International Journal of Medical Science and Research
www.medicaljournals.in
Online ISSN: 2664-8768, Print ISSN: 2664-875X
Received: 15-05-2022, Accepted: 30-05-2022, Published: 16-06-2022
Volume 4, Issue 1, 2022, Page No. 65-70

Advanced computational oncology: A systematic review of deep learning methods for carcinoma detection

Priyadarshini SRP[1], Sivaramarajalu Ramadurai Venkataraajalu[2]
[1] Department of General Medicine, Tbilisi State Medical University, Georgia
[2] Department of Computer Science, State University of New York, United States

Abstract
Cancer remains a leading cause of morbidity and mortality globally, necessitating advancements in diagnostic techniques for timely and accurate identification of cancer cells. Deep learning, a subset of artificial intelligence, has emerged as a revolutionary approach in medical imaging, offering significant improvements in the detection and classification of cancerous cells. This systematic review aims to comprehensively evaluate the current literature on the application of deep learning techniques in the identification of cancer cells. We analyze various deep learning methodologies, datasets, performance metrics, and the associated challenges. By synthesizing findings from recent studies, we highlight the potential of deep learning to transform cancer diagnostics, discuss the robustness and interpretability of these models, and identify areas for future research. Our review underscores the promise of deep learning in enhancing diagnostic accuracy, reducing workload for pathologists, and ultimately improving patient outcomes.

Keywords: Deep learning, cancer cells, identification, systematic review, medical imaging, artificial intelligence

Introduction

Cancer remains one of the most challenging and prevalent diseases worldwide, characterized by uncontrolled cell growth and metastatic potential. Early and accurate detection of cancer cells is crucial for effective treatment and improved survival rates [1, 2]. Traditional methods of cancer cell identification, such as histopathological examination, are often labor-intensive, time-consuming, and subject to inter-observer variability [3, 4]. These limitations highlight the need for innovative diagnostic tools that can provide reliable and reproducible results. Deep learning, a branch of artificial intelligence (AI) that involves training neural networks on large datasets, has shown remarkable success in various fields, including medical imaging [5, 6]. In the context of cancer detection, deep learning techniques, particularly convolutional neural networks (CNNs), have demonstrated exceptional capabilities in automatically extracting and analyzing complex features from medical images [7, 8]. This has opened new avenues for automating the identification of cancer cells, potentially enhancing diagnostic accuracy and efficiency [9, 10]. Recent studies have leveraged deep learning for various cancer types, including skin cancer, breast cancer, lung cancer, and colorectal cancer [11, 12]. These studies have utilized diverse datasets and have reported high performance metrics, indicating the robustness of deep learning models in cancer detection. However, challenges such as the need for large, annotated datasets, variability in image quality, and the interpretability of deep learning models remain significant barriers to clinical adoption [13, 14].

The application of deep learning in medical imaging extends beyond cancer detection to various other medical conditions and diagnostic tasks. For instance, deep learning models have shown promise in detecting pneumonia from chest X-rays, showcasing the potential of AI to assist in rapid and accurate diagnosis of respiratory conditions [15]. In the field of histopathology, researchers have explored the use of deep learning to enhance the accuracy and efficiency of histopathological diagnosis [16]. These studies highlight how deep learning tools could aid pathologists in analyzing tissue samples, potentially reducing workload and improving diagnostic precision. As the field of AI in medical imaging continues to evolve, it is crucial to address the clinical challenges and explore the full spectrum of applications [13]. The convergence of human expertise and artificial intelligence holds great promise for high-performance medicine [14]. These developments underscore the broader impact of deep learning across various domains of medical imaging and diagnosis, suggesting its transformative potential in healthcare beyond cancer detection [17, 18].

This systematic review aims to provide a comprehensive overview of the current state of deep learning applications in cancer cell identification. We systematically search and analyze the literature to identify key trends, methodologies, datasets, and performance metrics. Additionally, we discuss the challenges and future directions in this rapidly evolving field. By synthesizing the findings from recent research, we aim to highlight the potential of deep learning to revolutionize cancer diagnostics and improve patient care.

