsea pipeline: **a** Source image with an anomaly from which a mask **b** and seamlessly interpolated onto destination image **c**, resulting in **d**

$$\triangle f = \triangle\, g \; over\, \Omega,\; with f \mid \delta\Omega = f * \mid_{\delta\Omega} \tag{8}$$

Perez et al. [35] described and proposed a method for seamless object blending. The seamless blending method is based on a Poisson partial differential equation with Dirichlet boundary conditions that specify the Laplacian of an unknown function over the domain of interest and the unknown function values at the domain's boundary. This allows an object to be seamlessly interpolated onto another object. Figure 2, described by Eqs. 7 and 8, illustrates a guided interpolation in terms of a function *f* that interpolates in domain Ω the destination function *f** within a closed subset *S* with boundary $\delta\Omega$, guided by vector **v**, as a gradient field of a source function *g* [35].

A detailed mathematical description of the process is offered in [35]. Seamless cloning and insertion of an object relies on importing the gradients where the most common option for the guidance field **v** is a gradient field extracted directly from the image source (i.e, color information from the source image). Gradient field performs non-linear mixing or seamless blending, between source and destination images and selects the more dominant features for blending (color, texture, etc.). Equation 7 is used to guide the interpolation of this source image, which is denoted by *g*, after which the final reading for the function *f* is described by Eq. 8.

Figure 3 shows the process of seamless blending on an image of a subsea pipeline. A source image (Fig. 3a) has a mechanical damage anomaly on the pipeline that is masked off using an open-source annotation tool for machine learning and image analysis applications. We used Label Studio [36] for this purpose to achieve a precise mask image as shown in Fig. 3b. Annotation or labeling of images with Label Studio [36] was performed by marking a local region on the image. The marked region contains the bounding box and is assigned a label. Exported labels of the labeled regions are then exported as mask images. The source and mask, along with the position of the local region (i.e. position on the pipeline surface) on the destination image (Fig. 3c) where the blending will occur (other changes such as resizing and reshaping of source/mask object can be made at this point) are provided for seamless blending. Finally, the resulting image is obtained as a synthetic anomaly, as depicted in Fig. 3d. Figure 4 shows other images with synthetic anomalies.

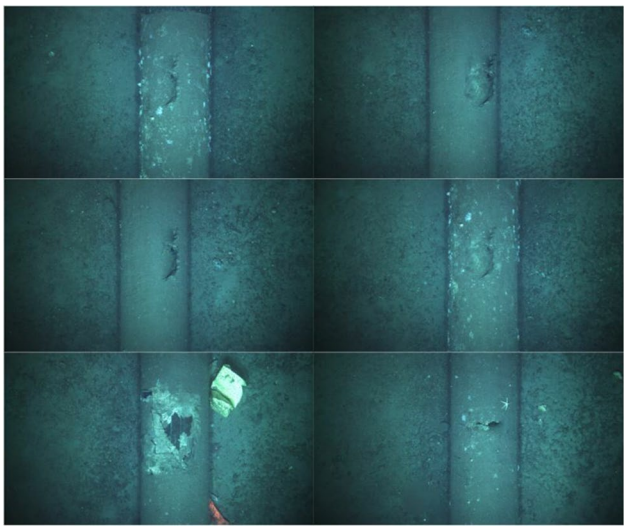

Fig. 4 Other examples of synthetic anomalies

Obtaining the mask images, which requires hand-labeling of anomalies with the knowledge of anomalies as risk factors, is the most labor-intensive aspect of creating synthetic anomaly images. However, once the masks have been obtained, the remaining steps are automated to produce batches of synthetic images. The reshaping and placement of the anomalies are randomized so that they do not appear in identical or similar forms. Nonetheless, generated synthetic anomaly images are manually inspected to identify any unrealistic or incorrect results.

5.4 Image classification without synthetic training data

CNN Global Average Pooling and Maximum Pooling on two-dimensional images have been implemented through Keras, a Python-based application programming interface for deep learning that runs on the machine learning platform TensorFlow [37, 38].

We analyzed the available data without added synthetic anomaly images to test the level at which CNN can classify the normal from anomalous images. The total number of images in the dataset without added synthetic mechanical damage is 164, out of which there are 126 normal images, and 38 of anomalous images with mechanical damage. We split the dataset into 80% for training, and 20% for testing. For the training, we have set the CNN to train over 30 epochs. During each epoch, one cycle of CNN training, all images are processed forward and backward to the CNN. Figure 5 shows the training and validation losses. Training loss measures how well the model fits the training data, while validation loss

