INTERNATIONAL JOURNAL of MEDICAL STUDENTS

Review

Navigating the Digital Frontier: A Review on the Clinical Applications of Convolutional Neural Networks and Emerging AI Models in Medicine and Surgery

Sri Sai Rohit Kosuri,[1] David Sunnucks.[1]

Abstract

Artificial intelligence (AI) is being integrated into several fields worldwide due to its impressive capabilities in completing tasks, sometimes autonomously. Research by several groups worldwide has shown that AI could similarly be incorporated into clinical practice. Convolutional neural network (CNN) models have an inherent capability of recognising and classifying patterns, allowing them to be used in imaging and other diagnostic techniques in various clinical specialities. With some AI systems already in use, it is anticipated that several other AI models will come into clinical practice in the coming years to improve healthcare and patient outcomes. Hence, it is paramount that current medical students and practising doctors keep up with these recent advances in AI to provide the best standard of care for patients. This narrative review explores the basis of deep learning CNN models and summarises extensive literature to provide an overview of some of the recent applications of CNN models to various clinical specialities in medicine and surgery.

Introduction

Artificial intelligence (AI) has shown incredible promise in clinical medicine, with key advances including enhanced diagnostic accuracy, better disease detection and improved workflow efficiency.[1] Moreover, deep learning models, such as convolutional neural networks (CNN), have the extraordinary ability to constantly learn and develop reasoning from provided datasets. This allows CNNs to perform complex tasks such as recognizing and classifying patterns of disease from unorganized datasets into different categories.[2-3] Such capability makes deep learning models great candidates for use in various clinical specialties to aid in imaging and other diagnostic techniques, helping to increase accuracy rates and improve patient outcomes.[3] Several CNN models are being trialed by numerous research groups and are anticipated to be introduced into mainstream clinical practice in several specialties within medicine and surgery, including the ones illustrated in ***Figure 1.*** However, several gaps in the literature persist, such as the lack of real-world prospective studies, limited models showcasing generalizability, "blackbox" models with inadequate interpretability, and insufficient studies with external validation.[4-5] These limitations are currently preventing the adoption of several AI models into mainstream clinical workflow.

This narrative review aims to evaluate the current and potential applications of CNNs and other AI models in several clinical specialties, answering the following research question: How can CNN models and other AI-based models be utilized across various clinical specialties, and what are the current challenges and ethical implications hindering their widespread integration into clinical practice? Given the variability in assessment methods of AI models used in different studies in the literature, this review presents the latest advances through various evaluation metrics, including, but not limited to, accuracy percentages, specificity, sensitivity, F1 scores, and direct comparison of model performance with physicians. Where applicable, other metrics such as data quality and utility for real-world clinical integration are also discussed. It is crucial that not only currently practicing physicians, but also medical students and future doctors are aware of these recent advances in AI, which are expected to change the clinical landscape in the coming years.

Methods

The main aim of this narrative review is to evaluate relevant literature and analyze recent advances in AI in various clinical specialties. An extensive literature search was conducted using multiple appropriate databases, including MEDLINE, PubMed, Google Scholar, Web of Science, and Embase, to identify relevant original research studies and review articles for this narrative review. In addition, certain websites were used for technical

1 Third-Year Medical Student. Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.

2 MBBCh. Cardiff University, Cardiff, UK.

About the Author: *Sri Sai Rohit Kosuri is currently a third-year medical student at Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK of a 5-year program. Prior to medical school, Sri graduated with a First Class Honours in BSc Medical Physiology from the University of Leicester, Leicester, UK. Dr. David Sunnucks is the Head of Anatomy and Head of Year 3 MBBS Medicine at Queen Mary University of London, Malta Campus, Victoria, Gozo, Malta – part of Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.*

Correspondence:

Sri Sai Rohit Kosuri.

Address: Garrod Building, Turner St, London E1 2AD, United Kingdom.

