International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 6, Nov-Dec-2026, ISSN (Online): 2395-566X

Artificial Intelligence in Cancer Detection: A Deep Learning-Based Approach for Early Diagnosis

Sakshi Yadav
Department Of Artificial Intelligence And Data Science Sahyadri Valley College Of Engineering And Technology Rajuri, Maharashtra, India
Email: Sakshiyadav6503@Gmail.Com

Abstract—Cancer diagnosis increasingly relies on imaging and pathology data, yet interpreting large volumes of medical images can be time-consuming and subject to variation between observers. Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), offers tools for assisting clinicians by identifying patterns that may be difficult to recognize consistently. This paper presents a literature-based technical review of AI applications in cancer detection and proposes a conceptual image-analysis framework. The framework includes data acquisition, image-quality assessment, preprocessing, region-of-interest analysis, feature learning, classification, and interpretable output. Convolutional Neural Networks (CNNs), transfer learning, traditional machine-learning classifiers, segmentation methods, and explainability techniques are discussed in relation to breast, lung, brain, skin, and other cancers. The paper highlights the importance of carefully separated training, validation, and test datasets; patient-level data splitting; class-imbalance handling; and evaluation using sensitivity, specificity, precision, F1-score, and Area Under the ROC Curve (AUC). It also examines dataset diversity, bias, privacy, interpretability, external validation, and clinical workflow integration as barriers to deployment. The framework is conceptual and does not report results from a newly trained model. It emphasizes that AI should be validated as a clinical decision-support tool, with final diagnosis and care decisions remaining under qualified medical professionals.

Keywords— Artificial Intelligence, Cancer Detection, Deep Learning, CNN, Medical Imaging, Transfer Learning, Explainable AI.

I. INTRODUCTION

Cancer comprises diseases in which abnormal cells grow uncontrollably and may invade surrounding tissues or spread to distant organs. Early and accurate detection can help clinicians select appropriate diagnostic tests and treatment pathways. However, screening and diagnosis often involve large volumes of imaging and pathology data, and interpretation may be affected by workload, image quality, clinical context, and observer variability. These factors motivate computer-assisted methods that support, rather than replace, clinical assessment.

AI methods can learn patterns from labelled examples and use them to identify suspicious regions, classify images, or prioritize cases for review. Traditional machine learning usually depends on features extracted from images, whereas deep-learning models such as CNNs can learn representations directly from image data. Transfer learning can reuse representations learned from large datasets and adapt them to medical-imaging tasks, although careful validation remains essential.

The objective of this paper is to review common AI approaches used in cancer image analysis and outline a conceptual workflow for an AI-assisted detection system. It discusses imaging modalities, algorithm families, evaluation metrics, applications, limitations, and future directions. This is a literature-based review; no original dataset was collected and no model was trained for this paper.

The interpretation of medical images requires expertise and can involve large volumes of data. Artificial Intelligence (AI) methods are being investigated to assist

with image analysis by detecting patterns, highlighting suspicious regions, and supporting consistent review. Machine Learning (ML) learns patterns from data, while Deep Learning (DL) uses multi-layer neural networks to learn representations from complex inputs such as images.

This paper reviews AI-based cancer detection methods with a focus on image analysis. It explains a conceptual workflow, summarizes common algorithms and application areas, and discusses limitations and research directions. The paper is literature-based and does not present experimental results from a newly implemented model.

II. RELATED WORK

Recent surveys report the use of deep learning for multiple cancer types and imaging modalities. Ahmad and Alqurashi reviewed cancer detection research published from 2020 to 2024, covering data, preprocessing, segmentation, feature extraction, deep learning, transfer learning, evaluation metrics, and challenges across several cancer types [1].

© 2026 IJSRET

Tandon et al. reviewed deep-learning-based automated cancer diagnosis models, including CNN architectures and pretrained networks, and discussed limitations and future research directions [2]. Sriraman et al. examined real-time deep-learning methods for image-based cancer diagnosis and emphasized issues such as data variability, generalization, and explainability [3].

An umbrella review by Xu et al. synthesized systematic reviews of AI-based imaging diagnosis across cancer types. The review reported promising findings but also noted that evidence quality and diagnostic performance

varied, underscoring the need for careful validation before clinical use [4].

1. Research Gap And Objectives

Although recent reviews report promising results for AI-assisted cancer image analysis, performance varies by cancer type, dataset, imaging modality, and study design. Many systems are evaluated using curated or retrospective datasets that may not represent differences in scanners, clinical populations, image protocols, or disease prevalence in routine practice. Further concerns include limited external validation, imbalanced datasets, unclear model explanations, and inconsistent reporting. A high score on a test dataset alone does not establish clinical benefit or readiness for deployment.