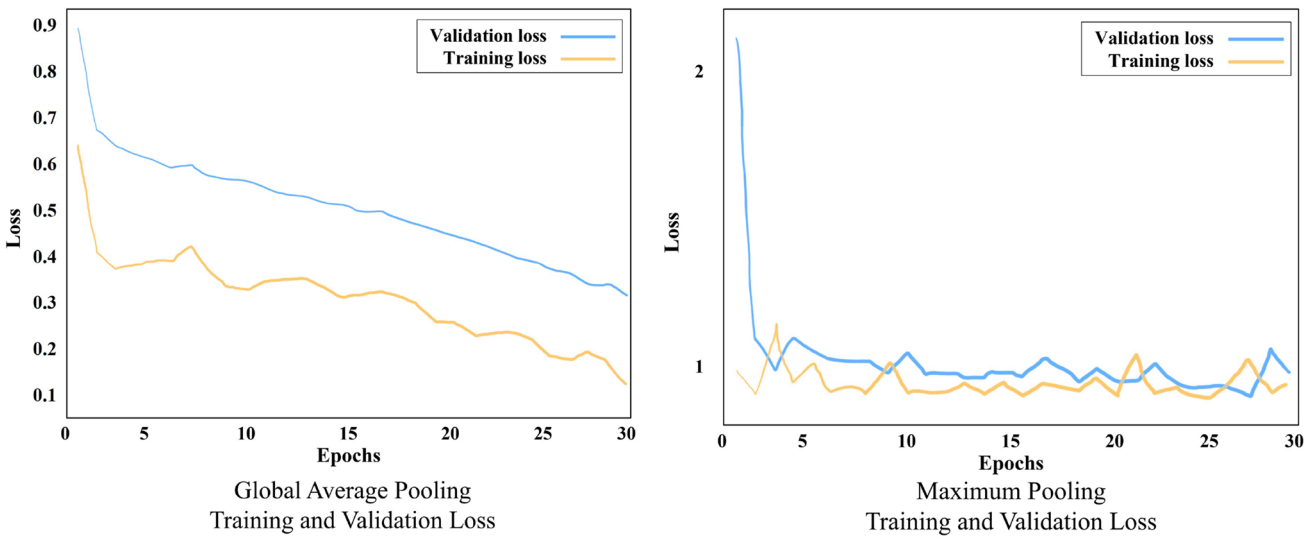

Fig. 5 Training and validation loss by global average and max pooling, without synthetic training data

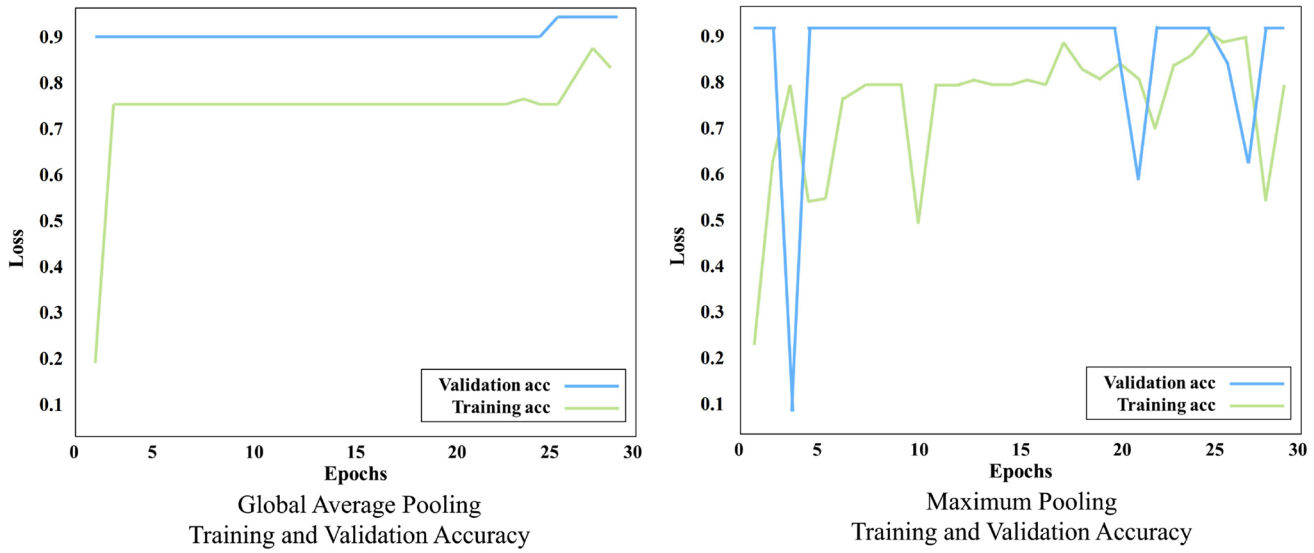

Fig. 6 Training and validation accuracy by global average and max pooling, without synthetic training data

