This is the author accepted manuscript of: Young, A., Zabalza, J., Fei, Z., West, G., Murray, P., & McArthur, S. (2025). Knowledge-Driven Explainable AI for Automated Defect Detection in Nuclear Reactor Components. Paper presented at 14th International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies (NPIC&HMIT 2025), Chicago, Illinois, United States.

Knowledge-Driven Explainable AI for Automated Defect Detection in Nuclear Reactor Components

Young A[1]*, Zabalza J[2], Fei Z[1], West G[1], Murray P[2], McArthur S[1]

[1]Institute for Energy and Environment, University of Strathclyde, Glasgow, UK
[2]Institute for Sensors, Signals & Communications, University of Strathclyde, Glasgow, UK

ABSTRACT

Critical assets in nuclear power plants, such as reactor vessels, coolant systems, and containment structures, must be inspected to ensure safe and reliable operations. However, these inspections are often complicated by the harsh environments in which they are conducted, with high radiation levels creating significant visual noise in inspection footage. Identifying defects, such as cracks or surface anomalies, is vital for preventing failures, but manual review by engineers can be time-consuming due to challenging visual conditions. To address these challenges, a novel, automated defect detection algorithm has been developed, integrating image processing techniques with a knowledge-driven framework. The approach uses frame differencing to detect temporal changes in video frames, thresholding to isolate potential defects, and morphological operations to eliminate noise. The main contribution of this work is a rule-based filtering process that incorporates domain-specific knowledge, including factors such as anomaly size, persistence across multiple frames, and proximity to critical surfaces. A key feature of this approach is the emphasis on explainability. Unlike black-box machine learning models, this method provides clear, rule-based justifications for each detected anomaly. Such transparency is crucial in the highly regulated nuclear industry, where every decision must be traceable and defensible. To validate the approach, it was applied to a case study involving Calandria Tubesheet Bore (CTSB) inspection videos, a particularly challenging dataset due to the visual noise caused by radiation. The knowledge-based rules tailored to this inspection process helped filter out irrelevant anomalies and generated detailed reports with visualisations to assist engineers in their final assessments.

Keywords: Image processing, in-core reactor inspections, defect detection, major component replacement, nuclear power plant

1. INTRODUCTION

Visual inspection is a vital process in nuclear power, allowing for evaluating components, systems, and structures within nuclear plants [1]. Several techniques, such as in-situ visual inspections, remote boroscopic methods [2], and robotic inspections [3], are utilised for these assessments. However, all these approaches encounter a shared issue: radiation-induced noise affecting imaging systems. This noise makes it challenging to detect and evaluate defects in critical components accurately [4]. As a result, it often causes false positives and reduces the efficiency of industrial operations. Obtaining precise and reliable visual data in high-radiation settings is crucial for maintaining the safety, dependability, and performance of nuclear facilities.

*Email address: andrew.young.101@strath.ac.uk

Traditionally, these inspections are performed manually, relying on human expertise to identify and address potential issues. However, manual methods are time-intensive and prone to subjectivity. Artificial Intelligence (AI) offers significant potential to enhance visual inspections through automation and advanced analysis. AI encompasses diverse methodologies, including data-driven approaches, symbolic reasoning, and generative AI. Data-driven approaches, often called ”black box” systems, excel at anomaly detection but provide limited explainability. Symbolic approaches, while interpretable, have historically suffered from brittleness and complexity in design. Recently, generative AI, including large language models (LLMs), has shown promise in summarising, and searching large volumes of textual data. However, for anomaly detection in visual inspection, generative AI lacks suitability due to its textual focus and limited applicability to image-based problems.

In most inspection applications, the collected visual data is analysed to detect defects in nuclear components, assess their severity, and prioritise necessary maintenance or repair activities. These inspections play a critical role in ensuring the safe and reliable operation of nuclear facilities by identifying potential issues before they escalate into serious failures. Examining the interior of a nuclear reactor core presents unique challenges due to the extreme radiological environment, which complicates the detection of defects. Artefacts, distortions, and brightness fluctuations in the footage can obscure critical features, making defect identification more difficult. Additionally, prolonged radiation exposure can damage camera components, further reducing the quality of the footage [5]. To overcome these obstacles, advanced defect detection algorithms and robust analysis methodologies extract meaningful insights from the inspection footage, enabling engineers to accurately identify and address anomalies while maintaining operational safety and efficiency [6].