Methodology

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines.

1. PICOTT Table

The methodology comprised three main components: a PICOTT (Population, Intervention, Comparison, Outcome, Time Type) table to define the scope of the review (as depicted in Table 1), a comprehensive set of search keywords to identify relevant literature, and the PRISMA flow diagram to document the study selection process. All these elements were carefully designed and implemented to ensure a thorough and transparent review of the current literature on deep learning applications in cancer cell identification.

Additional criteria

- **Language:** Include only English language publications
- **Sample size:** Include studies with a minimum of 100 cases/images
- **Data availability:** Include studies using publicly available datasets or clearly described proprietary datasets

Table 1: PICOTT table that defines the scope of the review

Criterion	Inclusion	Exclusion
Population	Adult patients (≥18 years old) with suspected or confirmed carcinoma	Pediatric patients, animal studies
Intervention	Deep learning or artificial intelligence algorithms applied to medical imaging or histopathology for carcinoma detection/diagnosis	Traditional machine learning methods not involving deep learning
Comparison	Standard diagnostic methods (e.g., manual interpretation by radiologists/pathologists), other AI methods, or no comparison	Studies without a clear comparison or baseline
Outcome	Diagnostic accuracy metrics (e.g., sensitivity, specificity, AUC-ROC), classification performance, detection rates	Studies not reporting quantitative performance metrics
Timing	Studies published from January 2010 onwards	Studies published before 2010
Type of study	Original research articles, including retrospective and prospective studies	Review articles, case reports, conference abstracts, editorials

2. Search Strategy and Initial Screening

Searches were conducted using keywords and search terms in PubMed and Google Scholar. Date filters were applied to limit the search results to articles published after January 2010 as the computing power and efficient deep learning algorithms advanced after around this time. The initial screening involved reviewing the titles and abstracts to exclude studies that did not meet the inclusion criteria. We used different combinations of keywords, including deep learning, artificial intelligence, cancer, carcinoma, medical imaging, histopathology, image processing and radiology to search the mentioned databases. MeSH keywords were used in PubMed to further refine search results (as depicted in Table 2).

Table 2: Search strategy and keywords used in the study

Search Strategy	Keywords
Regular keywords	Deep Learning, Artificial Intelligence, Cancer, Carcinoma, Medical Imaging, Histopathology, Image Processing, Radiology
MeSH (Medical Subject Headings) keywords	(("deep learning"[MeSH Terms] OR "deep learning"[All Fields] OR "artificial intelligence"[MeSH Terms] OR "artificial intelligence"[All Fields]) AND ("carcinoma"[MeSH Terms] OR "carcinoma"[All Fields] OR "neoplasms"[MeSH Terms] OR "cancer"[All Fields]) AND ("diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "detection"[All Fields]) AND ("diagnostic imaging"[MeSH Terms] OR "medical imaging"[All Fields] OR "radiology"[MeSH Terms] OR "radiology"[All Fields])) AND ("2010/01/01"[Date - Publication] : "2025"[Date - Publication])

3. PRISMA 2020 Flow Diagram

The PRISMA 2020 flow diagram is illustrated below

Discussion

The reviewed studies utilized a diverse range of datasets with large image collections and different types of deep learning models with varying degrees of customization in the training mechanisms to study the efficiency of carcinoma detection. Esteva *et al.* [19] employed a dataset of 129,450 clinical images for skin cancer classification. Hu *et al.* [20] used the CAMELYON16 dataset containing whole-slide images of lymph node sections for breast cancer analysis. Cao *et al.* [21] analyzed 417 patients with fully annotated mp-MRI studies for prostate cancer detection and Gleason score prediction. Coudray *et al.* [22] leveraged The

