THE BENCHMARKING VOID: A ROADMAP FOR DOMAIN-ADAPTED COMPUTER VISION IN FUEL CELL DEFECT DETECTION

Rahul D. Ray
Department of Electrical and Electronics Engineering
Birla Institute of Technology and Science (BITS) Pilani, Hyderabad Campus
Hyderabad, Telangana, India
f20242213@hyderabad.bits-pilani.ac.in

Utkarsh Srivastava
Department of Economics
Birla Institute of Technology and Science (BITS) Pilani, Hyderabad Campus
Hyderabad, Telangana, India
f20240633@hyderabad.bits-pilani.ac.in

Rishi Mohapatra
Department of Chemical Engineering
Birla Institute of Technology and Science (BITS) Pilani, Hyderabad Campus
Hyderabad, Telangana, India
f20240796@hyderabad.bits-pilani.ac.in

Naveen K. Shrivastava
Supervisor, Assistant Professor
Department of Mechanical Engineering
Birla Institute of Technology and Science (BITS) Pilani, Hyderabad Campus
Hyderabad, Telangana, India
naveenks@hyderabad.bits-pilani.ac.in

ABSTRACT

Object detection has achieved remarkable success in natural-image domains, yet its deployment in scientific microstructural analysis remains fundamentally constrained by data scarcity, domain mismatch, and a lack of standardized evaluation. This survey systematically reviews a corpus of 129 studies on computer vision–based fuel cell defect detection and conducts in-depth comparative analysis on a representative subset, identifying the absence of benchmarking infrastructure as the primary bottleneck to progress. We demonstrate that direct reuse of general-purpose detection architectures such as Faster R-CNN and YOLO fails to address domain-specific challenges including extreme class imbalance, multi-scale defects spanning orders of magnitude, and low- contrast anomalies embedded in complex material textures. Beyond cataloging architectures, we synthesize the literature around three structural gaps: fragmented and proprietary datasets, architectural mismatch with fuel cell degradation physics, and a deployment chasm between laboratory validation and production environments. We outline actionable research directions centered on standardized multi-modal benchmarks, detection-specific architectures, few-shot learning, and explainable AI to enable trustworthy industrial adoption. Fuel cell inspection is presented as a representative case study for applying computer vision in data-poor, scientifically complex domains.

1 INTRODUCTION

Fuel cell technologies are central to global decarbonization strategies for transportation, backup power, and distributed energy systems. Among available architectures, polymer electrolyte membrane fuel cells (PEMFCs) are particularly attractive due to their high power density, rapid start-up, and low operating temperature. Despite decades of development, large-scale commercialization remains constrained by durability, reliability, and manufacturing consistency challenges.

Microstructural degradation within the membrane electrode assembly (MEA) is a dominant driver of performance loss and premature failure. Key degradation modes include catalyst layer cracking, interfacial delamination, catalyst agglomeration, platinum migration, and membrane thinning, all of which impair proton transport, electronic conductivity, and reactant accessibility. Prior experimental and modeling studies have established that load cycling induces competing crack initiation and delamination processes within catalyst layers, accelerating electrochemical degradation Rong et al. (2008a;b), while dynamic driving conditions further exacerbate platinum redistribution and membrane damage, directly linking microstructural evolution to macroscopic performance decay Lin et al. (2009).

Advanced imaging modalities have become indispensable for characterizing these degradation mechanisms. Scanning electron microscopy (SEM) enables nanometer-scale surface analysis of catalyst layers and interfaces, while X-ray computed tomography (X-CT) provides non-destructive three-dimensional reconstruction of porous electrode microstructures and transport pathways Epting et al. (2012). In situ and four-dimensional X-CT further allow direct observation of crack propagation and catalyst layer thinning during accelerated degradation White et al. (2017). However, interpretation of such high-resolution imaging data remains largely manual, time-consuming, and subjective, limiting scalability for high-throughput diagnostics and manufacturing quality control.

In parallel, computer vision (CV) and deep learning have transformed automated visual analysis across materials science and industrial inspection. Prior work has demonstrated the potential of CV-based approaches for microstructural characterization and defect analysis DeCost & Holm (2015); Holm et al. (2020), while deep learning based defect detection systems have shown strong performance in manufacturing contexts, particularly when combined with transfer learning Ferguson et al. (2018); Mei et al. (2018). To address limited labeled data, recent studies have explored few-shot learning, synthetic data augmentation, and multi-scale feature representations Wu et al. (2020); Jain et al. (2022); Song et al. (2023); Wang & Hong (2025); Yang et al. (2019); Lin et al. (2022); Rudolph et al. (2022).