In this work, we present a novel automated defect detection algorithm designed to tackle the challenges of radiation-induced noise in nuclear power plant inspections. By integrating advanced image processing techniques with a knowledge-driven framework, our approach effectively identifies and isolates potential defects, even in high-noise environments. The algorithm employs frame differencing, thresholding, and morphological operations, combined with a rule-based filtering process that leverages domain-specific knowledge to reduce false positives and enhance detection accuracy. Emphasising explainability, our method provides clear, rule-based justifications for each detected anomaly, to aid compliance with regulatory standards. Validated through a case study involving Calandria Tubesheet Bore (CTSB) inspection videos, gathered by Bruce Power during their Major Component Replacement (MCR) program, our approach significantly reduces inspection time while maintaining accuracy, demonstrating its potential to streamline and improve the reliability of nuclear plant inspections.

2. BACKGROUND

Anomaly detection in visual inspection, particularly in high-radiation environments like nuclear power plants, has been extensively studied [4]. Traditional methods often rely on manual inspection, which is time-consuming and prone to human error, especially under challenging visual conditions. To address these limitations, various automated techniques have been developed [7] [8].

One common approach is frame differencing, which detects temporal changes between consecutive video frames. This technique has been widely used in surveillance and industrial inspection applications due to its simplicity and effectiveness in highlighting regions of interest, i.e., regions that have changed significantly or are showhing something unexpected. However, frame differencing alone is insufficient in high-noise environments, as it can generate numerous false positives [9].

Domain expert knowledge plays a crucial role in enhancing the accuracy and reliability of automated defect detection systems. In the context of nuclear power plant inspections, this knowledge is vital for developing algorithms and processes to the unique challenges posed by high-radiation environments. Experts contribute to defining key parameters, such as defect thresholds, anomaly size limits, and critical surface areas, which are used to guide the detection process [10].

Knowledge graphs have emerged as powerful tools for representing and organising complex information in a structured and interconnected manner [11]. In the context of defect detection, knowledge graphs can enhance the explainability and traceability of the inspection process by providing a clear representation of detected anomalies and their attributes. Knowledge graphs have been widely used in various domains, including natural language processing, biomedical research, and industrial applications [12]. They enable the integration of heterogeneous data sources and facilitate the extraction of meaningful insights through their interconnected structure [13]. In visual inspection, knowledge graphs can represent relationships between detected anomalies, their characteristics, and domain-specific knowledge, such as known defect patterns and critical surface areas [12].

3. METHODOLOGY

The proposed automated defect detection algorithm integrates several image processing techniques within a knowledge-driven framework to enhance the accuracy and efficiency of inspections in nuclear power plants but also to imitate the human decision making process. The key components of the algorithm are as follows.

3.1. Anomaly Detection Algorithm

The anomaly detection process, shown in Figure 1, begins with frame differencing, a technique employed to detect temporal changes between consecutive video frames. By comparing each frame with the average of the ten previous frames, the algorithm identifies regions where significant changes occur, which may indicate potential defects. This is because it is expected that there will be a consistent/homogenous surface and any different should be highlighted as a potential defect. Once potential defect regions are identified through this frame differencing, thresholding is applied to isolate these regions from the background. The pixel intensity threshold is set based on an analysis of historical data, ensuring that the threshold accurately distinguishes defects from surrounding noise.

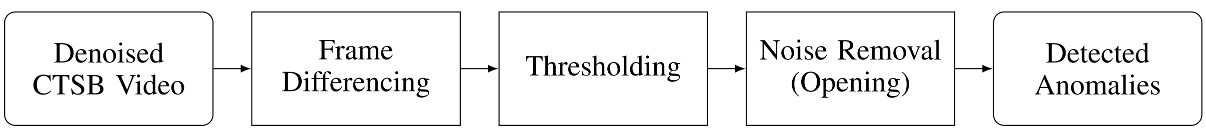


Figure 1: Diagram of the video processing pipeline.