Cancer Genome Atlas (TCGA) for lung cancer subtype classification. Aresta *et al.* [23] utilized the BACH (Breast Cancer Histology) dataset for their grand challenge in breast cancer detection. Lu *et al.* [24] used a dataset of whole-slide images for colorectal cancer classification from The Cancer Genome Atlas (TCGA). The variation in dataset sizes and types underscores the adaptability of deep learning models across different cancer types and imaging modalities. A common thread among these studies is the use of large, diverse datasets, which is crucial for training robust AI models. The datasets ranged from clinical images in dermatology to whole-slide images in pathology and multi-parametric MRI in radiology, showcasing the versatility of deep learning approaches. A summary of these details are depicted in Table 3 below.

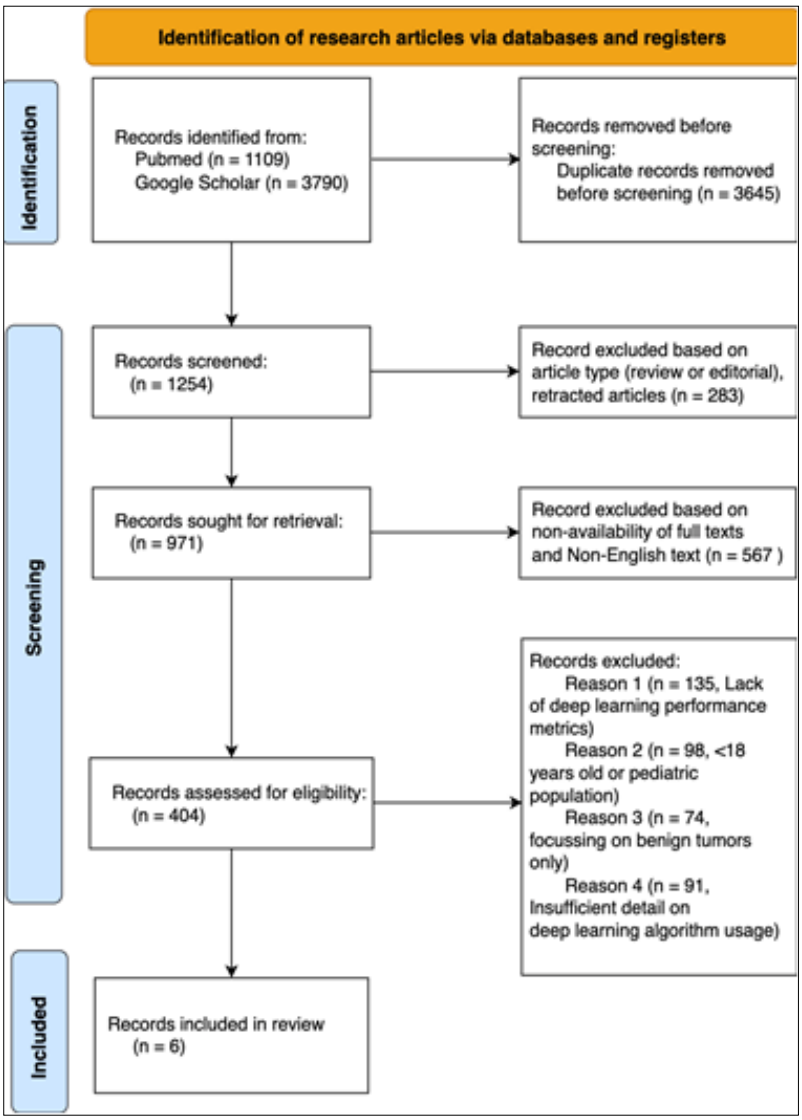

Fig 1: Prisma 2020 flow diagram for systematic review following inclusion and exclusion criteria

Table 3: Datasets and Artificial Intelligence Models Employed in Selected Studies for Various Cancer Types

