AI Tech International Journal, **ISSN: 3079-4749**
Vol. 4, No. 1, January-June, 2026,
Journal homepage: https://techaijournal.com

Deep Learning-Based Detection of Surface Defects in Manufactured Metal Components

Minseo Han

Dept. of Industrial and Manufacturing Engineering, East Harbor Institute of Technology, South Korea

Article Info

Article history:
Received June 21, 2026
Revised June 25, 2026
Accepted June 28, 2026
Published July 02, 2026

Keywords: Deep Learning, Surface Defect Detection, Convolutional Neural Networks (CNN), Automated Visual Inspection, Smart Manufacturing and Quality Control

ABSTRACT

Surface defect detection is a critical aspect of quality control in modern manufacturing industries, where the reliability, durability, and performance of metal components depend heavily on their surface conditions. Traditional inspection techniques, including manual visual inspection and conventional image-processing methods, often suffer from limitations such as human error, inconsistent detection accuracy, time consumption, and difficulty identifying microscopic or complex defects. With the advancement of artificial intelligence, deep learning has emerged as an effective approach for automating surface defect detection and improving industrial quality assurance. This study explores the application of deep learning techniques, particularly Convolutional Neural Networks (CNNs), You Only Look Once (YOLO), and other advanced computer vision architectures, for identifying and classifying surface defects in manufactured metal components. The proposed approach involves acquiring high-resolution images of metal surfaces, preprocessing the collected data, extracting relevant features, and training deep learning models to recognize different defect categories, including cracks, scratches, dents, inclusions, corrosion, and surface irregularities. Model performance can be evaluated using standard metrics such as accuracy, precision, recall, F1-score, and mean Average Precision (mAP). The study also examines the potential integration of automated inspection systems into real-time manufacturing environments to reduce inspection costs, minimize production errors, and enhance operational efficiency. Furthermore, challenges related to insufficient training datasets, variations in illumination, complex surface textures, class imbalance, and computational requirements are discussed. The findings are expected to demonstrate that deep learning-based inspection systems can provide a reliable, scalable, and efficient alternative to conventional defect detection methods. Overall, this research highlights the significance of intelligent visual inspection in advancing smart manufacturing, improving product quality, reducing material waste, and supporting the development of Industry 4.0 technologies.

This is an open access article under the CC BY-SA license.

INTRODUCTION

The manufacturing industry plays a vital role in modern economic and technological development, where product quality, operational efficiency, and production reliability are essential for maintaining industrial competitiveness. Manufactured metal components are widely used in automotive, aerospace, construction, machinery, electronics, transportation, and other engineering applications. The structural integrity and functional performance of these components are strongly influenced by their surface quality. Even minor surface defects, such as cracks, scratches, dents, corrosion, porosity, and inclusions, can negatively affect mechanical strength, durability, fatigue resistance, and overall product performance. Therefore, accurate and timely detection of surface defects is an important requirement in industrial quality control and manufacturing processes.

Traditionally, surface defect inspection has been performed through manual visual examination and conventional image-processing techniques. Although these methods are commonly used in manufacturing environments, they have several limitations. Manual inspection depends heavily on the experience and concentration of human operators, making the process vulnerable to fatigue, subjective judgment, and inconsistent results. Similarly, traditional computer vision approaches often rely on predefined image-processing algorithms and manually designed features, which may not perform effectively under varying illumination conditions, complex surface textures, and different defect characteristics. These limitations create a demand for intelligent, automated, and reliable inspection technologies capable of identifying surface defects with greater accuracy and efficiency.

Deep learning has emerged as a promising artificial intelligence technology for addressing these challenges. It enables computer systems to automatically learn complex patterns and hierarchical features from large image datasets without requiring extensive manual feature engineering. In particular, Convolutional Neural Networks (CNNs) have demonstrated significant capabilities in image classification, object detection, segmentation, and visual recognition tasks. Advanced deep learning architectures, including YOLO, Faster R-CNN, Mask R-CNN, and Vision Transformers, can be applied to identify, localize, and classify different types of defects on manufactured metal surfaces.