To further refine the detected anomalies and eliminate noise, morphological operations such as opening are used. A morphological opening [14], which involves erosion followed by dilation, helps in removing small, irrelevant artefacts and enhancing the shape of the detected defects. The parameters for these operations, such as the kernel size, are also determined through historical analysis, ensuring that the morphological operations are tailored to the specific characteristics of the inspection environment.

3.2. Rule-Based Filtering Process

A significant contribution of this work is the incorporation of a rule-based filtering process that leverages domain-specific knowledge, shown in Figure 2. This process involves several steps to enhance the accuracy and reliability of the anomaly detection. First, anomalies must exceed a minimum size threshold to be considered significant, as small anomalies are often noise or irrelevant artefacts. By setting a size threshold, the algorithm focuses on defects that are more likely to impact the integrity of the inspected component. Similarly, the algorithm checks that the width of the anomaly is greater than a specified pixel count. Anomalies with dimensions smaller than this are also likely to be insignificant features rather than meaningful defects. The next set of rules evaluates the contrast and variability of the anomaly compared to the surrounding area. If the average difference in pixel values between the anomaly and the background is low, it suggests the anomaly is not sufficiently distinct and may be indistinguishable from the asset surface, and therefore not a defect. Likewise, if the maximum difference in pixel values is also low, it indicates the anomaly does not have a significant contrast, which could lead to false positives.

Additionally, the algorithm checks the variance of pixel values within the anomaly region. If the variance is low, it implies the anomaly does not have sufficient textural or intensity variations to be considered a genuine defect. Finally, the approach verifies the location of the anomaly, ensuring that it is within the asset

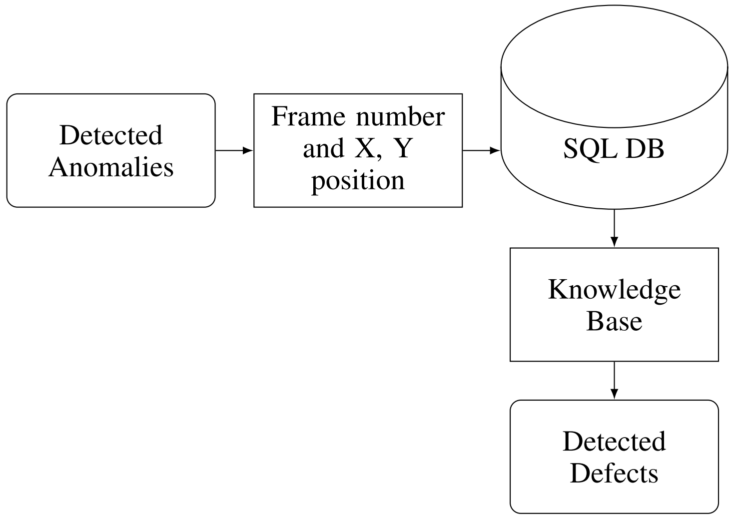


Figure 2: Diagram of the video processing pipeline.

surface. Anomalies positioned outside of this surface are more likely to be background features or irrelevant artefacts rather than defects within the inspected component. An example of the rule base is shown in Figure 3.

IF Area $< A$	THEN:	Small Area	$=$ True,
IF Width $< W$	THEN:	Small Dimension	$=$ True,
IF Average Difference $< D_{\text{avg}}$	THEN:	Low Average Difference	$=$ True,
IF Max Difference $< D_{\text{max}}$	THEN:	Low Max Difference	$=$ True,
IF Variance $< V$	THEN:	Low Variance	$=$ True,
IF Anomaly Position $\notin P_{\text{CTSB}}$	THEN:	Background	$=$ True.

Figure 3: Rule based filtering approach

By applying these rule-based filters, the algorithm can effectively distinguish genuine, significant defects from noise, small irrelevant features, and anomalies that are unlikely to have a meaningful impact on the inspected component, thereby enhancing the accuracy and reliability of the anomaly detection process.

3.3. Knowledge Graphs for Traceability

After anomalies are highlighted and processed using the rule-based filtering process, all relevant information is stored in a knowledge graph. This approach enhances the explainability of the defect detection algorithm by providing a structured and interconnected representation of the detected anomalies and their attributes. The knowledge graph is constructed by mapping each detected anomaly to a set of nodes and edges that represent its characteristics and relationships. For instance, the schema incorporates nodes such as the anomaly identifier, associated defect attributes (e.g., location, severity), timestamps, and relevant inspection metadata. These nodes are interconnected through relationships that trace the origin, propagation, and resolution of each anomaly.