Despite this progress, existing surveys remain largely domain agnostic or focused on surface-level industrial inspection. Reviews of transfer learning, anomaly detection, and explainable artificial intelligence (XAI) primarily target time-series data or generic manufacturing systems rather than imaging physics–driven microstructures Yan et al. (2024); Jin & Chen (2022); Hoffmann & Reich (2023); Das & Rad (2020); Senoner et al. (2022). Although explainable and Bayesian deep learning frameworks have been proposed for industrial quality assurance Sinha et al. (2021); Aboulhosn et al. (2024); Lin et al. (2025), their relevance to PEMFC microstructural degradation remains largely unexplored.

This survey addresses this gap by providing a detection-centric review of computer vision methods for PEM fuel cell microstructural defect analysis. By bridging fuel cell physics, advanced imaging, and modern deep learning, we identify current limitations and outline practical research directions toward scalable, interpretable, and industrially deployable fuel cell diagnostics. A basic outline of this survey paper is explained in Figure 1

2 SURVEY METHODOLOGY

This survey follows a structured and reproducible literature review methodology to ensure comprehensive coverage of computer vision–based approaches for fuel cell microstructural defect analysis. The review process was designed to balance breadth across interdisciplinary domains with sufficient depth for methodological comparison. The methodology consists of four sequential phases: literature identification, screening and selection, standardized data extraction, and thematic synthesis.

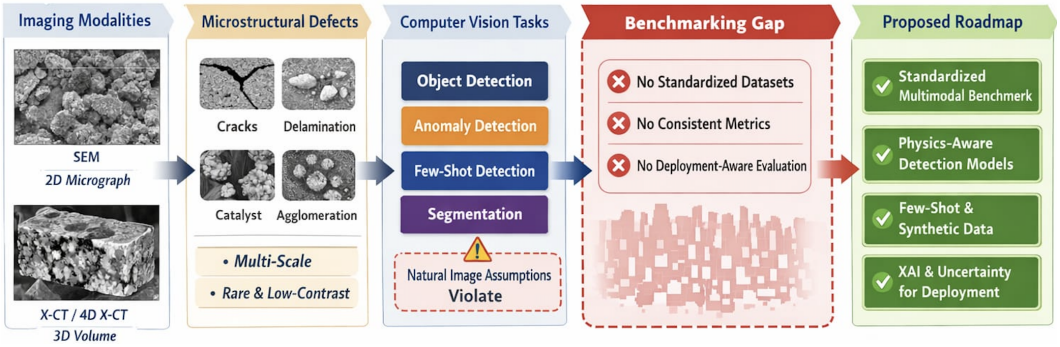

Figure 1: Overview of fuel cell defect detection, linking imaging modalities and microstructural defects to computer vision tasks, key benchmarking gaps, and a roadmap toward standardized, physics-aware, and deployable inspection pipelines.

A systematic multi-database search was conducted covering publications from January 2000 to June 2025, with particular emphasis on work published after 2015 to capture the rapid emergence of deep learning– based methods. Searches were performed across IEEE Xplore, Scopus, Web of Science, and Google Scholar to ensure broad disciplinary coverage spanning fuel cell engineering, materials science, and computer vision. Search queries were constructed using structured Boolean combinations across four conceptual dimensions: (i) fuel cell technologies (e.g., PEMFC, SOFC, PCFC), (ii) imaging modalities (e.g., SEM, X-ray computed tomography), (iii) microstructural defects and degradation mechanisms, and (iv) computer vision and machine learning methods (e.g., object detection, segmentation, anomaly detection). To reduce omission bias, secondary search strategies included backward reference tracing, citation network exploration, and review of relevant conference proceedings in both computer vision and materials-focused venues.

A two-stage screening process was employed consisting of title and abstract screening followed by full-text review. Studies were included if they satisfied the following criteria: (1) direct analysis of fuel cell materials or components, with emphasis on microstructural defects or degradation; (2) application of computer vision or machine learning techniques for defect localization, detection, or characterization; and (3) sufficient methodological detail to enable comparative assessment. Peer-reviewed journal articles, conference papers, and technically validated studies were considered. Studies were excluded if they focused exclusively on system-level diagnostics or control without imaging-based microstructural analysis, relied solely on physics-based modeling without data-driven components, or lacked adequate methodological transparency. Duplicate publications and preliminary versions of subsequently published work were also excluded. Screening was conducted independently by the authors to ensure consistent application of the predefined criteria.

