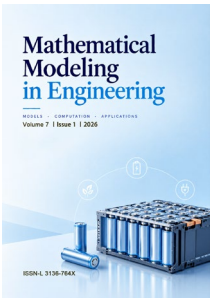
UNIVERSITY OF WEST ATTICA
eISSN: 3136-764X

Mathematical Modeling in Engineering

Vol. 7, No. 2, April, 2026



Topology-Preserving Feature Fusion with a DINOv3 Adapter for Corrosion-Under-Insulation Prediction in Pipelines

Draginja Petković[a,*], Obrada Vasić[a]

[a] *Faculty of Industrial Engineering, University of Kragujevac, Kragujevac, 34000, Serbia*

Received 15 January 2026; Revised 19 April 2026; Accepted 22April 2026; Published 28 April 2026

Abstract

Corrosion under insulation is difficult to predict because moisture ingress, temperature cycling, coating damage, and incomplete inspection records can conceal wall loss before visible symptoms emerge. This study proposes a topology-preserving feature-fusion model that adapts a frozen DINOv3 visual backbone to pipeline CUI assessment through lightweight residual adapters. Registered image tokens are coupled with thermal, ultrasonic, insulation, and operating descriptors on a segment graph that retains axial and circumferential neighborhoods while limiting message transfer across physical discontinuities. The model jointly predicts four corrosion-severity classes and maximum wall-loss depth, and its calibrated output supports maintenance ranking rather than replacing direct inspection. Experiments used 1,248 insulated carbon-steel pipe segments with grouped separation by physical line, class-balanced training, modality masks for incomplete records, and an untouched test distribution. The proposed model achieved 0.913 grade accuracy, 0.887 macro F1, 0.872 severe-CUI recall, and 0.42 mm mean absolute error for depth prediction. Relative to the non-topological DINOv3 Adapter baseline, topology-aware fusion reduced depth error by 18.6% and improved severe-case recall by 7.1 percentage points. Ablation and subgroup analyses indicate that adapters improve domain transfer, ultrasonic evidence anchors loss magnitude, and graph neighborhoods improve recognition of spatially clustered degradation. The resulting framework provides a parameter-efficient and auditable basis for prioritizing intrusive CUI inspection under heterogeneous field evidence.

Keywords: Corrosion Under Insulation, Topology Preserving Fusion, DINOV3 Adapter, Pipeline Integrity, Wall Loss Prediction

1. Introduction

Corrosion under insulation remains a problem for the integrity of process pipelines due to damaging electrochemical reactions that occur in the protected area away from regular inspection [1]. Moisture ingress through damaged jacketing, chloride-induced cracking, thermal gradients and intermittent wet-dry cycles can cause local damage to the wall before staining and insulation swelling or leakage [2]. Non-uniform corrosion in the vicinity of the elbow, support, weld and low-drainage area will also increase the engineering difficulty [3]. Therefore, the traditional screening program has been enhanced with visual walk-throughs, thermal anomalies, pulsed eddy current measurements, ultrasonic spot measurements and operating records, each having different spatial resolution, uncertainty and physical significance [4]. Therefore, the evaluation of CUI will need to be based on infrequent inspections and expert judgments for many insulated pipe sections at different temperatures and insulation systems.

Data driven CUI prediction has advanced from rule-based susceptibility matrices toward statistical learning models that integrate inspection histories and environmental variables. Early regression or tree-based models can rank large

* Corresponding author: draginja.pet@fim.kg.ac.rs

https://doi.org/10.68406/mme.2026.vol7iss2nm3:27-40

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

asset populations, but their representations often treat pipe segments as independent samples and therefore underuse the spatial continuity that exists along a pipeline [5]. Deep visual models offer a different route because field images, infrared patterns, coating states, and insulation damage cues can be embedded into high dimensional descriptors with better tolerance to lighting and texture variation [6]. Foundation vision models are especially attractive because their pretrained representations can capture edges, materials, local shapes, and defect patterns without requiring a large task specific image corpus [7]. However, directly fine tuning a large backbone is usually inefficient for industrial inspection data, where labels are expensive, class imbalance is strong, and the number of severe corrosion samples is limited [8].

The central research gap is the absence of a compact CUI predictor that can exploit foundation-model features while preserving the physical organization of pipeline inspection evidence. Corrosion observed at one grid location is rarely independent of adjacent axial and circumferential positions, and an identical wall-loss value can imply different risk when it belongs to a coherent band rather than an isolated measurement. This study makes three contributions. First, it adapts frozen DINOv3 representations through lightweight modules that limit task-specific training. Second, it introduces topology-preserving fusion that combines registered visual, thermal, ultrasonic, insulation, and operating evidence on a pipe-segment graph. Third, it couples grade classification, wall-loss regression, calibration, and inspection-budget analysis so that predictive performance can be interpreted in an integrity-management context.

The remainder of the paper follows the evidence path from inspection data to maintenance ranking. Section 2 reviews CUI prediction, multimodal inspection, foundation models, adapters, and spatial relation modeling. Section 3 defines the inspection graph, topology-preserving fusion, prediction targets, and loss terms. Section 4 describes DINOv3 adapter placement, multimodal gating, calibration, and the implementation protocol. Section 5 reports the dataset, grouped evaluation design, comparative results, ablation tests, subgroup behavior, and inspection-budget analysis using two tables and six data figures. Section 6 presents the conclusions, limitations, and deployment-oriented research directions.

2. Related Work

2.1. CUI Prediction and Multimodal Inspection Features

Traditionally, the prediction of CUI has used susceptibility rules that combine operating temperature, insulation type, cladding condition, coating age and environmental exposure. The above rules are suitable for pre-invasive tests; however, discrete scores may lose the continuous characteristic of physical evidence and be reduced to a general risk range [9]. Machine learning has also been employed in research to improve the system by learning the non-linear relationship among thermal history, moisture-exposure and wall-thickness data, especially when inspection records from multiple shutdown cycles are available [10]. However, many of the reported models use tabular descriptors extracted at the segment level and thus lose information about the local arrangement of corrosion cells along the pipe surface [11].

A number of tests can cover the entire process; infrared detection can identify an abnormal heat-transfer condition, ultrasonic thickness-grid methods can determine the reduced wall thickness, and external-image observation logs damage to the outer jacket, insulation sag, sealant cracks, etc., at the support joint. Firstly, the density of samples and the uncertainty in registration for these modes are not uniform. The area of the thermal anomaly is relatively large; the ultrasonic values are sparse and localized; only photographs are available, and the internal metal loss cannot be identified. Therefore, the effective fusion model should align the features without assuming that all modalities have the same spatial support.

Another defect of the existing CUI datasets is the mismatch between screening intent and modelling targets. Screening tools are used to quickly identify areas with defects, and prediction models are typically trained on confirmed wall loss data obtained after insulation removal. If the two stages are combined without proper annotation, the model may learn the visibility of a screening symptom instead of the actual severity of hidden corrosion. The definition of the label in this paper is related to the measured wall loss and engineering grade, and external symptoms and thermal signatures are used as evidence channels. Therefore, the prediction object is associated with pipeline condition rather than with the convenience of any single inspection device.

2.2. Foundation Models and Adapter Based Transfer

Vision foundation models have begun to attract attention in engineering inspection due to the transferability of pretrained transformer features to small, annotated industrial datasets [12]. DINOv3-style representations can be used for CUI analysis to capture texture, edges, object boundaries, and local changes in insulation jackets and corrosion stains, etc., in thermal images. Full fine-tuning may overfit to the inspection sites and operating conditions, especially if the source distribution contains more common scenes of industrial pipe surfaces [13]. Xin et al. [14] addressed this problem by adding a small trainable module to a frozen backbone and performing task adaptation solely through this compact parameter set while keeping the general visual prior.

Adapter learning is highly suitable for asset integrity problems that need to be updated with new inspection campaigns. The same frozen backbone can be used for several plants, and adapters store site-specific insulation materials and camera settings, etc. Most previous research on adapters has focused on image classification or segmentation and has not explicitly encoded pipe topology. Topology of the prediction for CUI is not an aesthetic trait but indicates whether the adjacent measurements are in line with a corrosion pattern or are isolated measurement errors [15].