Each detected anomaly is linked to the specific rules and domain knowledge that contributed to its identification, allowing engineers to understand and verify the decision-making process. For example, in the schema shown in Figure 4, nodes representing defects are connected to inspection results, timestamps, and contributing rules, forming a cohesive network of relationships. This rule-based justification ensures that the decision-making process is transparent and understandable. Additionally, visualisations of the knowledge graph, like the one shown (Figure 4), highlight the relationships between anomalies, inspection passes, and associated metadata (e.g., defect categories, inspection dates). These visualisations aid engineers in interpreting inspection results, and understanding defect context. Furthermore, the structured representation of information in the knowledge graph ensures that all decisions are traceable and defensible, supporting compliance with regulatory standards in the nuclear industry.

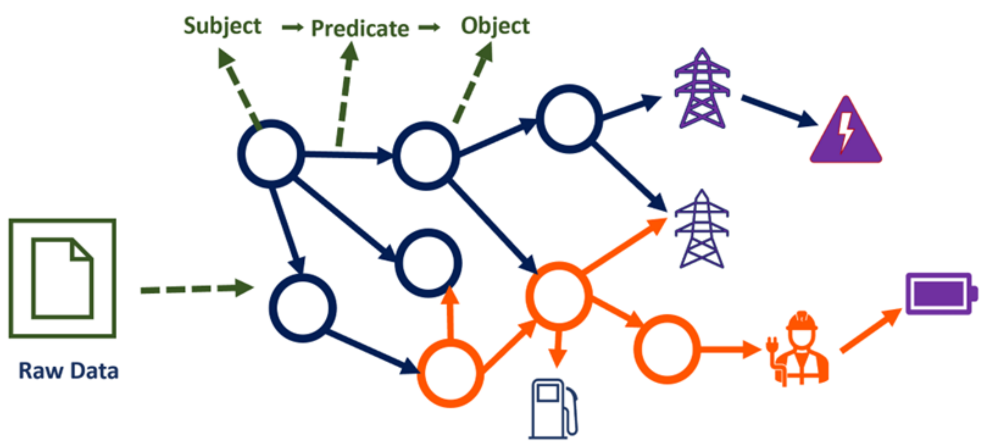


Figure 4: Example of knowledge graph showing details of a specific defect in a specific frame.

4. CASE STUDY

4.1. Calandria Tubesheet Bore (CTSB) Inspection Videos

This case study focused on a specific subset of data collected by Bruce Power (Ontario, Canada) during their Major Component Replacement program, which involved the inspection of Calandria Tubesheet Bores (CTSB) [15] [16]. The dataset included over 1,000 individual 360-degree scan videos, each capturing detailed footage of the interior surface of the CTSB (Figure 5). These scans were conducted using Ahlberg PTZ620 cameras, which are high-definition and specifically designed to withstand the harsh radiation levels present in nuclear inspection environments [17].

4.2. Data Preprocessing

The preprocessing of the CTSB inspection videos involved several steps to prepare the data for anomaly detection. First, the videos were converted into individual frames to facilitate frame-by-frame analysis. The frames are captured at 24 frames per second (fps) and digitally upscaled to 60 fps at a resolution of 1920×1080 during the capture process. Due to the upscaling process, duplicate frames were present and had to be removed to ensure the integrity and accuracy of the subsequent analysis. However, analysis confirmed that the duplicate frames were identical, ensuring that the upscaling process did not introduce any unintended artefacts to the data.

Each frame was then subjected to noise reduction techniques to mitigate the impact of high radiation levels and other sources of noise. This included the application of the motion-based video denoising algorithm discussed in [16]. The effectiveness of the noise reduction technique is being evaluated by domain experts. While earlier concerns were raised about anomaly blurring, these were linked to non-standard rotation speeds. No issues have been observed at expected speeds, and validation is planned to confirm that denoising does not affect detection accuracy. The approach is under review for future MCR deployment.