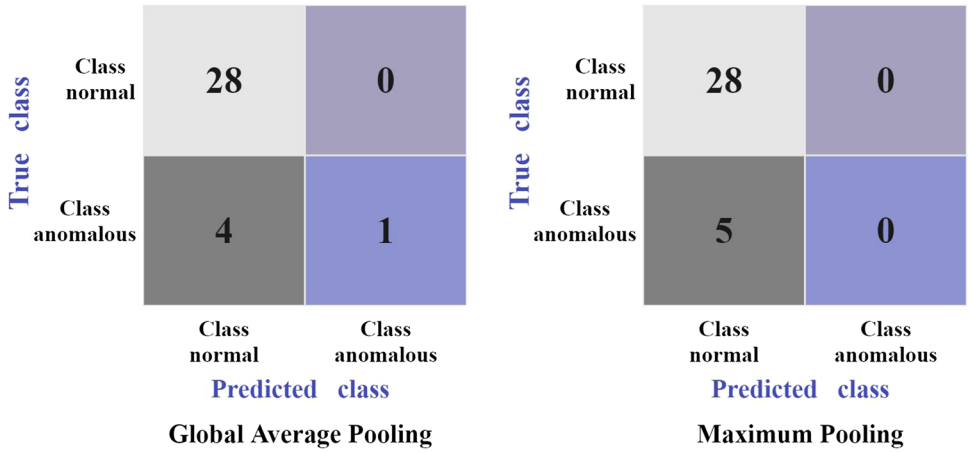

Fig. 7 Confusion matrix by global average and max pooling, without synthetic training data

measures how well the model fits new data. The left graph in Fig. 5 shows the training and validation loss lowering with the epochs, indicating that the model is getting better with learning. However, the graph on the right in Fig. 5, displaying losses for The Maximum Pooling model, shows a mismatched pattern for training and validation, indicating that as the model is struggling to learn the pattern with epochs. These trends are also visible through the accuracy, Fig. 6, and particularly when observed in the resulting confusion matrix, Fig. 7. The confusion matrix in Fig. 7 shows that Global Average Pooling resulted in four incorrectly classified anomalies and only one correctly classified anomaly. Maximum Pooling, however, was not able to learn the trends of anomalous class and did not classify any images as anomalies.

5.5 Image classification with synthetic training data

This section describes the results achieved with added synthetic anomalies through analysis with CNN Global Average Pooling and Maximum Pooling on two-dimensional images [37, 38].

Total number of images in the dataset with added synthetic mechanical damage, is 199, out of which there are 126 normal images and 73 anomalous images with mechanical damage where original and synthetic images are mixed. We split the dataset into 80% for training, and 20% for testing and set the CNN to train over 30 epochs. Figure 8 shows the training and validation loss for Global Average Pooling, and Maximum Pooling, with added synthetic data. Unlike Maximum Pooling the loss for Global Average Pooling shows a good result, with a promising learning trend with the epochs. This is also observed in Figs. 9 and 10 where the accuracy improves for both, training and validation over the epochs, in

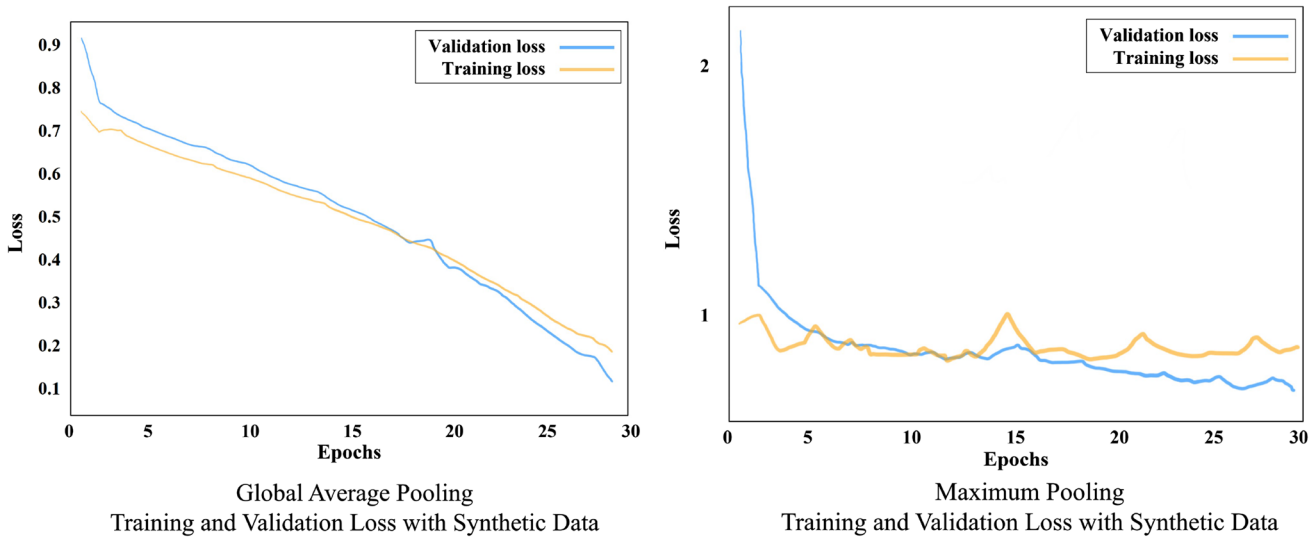

Fig. 8 Training and validation loss by global average and max pooling, with added synthetic training data