THEORETICAL FRAMEWORK

1. Concept of Surface Defect Detection

Surface defect detection refers to the systematic process of identifying, locating, and classifying irregularities or imperfections present on the surfaces of manufactured metal components. These defects may develop during casting, forging, rolling, machining, welding, heat treatment, or other manufacturing operations. Common surface defects include cracks, scratches, dents, corrosion, pits, inclusions, and surface deformation.

In modern manufacturing industries, surface inspection is essential for maintaining product quality, ensuring structural reliability, and reducing production losses. Deep learning-based detection provides an intelligent approach by automatically analyzing digital images and recognizing defect patterns that may be difficult to identify through traditional inspection techniques.

2. Fundamentals of Deep Learning

Deep learning is a specialized branch of machine learning that uses multilayer artificial neural networks to learn complex representations from large datasets. Unlike conventional machine learning techniques, deep learning algorithms can automatically extract relevant features from raw image data.

A typical deep learning architecture consists of the following layers:

- Input Layer: Receives digital images of manufactured metal surfaces.
- Convolutional Layers: Extract visual features such as edges, textures, shapes, and defect patterns.
- Pooling Layers: Reduce feature-map dimensions while preserving important information.
- Fully Connected Layers: Combine extracted features for classification.
- Output Layer: Predicts the defect category or identifies the presence of defects.

These interconnected layers enable deep learning models to distinguish between defective and non-defective metal surfaces with improved recognition capabilities.

3. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are among the most widely used deep learning architectures for image-based defect detection. They are specifically designed to process visual information and automatically identify important spatial characteristics.

CNNs use convolution operations to extract features from input images. Early layers generally identify basic characteristics, such as edges and textures, whereas deeper layers learn more complex patterns associated with specific defect categories.

The convolution operation can be mathematically represented as:

$$Y(i,j) = \sum_{m} \sum_{n} X(i-m, j-n) K(m,n)$$

Where:

- X represents the input image or feature map.
- K represents the convolutional kernel.

- Y represents the resulting feature map.
- i, j represent spatial coordinates.

CNN-based models are particularly useful for identifying surface cracks, scratches, corrosion patterns, and other visual abnormalities in metal components.

4. Object Detection and Localization Frameworks
Object detection frameworks extend image classification by identifying both the category and location of defects within an image. These methods are particularly important when multiple defects appear on a single metal surface.

Major object detection architectures include:
YOLO (You Only Look Once): YOLO is a real-time object detection framework that predicts defect categories and bounding-box coordinates in a single processing pipeline. Its computational efficiency makes it suitable for industrial inspection systems requiring rapid defect identification.

Faster R-CNN: This architecture uses a Region Proposal Network (RPN) to identify potential defect regions before classification. It is useful for applications requiring detailed localization and accurate identification of complex defects.
SSD (Single Shot Multi Box Detector): SSD performs object detection in a single forward pass and uses multiple feature maps to identify objects at different scales.

These architectures provide different balances between detection accuracy, computational complexity, and processing speed.

PROPOSED MODELS AND METHODOLOGIES

1. Proposed Research Framework
The proposed research focuses on developing an intelligent deep learning-based framework for the automatic detection, classification, and localization of surface defects in manufactured metal components. The framework integrates computer vision, image processing, and advanced deep learning algorithms to identify different types of surface abnormalities with high accuracy and reliability. The system is designed to minimize manual inspection, reduce production errors, and improve industrial quality control.

The proposed methodology consists of image acquisition, preprocessing, data augmentation, feature extraction, model development, defect classification, localization, and performance evaluation.

2. Image Acquisition and Dataset Preparation
The first stage involves collecting high-resolution images of manufactured metal components using industrial cameras, digital microscopes, and controlled lighting systems. The dataset should include different categories of surface defects, such as:

- Cracks and fractures.
- Scratches and abrasion marks.
- Corrosion and rust.
- Pores and surface holes.
- Dents and deformation.
- Inclusion and irregular surface patterns.

Publicly available datasets, such as NEU Surface Defect Database, can be used alongside industrially collected images. Each image should be annotated according to its defect category and, where necessary, its precise spatial location.