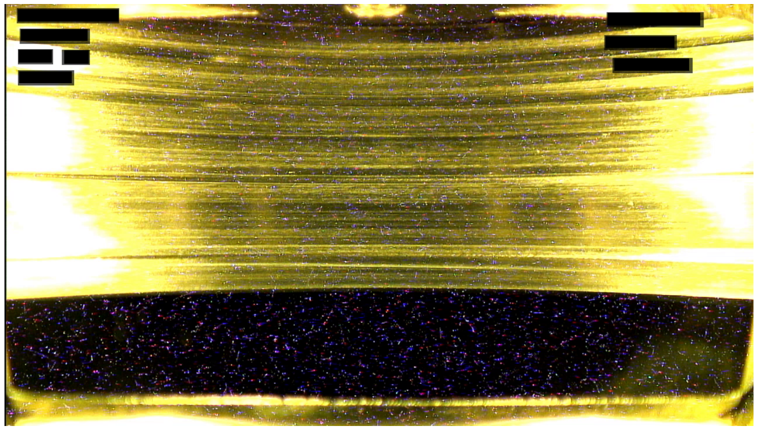
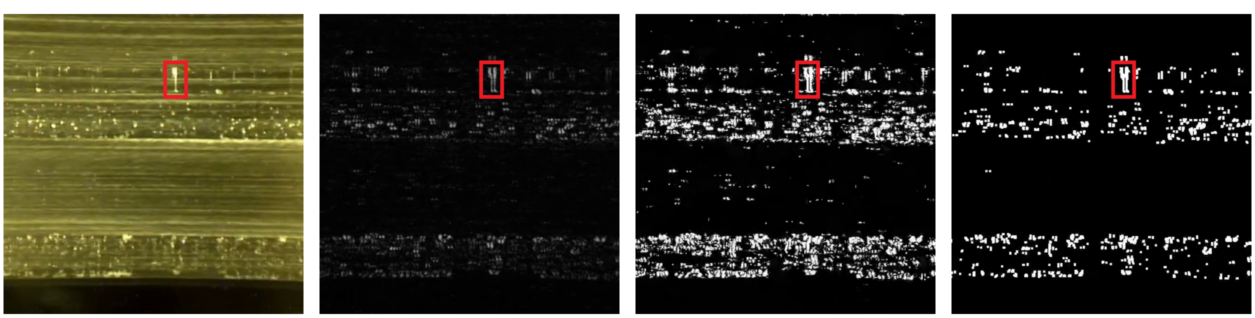


Figure 5: Representative frame gathered from CTSB visual inspection with overlay text anonymised.

Next, the saturated regions identified on the left and right sides of the frames were masked to prevent them from interfering with the anomaly detection process. Similarly to the denoising approach, the masking of saturated regions was assessed by engineers, who verified that no relevant information was lost, ensuring that the anomaly detection process remained unaffected.

4.3. Anomaly Detection

The anomaly detection process was applied to the preprocessed CTSB inspection videos, producing step-by-step results that demonstrate the progression through each stage of the algorithm. Figures 6a to 6d show the individual steps for one frame from an example inspection video. Initially, frame differencing was applied to the preprocessed frames, this was performed on a grayscale version of the image because colour was determined to not be an important feature for the defect detection. This highlights regions of significant temporal change between consecutive video frames. This step effectively identified areas where potential defects may exist by focusing on regions with high-intensity variation (Figure 6b). Next, thresholding was applied to isolate these regions from the background (Figure 6c). By setting a pixel intensity threshold, the algorithm effectively separated anomalies from surrounding noise. This step produced binary masks where potential anomalies were more clearly refined. To enhance the detection accuracy and remove irrelevant artefacts, a morphological opening was performed, as shown in Figure 6d. This operation removed smaller, disconnected noise regions while preserving the structural integrity of larger defect regions. At this stage,



(a) **(b)** **(c)** **(d)**

Figure 6: (a) Original image, (b) Frame differencing, (c) Thresholding, (d) Noise removal (opening). The red box in each image denotes the known location of a defect.

the anomalies were refined into distinct shapes that corresponded to potential physical defects on the CTSB surface.

4.4. Defect Classification

Once anomalies were detected, they were classified based on characteristics such as size, shape, intensity, and location. The classification process involved setting thresholds for each characteristic to determine whether the detected anomaly was likely a significant defect. Since labelled ground truth data was unavailable, the thresholds were determined using expert knowledge and iterative feedback.