Email: rohitkosuri24@gmail.com

Submission: Mar 18, 2024
Revisions: Jun 12, 27, 2025
Responses: Jun 10, Jul 10, 2025
Acceptance: Jul 13, 2025
Publication: Dec 16, 2025
Process: Peer-reviewed

information on AI products that are currently in use and where relevant articles were not available. Searches were made using keywords, including "artificial intelligence", "AI", "deep learning", "CNN models" and "large language models" for the range of clinical specialties covered in this narrative review. Studies were selected based on their inclusion of relevant recent advances using CNN models, deep learning algorithms or other AI models in the appropriate specialties discussed. The time frame of the included studies was 2017-2025. The scale for the quality assessment of narrative review articles (SANRA) guidelines was considered and used when reviewing literature.

Discussion

The Basis of AI Learning Algorithms

Initially, AI algorithms were designed and developed using machine learning (ML), an AI learning method that uses pattern identification to learn from presented data and minimize errors.[3] However, this type of learning requires large amounts of structured data for pattern recognition. Deep learning (DL), a subset of ML, can eliminate the manual task of data mining by using unstructured data to effectively group similarities and enhance pattern recognition.[2]

DL and ML are becoming the most common forms of AI learning that many organizations across the world are adopting, including organizations in the healthcare sector. For example, UC San Diego Health has adopted an AI model supported by Amazon Web Services (AWS) to analyze chest X-rays and assist radiologists in the detection of pneumonia in COVID-19 patients.6 Moreover, the AI model helped diagnose pneumonia in COVID-19 patients where the typical symptoms were absent.6 Another example includes the use of the Targeted Real-Time Early Warning System (TREWS) to identify patients at risk of developing sepsis by Johns Hopkins Medicine.

As part of a study, TREWS was used by over 4,000 clinicians across 5 hospitals, where the tool was used to treat 590,000 patients.[7] In contrast to previously tested electronic tools, which could correctly predict sepsis only 2-5% of the time, this AI model accurately predicted almost 40% of the sepsis cases among the 82% presented.[7]

Additionally, DL algorithms can be used to develop an artificial neural network (ANN) where there is an input layer, middle hidden layer(s), and an output layer to broadly function as neurons in a human brain. Data can be fed into the input layer, and information can be passed onto the next layer and receive an output, much like a brain.[8] Now, using such algorithms, researchers are developing a type of ANN called convolutional neural networks (CNN), which relies on computer vision (CV), where images and videos are fed into the neural network.[3] CNNs employ convolutional layers consisting of learnable filters, which are applied to the input image to detect specific features.9 These features can be associated with clinically meaningful entities, facilitating classification, detection and segmentation tasks, as shown in ***Figure 1.***[9]

With imaging investigations being central to the diagnosis and management of patients in several medical and surgical specialties, AI models, particularly CNNs, could aid medical teams in image analysis, allowing for better pathology detection.

If implemented correctly into clinical practice, these CNN models can allow for more accurate and faster diagnoses, leading to better healthcare outcomes.[10]

AI Use in Cardiology: Using AI models for ECGs, echocardiography and intravascular imaging