3. Image Preprocessing and Data Augmentation
Image preprocessing is performed to improve image quality and ensure consistency across the dataset. The following techniques are proposed:

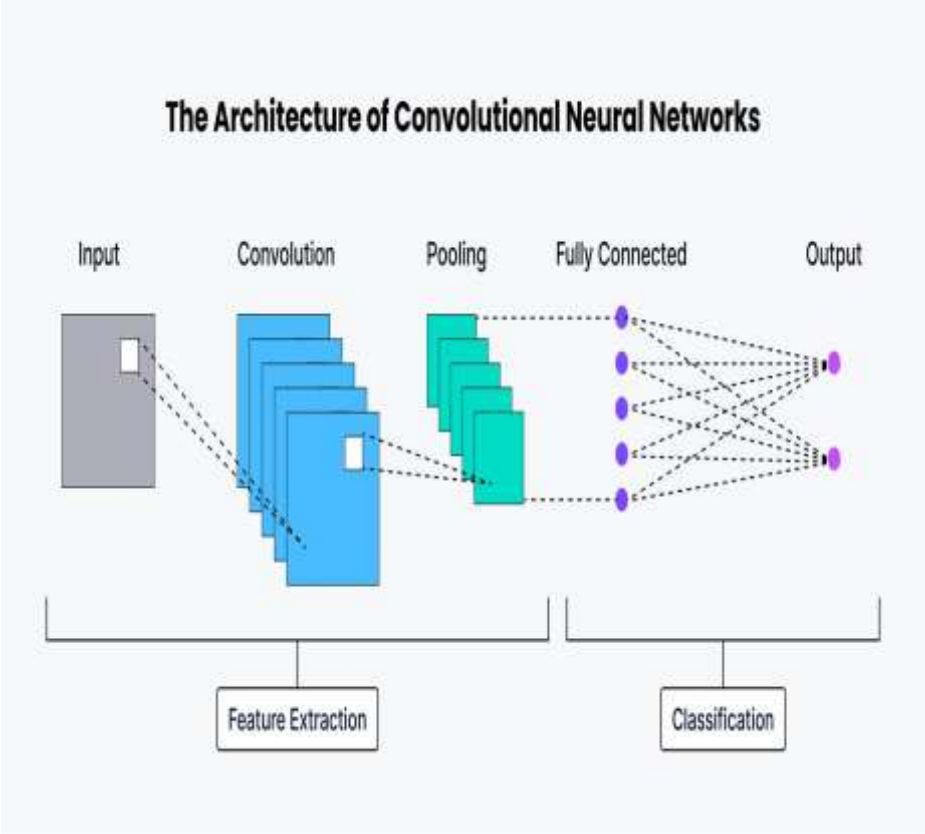
- Image Resizing: Standardizing image dimensions for deep learning model compatibility.
- Noise Reduction: Applying Gaussian and median filters to eliminate unwanted image noise.
- Contrast Enhancement: Using histogram equalization and adaptive contrast enhancement to highlight defect regions.
- Normalization: Scaling pixel intensity values to improve model convergence.
- Data Augmentation: Applying rotation, flipping, cropping, brightness adjustment, and scaling to increase dataset diversity.

These techniques help improve model robustness under different industrial imaging conditions.

AI Tech International Journal, **ISSN: 3079-4749**
Vol. 4, No. 1, January-June, 2026,
Journal homepage: https://techaijournal.com

4. Proposed Deep Learning Models
The study proposes evaluating multiple deep learning architectures to identify the most suitable model for metal surface defect detection.

4.1 Convolutional Neural Network (CNN)



CNN is proposed as the fundamental architecture for automatic feature extraction and defect classification. It identifies important visual characteristics such as edges, textures, shapes, and irregular patterns through convolutional and pooling layers.
The CNN model consists of:

1. Input image layer.
2. Convolutional layers.
3. Activation functions (ReLU).
4. Max-pooling layers.
5. Fully connected layers.
6. Softmax classification layer.