The classification of anomalies was based on several key characteristics: size, shape, intensity, and location. The area and dimensions of an anomaly were considered, with experts identifying specific size ranges associated with significant defects such as cracks, pitting, or other forms of damage. The width and height of the anomalies were also examined, with domain experts providing insight into the typical shapes of critical defects. Intensity differences between the anomaly and the surrounding surface were calculated, with experts noting that significant defects often exhibited higher intensity contrasts compared to background features. The thresholds for each variable were initially set based on typical defect sizes and intensities observed by experts in previous inspections.

Since ground truth labels were unavailable, the validation of the thresholds relied on iterative feedback from the domain experts. This process involved detecting anomalies, classifying them based on the established thresholds, and then having the experts review the results. This iterative process helped to fine-tune the classification thresholds, ensuring that only significant defects were flagged for further review. As a result, the thresholds were gradually optimised without the need for ground truth data. In addition, inspection reports were used to further define the thresholds, with all highlighted defects from these reports correctly detected in the inspection video. However, this validation was conducted on a small sample size of 14 videos.

To ensure no defects were missed, the approach was deliberately conservative, prioritising recall over precision. This meant that while more anomalies were flagged for review, the focus was on capturing all potential defects, even at the cost of some false positives. The goal was to ensure that any defects, which could affect the CTSB surface, were not missed.

Figure 7 provides a detailed visual representation of this classification process for a specific frame from

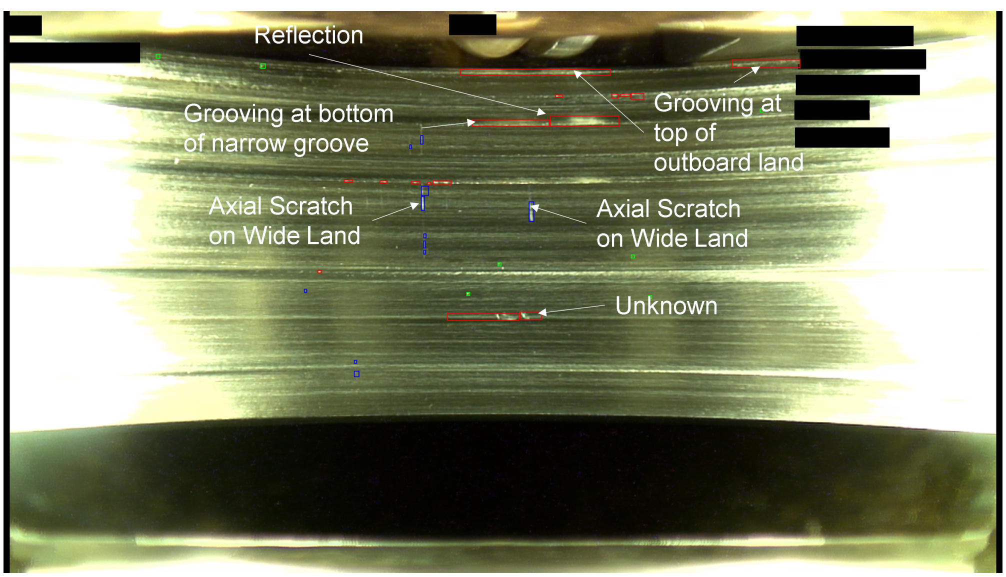


Figure 7: Example of defect classification for a specific frame in an inspection video.

the inspection video. The figure demonstrates how individual anomalies are segmented, colour-coded, and labelled based on their severity and type. For instance, a red-outlined region indicates a defect with a larger width than height, while blue regions represent anomalies with larger height than width, and finally green regions have equal width and height. This colour-coding scheme highlights defects spanning across lands and grooves, with these vertically-oriented (height > width) anomalies being more likely to have an impact on the asset than the other two types.

4.5. Detection Traceability

All the information relating to anomalies, defects and noise was then stored in a knowledge graph, which provided a structured and interconnected representation of the detected anomalies and their attributes. This knowledge graph, see Figure 8, will facilitate the explainability and traceability of the defect detection process, allowing engineers to understand and verify the decision-making process.