For each included study, a standardized data extraction framework was applied to capture key technical and practical attributes. Extracted information included imaging modality, defect type and scale, dataset availability and annotation practices, computer vision architecture, training strategies, evaluation metrics, and reported deployment considerations. These attributes were selected to support both methodological comparison and assessment of industrial relevance across diverse fuel cell technologies and imaging conditions.

Following extraction, a structured thematic synthesis was performed to identify dominant methodological trends, recurring limitations, and underexplored research directions. The synthesis emphasizes detection- centric approaches, data scarcity challenges, multi-scale defect representation, and the gap between laboratory demonstrations and deployable inspection pipelines. Comparative analysis further examines methodological evolution over time, inter-domain transfer of computer vision techniques, and practical constraints affecting industrial adoption.

The systematic review process yielded 129 unique records across all search sources. After title and abstract screening, 86 publications were retained for full-text assessment. Application of the inclusion and exclusion criteria resulted in 40 studies undergoing detailed evaluation, with 24 publications forming the final systematic review corpus. The excluded publications mainly failed to meet the methodological specific requirements or focused exclusively on system-level monitoring without

microstructural defect analysis. The included publications represent the comprehensive field of study of computer vision applications for fuel cell defect detection, providing sufficient coverage for robust trend analysis and methodological assessment across the diverse research domain. The systematic review process is summarized in the Table 1

Table 1: Systematic review inclusion statistics.

Category	**Initial**	**Screened**	**Full-Text**	**Included**
Computer vision for PEMFC defects	50	29	10	7
Imaging modalities and microstructures	30	21	8	5
Industrial inspection applications	20	16	11	6
Cross-domain methodological studies	29	20	11	6
Total (Unique)	**129**	**86**	**40**	**24**

3 Imaging Modalities and Detection Implications

SEM and X-ray computed tomography (X-CT) dominate fuel cell defect visualization due to their complementary spatial and structural capabilities. SEM provides nanometer scale surface resolution essential for analyzing catalyst layers, membrane interfaces, and agglomeration phenomena, but is inherently limited by small field-of-view, charging artifacts, and preparation induced variability Rong et al. (2008a;b); DeCost & Holm (2015). These constraints frequently produce low-contrast, texture-dominated images in which defect boundaries are weakly defined, confounding conventional convolutional feature extractors Holm et al. (2020).

X-CT enables non-destructive three-dimensional imaging of internal defects such as pore blockage, delamination, and crack networks, providing volumetric context inaccessible to surface microscopy. Nanometer-scale X-CT systems have achieved resolutions below 100 nm for PEM fuel cell electrodes Epting et al. (2012), while in situ and operando 4D X-CT allows direct observation of crack propagation and catalyst layer thinning during accelerated degradation White et al. (2017). From a detection perspective, however, X-CT introduces challenges related to extreme scale heterogeneity, high dimensional volumetric data, and resolution–field-of-view trade-offs, which complicate both annotation and model training Holm et al. (2020); Jin & Chen (2022).

These modality-specific characteristics directly shape detection algorithm requirements. Effective models must operate under severe foreground–background imbalance, detect low contrast and morphologically ambiguous anomalies, and resolve multi-scale defects spanning orders of magnitude conditions under which standard industrial inspection pipelines and generic defect detectors often fail Mei et al. (2018); Yang et al. (2019); Lin et al. (2022); Rudolph et al. (2022).

4 Object Detection for Fuel Cell Defects

Object detection provides a natural framework for automated fuel cell inspection by jointly localizing defects and assigning semantic labels. Most existing studies rely on off-the-shelf detectors adapted from natural-image benchmarks or industrial surface inspection pipelines Ferguson et al. (2018); Mei et al. (2018); Holm et al. (2020). In broader industrial contexts, transfer learning from large-scale datasets such as ImageNet or COCO has proven effective for defect detection under limited data availability Ferguson et al. (2018); Wu et al. (2020). However, fuel cell microstructures violate key assumptions underlying natural-image detection: defects frequently occupy less than 1% of the image area, exhibit weak contrast, and lack consistent shape priors DeCost & Holm (2015); Jin & Chen (2022).