Adapter based learning also offers a practical advantage for lifecycle management. Industrial inspection data are not static: new insulation repairs, camera models, temperature regimes, and labeling policies appear as plants update their integrity programs. A fully fine-tuned model may require heavy retraining when these conditions change. A frozen backbone with replaceable adapters can be updated with a smaller computational burden and can preserve a common representation across several asset groups. This property is valuable when a company needs consistent risk ranking across lines while still allowing site specific adaptation.

2.3. Research Gap

The two missing links are shown above. First, the prediction of CUI should not be restricted to feature concatenation; instead, the hidden corrosion field can be organized spatially based on pipe geometry, drainage paths, supports, thermal gradients, etc. Second, a method is needed to connect visual representations with inspection grids and operating descriptions for foundation model adaptation without full backbone retraining. A topology-preserving fusion layer is introduced to connect the DINOv3 adapter features with a multi-task corrosion prediction head in this paper.

Another deficiency is the form of evidence provided to integrity engineers. Most image classifiers only output a category; however, the decision of inspection needs to consider the relationship among visible conditions, measured wall loss, uncertainty, and the physical location of a segment. Concatenate directly to combine the variables; however, in this case, the information of neighboring observations or discontinuities caused by flanges, supports and insulation breaks will be lost. Graph models can remember relationships, but if only tabular data is used to construct a graph, the extended visual representations offered by foundation models cannot be realized. Therefore, a good engineering architecture should combine parameter-efficient visual adaptation with an explicit inspection graph, provide both categorical and continuous outputs, and keep some transparency in how the neighbors of the evidence affect the inspection. The above requirements specify the method scope of the topology-preserving DINOv3 adapter introduced in the following sections.

3. Topology Preserving Feature Fusion Framework

3.1. Problem Definition and Inspection Graph

The inspected pipeline is represented as an ordered graph of pipe segments rather than an unordered sample table. Each node corresponds to a registered axial interval and circumferential band and stores visual, thermal, ultrasonic, insulation, and operating descriptors. Axial edges connect consecutive stations on the same circumferential band, whereas circumferential edges connect bands within the same axial station. Shortcut edges are permitted only for documented support, elbow, or insulation-termination contexts and are removed when a flange, valve, insulation break, or incompatible line identifier interrupts physical continuity. Figure 1 shows how field evidence is registered before graph construction, topology-preserving fusion, and corrosion prediction.

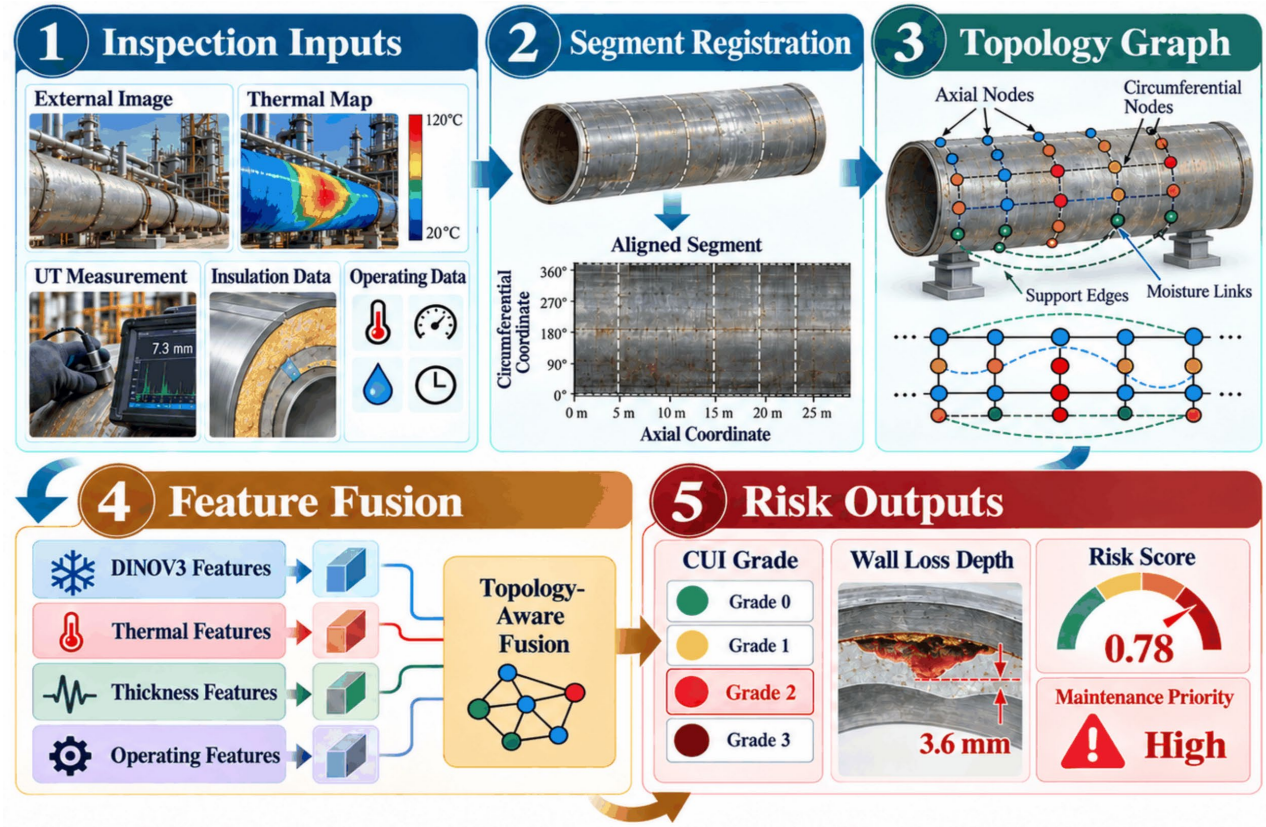


Figure 1. Topology preserving feature fusion workflow for CUI prediction

For a segment node i, the fused input vector combines the DINOV3 visual token, the pooled thermal response, the ultrasonic thickness statistics, and operating descriptors. The following feature-definition equation makes this composition explicit:

$$\boldsymbol{x}_i = [\boldsymbol{v}_i; \boldsymbol{t}_i; \boldsymbol{u}_i; \boldsymbol{s}_i; \boldsymbol{o}_i] \tag{1}$$

The first stage of graph construction is registration, and all visual and thermal observations are linked to an inspection area. Continuous survey of the surface of a pipe obtains the axial coordinate via line isometrics or inspection-distance markers, and the circumferential coordinate is approximated by camera orientation or ultrasonic grid location. No ideal geometry reconstruction is needed for this. The location of the measurement should be reasonably close to the object and a fitting should be employed; otherwise, the measurement result will be unreliable. Therefore, the resulting graph shows engineering adjacency rather than an image-based mesh.

The three reasons for the change in the weight of an edge are: distance, insulation continuity and similarity of moisture evidence. A small axial distance increases the edge weight; otherwise, if there is a flange, a valve or insulation break that interrupts the continuity, the weight will be reduced. Thermal similarity is used only as an auxiliary factor because the heat response can also be influenced by wind, emissivity and operating conditions. The design of the edge will not lead to unrelated parts of the graph sharing prediction results merely because they are similar. It can give the field engineer a specific direction for the investigation of why one part interferes with another during inference.

3.2. Topology Preserving Feature Fusion

The fusion layer performs two related but distinct operations. Weighted message passing aggregates evidence from physically admissible neighbors, while topology regularization penalizes inconsistent predictions only when connected segments share comparable geometry and exposure. Edge weights combine distance, insulation continuity, and moisture similarity and are normalized before aggregation. A residual path retains the original node descriptor, preventing strong local ultrasonic evidence from being erased by neighboring observations. This separation allows clustered wet-insulation patterns to reinforce one another without forcing isolated pits or measurement artifacts to become spatially smooth.