The construction of the knowledge graph involved defining entities representing key attributes of the inspection process, including defect identifiers, timestamps, inspection passes, and related metadata. Each node in the graph corresponds to a specific aspect of the inspection, such as individual defect detections, frame number, or anomaly attributes, while edges capture the relationships between these elements. This structured approach enables engineers to trace detected anomalies back to their originating inspection data, ensuring that each detection is explainable and verifiable. Additionally, the knowledge graph provides a persistent record of detected anomalies, allowing for historical comparisons and aiding in long-term monitoring of defect evolution. By utilising this structured representation, engineers can efficiently navigate large datasets, validate detection outcomes, and enhance decision-making processes while maintaining full traceability of each detection.

5. CONCLUSIONS

In this paper, a novel automated defect detection algorithm was presented to enhance the inspection processes of critical assets in nuclear power plants. The algorithm integrates advanced image processing tech-

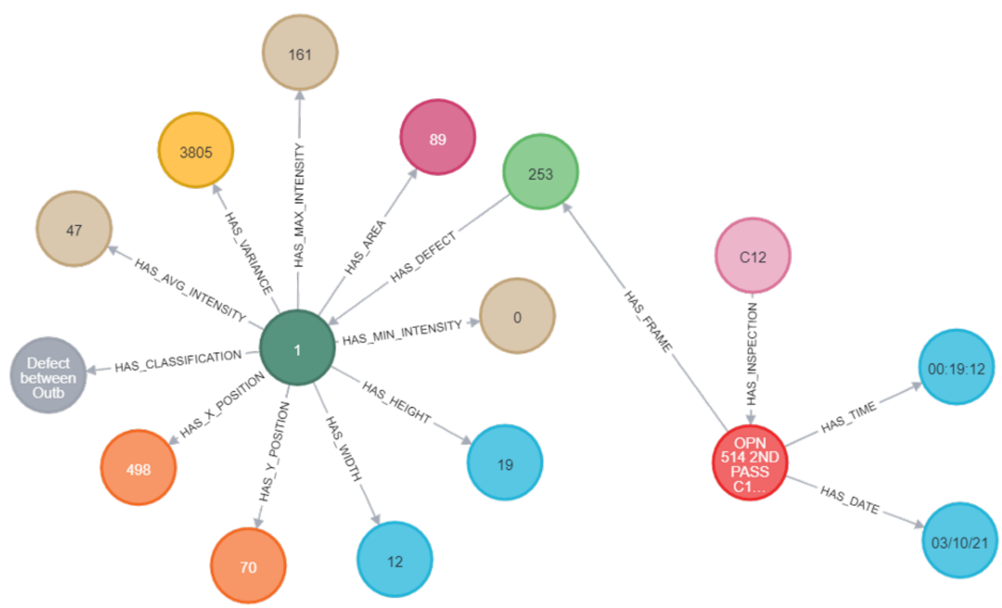


Figure 8: Example of knowledge graph showing details of a specific defect in a specific frame. Defect number (1) has a variety of properties, and is present in frame (253) which is within inspection video (OPN 514 2ND PASS).

niques with a knowledge-driven framework, addressing the challenges posed by high radiation levels and other sources of visual noise in inspection footage.

The methodology involved the use of frame differencing, thresholding, and morphological operations to detect and refine potential defects. A rule-based filtering process, incorporating domain-specific knowledge, was employed to further enhance detection accuracy by reducing false positives. This approach not only improves the reliability of defect detection but also ensures that the process is transparent and explainable, which is crucial in a highly regulated environment such as the nuclear industry.

A case study involving Calandria Tubesheet Bore (CTSB) inspection videos demonstrated the effectiveness of the proposed algorithm. The knowledge-based rules tailored to the CTSB inspection process proved effective in filtering out irrelevant anomalies and generating detailed reports with visualisations to assist engineers in their final assessments.

The integration of a knowledge graph provided a structured and interconnected representation of the detected defects and their attributes, enhancing the explainability and traceability of the defect detection process. This transparency is essential for ensuring compliance with strict safety and regulatory standards in the nuclear industry.

Future work will focus on addressing the limitations of the current study by incorporating a labelled dataset to enable quantitative evaluation of the approaches performance. This will allow for a more comprehensive comparison with existing methods and provide insights into its accuracy and reliability. Additionally, the defect detection and classification algorithms will be refined to enhance their adaptability to different light levels and defect types.