Study	Cancer Type	Dataset	Artificial Intelligence Model
Esteva *et al.* [19]	Skin cancer	Clinincal images	Inception v3 CNN
Hu*et al.* [20]	Metastatic breast cancer	CAMELYON16	GoogLeNet CNN + Random Forest
Cao *et al.* [21]	Prostate cancer	Multiparametric MRI	FocalNet (Specialized CNN)
Coudray *et al.* [22]	Lung cancer (NSCLC)	The Cancer Genome Atlas (TCGA)	Inception v3 CNN
Aresta *et al.* [23]	Breast cancer	BACH (Breast Cancer Histology)	Primarily CNN (various models were used as part of the challenge)

The reviewed studies utilized a diverse range of datasets with large image collections and different types of deep learning models with varying degrees of customization in the training mechanisms to study the efficiency of carcinoma detection. Esteva *et al.* [19] employed a dataset of 129,450 clinical images for skin cancer classification. Hu *et al.* [20] used the CAMELYON16 dataset containing whole-slide images of lymph node sections for breast cancer analysis. Cao *et al.* [21] analyzed 417 patients with fully annotated mp-MRI studies for prostate cancer detection and Gleason score prediction. Coudray *et al.* [22] leveraged The Cancer Genome Atlas

(TCGA) for lung cancer subtype classification. Aresta *et al*. [23] utilized the BACH (Breast Cancer Histology) dataset for their grand challenge in breast cancer detection. Lu *et al*. [24] used a dataset of whole-slide images for colorectal cancer classification from The Cancer Genome Atlas (TCGA). The variation in dataset sizes and types underscores the adaptability of deep learning models across different cancer types and imaging modalities. A common thread among these studies is the use of large, diverse datasets, which is crucial for training robust AI models. The datasets ranged from clinical images in dermatology to whole-slide images in pathology and multi-parametric MRI in radiology, showcasing the versatility of deep learning approaches. A summary of these details is depicted in Table 3.

The studies utilized various deep learning architectures, predominantly employing widely available Convolutional Neural Networks (CNNs) or customized CNN variants. The studies by Esteva *et al*. [19] and Hu *et al*. [20] showcase the versatility of deep learning architectures in addressing diverse medical imaging challenges, while highlighting the effectiveness of transfer learning and hybrid approaches. Esteva's team adapted an Inception v3 CNN pre-trained on ImageNet for skin lesion classification, achieving performance on par with 21 board-certified dermatologists with AUCs of 0.96 for carcinoma and 0.94 for melanoma classification. In contrast, Hu *et al*. employed a hybrid model combining GoogleNet CNN with a Random Forest classifier for metastatic breast cancer detection in lymph nodes, achieving an impressive AUC-ROC of 0.993 on the CAMELYON16 dataset. This performance suggests potential for significant improvements in pathology workflow, particularly in detecting micro-metastasis that might be overlooked in routine examinations. While Esteva's model excelled in classifying a wide range of skin lesions, Hu's approach demonstrated superior performance in detecting subtle metastases, significantly outperforming human pathologists (AUC 0.993 vs 0.884 for a pathologist without time constraint). Both studies underscore the potential of AI in medical imaging, but their methodologies differ notably. Esteva's use of transfer learning efficiently leveraged general image recognition capabilities for a specialized task, whereas Hu's hybrid approach effectively addressed the computational challenges of processing whole-slide images. These contrasting yet successful approaches highlight the adaptability of deep learning in medical imaging. Esteva's work suggests AI's potential as a valuable second opinion in dermatological practice, while Hu's study points to AI's capability to reduce pathologist workload and improve cancer staging accuracy, particularly in identifying small metastases that human observers might miss. Together, these studies demonstrate how different AI architectures can achieve state-of-the-art performance in their respective domains, paving the way for AI to assist in comprehensive cancer screening programs and potentially revolutionize diagnostic workflows by enhancing both efficiency and accuracy.