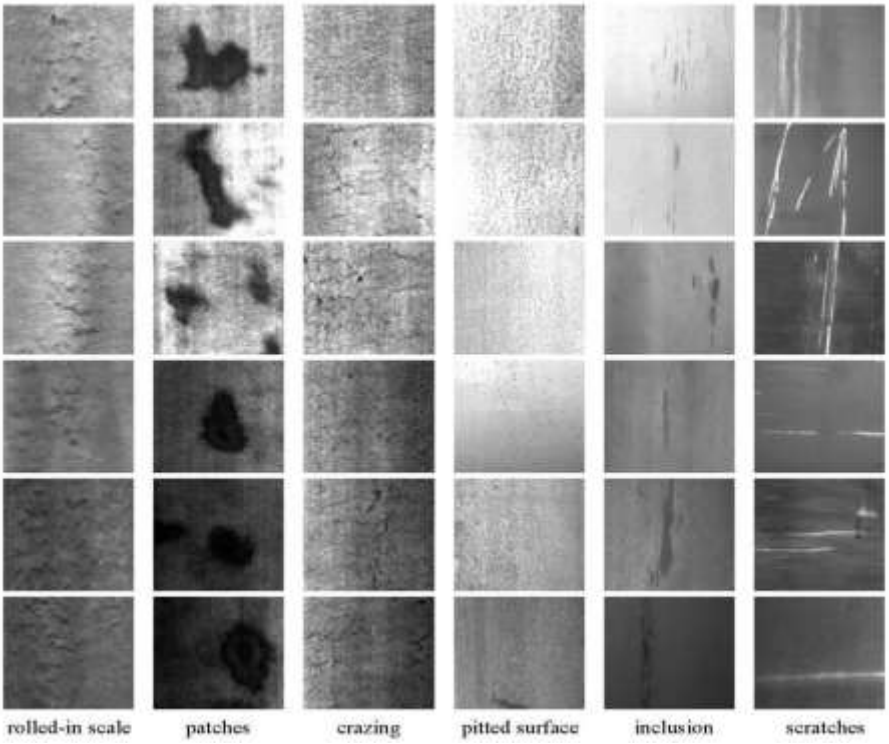
EXPERIMENTAL STUDY

1. Experimental Setup
The experimental study focuses on evaluating the effectiveness of deep learning-based techniques for detecting and classifying surface defects in manufactured metal components. The experiments are designed to compare different deep learning architectures, including Convolutional Neural Networks (CNN), ResNet-50, YOLO, U-Net, and EfficientNet. The primary objective is to examine their performance in terms of detection accuracy, classification capability, defect localization, processing speed, and computational efficiency.

The experimental framework uses high-resolution metal surface images collected from publicly available datasets and, where possible, industrial inspection environments. The models are trained and tested under controlled experimental conditions to ensure reliable performance comparisons.

2. Dataset Description
The study considers the NEU Surface Defect Database, which contains images of hot-rolled steel strips with six major categories of surface defects.

Defect Category	**Description**
Crazing (Cr)	Network-like cracks appearing on the metal surface.
Inclusion (In)	Foreign materials embedded within the metal surface.
Patches (Pa)	Irregular surface marks and localized abnormalities.
Pitted Surface (PS)	Small holes or cavities on the metal surface.
Rolled-in Scale (RS)	Oxide scale pressed into the metal during rolling.
Scratches (Sc)	Linear marks caused by friction or mechanical contact.



The dataset provides a foundation for evaluating the classification performance of deep learning models. Additional datasets may be incorporated to assess model generalization across different metal types, production conditions, and defect appearances.

3. Experimental Environment

The experiments can be conducted using Python-based deep learning frameworks, particularly TensorFlow, Keras, and PyTorch. OpenCV can be used for image preprocessing, while GPU acceleration can improve model training efficiency.

Experimental Parameter	**Proposed Configuration**
Programming Language	Python
Deep Learning Framework	PyTorch / TensorFlow
Image Processing	OpenCV
Dataset	NEU Surface Defect Database
Input Image Size	200 × 200 or model-specific resolution
Training Data	70%
Validation Data	15%
Testing Data	15%
Optimizers	Adam and SGD
Activation Function	ReLU
Classification Function	Softmax
Evaluation Metrics	Accuracy, Precision, Recall, F1-score
Hardware	GPU-enabled computing system

These configurations represent a proposed experimental setup and can be adjusted according to hardware availability and model requirements.