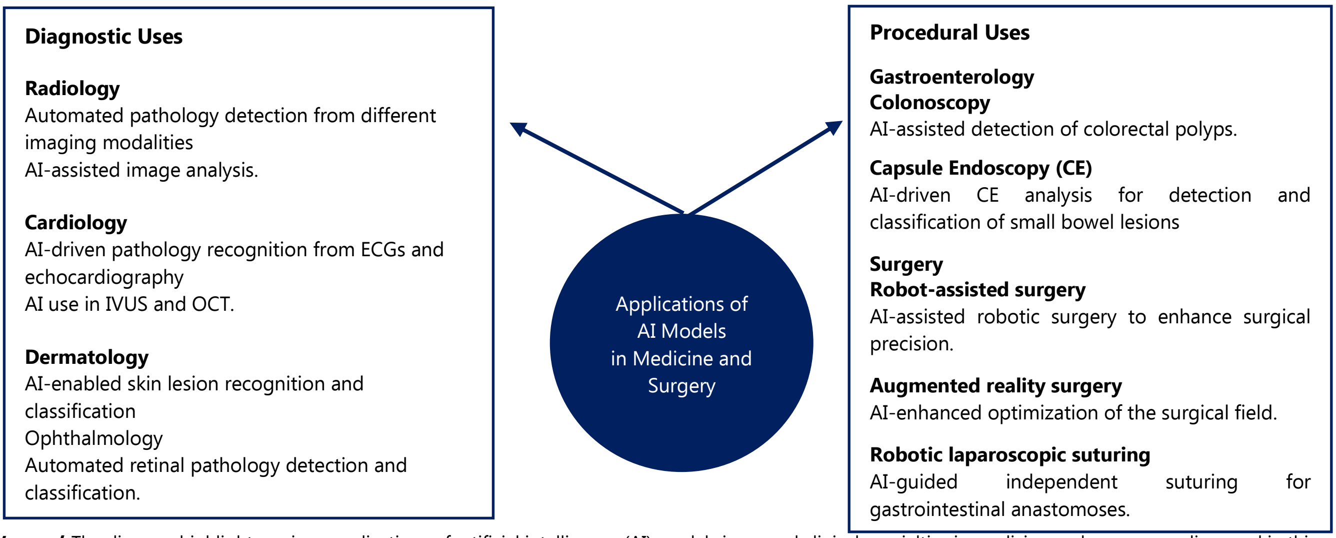
ECGs and Echocardiography

Cardiovascular medicine encompasses several serious heart conditions, including myocardial infarction (MI), heart failure (HF), or fatal arrhythmias requiring immediate medical attention. Several diagnostic techniques are used in diagnosis, with electrocardiograms (ECG) and echocardiography being some of the most requested investigations in cardiology, in addition to others. Analysis of an ECG accurately and efficiently is crucial for a diagnosis to allow the medical team to proceed with the appropriate management for the patient. However, ECGs may not always be correctly interpreted, with one meta-analysis study showing that the cardiologists' interpretation accuracy ranged from 49% to 92%.[12]

CNN models hold strong potential for enhancing the interpretation of ECGs with more accuracy, leading clinicians towards a more accurate diagnosis. Hughes et al. were able to train a CNN model to interpret a range of ECGs and even perform on par with cardiologists, and also exceed the standards of the currently in-use automated ECG detection system in 4 out of 5 diagnostic classes.[13] Further research has yielded more impressive results in not only recognizing abnormal ECG patterns but also classifying them into various cardiovascular diseases.[14-15] Yoon et al. used ECG graphs and converted them into grayscale images to train the CNN to recognize ECG patterns with a remarkable accuracy of 95.1%.[14] Going further, Makimoto et al. were able to train a CNN model to detect an MI from ECG images and outperformed 10 physicians with a higher F1 score (83% vs 81%) and higher accuracy (70% vs 67%).

However, while these results are impressive, it is important to note that generalizability holds significant weight in determining whether such models can be implemented into mainstream clinical practice. For instance, the methods used to train the CNN were different in these studies, with Yoon et al. using ECG graphs and converting them into grayscale images, while Makimoto et al. were able to train their model directly from ECG images.[14-15] Such differences indicate the need for further external validation, and importantly, research groups adopting a universal training protocol to prevent any overfitting to specific protocols and datasets.

Figure 1. Overview of the Applications of AI Models Across Various Specialties in Medicine and Surgery.



Legend: The diagram highlights various applications of artificial intelligence (AI) models in several clinical specialties in medicine and surgery, as discussed in this article.

In addition to ECGs, echocardiography is also a highly useful diagnostic tool for various cardiac pathologies. Echocardiography can be used to determine any abnormalities in cardiac size and shape, pumping strength via the ejection fraction, valvular disorders, cardiac muscle damage, congenital heart defects, as well as many others.[17] As with ECGs, there is a potential for misinterpretation of echocardiography images with inaccuracies up to 30% for transthoracic echocardiography (TTE). Using deep learning algorithms could help reduce interpretation inaccuracies as well as reduce the time required for echocardiogram interpretation.[18]

In one study, Madani et al. developed CNN models for TTE analysis through videos and images using over 250 echocardiograms, with the CNN model achieving an impressive overall accuracy of 97.8%.[19] Additionally, analysis on single images provided an accuracy of 91.7%, significantly higher than the accuracies of 70.2-84.0% achieved by board-certified echocardiographers.19 In a similar study, Naser et al. trained two-dimensional and three-dimensional CNN models to classify cardiac views obtained from TTE.20 The two-dimensional CNN achieved an overall accuracy of 96.8%, while the three-dimensional CNN had an overall accuracy of 96.3%.[20] These high diagnostic values emphasize the great potential for deep learning models to improve the diagnoses of cardiac diseases through echocardiography.