An interesting trend in the application of deep learning for cancer detection is the evolution from single task to multi-task and multi-modal approaches. While the above studies focused on classification tasks, studies like Cao *et al*. [21], Coudray *et al*. [22] and Lu *et al*. [24] incorporated multiple tasks or data types. This progression reflects the increasing sophistication of AI models and their ability to mimic complex clinical decision-making processes. Cao's FocalNet simultaneously performs prostate cancer detection and Gleason score prediction from multi-parametric MRI, leveraging different MRI sequences (T2-weighted, diffusion-weighted, and dynamic contrast-enhanced) to provide comprehensive tissue characterization. Lu's model tackles the challenging problem of identifying the origin of cancers of unknown primary (CUP) by incorporating both global and local image features from whole-slide histopathology images, using a multi-scale approach that captures tissue architecture and cellular details. Coudray's CNN goes a step further by not only classifying non-small cell lung cancer (NSCLC) subtypes but also predicting common genetic mutations directly from histopathology images, utilizing an Inception v3 architecture pre-trained on ImageNet and fine-tuned on tumor tiles. These studies highlight a key advantage of deep learning: the ability to automatically extract and synthesize relevant features at different levels of abstraction, integrating multiple levels of information to provide more comprehensive diagnostic insights.

The multi-task and multi-modal approaches demonstrated in these studies have significant implications for clinical practice and personalized medicine, with each model achieving impressive performance in their respective, challenging domains. Cao's FocalNet, achieving sensitivities of 89.7% for index lesions and 87.9% for clinically significant lesions, shows remarkable accuracy in prostate cancer diagnosis and grading, a task that typically requires extensive expertise and is prone to inter-observer variability. Lu's model tackles the even more complex challenge of identifying primary sites across diverse cancer types, a problem that often confounds experienced pathologists. Its top-1 accuracies of 0.83 and 0.80 on held-out and external test sets represent a significant advancement in addressing this long-standing oncological dilemma. Coudray's work sets a new benchmark in lung cancer classification, reporting exceptionally high accuracies of 0.97, 0.98, and 0.99 for adenocarcinoma, squamous cell carcinoma, and normal lung tissue classification respectively. These results are particularly impressive given the subtle histological differences between cancer subtypes and the additional complexity of predicting genetic mutations from image data alone. The performance of these models becomes even more significant when considering the multifaceted nature of their tasks - from integrating multiple MRI sequences in FocalNet to analyzing both global and local features in Lu's model and combining histopathological and genetic insights in Coudray's approach. By successfully performing multiple related tasks simultaneously and integrating various data types, these AI models are demonstrating a level of comprehensive analysis that closely mirrors the nuanced, multi-factor assessments required in clinical oncology practice. This convergence of AI capabilities with clinical needs suggests a potential revolution in cancer diagnostics, where AI could serve not

just as a supportive tool, but as a powerful partner in enhancing the accuracy, efficiency, and depth of cancer diagnosis and characterization.

While Cao, Coudray, and Lu demonstrated the power of multi-task and multi-modal approaches in individual models, Aresta *et al*. [23] took a broader, community-driven approach by organizing the BACH (Breast Cancer Histology) grand challenge. This innovative initiative evaluated multiple AI models on a standardized breast cancer histopathology dataset, providing valuable insights into the current state-of-the-art in AI-based histopathology analysis. The challenge encompassed various tasks, from high-resolution image classification to whole slide image labeling, mirroring the complexity of real-world pathology workflows. The best-performing methods achieved an accuracy of 0.87 in classifying high-resolution microscopy images into four breast tissue categories and a score of 0.69 in labeling entire whole slide images. These results, while not as high as some individual studies, are significant given the standardized and challenging nature of the dataset. The BACH challenge complements the other studies by providing a benchmark for comparing different AI approaches and highlighting areas for improvement. It also underscores the potential of AI to assist pathologists in routine diagnostic tasks, potentially enhancing efficiency and consistency in breast cancer assessment. By fostering collaboration and competition in the AI community, initiatives like BACH play a crucial role in advancing the field of AI in cancer diagnostics, driving innovation in multi-task and multi-modal approaches that can be applied across various cancer types and imaging modalities.