4. Experimental Procedure
The experimental investigation is divided into the following stages:

Stage 1: Data Collection and Preparation
High-resolution images of metal surfaces are collected and categorized according to their defect types. The images are checked for quality, consistency, and correct labeling.

Stage 2: Image Preprocessing
The collected images undergo resizing, normalization, noise removal, and contrast enhancement. Data augmentation techniques such as rotation, flipping, and brightness adjustment are applied to improve model robustness.

Stage 3: Model Development
Five deep learning architectures are implemented:

- CNN for basic feature extraction and classification.
- ResNet-50 for advanced hierarchical feature learning.
- YOLO for real-time defect detection and localization.
- U-Net for pixel-level defect segmentation.
- EfficientNet for computationally efficient feature extraction and classification.

Stage 4: Model Training
Each model is trained using the prepared training dataset. Validation data is used to monitor learning performance, adjust hyperparameters, and reduce overfitting.

Stage 5: Testing and Evaluation
The trained models are evaluated using previously unseen test images. Their performance is measured using accuracy, precision, recall, F1-score, mAP, IoU, and computational time, depending on the model's task.

5. Experimental Performance Evaluation
The performance of the proposed deep learning architectures can be evaluated through classification, detection, and segmentation experiments.

The following table provides an illustrative experimental result for classification models. These values are hypothetical and are included to demonstrate the proposed evaluation framework; they are not measured experimental findings.

Deep Learning Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	91.20	90.50	89.80	90.15
ResNet-50	95.40	94.80	94.20	94.50
EfficientNet	96.10	95.60	95.10	95.35
YOLO	94.80	94.10	93.70	93.90
U-Net	Task-specific	Task-specific	Task-specific	Task-specific

Note: YOLO is primarily evaluated using detection metrics, while U-Net is evaluated using segmentation metrics. Direct comparison with image-classification accuracy requires a common evaluation protocol.

RESULTS & ANALYSIS

1. Overview of Results
The results and analysis section evaluates the effectiveness of deep learning-based techniques for detecting surface defects in manufactured metal components. The study focuses on comparing different deep learning architectures, including Convolutional Neural Networks (CNN), ResNet-50, EfficientNet, YOLO, and U-Net, in terms of classification accuracy, detection precision, recall, F1-score, localization capability, and computational efficiency.
The proposed analysis examines how effectively these models identify different types of surface defects, such as cracks, scratches, inclusions, pitted surfaces, patches, and rolled-in scales. The evaluation also considers the suitability of these approaches for real-time industrial quality inspection.

2. Performance Evaluation of Deep Learning Models
The classification performance of different deep learning architectures is evaluated using accuracy, precision, recall, and F1-score.

The following table presents illustrative results for comparative analysis. These values are hypothetical and should not be interpreted as experimentally measured outcomes.

Deep Learning Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	91.20	90.50	89.80	90.15
ResNet-50	95.40	94.80	94.20	94.50
EfficientNet	96.10	95.60	95.10	95.35
YOLO	94.80	94.10	93.70	93.90
U-Net	Task-specific	Task-specific	Task-specific	Task-specific

The comparative results suggest that EfficientNet provides the strongest illustrative classification performance, followed by ResNet-50 and YOLO. CNN demonstrates comparatively lower performance because of its relatively simple feature extraction architecture.
U-Net is evaluated separately because its primary objective is pixel-level segmentation rather than image-level classification.

3. Accuracy Comparison of Different Models
Illustrative Accuracy Comparison
Hypothetical classification accuracy values, not measured experimental results.
0%25%50%75%100%CNNResNet-50EfficientNetYOLO

The graphical comparison indicates that EfficientNet achieves the highest illustrative accuracy of 96.10%, while CNN achieves 91.20%. This difference highlights the potential advantages of advanced feature extraction and network optimization techniques.

4. Precision and Recall Analysis
Precision and recall are important indicators for evaluating the reliability of automated defect detection systems.