Recent advances in defect detection for textured and low-contrast surfaces emphasize the importance of multi-scale modeling and background aware representations. Cross-scale normalizing flows explicitly capture unlikely feature distributions across scales Rudolph et al. (2022), while multiscale autoencoder-based methods reconstruct complex textured backgrounds to isolate anomalous regions Yang et al. (2019). Attention-driven architectures such as EMRA-Net further demonstrate strong performance for tiny, low-contrast defects by fusing local edge cues with global contextual features Lin et al. (2022).

Despite these methodological advances, evaluation practices remain inconsistent. Mean average precision (mAP) is frequently reported without standardized intersection-over-union thresholds, dataset splits, or latency measurements, limiting reproducibility, cross-study comparison, and deployment relevance Hoffmann & Reich (2023); Jin & Chen (2022).

5 Structural Gaps and Future Directions

The dominant barrier to progress in fuel cell defect detection is not algorithmic capability but the absence of standardized benchmarking infrastructure. Existing datasets remain fragmented, predominantly proprietary, and inconsistently annotated, preventing reproducible evaluation, cross-study comparison, and cumulative methodological progress Holm et al. (2020); Jin & Chen (2022); Hoffmann & Reich (2023).

Severe data scarcity further motivates few-shot and transfer learning approaches. Meta learning and few-shot transfer frameworks have demonstrated strong performance in industrial fault diagnosis under limited labeled samples Wu et al. (2020) and are increasingly recognized as critical enablers for defect detection in data constrained environments Song et al. (2023); Wang & Hong (2025). Synthetic data generation, particularly using generative adversarial networks, offers a complementary strategy for expanding training distributions and improving generalization in surface and microstructural defect detection tasks Jain et al. (2022).

Explainable artificial intelligence (XAI) is essential for industrial deployment, enabling trust, regulatory acceptance, and root cause analysis. Recent surveys highlight the limited adoption of XAI in manufacturing quality assurance despite its demonstrated potential Hoffmann & Reich (2023); Das & Rad (2020). Techniques such as Grad-CAM, SHAP, and Bayesian interpretability frameworks provide practical pathways toward transparent and uncertainty-aware inspection systems Sinha et al. (2021); Aboulhosn et al. (2024); Senoner et al. (2022).

6 Conclusion

This survey demonstrates that computer vision–based fuel cell defect detection is constrained less by architectural limitations than by structural deficiencies in data availability, benchmarking practice, and deployment integration. While modern detection architectures provide a strong foundation, their impact remains limited without coordinated community efforts toward standardized datasets, physics-aware modeling, and industrial relevance. By reframing fuel cell inspection as a benchmark-driven, detection-centric problem, the field can transition from fragmented laboratory demonstrations to scalable quality assurance systems. The lessons identified here extend beyond fuel cells, offering broader insights for applying computer vision in data-poor, scientifically complex domains central to the clean energy transition.

References

Zeina Aboulhosn, Ahmad Musamih, Khaled Salah, Raja Jayaraman, Mohammed Omar, and Zeyar Aung. Detection of manufacturing defects in steel using deep learning with explainable artificial intelligence. *IEEE Access*, 2024.

Arun Das and Paul Rad. Opportunities and challenges in explainable artificial intelligence (xai): A survey. *arXiv preprint arXiv:2006.11371*, 2020.

Brian L. DeCost and Elizabeth A. Holm. A computer vision approach for automated analysis and classification of microstructural image data. *Computational Materials Science*, 110:126–133, 2015.

William K. Epting, Jeff Gelb, and Shawn Litster. Resolving the three-dimensional microstructure of polymer electrolyte fuel cell electrodes using nanometer-scale x-ray computed tomography. *Advanced Functional Materials*, 22(3):555–560, 2012.

Max Ferguson, Ronay Ak, Yung-Tsun Tina Lee, and Kincho H. Law. Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning. *Smart and Sustainable Manufacturing Systems*, 2(1):137–164, 2018.

Rudolf Hoffmann and Christoph Reich. A systematic literature review on artificial intelligence and explainable artificial intelligence for visual quality assurance in manufacturing. *Electronics*, 12 (22):4572, 2023.

Elizabeth A. Holm, Ryan Cohn, Nan Gao, Andrew R. Kitahara, Thomas P. Matson, Bo Lei, and Srujana Rao Yarasi. Overview: Computer vision and machine learning for microstructural characterization and analysis. *Metallurgical and Materials Transactions A*, 51(12):5985–5999, 2020.