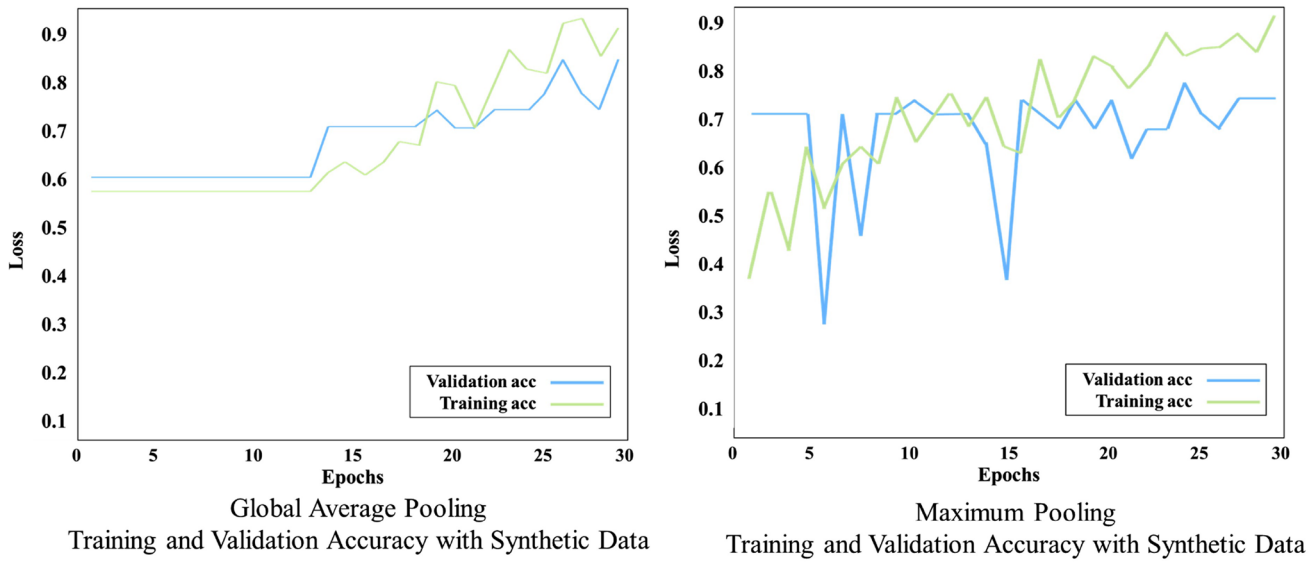

Fig. 9 Training and validation accuracy by global average and max pooling, with added synthetic training data

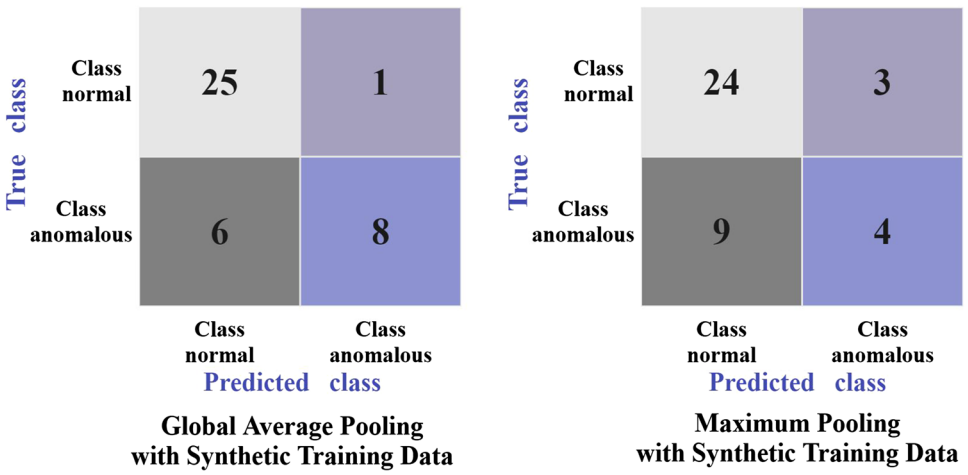

Fig. 10 Confusion matrix by global average and max pooling, with added synthetic training data