The objectives of this paper are: (1) to summarize AI and deep-learning techniques used in cancer image analysis;

(2) to describe a conceptual end-to-end workflow; (3) to compare traditional ML, CNNs, transfer learning, and segmentation approaches; (4) to discuss evaluation metrics and validation considerations; and (5) to identify technical, ethical, and clinical challenges for future research.

2. Proposed Ai-Assisted Workflow

The conceptual framework begins with medical image acquisition and proceeds through preprocessing, feature learning, classification, and output presentation. The stages are intended to support an AI-assisted analysis workflow; the specific model and classes must be selected according to the cancer type and dataset.

3. Image Acquisition

Images may be obtained from approved clinical sources or public research datasets. Depending on the task, input data can include mammograms, CT scans, MRI scans, dermoscopic images, or histopathology slides. Dataset labels and patient-level separation should be checked to avoid leakage between training and testing sets.

Preprocessing

Preprocessing may include resizing, intensity normalization, noise reduction, contrast adjustment, and image quality screening. Augmentation techniques such as modest rotations, flips, or zooms may be applied to training images when medically appropriate. Augmentation must not create misleading anatomical changes.

4. Feature Learning

Traditional ML pipelines may extract handcrafted features such as texture, shape, intensity, and statistical descriptors. CNN-based approaches learn feature representations from image pixels through convolutional layers and nonlinear activations, with pooling or other operations used to manage spatial dimensions.

5. Classification And Output

A classifier can produce probabilities or scores for dataset-defined categories such as normal, benign, or malignant. These labels are task-specific and should not be interpreted as a definitive clinical diagnosis. The output

may include a class, a calibrated confidence score, and a visualization of regions contributing to the prediction.

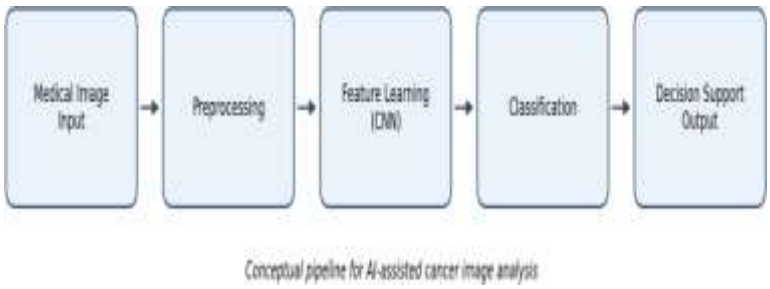


Fig. 1. Conceptual workflow for AI-assisted cancer image analysis.

For practical implementation, data should be split at the patient level to reduce leakage between training and test sets. Model selection should use validation data, while the final test set should remain untouched until evaluation.

External datasets and prospective evaluation, where feasible, are important for assessing generalizability.

6. AI Techniques For Cancer Detection

Table II compares common AI approaches and their typical roles in cancer image-analysis workflows.

Approach	Typical use	Key consideration
Traditional ML (SVM, Random Forest)	Classification using engineered image features	Feature design and data quality strongly influence results

© 2026 IJSRET

Approach	Typical use	Key consideration
CNN	Image classification and feature learning	Requires representative labelled data and careful validation
Transfer learning	Adapting pretrained networks to medical images	Pretraining domain and fine-tuning strategy matter
Segmentation models	Locating lesion or tumour regions	Pixel-level annotations may be costly and variable
Explainable AI	Visualizing regions influencing a model output	Visual explanations are not proof of clinical reasoning

1. **Machine Learning Algorithms**
Algorithms such as Support Vector Machines (SVM), Logistic Regression, Decision Trees, Random Forest, and K-Nearest Neighbour (KNN) can be used for classification when suitable features are available. Their performance depends on feature quality, dataset size, preprocessing, and parameter selection.

2. **Convolutional Neural Networks**
CNNs are widely used for image classification. Convolution layers learn local patterns such as edges and textures; activation functions introduce nonlinearity; pooling or downsampling layers reduce feature-map size; and later layers combine learned features for classification.

CNNs can learn features directly from images, reducing reliance on handcrafted descriptors.

6. Transfer Learning
Transfer learning adapts a model pretrained on a large dataset to a target medical-imaging task. Common model families include VGG, ResNet, MobileNet, and EfficientNet. Transfer learning may reduce training requirements when medical data are limited, but careful fine-tuning and validation are still required.

7. Segmentation And Explainability
Image segmentation identifies regions of interest, such as a lesion or tumour boundary, and can support measurement or downstream classification. Explainability tools such as Grad-CAM generate visual indications of image regions influencing a model output. Such visualizations can assist review but do not independently establish correctness

III. APPLICATIONS OF AI IN CANCER DETECTION

1. Breast Cancer
AI can analyse mammography, ultrasound, MRI, and pathology images to assist in detecting suspicious findings and classifying lesions. AI may be used as a second-reader or triage aid, subject to clinical validation.