While a key strength observed across studies was the rigorous validation strategies employed, including direct comparisons of AI performance against human experts, several challenges persist in the application of deep learning for cancer detection. These challenges include the need for large, annotated datasets, variability in image quality, and the presence of artifacts affecting model performance. The "black box" nature of many deep learning models poses a significant hurdle for clinical adoption, as healthcare providers require transparent and interpretable decision-making processes. Although some studies, like Xie *et al*. [25], have explored knowledge-based collaborative approaches to enhance interpretability, this area requires further research to develop explainable AI crucial for gaining trust from healthcare professionals and meeting regulatory requirements. Looking forward, there's a pressing need for prospective, multi-center studies to validate these AI models in real-world clinical settings, focusing on integrating AI tools into existing workflows and assessing their impact on patient outcomes and healthcare economics. Future work should prioritize developing more robust and interpretable models, improving data quality, and exploring novel techniques for seamless integration into diagnostic workflows. Ultimately, successful translation of these technological advancements into clinical practice will hinge on close collaboration between AI researchers and healthcare professionals, with the goal of significantly improving cancer diagnosis and patient care.

Ethical considerations and potential biases in AI models were not extensively addressed in most studies, highlighting a critical area for future research. Ensuring diverse representation in training datasets is crucial to mitigate potential biases and improve generalizability across different patient populations. Additionally, our analysis revealed that traditional quality assessment tools, such as the Newcastle-Ottawa Scale (NOS), may not be fully applicable to AI-based cancer detection studies. While the reviewed studies scored highly in domains like cohort selection and outcome assessment, follow-up aspects were often limited or not applicable due to the cross-sectional nature of imaging-based diagnostic research. This underscores a significant gap in current quality assessment methodologies for AI in medical imaging. Moving forward, the field should not only focus on advancing AI technologies but also on developing tailored quality assessment tools that can adequately evaluate the unique aspects of AI-based diagnostic accuracy studies in cancer detection, including their ethical implications and potential biases. Such tools would be invaluable for ensuring rigorous evaluation of future research in this rapidly evolving field, ultimately contributing to more robust, unbiased, and ethically sound AI applications in cancer detection.

Conclusion

As we have explored throughout this review, deep learning has emerged as a powerful tool in the fight against cancer, offering promising results across various medical specialties. From dermatology to pathology and radiology, the studies we've examined demonstrate the potential of AI to enhance diagnostic accuracy and efficiency. However, significant challenges remain, particularly in terms of data quality, model interpretability, and clinical integration.

By bringing together AI researchers, healthcare professionals, and policymakers, we can work towards developing more robust, interpretable models and integrating them effectively into clinical workflows. Also, as these technologies evolve, it is crucial to establish standardized protocols and ethical guidelines to ensure the safe and effective use of AI in healthcare. While the journey ahead looks challenging, the potential rewards - improved patient outcomes, optimized healthcare resources, and a reduced global cancer burden - make it a worthy endeavor. It's evident that the future of cancer detection lies at the intersection of human expertise and artificial intelligence, promising a new era in personalized, efficient, and effective cancer care.