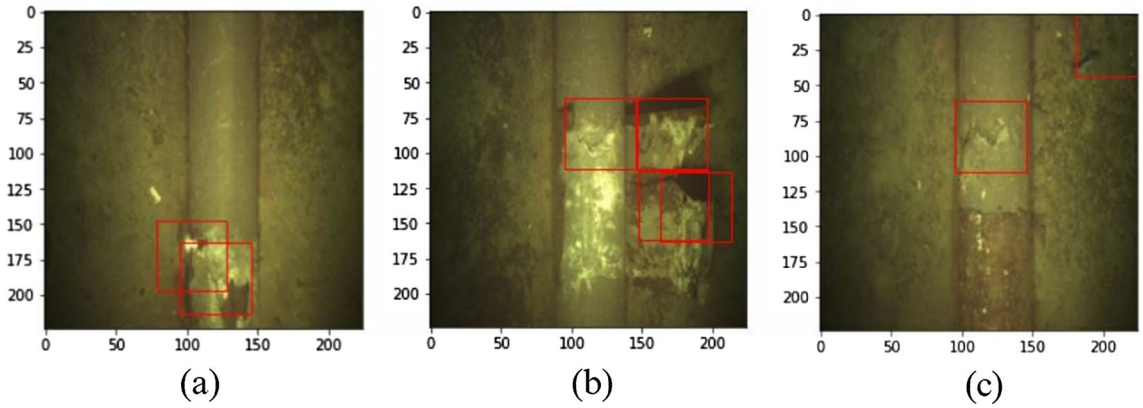

Fig. 11 Localized mechanical damage

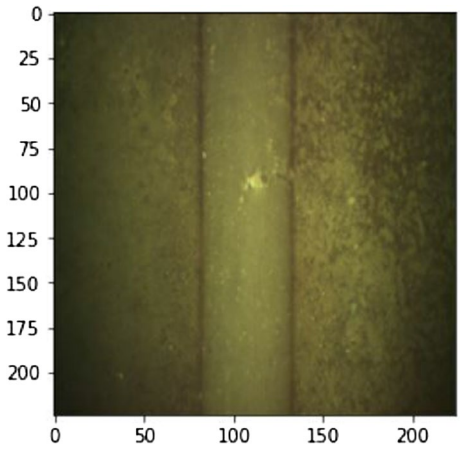

Fig. 12 Inaccurate classification of undersized anomalies. True label: Anomaly; Predicted: Non-anomalous (normal)

both cases Global Average, and Maximum Pooling CNN. In the confusion matrices, Fig. 10, both approaches show that the network was able to learn patterns of anomalous images. With additional synthetic training data, the CNN model has learned the pattern of anomalies more successfully, which is the most optimistic result. In the case of Global Average Pooling, eight anomalies were classified correctly, and six incorrectly. For Maximum Pooling, four anomalies were classified correctly, and nine incorrectly. Maximum Pooling showed difficulty to classify small-sized anomalies (such as the anomaly in Fig. 12). In both cases, there is a high accuracy rate for classifying images without anomalies. When synthetic anomalies are added to the training data, the normal and anomaly classes become more balanced, and the CNN model has more anomaly data to learn from.

5.6 Localized anomaly detection

Localized anomaly detection highlights the anomaly on the evaluated image. The highlighted part of the image illustrates how CNN classified the image into noromal and anomalous regions. Figure 11 illustrates the examples of localized mechanical damage on accurately classified subsea pipeline anomalies, we see three different regions highlighted with red boxes:

(a) Localized damage on the pipeline without any noise.
(b) Localized damage on the pipeline surface, and dislocated anode cover on the sides of the pipeline.
(c) Localized damage on the pipeline surface, and noise in the corner of the image.

Figure 11a shows a clean image of highlighted damage as the most desirable outcome. However, two cases Fig. 11b and c have resulted in additional highlighted regions that do not represent mechanical damage. The highlighted regions

give insight into possible noise levels that result in inaccurately classified anomalies. Similarly, Fig. 12 shows one of the inaccurately classified images where an undersized anomaly is not recognized and captured by CNN.

6 Resulting methodology

The case study and its objectives are summarized in the proposed resulting methodology presented in Fig. 13.

The resulting methodology proposes the eight-task data analysis lifecycle for pipeline damage detection on images of imbalanced subsea pipelines. Tasks 3, 4, and 8 are the most novel contributions to a traditional data analysis lifecycle:

1. The first task is to understand the objective, the problem and gather the data.
2. As the objective is to detect pipeline damage, the second task is to observe the data, identify the anomalies that are pipeline damage within the data, and determine the imbalances between the anomaly and no-anomaly data classes.
3. Once the anomaly and no-anomaly classes have been determined, the third step is to prepare the data by extracting images with pipeline damage from the dataset, masking, and annotating images in preparation for the next step.
4. The fourth task is processing the data which entails generating synthetic damage by seamless blending and image manipulation. This step allows us to expand the training data with additional evidence of pipeline damage.
5. Once the training data is complete, the fifth task consist of training the classification models.
6. After the training is complete, the sixth task is testing the classification models.
7. Utilizing appropriate evaluation metrics, the seventh task is the validation of classification outcomes.
8. Finally, the eighth task is to communicate and interpret the classification results. One of the efforts at interpretation is the application of localized anomaly detection that provides more precise insight into damage detection and possible errors. The last task is particularly important for complex image analysis algorithms that are challenging to explain.