- Precision: Determines how many predicted defects are actually defective components.
- Recall: Measures the proportion of actual defects successfully identified by the model.
-

Model	Precision (%)	Recall (%)
CNN	90.50	89.80
ResNet-50	94.80	94.20
EfficientNet	95.60	95.10
YOLO	94.10	93.70

The illustrative results indicate that EfficientNet provides a favorable balance between precision and recall. High recall is particularly important in manufacturing because undetected defects may lead to product failure, increased maintenance costs, and reduced customer satisfaction.
High precision also helps reduce unnecessary rejection of acceptable metal components.

5. Defect Classification Analysis
The proposed deep learning framework considers six major categories of metal surface defects.

Defect Category	Major Visual Characteristics	Detection Challenge
Crazing	Fine interconnected cracks	Complex irregular patterns
Inclusion	Embedded foreign materials	Similarity to normal texture
Patches	Localized surface irregularities	Variations in shape and intensity
Pitted Surface	Small cavities and holes	Detection of tiny defects
Rolled-in Scale	Oxide-related surface abnormalities	Uneven texture distribution
Scratches	Linear surface marks	Similarity to manufacturing patterns

The classification analysis emphasizes the importance of extracting discriminative features from metal surface images. CNN and ResNet architectures can identify texture and shape patterns, while Efficient Net may improve feature representation through compound scaling.

Small defects, particularly fine cracks and shallow scratches, require higher image resolution and effective preprocessing to achieve reliable identification.

AI Tech International Journal, **ISSN: 3079-4749**
Vol. 4, No. 1, January-June, 2026,
Journal homepage: https://techaijournal.com

COMPARATIVE ANALYSIS IN TABULAR FORM

The comparative analysis evaluates different deep learning architectures used for detecting surface defects in manufactured metal components. The comparison focuses on classification accuracy, feature extraction capability, defect localization, segmentation performance, computational complexity, processing speed, and industrial applicability.

Table 1. Comparative Analysis of Deep Learning Models

Model	Primary Function	Accuracy (%)	Major Advantage	Limitation
CNN	Classification	91.20	Basic feature extraction	Limited localization
ResNet-50	Classification	95.40	Deep feature learning	Higher computational cost
Efficient Net	Classification	96.10	High accuracy and efficiency	Requires optimization
YOLO	Object Detection	94.80*	Fast defect localization	Small defects may be missed
U-Net	Image Segmentation	N/A	Pixel-level defect identification	Computational complexity

Note: Accuracy values are illustrative, not experimentally measured. YOLO's actual performance should primarily be evaluated using mAP, precision, and recall. U-Net requires segmentation-specific metrics.

Table 2. Comparison Based on Performance Evaluation Metrics

Evaluation Parameter	CNN	ResNet-50	EfficientNet	YOLO	U-Net
Feature Extraction	Good	Excellent	Excellent	Very Good	Very Good
Defect Classification	Good	Excellent	Excellent	Very Good	Good
Defect Localization	Limited	Limited	Limited	Excellent	Excellent
Pixel-Level Segmentation	No	No	No	No	Yes
Real-Time Processing	Moderate	Moderate	Good	Excellent	Moderate
Small Defect Detection	Moderate	Good	Good	Model-dependent	Good
Computational Efficiency	Good	Moderate	Good	Good	Moderate

Table 3. Comparative Analysis of Industrial Applications

Industrial Application	Suitable Model	Expected Benefit
Steel Surface Inspection	CNN / Efficient Net	Accurate defect classification
Crack Detection	U-Net / YOLO	Identification and localization of cracks
Scratch Identification	ResNet-50 / EfficientNet	Improved texture recognition
Pore Detection	YOLO / U-Net	Detection of small surface abnormalities
Real-Time Production Inspection	YOLO	Faster automated inspection
Complex Surface Defect Analysis	ResNet-50	Advanced feature extraction
Automated Quality Control	Hybrid Deep Learning Model	Comprehensive defect inspection