ACKNOWLEDGEMENTS

The authors would like to thank Bruce Power (Ontario, Canada) for providing the video footage and related information used in this work.

REFERENCES

[1] Y. Xu, Y. Cai, and L. Song. "Condition Assessment of Nuclear Power Plant Equipment Based on Machine Learning Methods: A Review." *Nuclear Technology*, **volume 209**(7), pp. 929–962 (2023).

[2] Z. Fei, G. West, P. Murray, and G. Dobie. "Enhanced video-level anomaly feature detection for nuclear power plant component inspections using the latency mechanism." In *13th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies* (2023).

[3] R. Bogue. "Robots in the nuclear industry: a review of technologies and applications." *Industrial Robot: An International Journal*, **volume 38**(2), pp. 113–118 (2011).

[4] M. G. Devereux, P. Murray, and G. West. "Automated Object Detection for Visual Inspection of Nuclear Reactor Cores." *Nuclear Technology*, **volume 208**(1), pp. 115–128 (2022).

[5] C. Wang, S. Hu, C. Gao, and C. Feng. "Nuclear Radiation Degradation Study on HD Camera Based on CMOS Image Sensor at Different Dose Rates." *Sensors*, **volume 18**(2) (2018).

[6] S. Xu, H. Dong, Z. Qin, Y. Han, D. Gong, S. Zou, C. Wei, and F. Zhao. "Parallel processing of radiation measurements and radiation video optimization." *Opt Express*, **volume 30**(26), pp. 46870–46887 (2022).

[7] Z. Guo, Z. Wu, S. Liu, X. Ma, C. Wang, D. Yan, and F. Niu. "Defect detection of nuclear fuel assembly based on deep neural network." *Annals of Nuclear Energy*, **volume 137**, p. 107078 (2020). URL https://www.sciencedirect.com/science/article/pii/S0306454919305808.

[8] S. J. Schmugge, L. Rice, N. R. Nguyen, J. Lindberg, R. Grizzi, C. Joffe, and M. C. Shin. "Detection of cracks in nuclear power plant using spatial-temporal grouping of local patches." In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–7. IEEE (2016).

[9] D. S. Kumar, P. Tejaswini, S. Abirame, N. Gopinath, M. Meena, and S. Gurusubramani. “Analyze and Highlight the Abnormal Activities While Recording in CCTV.” In *2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, pp. 1–5. IEEE (2024).

[10] D. Martin, S. Heinzel, J. K. von Bischhoffshausen, and N. Kühl. “Deep learning strategies for industrial surface defect detection systems.” *arXiv preprint arXiv:210911304* (2021).

[11] X. Chen, S. Jia, and Y. Xiang. “A review: Knowledge reasoning over knowledge graph.” *Expert systems with applications*, **volume 141**, p. 112948 (2020).

[12] X. Zou. “A survey on application of knowledge graph.” In *Journal of Physics: Conference Series*, volume 1487, p. 012016. IOP Publishing (2020).

[13] J. Z. Pan, S. Razniewski, J.-C. Kalo, S. Singhania, J. Chen, S. Dietze, H. Jabeen, J. Omeliyanenko, W. Zhang, M. Lissandrini, et al. “Large language models and knowledge graphs: Opportunities and challenges.” *arXiv preprint arXiv:230806374* (2023).

[14] J. Serra and L. Vincent. “An overview of morphological filtering.” *Circuits, Systems and Signal Processing*, **volume 11**, pp. 47–108 (1992).

[15] H. Barpugga, R. Cluett, and B. Singh. “Bruce Power major component replacement automated inspection and Calandria tube install innovation.” *42nd Annual CNS conference*, **volume 54** (2023).

[16] A. Young, Z. Fei, J. Zabalza, G. M. West, P. Murray, and S. D. McArthur. “Real-World Video Denoising for Visual Inspection in High-Dose Radiological Environments.” *Nuclear Technology*, pp. 1–11 (2024).

[17] Ahlberg Cameras AB. “PTZ620 - HD PTZ Camera with mirror - Nuclear Visual Inspection.” https://www.ahlbergcameras.com/products/cameras/ptz620/.