The proposed methodology is based on the case study presented in this paper and the primary challenges identified in image analysis and damage detection, such as a lack of training data and the difficulty explaining Black-box algorithms.

7 Discussion

Despite the small data size, the resulting methodology that includes generating synthetic anomalies to balance the heavily imbalanced data and employing localized anomaly detection has proven to be a promising strategy for addressing the lack of training data, imbalance, and explainability issues that are commonly encountered in image analysis. The subsea images present additional difficulties with visibility, color, and resizing which is especially evident in cases of small and less evident anomalies that are challenging to detect. The resizing of the images has contributed to loss of information resulting in small and less evident anomalies to be less visible. However, resizing of the images is necessary

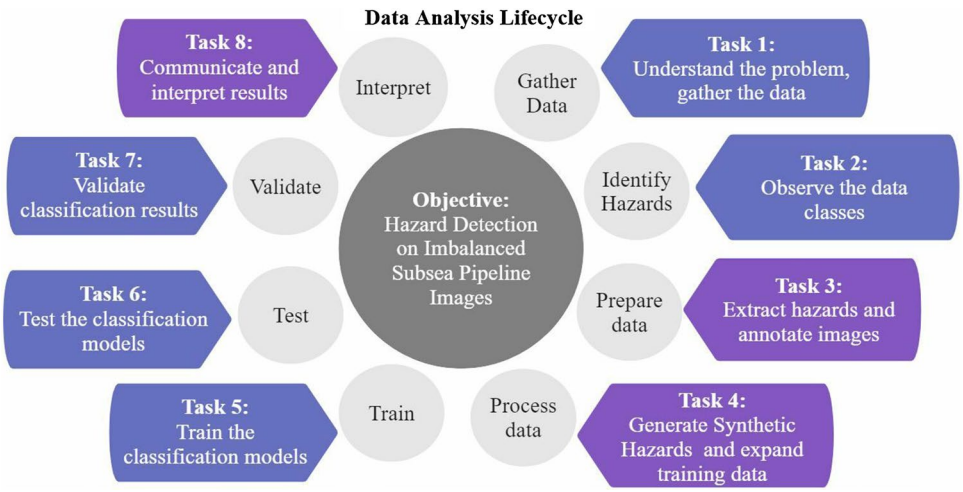

Fig. 13 Resulting methodology

because the computational requirement is a critical challenge. Analysis of large, high quality images requires significant computational resources. Therefore, resizing of images is necessary and during this process, information may be lost. Despite considerable image compression, seamless blending, manipulation, and generation of anomalies allow for the realistic and straightforward expansion of data as required. Moreover, since there is a general absence of high-quality data on subsea pipelines, this method of creating synthetic images may prove useful in industry for generating new data with minimal effort and sharing the data openly and anonymously, while maintaining the realism of the images.

8 Conclusion and future work

As one of the most important assets in the transportation of oil and gas products, subsea pipelines are vulnerable to environmental hazards that can compromise their structural integrity and result in catastrophic environmental damage and financial loss. Autonomous underwater systems (AUS) are expected to assist subsea pipeline inspection and enhance damage detection. However, image-based inspections with computer vision and anomaly detection methods for detecting anomalies, such as pipeline damage, continue to face numerous obstacles that reduce their reliability. These obstacles include visibility, color reconstruction, and shape reconstruction. The lack of training data for image analysis impedes reliable subsea pipeline inspection. In this paper, we analyzed images of subsea pipelines provided by the industry and generated a set of synthetic images using seamless blending techniques. We compared the outcomes of convolutional neural networks trained on data with and without synthetic anomalies. In addition, localized anomaly detection during CNN training and validation increases explainability by highlighting regions of classification impact. Finally, we demonstrated the potential of our approach of augmenting the data with synthetic anomalies and presented the tasks in a new methodology that expands the traditional data analysis lifecycle. The proposed methodology shows a potential in training AUS for more reliable damage detection, and assisting pipeline inspection tasks.

We plan to generate additional anomalies as risk factors, such as misplaced objects and boulders, and test multivariate classification and semantic segmentation in our future research. With additional data and evidence of pipeline damage, we intend to further test the proposed methodology. In addition, we plan to investigate methods that can analyze large images without resizing or information loss, such as moving window methods that analyze large images in batches.

Acknowledgements We would like to express our appreciation to Tor Arne Nordberg Bergset for sharing insightful knowledge on image editing methods, and Assist. Prof. Simo Makiharju for sharing knowledge on corrosion and materials in offshore structures. This research is part of BRU21—NTNU Research and Innovation Program on Digital and Automation Solutions for the Oil and Gas Industry (www.ntnu.edu/bru21) and supported by Equinor, who also provided the subsea pipeline images for the case study. This research is also a part of a project at the University of California, Berkeley, funded by Peder Sather Center for Advanced Studies.