The normalized adjacency matrix controls message passing through a residual graph operator. Message propagation for node i at layer l follows the graph-update equation below:

$$\boldsymbol{h}_i^{(l+1)} = \sigma\left(W_s \boldsymbol{h}_i^{(l)} + \sum_{j \in \mathcal{N}(i)} a_{ij} W_n \boldsymbol{h}_j^{(l)} \right) \tag{2}$$

A topology regularization term further encourages neighboring predictions to remain consistent when their inspection context is similar. That constraint is formalized by the topology equation that follows:

$$\mathcal{L}_{topo} = \sum_{i,j} a_{ij} \left\| \boldsymbol{p}_i - \boldsymbol{p}_j \right\|_2^2 \tag{3}$$

A fusion layer is added to provide a residual connection and keep local evidence in the absence of neighboring information. CUI can be applied at the site of the damage for a local pit, and otherwise, strong smoothing operations would reduce the signal strength. The residual path preserves the original node descriptor and appends structured context to the graph path. Based on the above validation results, as shown in the graphs, clustered wet insulation patterns are more strongly correlated with hidden corrosion than isolated visual blemishes; this is consistent with the actual occurrence of concealed corrosion.

Topology preservation can also reduce the instability caused by sparse ultrasonic grids. The accuracy of ultrasonic measurement at the location is relatively high; however, if the grid spacing is too large, a peak in local wall loss may be missed. An adjacent Area with a similar degree of moisture and visual conditions can be used to extend the sparse data and determine whether the local corrosion rate has been underestimated. Conversely, a single low reading that is not supported by neighboring evidence can be considered unreliable, and thus the proportion of measurement artifacts in the segment risk score will be reduced.

3.3. Prediction Targets and Loss Function

The model predicts two coupled targets: a corrosion grade and a maximum wall loss depth. The grade target supports inspection prioritization, while the depth target supports engineering assessment of remaining wall thickness. Cross entropy, robust absolute error, and topology regularization are coupled in the multitask objective equation:

$$\mathcal{L} = \mathcal{L}_{ce}(y, \hat{y}) + \lambda_d |d - \hat{d}| + \lambda_t \mathcal{L}_{topo} \tag{4}$$

The equation deliberately uses moderate mathematical complexity because the engineering value lies in making heterogeneous inspection evidence coherent. The feature fusion layer is therefore designed to be auditable: every input group can be traced to a field measurement, and the graph edges follow pipe geometry or documented exposure similarity rather than arbitrary latent proximity.

The classification and regression targets are trained simultaneously because the maintenance decision is rarely a single output. A grade label is used to show the risk quickly and provide a depth estimate for comparison with the remaining wall requirements of engineers. Jointly train the two targets to have the model share representations of categorical severity and measured wall loss. Based on the validation set, loss weights have been assigned to increase the weight of high-recall for the severe class and depth stability; these two features are more important for CUI screening than a small increase in the accuracy of the majority class.

The loss function also supports uncertainty review through the distribution of grade probabilities. When the model assigns similar probabilities to mild and moderate classes, the output can be flagged as a boundary case rather than presented as an overly precise decision. This behavior is useful near grade thresholds, where measurement noise and labeling practice can change the assigned class. The model is therefore designed to support inspection planning, not to replace direct engineering judgment or code based remaining life calculation.

4. DINOV3 Adapter Prediction Model

4.1. Adapter Insertion in DINOV3

The DINOv3 backbone serves as a frozen visual feature extractor, while trainable adapters provide task-specific residual corrections. Adapters are inserted after selected intermediate and late transformer blocks; the early blocks remain unchanged to preserve general edge, texture, and material cues. Each adapter applies a bottleneck down-projection, nonlinear activation, up-projection, and residual update to the incoming token sequence. Backbone gradients are disabled, so only adapters, multimodal fusion weights, and prediction heads are optimized. Figure 2 traces the resulting path from registered inspection inputs to calibrated grade and wall-loss outputs.

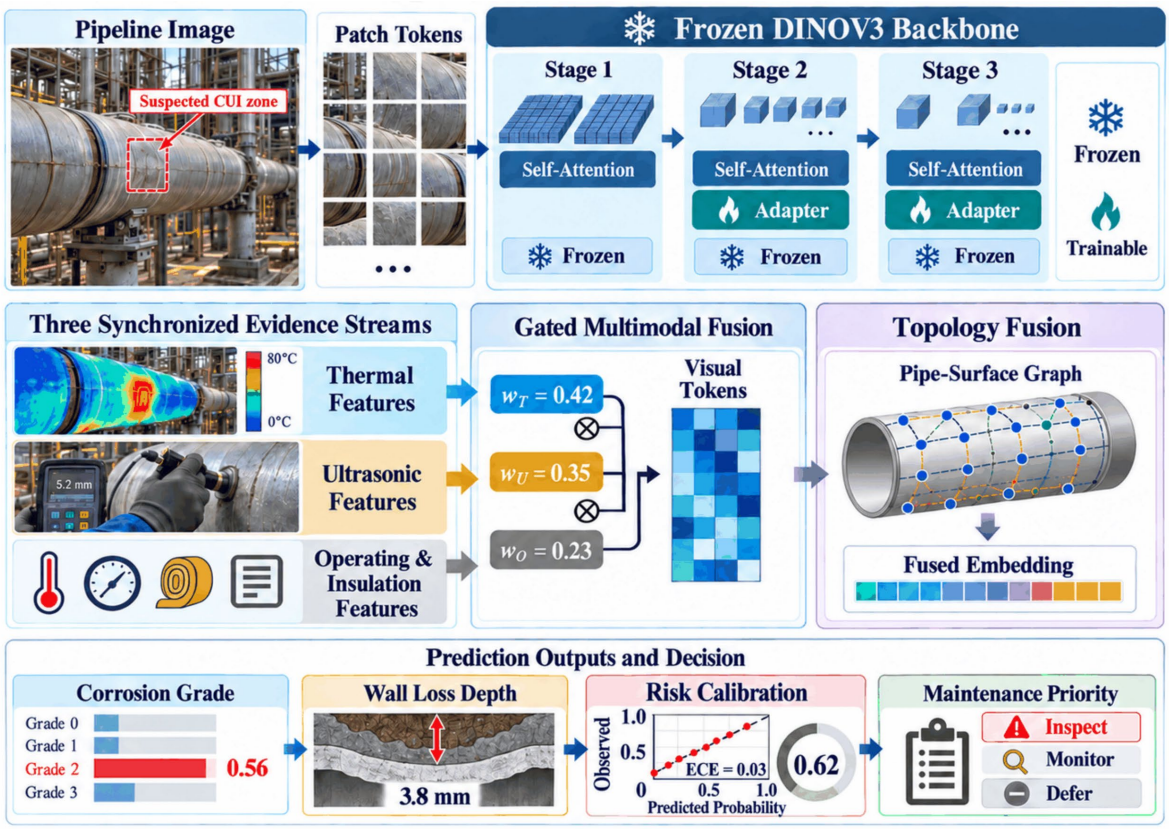


Figure 2. DINOV3 Adapter architecture with multimodal coupling and calibrated CUI outputs

For token matrix X at a selected transformer block, the residual bottleneck operation is specified by the adapter equation below:

$$Adapter(X) = X + W_{up}\,\phi(W_{down}\,LN(X)) \tag{5}$$

Adapter locations are selected after the intermediate and late transformer blocks. The first few visual blocks are not altered because they are general edges, colors and textures that are not particular to CUI. The adapter in the middle section learns patterns of insulation damage, jacket seams, moisture stains and thermal boundary changes. Late blocks are connected to map visual evidence with engineering labels. The distribution of adapters keeps the trainable part small and can move the representation to areas with more specific task semantics.

The Bottleneck dimension is relatively small compared to the backbone hidden dimension. A small bottleneck can be used as a regularize to prevent overfitting to the plant-specific background or camera artefacts in the adapter. If the bottleneck is too small, we will not be able to learn the connections between thermal anomalies, coating conditions and geometry in the model. Therefore, several widths will be selected for the sensitivity analysis of the experiment to determine an appropriate one.