References

1. Siegel RL, Miller KD, Fuchs HE, *et al*. Cancer statistics 2021. CA Cancer J Clin,2021:71:7-33. 10.3322/caac.21654

2. Bray F, Ferlay J, Soerjomataram I, *et al*. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin,2018:68:394-424. 10.3322/caac.21492
3. Gurcan MN, Boucheron LE, Can A, *et al*. Histopathological image analysis: A review. IEEE Rev Biomed Eng,2009:2:147-171. 10.1109/RBME.2009.2034865
4. Veta M, Pluim JPW, van Diest PJ, *et al*. Breast cancer histopathology image analysis: A review. IEEE Trans Biomed Eng,2014:61:1400-1411. 10.1109/TBME.2014.2303852
5. Litjens G, Kooi T, Bejnordi BE, *et al*. A survey on deep learning in medical image analysis. Med Image Anal,2017:42:60-88. 10.1016/j.media.2017.07.005
6. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng,2017:19:221-248. 10.1146/annurev-bioeng-071516-04420442
7. Liu Y, Kohlberger T, Norouzi M, *et al*. Artificial intelligence-based breast cancer nodal metastasis detection: Insights into the black box for pathologists. Arch Pathol Lab Med,2019:143:859-868. 10.5858/arpa.2018-0147-OA
8. McKinney SM, Sieniek M, Godbole V, *et al*. International evaluation of an AI system for breast cancer screening. Nature,2020:577:89-94. 10.1038/s41586-019-1799-6
9. Tschandl P, Codella N, Akay BN, *et al*. Comparison of the accuracy of human readers versus machine learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. Lancet Oncol,2019:20:938-947. 10.1016/S1470-2045(19)30333-X
10. Bejnordi BE, Veta M, van Diest PJ, *et al*. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer. JAMA,2017:318:2199-2210. 10.1001/jama.2017.14585
11. Ardila D, Kiraly AP, Bharadwaj S, *et al*. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med,2019:25:954-961. 10.1038/s41591-019-0447-x
12. Kather JN, Krisam J, Charoentong P, *et al*. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. PLOS Medicine,2019:16:1002730. 10.1371/journal.pmed.1002730
13. Bi WL, Hosny A, Schabath MB, *et al*. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA Cancer J Clin,2019:69:127-157. 10.3322/caac.21552
14. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med,2019:25:44-56. 10.1038/s41591-018-0300-7
15. Rajpurkar P, Irvin J, Zhu K, *et al*. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLOS Medicine,2018:15:10. 1371/journal.pmed.1002686
16. Litjens G, Bejnordi BE, Setio AAA, *et al*. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep,2016:6:26286. 10.1038/srep26286
17. Hosny A, Parmar C, Quackenbush J, *et al*. Artificial intelligence in radiology. Nat Rev Cancer,2018:18:500-510. 10.1038/s41568-018-0016-5
18. Jiang F, Jiang Y, Zhi H, *et al*. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol,2017:2:230-243. 10.1136/svn-2017-000101
19. Esteva A, Kuprel B, Novoa RA, *et al*. Dermatologist-level classification of skin cancer with deep neural networks. Nature,2017:542:115-118. 10.1038/nature21056
20. Hu Y, Wei J, Zhao W, *et al*. End-to-End Deep Learning for Detecting Metastatic Breast Cancer in Axillary Lymph Node from Digital Pathology Images. Gastric Cancer,2021:24:868-877. 10.1007/s10120-021-01158-9
21. Cao R, Bajgiran AM, Mirak SA, *et al*. Joint prostate cancer detection and Gleason score prediction in mp-MRI via FocalNet. IEEE Trans Med Imaging,2019:38:2496-2506. 10.1109/TMI.2019.2901928
22. Coudray N, Ocampo PS, Sakellaropoulos T, *et al*. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med,2018:24:1559-1567. 10.1038/s41591-018-0177-5
23. Aresta G, Araujo T, Kwok S, *et al*. BACH: Grand challenge on breast cancer histology images. Med Image Anal,2019:56:122-139. 10.1016/j.media.2019.05.010
24. Lu MY, Chen TY, Williamson DFK, *et al*. AI-based pathology predicts origins for cancers of unknown primary. Nature,2021:594:106-110. 10.1038/s41586-021-03512-4
25. Xie Y, Xia Y, Zhang J, *et al*. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. IEEE Trans Med Imaging,2020:39:1585-1595. 10.1109/TMI.2018.2876510