Author contributions All authors contributed to the research conception. Rialda Spahic performed material preparation, literature and data analysis, and manuscript writing. Kameshwar Poolla and Mary Ann Lundteigen performed writing reviews and supervision of all prior drafts of the manuscript. Vidar Hepsø contributed to the concept visualization of the research. All authors reviewed and commented on prior manuscript versions. All authors read and approved the final manuscript.

Funding Open access funding provided by Norwegian University of Science and Technology. This research is a part of BRU21—NTNU Research and Innovation Program on Digital and Automation Solutions for the Oil and Gas Industry (www.ntnu.edu/bru21), supported by Equinor, and a part of a project at the University of California, Berkeley, funded by Peder Sather Center for Advanced Studies.

Data availibility Data sharing is available per request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests There are no financial or non-financial interests to disclose by the authors.

References

1. ISO:51: Safety aspects—Guidelines for their inclusion in standards ISO/IEC Guide.2014;51:2014(E).
2. Ho M, El-Borgi S, Patil D, Song G. Inspection and monitoring systems subsea pipelines: a review paper. SAGE Publications Ltd; 2020. https://doi.org/10.1177/1475921719837718.
3. Yasseri S. Selection of leak detection systems by aggregation of experts' judgment. In: Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering—OMAE, vol. 6A. American Society of Mechanical Engineers (ASME), 2014. https://doi.org/10.1115/OMAE2014-23111.
4. Yasseri S, Bahai H. Safety in marine operations. Int J Coastal Offshore Eng. 2018;2(3):29–40. https://doi.org/10.29252/ijcoe.2.3.29.
5. Abicht D, Torvestad JC, Solheimsnes PA, Stenevik KA. Underwater intervention drone subsea control system. Proc Annu Offshore Technol Conf. 2020. https://doi.org/10.4043/30701-ms.
6. Oxford University Press: Oxford Learner's Dictionaries. 2021. https://www.oxfordlearnersdictionaries.com/.
7. Zhou Y, Li B, Wang J, Rocco E, Meng Q. Discovering unknowns: context-enhanced anomaly detection for curiosity-driven autonomous underwater exploration. Pattern Recognition. 2022. https://doi.org/10.1016/j.patcog.2022.108860.
8. Spahic R, Hepsø, Vidar, Lundteigen MA. Using Risk Analysis for Anomaly Detection for Enhanced Reliability of Unmanned Autonomous Systems. In: Leva, M.C., Patelli, E., Podofillini, L., Wilson, S. (eds.) Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022), 2022;pp. 273–280. Research Publishing, Singapore, Singapore. https://doi.org/10.3850/978-981-18-5183-4_R08-03-390.
9. Rausand M. Risk assessment theory, methods, and applications. John Wiley and Sons Inc, Hoboken. 2011. https://doi.org/10.10029781118281116.
10. Zhu H, Xie W, Li J, Shi J, Fu M, Qian X, Zhang H, Wang K, Chen G. Advanced computer vision-based subsea gas leaks monitoring: a comparison of two approaches. Sensors. 2023;23(5):2566. https://doi.org/10.3390/s23052566.
11. González-Sabbagh SP, Robles-Kelly A. A survey on underwater computer vision. ACM Comput Surv. 2023. https://doi.org/10.1145/3578516.
12. Rumson AG. The application of fully unmanned robotic systems for inspection of subsea pipelines. Ocean Eng. 2021. https://doi.org/10.1016/j.oceaneng.2021.109214.
13. Wang J, Fu P, Gao RX. Machine vision intelligence for product defect inspection based on deep learning and Hough transform. J Manuf Syst. 2019;51:52–60. https://doi.org/10.1016/j.jmsy.2019.03.002.
14. Jacobi M, Karimanzira D. Underwater pipeline and cable inspection using autonomous underwater vehicles. In: OCEANS 2013 MTS/IEEE Bergen: The Challenges of the Northern Dimension. 2013. https://doi.org/10.1109/OCEANS-Bergen.2013.6608089.
15. Nash WT, Powell CJ, Drummond T, Birbilis N. Automated corrosion detection using crowdsourced training for deep learning. Corrosion. 2020;76(2):135–41. https://doi.org/10.5006/3397.
16. Carlevaris-Bianco N, Mohan A, Eustice RM. Initial results in underwater single image dehazing. In: MTS/IEEE Seattle, OCEANS 2010. 2010. https://doi.org/10.1109/OCEANS.2010.5664428.
17. Ren Z, Fang F, Yan N, Wu Y. State of the art in defect detection based on machine vision. Korean Soc Precision Eng. 2022. https://doi.org/10.1007/s40684-021-00343-6.
18. ...Alyamkin S, Ardi M, Berg AC, Brighton A, Chen B, Chen Y, Cheng HP, Fan Z, Feng C, Fu B, Gauen K, Goel A, Goncharenko A, Guo X, Ha S, Howard A, Hu X, Huang Y, Kang D, Kim J, Ko JG, Kondratyev A, Lee J, Lee S, Lee S, Li Z, Liang Z, Liu J, Liu X, Lu Y, Lu YH, Malik D, Nguyen HH, Park E, Repin D, Shen L, Sheng T, Sun F, Svitov D, Thiruvathukal GK, Zhang B, Zhang J, Zhang X, Zhuo S. Low-power computer vision: status, challenges, and opportunities. IEEE J Emerg Selected Topics Circuits Syst. 2019;9(2):411–21. https://doi.org/10.1109/JETCAS.2019.2911899.
19. Spahic R, Lundteigen MA, Hepsø V. Context-based and image-based subsea pipeline degradation monitoring. Discover Artif Intel. 2023;3(1):17. https://doi.org/10.1007/s44163-023-00063-7.
20. Mishra D, Singh SK, Singh RK. Deep architectures for image compression: a critical review. Elsevier B.V. 2022. https://doi.org/10.1016/j.sigpro.2021.108346.
21. Yin F. Inspection Robot for Submarine Pipeline Based on Machine Vision. In: Journal of Physics: Conference Series, vol. 1952. IOP Publishing Ltd, 2021. https://doi.org/10.1088/1742-6596/1952/2/022034.
22. Syamsul Amri SQ, Abdul Ghani AS, Kamarul Baharin MAS. Implementation of Underwater Image Enhancement for Corrosion Pipeline Inspection (UIECPI). In: 2023 19th IEEE International Colloquium on Signal Processing and Its Applications, CSPA 2023—Conference Proceedings, pp. 195–200. Institute of Electrical and Electronics Engineers Inc., 2023. https://doi.org/10.1109/CSPA57446.2023.10087382.
23. Samnejad M, Aboelatta M, Vu C, Wood D. Asset Inspection Powered by Computer Vision: The Use of Deep Neural Networks for Automating the Detection and Classification of Pipeline External Damage. In: PRCI Virtual Research Exchange (VREX2021). 2021.
24. Bastian BT, N J, Ranjith SK, Jiji CV. Visual inspection and characterization of external corrosion in pipelines using deep neural network. NDT E Int. 2019. https://doi.org/10.1016/j.ndteint.2019.102134.
25. Khan A, Ali SSA, Anwer A, Adil SH, Meriaudeau F. Subsea pipeline corrosion estimation by restoring and enhancing degraded underwater images. IEEE Access. 2018;6:40585–601. https://doi.org/10.1109/ACCESS.2018.2855725.
26. Vachtsevanos G, Lee B, Oh S, Balchanos M. Resilient design and operation of cyber physical systems with emphasis on unmanned autonomous systems. J Intel Robotic Syst Theory Appl. 2018;91(1):59–83. https://doi.org/10.1007/s10846-018-0881-x.
27. Jaiswal A, Babu AR, Zadeh MZ, Banerjee D, Makedon F. A survey on contrastive self-supervised learning. Technologies. 2020;9(1):2. https://doi.org/10.3390/technologies9010002.