4.2. Multimodal Coupling and Calibration

Visual tokens from the adapter are coupled with thermal and ultrasonic descriptors through gated fusion. The gate learns whether a segment prediction should rely more on external visual symptoms, thermal moisture evidence, or direct thickness readings. This behavior is useful when the external jacket appears intact but infrared response and ultrasonic thinning indicate hidden degradation. Grade confidence and normalized depth are combined through the risk-calibration equation:

$$R_i = \alpha \max_k P(y_i = k) + (1 - \alpha)\frac{\hat{d}_i}{d_{crit}} \tag{6}$$

Calibration is applied to grade probabilities after model fitting, whereas the maintenance risk score combines calibrated severity confidence with normalized wall-loss depth. The two quantities therefore serve different purposes: calibrated probabilities describe predictive confidence, while the combined score ranks segments for follow-up. Class-balanced sampling is used only during training to reduce majority-class dominance. Validation macro F1 and severe-class recall determine checkpoint and threshold selection, while the original class prevalence is retained for test evaluation.

A modality mask prevents unavailable thermal maps or incomplete ultrasonic grids from being interpreted as measured evidence. When a channel is absent, its contribution is set to zero before gate normalization, and the remaining available channels are reweighted. The mask is supplied explicitly so that numerical placeholders cannot be confused with physical measurements. This design retains incomplete field records while reducing the risk that missingness alone becomes a shortcut for corrosion severity.

Calibration is handled after the prediction head because risk ranking must be stable across inspection batches. A model with high raw accuracy may still be difficult to use if its severe probabilities are poorly calibrated. The risk score combines grade confidence and normalized depth because a high probability of moderate corrosion and a high depth estimate can represent a practical maintenance concern even when the severe class is not the maximum probability. This combined score is used only for prioritization; grade and depth remain available as separate outputs.

4.3. Implementation Protocol

Input images are resized to the DINOv3 input resolution, thermal maps are normalized using local operating temperature, and ultrasonic grids are converted into minimum remaining wall, mean wall loss, and thickness-variance descriptors. Explicit masks distinguish unavailable channels from valid zero values. Data augmentation is restricted to photometric variation and small crops so that pipe geometry and defect scale remain physically interpretable. The split key combines service unit, line number, and contiguous inspection run, preventing adjacent segments from the same physical line from appearing in both training and evaluation partitions.

Training uses grouped sampling of pipeline lines and service units to reduce spatial leakage. Otherwise, the adjacent sections of the same wet insulation area may be included in both the training and test sets, and the reported accuracy will be falsely high. The grouped split is more restrictive and better in line with the purpose of use; otherwise, the model would have to be generalized to lines that were not used during training. Data augmentation is limited to photometric changes and small crops to maintain the physical meaning of pipe geometry and visible defect scale.

The optimizer updates only adapter parameters, fusion weights, and prediction heads. Backbone gradients are disabled, which reduces memory use and allows repeated training runs for ablation without excessive computational cost. Early stopping monitors macro F1 and severe recall on the validation set. The selected checkpoint is not necessarily the one with the lowest training loss, because a low training loss can coincide with reduced recall for rare severe samples. This selection rule reflects the operational purpose of CUI screening.

5. Experimental Data and Analysis

5.1. Dataset and Evaluation Setting

The dataset contained 1,248 insulated carbon-steel pipe segments collected from refinery and chemical-process service lines. Each segment included an external image, an available thermal response map, ultrasonic thickness readings, insulation age, jacket condition, operating temperature, nominal pipe size, support proximity, and exposure category. Ground truth comprised four CUI classes—486 no-CUI, 412 mild, 228 moderate, and 122 severe segments—and measured maximum wall loss from 0.00 to 4.85 mm, with a median of 0.74 mm. Large screening campaigns commonly produce imbalanced severity distributions because severe degradation is rarer and is often repaired soon after confirmation [16]. Severe cases represented 9.8% of the records, while no-CUI and mild cases together represented 72.0%. Class-balanced sampling was applied only within training batches; validation and test metrics were calculated on their original grouped prevalence.

Before division, duplicate photos, inconsistent segment IDs and impossible thickness values in the records were filtered out. Twenty-six records were deleted because the ultrasonic values exceeded the nominal wall thickness plus the recorded tolerance, and the corresponding images belonged to a different line. Nineteen records with unavailable thermal files were kept and explicitly modality-masked to preserve the incomplete-data conditions in field archives. The reliability of CUI labels is based on the traceability of screening evidence, exposed-surface examination and thickness confirmation [17]. Therefore, consistency checks were conducted on the measured depths and assigned grades of the eleven labels near the 1.00mm mild-moderate boundary, and no changes were made to the measured wall-loss values. Based on the audit, the same photograph had not been used for the neighboring identifier after file renaming, and the stated nominal diameter matched the visible geometry and line metadata. Missing values were not replaced by cohort means to prevent leakage of prevalence information into the model. Therefore, each missing modality generated a mask variable to reduce the impact of the fusion gate directly.

The data were divided by line group into 70% training, 15% validation, and 15% testing, preventing neighboring segments from the same physical line from entering different partitions. Grouped partitioning is necessary in spatial inspection data because random segment splitting can inflate performance through local similarity and shared exposure history [18]. The operating-temperature distribution covered three regimes: 31% of segments were below 65 C, 44%

were between 65 C and 120 C, and 25% were above 120 C. Nominal diameter ranged from 2 to 20 inches, and insulation types included mineral wool, calcium silicate, aerogel, and mixed repairs. Table 1 records the sample scale, modality coverage, grade definition, adapter width, graph depth, and training configuration used for all comparisons. The grouping key combined service unit, line number, and contiguous inspection run, so the split remained conservative even when short segment identifiers were duplicated across plants. Thermal availability was 81%, ultrasonic coverage was 94%, and external imagery was present for every retained record. These coverage rates provided enough complete cases for multimodal learning while preserving a meaningful subset for testing robustness to missing channels.

Table 1. Experimental dataset and model configuration

Item	Value	Engineering meaning
Pipe segments	1,248	Line-grouped samples from insulated process pipelines
Modalities	Image, thermal, UT, operating descriptors	Heterogeneous evidence for hidden CUI
Grade labels	No, mild, moderate, severe	Maintenance-oriented risk classes
Depth range	0.00-4.85 mm	Measured maximum wall loss
Trainable parameters	8.1 M	Adapter and prediction heads only
Primary metrics	Accuracy, macro F1, severe recall, MAE	Classification and depth performance

Evaluation used grade accuracy, macro F1, severe-class recall, wall-loss mean absolute error, and calibration error. Macro F1 limits dominance by the two majority classes, severe recall quantifies missed high-consequence cases, and depth error preserves the engineering magnitude hidden by discrete grades. Probability calibration supports interpretable risk ranking in imbalanced classification [19]. Grade boundaries were no observable CUI, mild loss below 1.00 mm, moderate loss from 1.00 to 2.50 mm, and severe loss above 2.50 mm. Each configuration was trained with five random seeds on the same grouped partitions. Seed variation describes optimization stability, while confidence intervals were estimated by bootstrap resampling at the line-group level to retain within-line dependence. No claim of formal statistical significance is made. Model selection used validation macro F1 with severe recall as a secondary criterion, and the test set remained sealed until architecture and threshold choices were fixed.