Saksham Jain, Gautam Seth, Arpit Paruthi, Umang Soni, and Girish Kumar. Synthetic data augmentation for surface defect detection and classification using deep learning. *Journal of Intelligent Manufacturing*, 33(4):1007–1020, 2022.

Qifan Jin and Li Chen. A survey of surface defect detection of industrial products based on a small number of labeled data. *arXiv preprint arXiv:2203.05733*, 2022.

Chung-Yin Lin, Jinsu Gim, Demitri Shotwell, Mong-Tung Lin, Jia-Hau Liu, and Lih-Sheng Turng. Explainable artificial intelligence and multi-stage transfer learning for injection molding quality prediction. *Journal of Intelligent Manufacturing*, 36(5):3587–3606, 2025.

Qiangqiang Lin, Jinzhu Zhou, Qiurui Ma, Yongji Ma, Le Kang, and Jianjun Wang. Emra-net: A pixel-wise network fusing local and global features for tiny and low-contrast surface defect detection. *IEEE Transactions on Instrumentation and Measurement*, 71:1–14, 2022.

R. Lin, B. Li, Y. P. Hou, and J. M. Ma. Investigation of dynamic driving cycle effect on performance degradation and micro-structure change of pem fuel cell. *International Journal of Hydrogen Energy*, 34(5):2369–2376, 2009.

Shuang Mei, Hua Yang, and Zhouping Yin. An unsupervised-learning-based approach for automated defect inspection on textured surfaces. *IEEE Transactions on Instrumentation and Measurement*, 67(6):1266–1277, 2018.

Feng Rong, Cheng Huang, Zhong-Sheng Liu, Datong Song, and Qianpu Wang. Microstructure changes in the catalyst layers of pem fuel cells induced by load cycling: Part i. mechanical model. *Journal of Power Sources*, 175(2):699–711, 2008a.

Feng Rong, Cheng Huang, Zhong-Sheng Liu, Datong Song, and Qianpu Wang. Microstructure changes in the catalyst layers of pem fuel cells induced by load cycling: Part ii. simulation and understanding. *Journal of Power Sources*, 175(2):712–723, 2008b.

Marco Rudolph, Tom Wehrbein, Bodo Rosenhahn, and Bastian Wandt. Fully convolutional cross-scale-flows for image-based defect detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 1088–1097, 2022.

Julian Senoner, Torbjørn Netland, and Stefan Feuerriegel. Using explainable artificial intelligence to improve process quality: Evidence from semiconductor manufacturing. *Management Science*, 68 (8):5704–5723, 2022.

Sumit Sinha, Pasquale Franciosa, and Dariusz Ceglarek. Building a scalable and interpretable bayesian deep learning framework for quality control of free form surfaces. *IEEE Access*, 9: 50188–50208, 2021.

Yisheng Song, Ting Wang, Puyu Cai, Subrota K. Mondal, and Jyoti Prakash Sahoo. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities. *ACM Computing Surveys*, 55(13s):1–40, 2023.

Ruibo Wang and Tao Hong. Few-shot defect detection in industrial scenarios: A comprehensive review of challenges, advances, and frontier trends. In *MATEC Web of Conferences*, volume 413, pp. 04005. EDP Sciences, 2025.

Robin T. White, Alex Wu, Marina Najm, Francesco P. Orfino, Monica Dutta, and Erik Kjeang. 4d in situ visualization of electrode morphology changes during accelerated degradation in fuel cells by x-ray computed tomography. *Journal of Power Sources*, 350:94–102, 2017.

Jingyao Wu, Zhibin Zhao, Chuang Sun, Ruqiang Yan, and Xuefeng Chen. Few-shot transfer learning for intelligent fault diagnosis of machine. *Measurement*, 166:108202, 2020.

Peng Yan, Ahmed Abdulkadir, Paul-Philipp Luley, Matthias Rosenthal, Gerrit A. Schatte, Benjamin F. Grewe, and Thilo Stadelmann. A comprehensive survey of deep transfer learning for anomaly detection in industrial time series: Methods, applications, and directions. *IEEE Access*, 12: 3768–3789, 2024.

Hua Yang, Yifan Chen, Kaiyou Song, and Zhouping Yin. Multiscale feature-clustering-based fully convolutional autoencoder for fast accurate visual inspection of texture surface defects. *IEEE Transactions on Automation Science and Engineering*, 16(3):1450–1467, 2019.