28. Diao Y, Cheng W, Du R, Wang Y, Zhang J. Vision-based detection of container lock holes using a modified local sliding window method. Eurasip J Image Video Proc. 2019. https://doi.org/10.1186/s13640-019-0472-1.
29. Tadjiev D. ANOMALY CRITERIA FOR GENERAL VISUAL INSPECTION OF SUBSEA FLEXIBLE PIPES. In: Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering. 2020. https://doi.org/10.1115/OMAE2020-19044. http://asmedigitalcollection.asme.org/OMAE/proceedings-pdf/OMAE2020/84355/V004T04A023/6606309/v004t04a023-omae2020-19044.pdf.
30. Funge WJ. ASCE Pipeline Division Specialty. In: Proceedings of the ASCE Pipeline Division Specialty Conference. 1979.
31. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intel. 2022. https://doi.org/10.1007/s12065-020-00540-3.
32. Lin M, Chen Q, Yan S. Network In Network. 2013.
33. Liu W, Li R, Zheng M, Karanam S, Wu Z, Bhanu B, Radke RJ, Camps O. Towards Visually Explaining Variational Autoencoders. 2019. arXiv. https://doi.org/10.48550/arXiv.1911.07389.
34. H, M, S, MN. A review on evaluation metrics for data classification evaluations. Int J Data Mining Knowledge Manag Process. 2015. https://doi.org/10.5121/ijdkp.2015.5201.
35. Perez P, Gangnet M, Blake A. Poisson Image Editing. Microsoft Research UK: Technical report; 2003.
36. Heartex I. Label Studio: Open Source Data Labeling Platform. 2023. https://labelstud.io.
37. Keras & TensorFlow 2: GlobalAveragePooling2D layer. 2023.
38. Keras & TensorFlow 2: MaxPooling2D layer Keras. 2023.