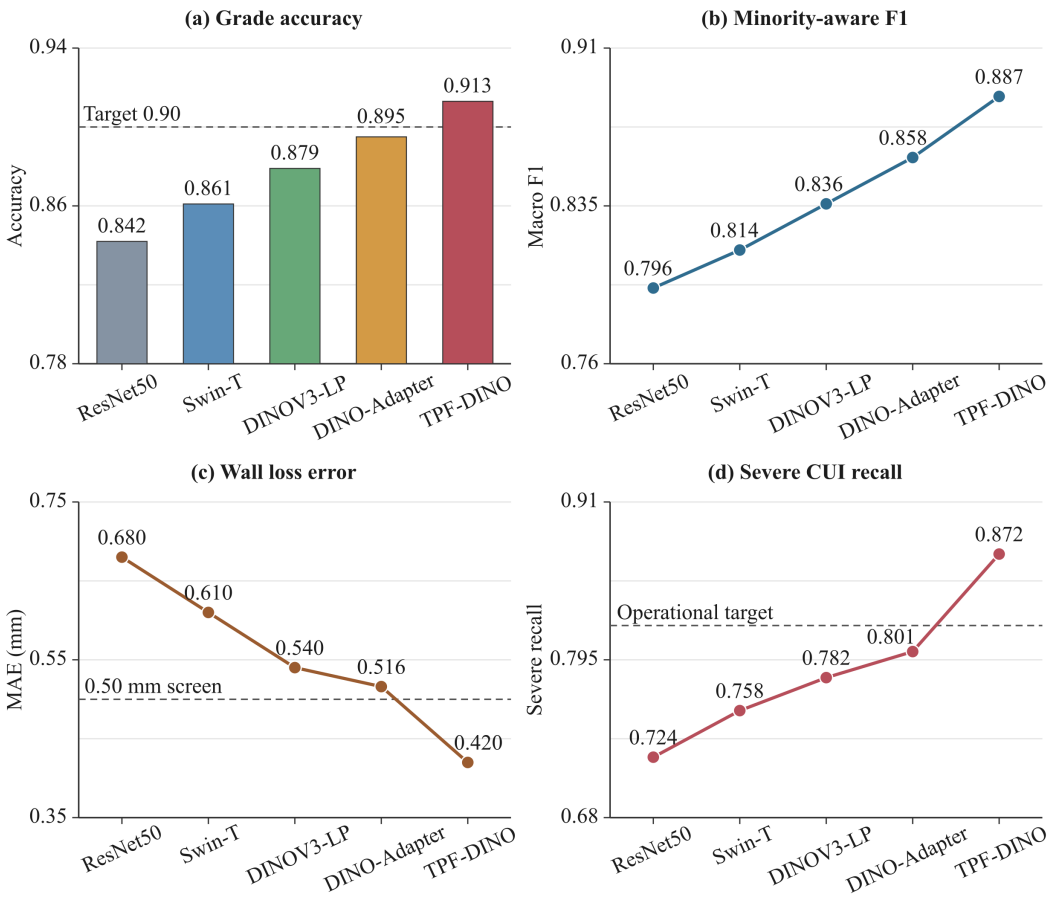


Figure 3. Comparative prediction performance: (a) grade accuracy, (b) macro F1, (c) depth MAE, and (d) severe CUI recall

5.2. Comparative Performance

All baselines used the same line-group split, records, and modality availability. ResNet50 and Swin Transformer received identical visual preprocessing and pooled engineering descriptors; DINOv3 linear probing trained only an output head; and the non-topological DINOv3 Adapter replaced graph fusion with direct concatenation. Foundation representations can improve data efficiency, but adaptation remains necessary when inspection imagery differs from pretraining data [20]. Figure 3(a) shows accuracy increasing from 0.842 for ResNet50 to 0.913 for the proposed model. Figure 3(b) shows macro F1 increasing from 0.858 for the non-topological adapter to 0.887. The gain is therefore concentrated in minority-class discrimination rather than only in the dominant no-CUI class. Across five seeds, the proposed model had standard deviations of 0.004 for accuracy and 0.006 for macro F1; these values indicate stable optimization but are not treated as a formal significance test.

Figure 3(c) reports wall-loss MAE values of 0.68, 0.61, 0.54, 0.516, and 0.42 mm, relative to the non-topological adapter, topology-preserving fusion reduced error by 18.6%, and severity-conditioned analysis showed mean errors of 0.24 mm for no CUI, 0.36 mm for mild CUI, 0.48 mm for moderate CUI, and 0.67 mm for severe CUI. The increasing error with severity reflects both the smaller high-loss sample count and wider physical variation, but the severe-case error remained below the 0.81 mm baseline value. Figure 3(d) gives a severe recall of 0.872, above the marked operational target of 0.82 and 7.1 points above the non-topological adapter. Severe-class recall deserves separate reporting because a high overall accuracy can coexist with unacceptable misses in an imbalanced integrity dataset [21]. To achieve a target recall of 0.82, the new model has a higher operating margin than the previous ones. A small margin will be left because of calibration drift in the inspection cycle, and thus, the actual recall will not meet the high standards for screening. Therefore, the topology-aware model had a higher mean and additional tolerance to threshold uncertainty.

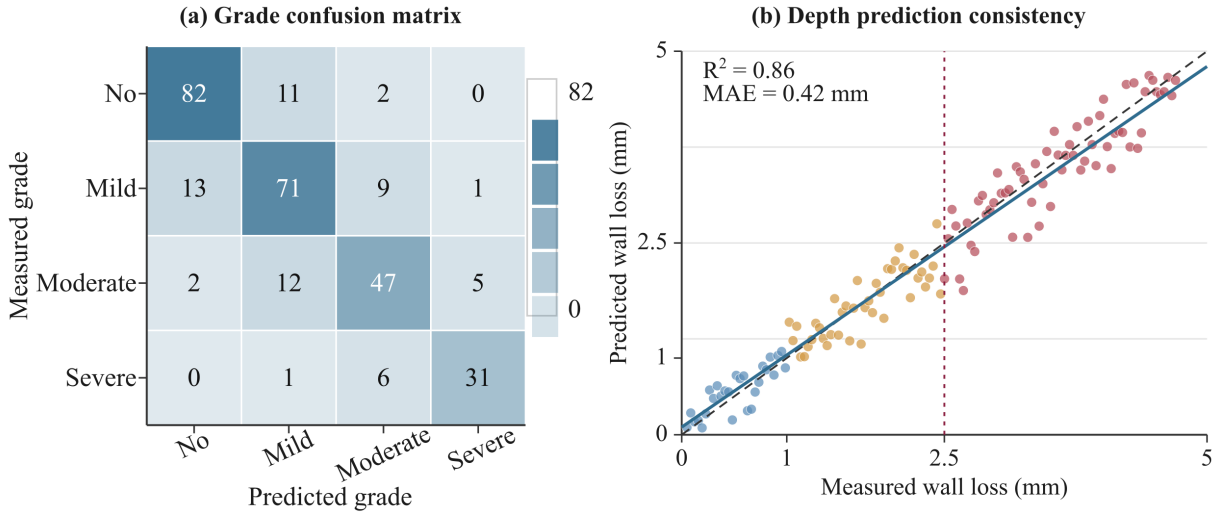


Figure 4. Prediction consistency: (a) confusion matrix for corrosion grade and (b) predicted versus measured wall loss

The confusion matrix in Figure 4(a) localizes the remaining classification errors. Mild and moderate cases formed the largest off-diagonal exchange, with 9 mild samples classified as moderate and 12 moderate samples classified as mild; this pattern is compatible with label sensitivity around the 1.00 mm boundary rather than a general collapse of class separation. Only 6 of 38 severe test cases were assigned to the moderate class, and none was assigned to no CUI. Figure 4(b) compares predicted and measured wall loss over the full 0-5 mm range. The fitted slope of 0.94, R squared of 0.86, and 0.42-mm MAE indicate limited systematic bias, while the visibly wider scatter above 3.5 mm identifies the range in which additional severe examples would provide the greatest benefit. Sparse observations at the high-damage tail are a recurrent limitation when confirmed severe defects are promptly repaired [22]. Residual inspection showed no pronounced bias below 2.5 mm: positive and negative errors were distributed on both sides of the identity line. Above that boundary, underprediction became somewhat more frequent, which is visible in points below the diagonal. The pattern supports targeted acquisition of high-loss examples rather than a global recalibration that would disturb the well-populated mild range.