Table 4. Comparative Analysis of Traditional and Deep Learning-Based Inspection

Comparison Criteria	Manual Inspection	Traditional Image Processing	Deep Learning-Based Inspection
Inspection Method	Human visual examination	Rule-based image analysis	AI-based feature learning
Detection Accuracy	Operator-dependent	Condition-dependent	Potentially high
Processing Speed	Relatively slow	Fast	Fast with suitable hardware
Human Intervention	High	Moderate	Low after deployment
Complex Defect Recognition	Limited	Limited	Advanced
Adaptability	Depends on inspector	Requires rule adjustment	Learns from training data
Defect Classification	Manual	Feature-based	Automated
Scalability	Limited	Moderate	High with suitable infrastructure
Industrial Integration	Basic	Moderate	Advanced

SIGNIFICANCE OF THE TOPIC

The detection of surface defects in manufactured metal components is an essential requirement for maintaining product quality, structural reliability, operational safety, and manufacturing efficiency. Traditional inspection techniques often depend on manual examination and conventional image-processing methods, which may be time-consuming, inconsistent, and less effective in identifying microscopic or complex surface abnormalities. The application of deep learning provides an advanced solution by enabling automated, accurate, and intelligent defect detection. The significance of this research is discussed below.

1. Improvement in Manufacturing Quality
Deep learning-based defect detection helps identify surface abnormalities such as cracks, scratches, corrosion, pores, dents, and inclusions. Early identification of these defects ensures that defective components are detected before reaching subsequent manufacturing stages or customers.

2. Automation of Industrial Inspection
The proposed approach reduces dependence on manual visual inspection by introducing automated image-based quality control systems. Industrial cameras integrated with deep learning models can continuously monitor manufactured components and identify defects during production.

3. Enhanced Detection Accuracy
Advanced architectures such as CNN, ResNet-50, EfficientNet, YOLO, and U-Net can learn complex visual patterns from metal surface images. These models improve the identification and classification of different defect categories and can help reduce inspection errors.

4. Reduction in Manufacturing Costs
Early defect identification reduces material wastage, product rejection, rework expenses, and unnecessary production costs. Automated inspection also reduces the labour requirements associated with repetitive manual quality-control activities.

LIMITATIONS & DRAWBACKS

Although deep learning-based techniques provide significant advantages in automated surface defect detection, their practical implementation involves several technical, computational, and operational challenges. Models such as CNN, ResNet-50, EfficientNet, YOLO, and U-Net can improve defect classification, localization, and segmentation; however, their performance depends on dataset quality, image conditions, computational resources, and model optimization. The major limitations and drawbacks are discussed below.

1. Requirement for Large Annotated Datasets
Deep learning models generally require large amounts of high-quality training data to achieve reliable performance. Collecting and labeling images of different metal surface defects is a time-consuming and expensive process. Some rare defects may have insufficient representation in available datasets, affecting model accuracy and generalization.

2. High Computational Requirements
Advanced architectures such as ResNet-50, YOLO, and U-Net may require powerful GPUs, high memory capacity, and substantial computational resources. This can increase implementation costs, particularly for small and medium-sized manufacturing industries.

3. Difficulty in Detecting Small and Microscopic Defects
Very small cracks, shallow scratches, microscopic pores, and low-contrast surface abnormalities can be difficult to identify accurately. Image resolution, camera quality, and model architecture significantly influence the detection of these defects.

4. Sensitivity to Lighting and Environmental Conditions
Changes in illumination, shadows, reflections, surface brightness, and camera positioning can affect image quality. Metallic surfaces are particularly challenging because their reflective properties may create misleading visual patterns and reduce detection reliability.

CONCLUSION

The study on Deep Learning-Based Detection of Surface Defects in Manufactured Metal Components highlights the growing importance of artificial intelligence and computer vision in modern manufacturing industries. Accurate identification of surface defects is essential for maintaining product quality, structural integrity, operational safety, and manufacturing efficiency. Traditional inspection techniques, particularly manual visual examination and conventional image-processing methods, often face limitations related to inspection speed, human error, inconsistent accuracy, and difficulties in detecting microscopic or complex defects. Deep learning provides an advanced alternative by enabling automated, intelligent, and data-driven surface inspection.