2. Lung Cancer
CT-based systems can be designed to identify and characterize pulmonary nodules. AI may assist with candidate detection and prioritization, but findings must be reviewed in clinical context.

3. Brain Cancer
MRI analysis can support tumour-region segmentation and classification tasks. Models may learn patterns in image sequences or modalities, but performance depends on tumour type, acquisition protocols, and labels.

4. Skin Cancer
Computer vision models can analyse dermoscopic images of skin lesions and estimate classes associated with benign or malignant conditions. Image quality, skin-tone diversity, and external validation are important considerations

5. Other Cancer Types
AI methods are also being investigated for colorectal, liver, cervical, prostate, and other cancers using imaging, pathology, endoscopy, or multimodal clinical data [1], [4].

6. Advantages, Challenges, And Evaluation
1. Potential Advantages
AI can automate repetitive image-analysis steps, support prioritization of suspicious cases, help standardize measurements, and assist clinicians in reviewing large image volumes. It may also support research by enabling reproducible analysis of defined datasets.

2. Challenges
Important limitations include limited or unrepresentative datasets, class imbalance, label quality,image variation across institutions, privacy and security concerns, bias, explainability, and the risk of over-reliance on predictions. A model trained on one dataset may not generalize to another hospital or population.

3. Evaluation Metrics
Classification systems are commonly evaluated using accuracy, precision, recall (sensitivity), specificity, F1-score, and the area under the receiver operating characteristic curve (AUC). For medical screening, sensitivity and specificity

© 2026 IJSRET

should be interpreted alongside prevalence, clinical workflow, and the consequences of false positives and false negatives. Results should be reported on data separated from training and tuning

Table I. Common evaluation metrics for cancer classification

Metric	Meaning	Use
Accuracy	Correct predictions / all predictions	Overall correctness
Sensitivity	True positives / actual positives	How many cases are detected
Specificity	True negatives / actual negatives	How many non-cases are correctly excluded
Precision	True positives / predicted positives	Reliability of positive predictions
F1-score	Harmonic mean of precision and recall	Balance of precision and recall

4. Future Scope

Future research can explore multimodal systems that combine imaging with clinical, laboratory, and molecular information; explainable models; privacy-preserving training such as federated learning; efficient architectures for edge devices; and prospective validation across diverse institutions. Researchers should also prioritize dataset transparency, calibration, reproducibility, fairness, and evaluation of clinical impact rather than relying only on retrospective accuracy.

IV. CONCLUSION

Artificial Intelligence has the potential to support cancer detection by analyzing medical images and helping identify patterns that warrant further review. CNNs, transfer learning, traditional ML, segmentation, and explainability methods provide tools for building image-analysis systems. However, their usefulness depends on data quality, representative datasets, robust validation, interpretability, privacy safeguards, and evaluation in realistic clinical settings. This paper presented a literature-based review and conceptual workflow rather than an experimentally validated model. Future research should focus on transparent reporting, multi-centre external validation, patient-centred outcomes, and responsible integration into clinical workflows. AI should complement healthcare professionals and not replace clinical diagnosis.

REFERENCES

1. I. Ahmad and F. Alqurashi, “Early cancer detection using deep learning and medical imaging: A survey,” Critical Reviews in Oncology/Hematology, vol. 204, Art. no. 104528, 2024. doi: 10.1016/j.critrevonc.2024.104528.
2. R. Tandon, S. Agrawal, N. P. S. Rathore, A. K. Mishra, and S. K. Jain, “A systematic review on deep learning-based automated cancer diagnosis models,” Journal of Cellular and Molecular Medicine, vol. 28, no. 6, e18144, 2024. doi: 10.1111/jcmm.18144.
3. H. Sriraman, S. Badarudeen, S. Vats, and P. Balasubramanian, “A systematic review of real-time deep learning methods for image-based cancer diagnostics,” Journal of Multidisciplinary Healthcare, vol. 17, pp. 4411–4425, 2024. doi: 10.2147/JMDH.S446745.
4. H.-L. Xu et al., “Artificial Intelligence Performance in Image-Based Cancer Identification: Umbrella Review of Systematic Reviews,” Journal of Medical Internet Research, vol. 27, e53567, 2025. doi: 10.2196/53567. AUTHOR PROFILE Sakshi Yadav is an engineering student in Artificial Intelligence and Data Science at Sahyadri Valley College of Engineering and Technology, Rajuri, Maharashtra, India. Her academic interests include artificial intelligence, machine learning, deep learning, medical image processing, and healthcare applications.

© 2026 IJSRET