Table 2 separates the effects of topology, adapter tuning, and inspection modalities. Removing topology regularization reduced macro F1 from 0.887 to 0.858 and increased depth MAE from 0.42 to 0.49 mm, whereas removing adapter tuning reduced accuracy to 0.884. The components therefore address different limitations: adapters align frozen visual tokens with site-specific inspection conditions, while topology constrains how evidence propagates among physically related segments. Removing thermal descriptors reduced severe recall from 0.872 to 0.813, and removing ultrasonic descriptors produced the largest depth error. Because thermal response is affected by load, emissivity, and environment, it should be interpreted together with direct thickness evidence [23]. The ablation results

do not make any single modality sufficient; they show that explicit neighbor aggregation, consistency regularization, and complementary measurements contribute different information.

Table 2. Ablation study of topology fusion, adapters, and modality groups

Model variant	Accuracy	Macro F1	Severe recall	Depth MAE mm
Full proposed model	0.913	0.887	0.872	0.42
Without topology regularizer	0.892	0.858	0.834	0.49
Without adapter tuning	0.884	0.842	0.821	0.53
Without thermal descriptors	0.887	0.849	0.813	0.50
Without ultrasonic descriptors	0.862	0.817	0.786	0.64

5.3. Detailed Data Analysis

Figure 5(a) resolves depth error by nominal pipe size. Median absolute error declined from 0.46 mm for small-bore segments to 0.41 mm for medium sizes and 0.37 mm for large-bore segments; the wider small-bore distribution coincided with stronger perspective distortion and fewer ultrasonic points per axial interval. Figure 5(b) examines adapter bottleneck width. Macro F1 increased from 0.846 at width 16 to 0.887 at width 64, then declined to 0.878 at width 128 even though trainable capacity increased. The response indicates underfitting at the narrowest setting and site-specific memorization at the widest setting, making width 64 a capacity compromise rather than an arbitrary default. Figure 5(c) shows that expected calibration error fell from 0.081 to 0.047 after topology fusion; clustered severe cases moved from confidence values near 0.55 toward the 0.70-0.85 interval when neighboring evidence supported the local prediction. Figure 5(d) places the minimum depth MAE at a topology weight of 0.15. Values above 0.20 increased error, consistent with oversmoothing of genuinely localized wall loss. The calibration curves remained close to the diagonal through the middle confidence bins, where most moderate cases occurred. Deviations were larger in the highest bin because it contained fewer observations, but the proposed curve still avoided the marked overconfidence of the adapter baseline. This behavior supports ranking without implying that the predicted probability is a substitute for direct inspection evidence.

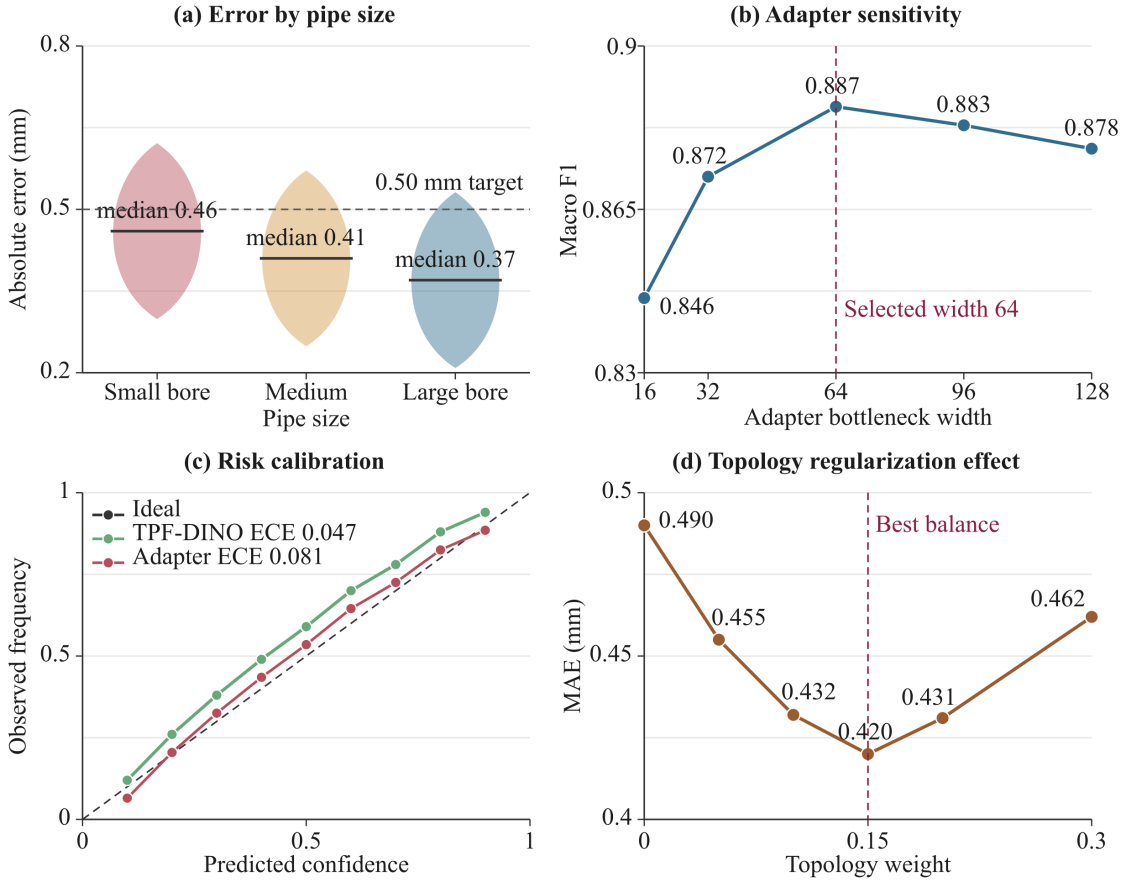


Figure 5. Model behavior under engineering variables: (a) depth error by pipe size, (b) adapter width sensitivity, (c) risk calibration, and (d) topology regularization effect

Figure 6(a) compares controlled modality combinations. Visual features alone achieved 0.802 macro F1 and 0.744 severe recall; adding ultrasonic descriptors raised the two metrics to 0.861 and 0.821, while the complete visual-thermal-ultrasonic-operating input reached 0.887 and 0.872. The visual-plus-thermal condition improved severe recall but did not match the depth stability obtained with ultrasonic evidence, demonstrating that thermal response located suspicious moisture patterns whereas thickness measurements anchored loss magnitude. Insulation material and jacket condition alter moisture retention and heat transfer beneath cladding [24]. Figure 6(b) therefore separates severe recall by material: mineral wool reached 0.891, aerogel 0.874, mixed repairs 0.858, and calcium silicate 0.846. The lower calcium-silicate value coincided with variable drying signatures and a greater proportion of repaired jackets, while the mixed category exposed the uncertainty created by incomplete maintenance metadata. Bubble size in the material comparison represents the number of available segments, making the uncertainty of the smaller aerogel and mixed-repair groups visible alongside recall. The material differences should therefore be interpreted as conditional performance within this dataset, not as intrinsic susceptibility rankings. Additional plant-specific examples are needed before changing inspection intervals by insulation type.