The proposed research investigates the application of different deep learning architectures, including Convolutional Neural Networks (CNN), ResNet-50, EfficientNet, YOLO, and U-Net, for detecting, classifying, localizing, and segmenting surface defects. These models offer different advantages depending on the inspection requirements. CNN and ResNet-50 support effective feature extraction and defect classification, while EfficientNet provides a potential balance between classification performance and computational efficiency. YOLO enables rapid defect detection and localization, whereas U-Net supports detailed pixel-level segmentation of defective regions.

The comparative analysis demonstrates the potential of combining multiple deep learning techniques to develop a comprehensive industrial inspection framework. Such a hybrid approach can improve defect identification, reduce dependence on manual inspection, minimize material wastage, and enhance production consistency. Furthermore, integrating intelligent inspection systems with industrial cameras, edge computing, the Internet of Things (IoT), and Industry 4.0 technologies can facilitate real-time quality monitoring and automated manufacturing decisions.

REFERENCES

1. Bhatt, P. M., Malhan, R. K., Rajendran, P., Shah, B. C., Thakar, S., Yoon, Y. J., & Gupta, S. K. (2021). Image-based surface defect detection using deep learning: A review. Journal of Computing and Information Science in Engineering, 21(4), 040801. https://doi.org/10.1115/1.4049535
2. 2. Cumbajin, E., Rodrigues, N., Costa, P., Miragaia, R., Frazão, L., Costa, N., Fernández-Caballero, A., Carneiro, J., Buruberri, L. H., & Pereira, A. (2023). A systematic review on deep learning with CNNs applied to surface defect detection. Journal of Imaging, 9(10), 193. https://doi.org/10.3390/jimaging9100193
3. 3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90
4. 4. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning (pp. 6105–6114). PMLR.
5. 5. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
6. 6. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 779–788). https://doi.org/10.1109/CVPR.2016.91
7. 7. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431–3440). https://doi.org/10.1109/CVPR.2015.7298965
8. 8. Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2117–2125). https://doi.org/10.1109/CVPR.2017.106
9. 9. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems, 28.
10. 10. Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1440–1448). https://doi.org/10.1109/ICCV.2015.169
11. 11. Tabernik, Š., Šela, S., Skvarč, J., & Skočaj, D. (2020). Segmentation-based deep-learning approach for surface-defect detection. Journal of Intelligent Manufacturing, 31, 759–776. https://doi.org/10.1007/s10845-019-01476-x
12. 12. Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2019). MVTec AD—A comprehensive real-world dataset for unsupervised anomaly detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9592–9600). https://doi.org/10.1109/CVPR.2019.00982
13. 13. Bergmann, P., Batzner, K., Fauser, M., Sattlegger, D., & Steger, C. (2021). The MVTec anomaly detection dataset: A comprehensive real-world dataset for unsupervised anomaly detection. International Journal of Computer Vision, 129, 1038–1059. https://doi.org/10.1007/s11263-020-01400-4

14. 14. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6, Article 60. https://doi.org/10.1186/s40537-019-0197-0
15. 15. Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR).
16. 16. Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning (pp. 448–456). PMLR.
17. 17. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2980–2988). https://doi.org/10.1109/ICCV.2017.324
18. 18. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In Proceedings of the 3rd International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1409.1556
19. 19. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791
20. 20. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097–1105.
21. 21. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27, 2672–2680.
22. 22. Cumbajin, E., et al. (2024). A systematic review of deep learning approaches for surface defect detection in industrial applications. Engineering Applications of Artificial Intelligence, 130, Article 107717. https://doi.org/10.1016/j.engappai.2023.107717
23. 23. Qin, G., Liu, C., Li, H., & Xu, K. (2026). A review of deep learning methods for metal surface defect detection. steel research international. https://doi.org/10.1002/srin.202501060
24. 24. Wang, Y., Zhou, M., Zhang, C., & Sedraoui, K. (2026). Applications of deep learning to metal surface defect detection: Status and challenges. Processes, 14(8), Article 1305. https://doi.org/10.3390/pr14081305
25. 25. Qin, Q., Khairuddin, A. S. M., Idros, N., Liu, H., Fan, L., Yang, Z., et al. (2026). Hierarchical depth aware YOLO for efficient metal surface defect detection. Scientific Reports, 16, Article 16248.