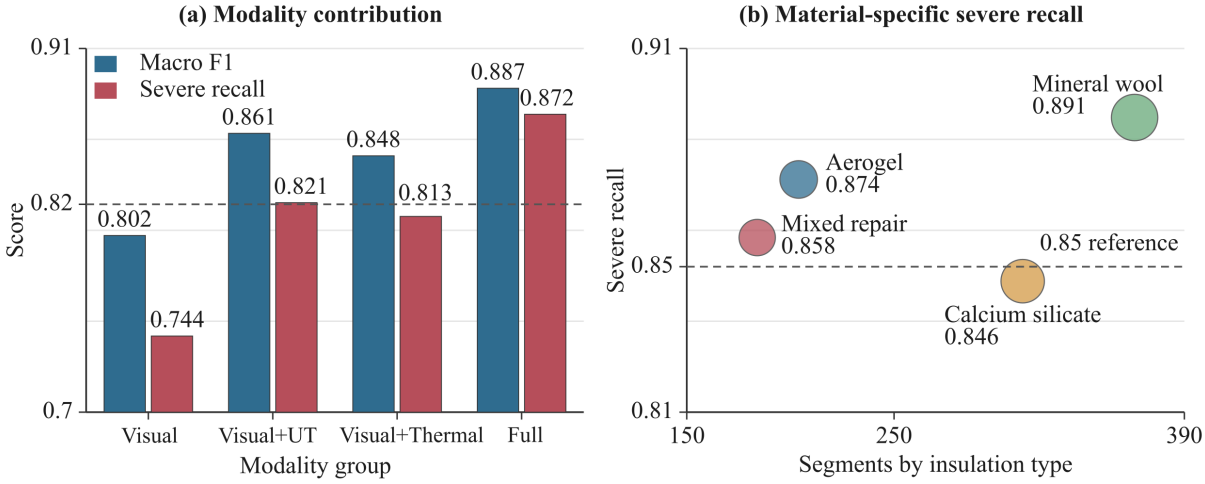


Figure 6. Modality and material effects: (a) contribution of modality groups and (b) severe recall by insulation type

Figure 7(a) converts calibration into an inspection decision. At a risk threshold of 0.62, severe recall was 0.872 while 23.4% of all segments were flagged for intrusive follow-up. Lowering the threshold to 0.50 increased recall to 0.918 but raised the flagged share to 41%, whereas a threshold of 0.70 reduced burden to 17% and recall to 0.812. The curve allows the operating point to change with outage duration and failure consequence instead of embedding a universal threshold. Supports can retain moisture, compress insulation, and restrict drainage, making their local context important in CUI screening [25]. Figure 7(b) shows a corresponding exposure gradient: mean calibrated risk was 0.66 within 0.25 m, 0.61 at 0.5 m, 0.45 at 1.5 m, and 0.34 at 3 m. The interquartile band nevertheless overlaps across distances, confirming that support proximity contributed context but did not override contradictory thermal or ultrasonic evidence. At distances beyond 2 m, a small number of segments still occupied the upper risk tail because their direct thickness evidence was strong. Conversely, several support-adjacent segments remained low risk when jackets, thermal response, and ultrasonic readings were all normal. These counterexamples demonstrate that the plotted distance trend arose from fused evidence rather than a deterministic support-distance rule.

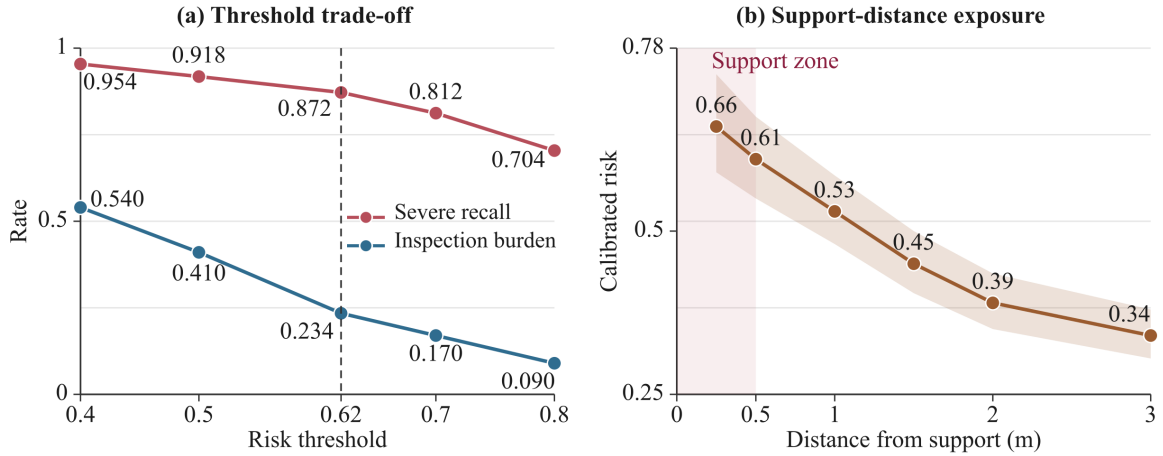


Figure 7. Risk threshold and spatial exposure analysis: (a) severe recall and inspection burden by threshold and (b) risk score by distance from support

Maintenance screening is constrained by insulation-removal access, scaffolding, outage labor, and confirmatory measurement capacity [26]. Figure 8 therefore evaluates ranking under an inspection budget rather than treating every high score as an automatic intervention. At a 20% budget, the proposed ranking captured 61.5% of severe cases, compared with 52.4% for the non-topological adapter and 45.1% for tabular gradient boosting; at 35%, the corresponding values were 82.8%, 73.0%, and 64.2%. In practice, the calibrated list should be reviewed by an integrity engineer, checked against data quality and consequence information, and followed by targeted ultrasonic testing or insulation removal. Missed severe cases were commonly associated with visually quiet jackets and sparse ultrasonic coverage, while false positives often combined damaged jacketing, elevated thermal response, and support proximity with limited measured loss. The model is therefore a prioritization tool, not a substitute for direct condition verification or code-based fitness-for-service assessment.

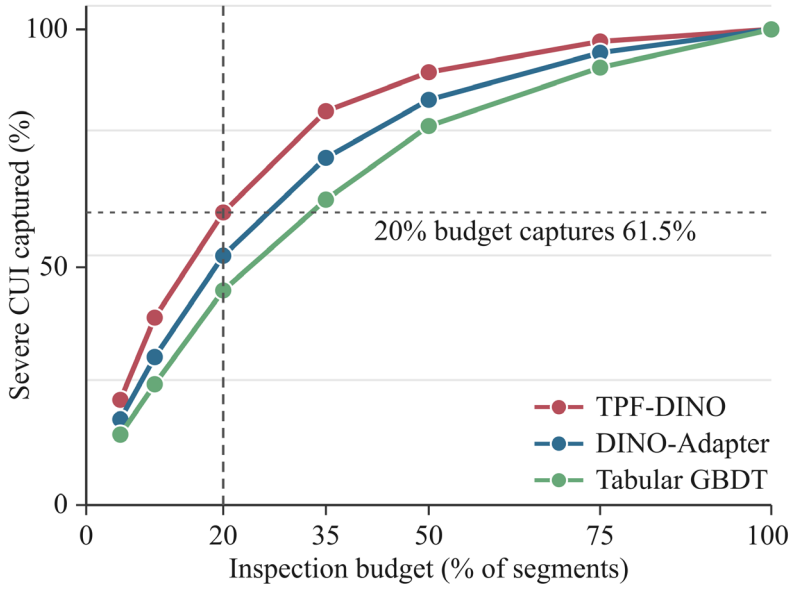


Figure 8. Maintenance prioritization curve comparing severe CUI captured under inspection budget

6. Conclusion

This paper presents a topology-preserving feature-fusion DINOv3 Adapter for pipeline corrosion-under-insulation prediction. The framework treats CUI assessment as multimodal spatial inference by combining frozen visual representations, lightweight adapter tuning, a pipe-segment graph, and coupled grade and wall-loss outputs. Grouped evaluation by physical line and calibration-oriented analysis connect predictive performance to the intended use of screening and maintenance prioritization.

The study remains limited by the available plants, climates, insulation materials, jacket repairs, coating systems, and pipe geometries. Severe cases are comparatively sparse, reflective surfaces can distort thermal evidence, and the graph depends on reliable registration of inspection locations. The method also assumes that consistent line identifiers and segment coordinates are available, which may not hold in older archives or repeatedly repaired insulation systems. These constraints limit claims of universal transfer and require human review of data quality and unusual operating conditions.

Future work should evaluate the method prospectively across additional plants and climates and incorporate time-series inspections that describe corrosion progression. Deployment research should connect calibrated risk to maintenance cost, shutdown constraints, consequence categories, and remaining-life assessment. Field use will also require uncertainty visualization, automatic checks of modality registration, minimum evidence rules for ultrasonic coverage, and controlled adapter updates after new inspection campaigns. Physics-informed corrosion-growth constraints may further improve data efficiency while preserving known moisture, temperature, and wall-loss relationships.

7. Declarations

7.1 Author Contributions

Draginja Petković contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Obrada Vasić contributes to validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

7.2 Funding and Acknowledgments

This research received no specific financial support from any funding agency.

7.3 Institutional Review Board Statement

Not applicable.

7.4 Informed Consent Statement

Not applicable.

7.5 Data Availability

Not applicable.

7.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7.7 Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

8. References

[1] Popova, K., Prošek, T., Šefl, V., Šedivý, M., Kouřil, M., & Reiser, M. (2024). Application of flexible resistometric sensors for real-time corrosion monitoring under insulation. Materials and Corrosion, 75(5), 625–635. https://doi.org/10.1002/maco.202314162

[2] Jiryaeisharahi, Z., De Keukeleere, D., Desmet, Y., Wouters, B., Hubin, A., & Terryn, H. (2026). A novel sensor-based approach for early detection of corrosion under insulation, validated by ORP-EIS technique. Progress in Organic Coatings, 219, Article 110335. https://doi.org/10.1016/j.porgcoat.2026.110335

[3] Bailey, J., Gouws, G. J., & Long, N. (2025). Numerical investigation of encircling eddy current signals for improved sizing of localised corrosion under insulation in steel pipelines. Structural Health Monitoring, 24(5), 2747–2761. https://doi.org/10.1177/14759217251371060

[4] Bailey, J., Gouws, G. J., & Long, N. (2024). Magnetic sensor angle adjustment to improve corrosion under insulation detection. Sensors, 24(3), Article 797. https://doi.org/10.3390/s24030797

[5] Dia, A. K., Bosca, A. G., & Ghazzali, N. (2024). Walk-through corrosion assessment of slurry pipeline using machine learning. International Journal of Corrosion, 2024, 1–11. https://doi.org/10.1155/2024/9427747

[6] Farooqui, M., Rahman, A., Alsuliman, L., Alsaif, Z., Albaik, F., Alshammari, C., Sharaf, R., Olatunji, S., Althubaiti, S. W., & Gull, H. (2024). A deep learning approach to industrial corrosion detection. Computers, Materials & Continua, 81(2), 2587–2605. https://doi.org/10.32604/cmc.2024.055262

[7] Moenck, K., Thieu, D. T., Koch, J., & Schüppstuhl, T. (2024). Industrial language-image dataset (ILID): Adapting vision foundation models for industrial settings. Procedia CIRP, 130, 250–263. https://doi.org/10.1016/j.procir.2024.10.084

[8] Xu, G., Ye, J., Liu, X., Wen, X., Li, Y., & Wang, J. (2024). Lv-Adapter: Adapting vision transformers for visual classification with linear-layers and vectors. Computer Vision and Image Understanding, 246, Article 104049. https://doi.org/10.1016/j.cviu.2024.104049

[9] Benz, A., Sjerve, E., Jonsson, K., Nichols, M., Townsend, M., & Sanchez, E. (2024). Corrosion under insulation—Data from thousands of circuits. Materials Performance, 63(1), 50–54. https://doi.org/10.5006/mp2024_63_1-50

[10] Cui, B., & Wang, H. (2024). Pipeline corrosion prediction and uncertainty analysis with an ensemble Bayesian neural network approach. Process Safety and Environmental Protection, 187, 483–494. https://doi.org/10.1016/j.psep.2024.05.011

[11] Liu, Y., Tan, X., & Bao, Y. (2024). Machine learning-assisted intelligent interpretation of distributed fiber optic sensor data for automated monitoring of pipeline corrosion. Measurement, 226, Article 114190. https://doi.org/10.1016/j.measurement.2024.114190

[12] Xiao, Y., Chen, X., Xie, L., Chen, S., Xiao, X., & Peng, L. (2026). Geometry-constrained open set recognition with frozen foundation model features for industrial inspection. Expert Systems with Applications, 320, Article 132025. https://doi.org/10.1016/j.eswa.2026.132025

[13] Zhou, Y., Yi, Z., & Yen, G. G. (2025). Efficient fine-tuning of vision transformer via path-augmented parameter adaptation. Information Sciences, 703, Article 121948. https://doi.org/10.1016/j.ins.2025.121948

[14] Xin, Y., Du, J., Wang, Q., Lin, Z., & Yan, K. (2024). VMT-Adapter: Parameter-efficient transfer learning for multi-task dense scene understanding. Proceedings of the AAAI Conference on Artificial Intelligence, 38(14), 16085–16093. https://doi.org/10.1609/aaai.v38i14.29541

[15] Yu, C., & Li, Y. (2024). Graph neural network incorporating time-varying frequency domain features with application in spatial wind speed field prediction. Journal of Wind Engineering and Industrial Aerodynamics, 253, Article 105875. https://doi.org/10.1016/j.jweia.2024.105875

[16] Velarde, G., Weichert, M., Deshmunkh, A., Deshmane, S., Sudhir, A., Sharma, K., & Joshi, V. (2024). Tree boosting methods for balanced and imbalanced classification and their robustness over time in risk assessment. Intelligent Systems with Applications, 22, Article 200354. https://doi.org/10.1016/j.iswa.2024.200354

[17] Yakubu, M. A., Caines, S., Adumene, S., & Khan, F. (2026). A data-driven model for pipeline corrosion under insulation-induced failure analysis. Reliability Engineering & System Safety, 266, Article 111826. https://doi.org/10.1016/j.ress.2025.111826

[18] Killeen, P., Kiringa, I., Yeap, T., & Branco, P. (2024). Corn grain yield prediction using UAV-based high spatiotemporal resolution imagery, machine learning, and spatial cross-validation. Remote Sensing, 16(4), Article 683. https://doi.org/10.3390/rs16040683

[19] Guilbert, T., Caelen, O., Chirita, A., & Saerens, M. (2024). Calibration methods in imbalanced binary classification. Annals of Mathematics and Artificial Intelligence, 92(5), 1319–1352. https://doi.org/10.1007/s10472-024-09952-8

[20] Wang, X., Zhang, W., Lin, Z., Li, H., Zhang, Y., Quan, W., Chen, Z., You, X., Zeng, Y., Wang, G., Luo, B., & Yu, Z. (2024). Current status of image recognition technology in the field of corrosion protection applications. Coatings, 14(8), Article 1051. https://doi.org/10.3390/coatings14081051

[21] Ding, Y., Zhao, Q., Li, T., Lu, C., Tao, L., & Ma, J. (2024). A rail defect detection framework under class-imbalanced conditions based on improved you only look once network. Engineering Applications of Artificial Intelligence, 138, Article 109351. https://doi.org/10.1016/j.engappai.2024.109351

[22] Miao, X., & Zhao, H. (2024). Maximum pitting corrosion depth prediction of buried pipeline based on theory-guided machine learning. International Journal of Pressure Vessels and Piping, 210, Article 105259. https://doi.org/10.1016/j.ijpvp.2024.105259

[23] Usamentiaga, R., Venegas, P., Sfarra, S., & Zhang, H. (2025). Enhanced temperature measurement using infrared thermography in dynamic environments through an automated robust detection-tracking approach. Quantitative InfraRed Thermography Journal, 22(5), 424–451. https://doi.org/10.1080/17686733.2025.2464972

[24] Wu, J., Li, W., Huang, F., Zhang, H., Zhang, B., Liu, F., & Jin, Y. (2025). Simulation of internal environmental conditions within rock wool insulation: Implications for corrosion under insulation in piping systems. Coatings, 15(5), Article 571. https://doi.org/10.3390/coatings15050571

[25] Manzoor, A., & Saleem, O. (2025). Investigating corrosion under insulation on hot oil piping. Journal of Failure Analysis and Prevention, 25(1), 301–313. https://doi.org/10.1007/s11668-025-02094-x

[26] Zheng, Y., Dong, Z., Zhang, X., & Shi, H. (2025). Random risk assessment model and risk-based maintenance decisions for natural gas pipelines. Journal of Loss Prevention in the Process Industries, 95, Article 105591. https://doi.org/10.1016/j.jlp.2025.105591