While these outcomes for the interpretation of ECGs and echocardiography imaging are encouraging, as indicated in ***Figure 1***, several limitations hinder their widespread integration into daily clinical practice. Many of the studies highlighting the ability of CNNs to assist in clinical decision making are often trained on existing, retrospective datasets. Such datasets may be useful for initial training, but CNNs must be trained with prospective data to account for the variability seen in real-world clinical practice. Further, generalizability remains a critical issue, as models for both ECG and echocardiography interpretation must perform similarly when presented with different, diverse data.[21]

AI use for arrhythmia detection from wearable devices

Additionally, AI models are being integrated into wearable devices, such as smartwatches and smart rings, to aid in the early detection of arrhythmias.[22] One of the most widely used features on these wearables includes the detection of atrial fibrillation (AF), namely via smartwatches.[22] These devices monitor heart rate (HR) and rhythm through either photoplethysmography (PPG) or a single-lead ECG, or both, where the time intervals between heartbeats are calculated and various algorithms are used to classify the heart rhythm.[22-23] PPG works by illuminating the skin with a light-emitting diode (LED) and detecting the amount of light reflected, which varies according to changes in blood volume during the cardiac cycle.23 A photodetector measures the intensity of the light reflected, building a pulse pressure waveform, which is not only used to calculate HR, but importantly, the time between each heartbeat corresponding to the R-R interval.[23] In AF, the pulse waveform is highly irregular, representing an irregular R-R interval, notifying the user of suggestive AF through an irregular pulse notification (IPN), which could be useful for paroxysmal AF.[23-24] Nevertheless, this technology is only suggestive and cannot confirm AF or other arrhythmias.[23] Devices capable of performing single-lead ECGs work by using two metal plates to create one positive electrode (often located on the back of the watch) and one negative electrode (often located on the digital crown), thereby allowing measurement of Lead I.[24] These devices can detect arrhythmias such as AF with more accuracy than those with only PPG capacity.[22] However, the main limitation of performing an ECG using such wearables is that only a 1-lead view is ever available.[24] Ultimately, this means that any abnormalities that would be seen in other leads are missed, and also increases the risk of artefacts such as poor sensor-skin contact and muscle motion, which could prevent accurate ECG readings from being recorded.[24]

To address these concerns, several wearables now incorporate DL detection algorithms to increase the accuracy of ECG recordings and PPG signals.[23] These algorithms aid in noise reduction, normalization of data and segmentation to increase the accuracy of PPG tachograms and ECG waveforms.[23] Moreover, CNNs can be employed not only to categorize unprocessed data, but also to detect complex patterns from PPG signals and spatial pattern recognition from ECG traces.[23] Such CNNs provide an additional advantage of detecting and analyzing dynamic changes throughout the day, allowing for continuous, passive ECG monitoring, provided that enough data is stored by the wearable.[23]

AI use in intravascular imaging

In addition to AI use for non-invasive investigations, there is potential for AI in invasive imaging modalities such as intravascular ultrasound (IVUS) and optical coherence tomography (OCT), as highlighted in ***Figure 1***.[25] Although both involve reconstructing images of intracoronary structures via a catheter inserted into coronary arteries, IVUS uses ultrasound, while OCT uses low-coherence light.[26] Additionally, IVUS offers deeper penetration of vessel walls (inclusive of adventitia) than OCT, making IVUS a useful investigation for arteries with increased plaque burden, albeit the low-resolution images.26 In contrast, OCT provides high-resolution images compared to IVUS, offering greater detection rates of thin-cap fibroadenoma (TCFA), arterial plaque rupture and stent malapposition.[25-26]

Currently, with OCT images, automation is primarily limited to the segmentation of atherosclerotic plaques, where the quantification and characterization of any detected plaques are performed.[25,27] Similarly, with IVUS, DL algorithms can assist in feature extraction to increase the detection rates of TCFA.[25] However, interpretation of these intravascular images requires a clinician with extensive training, and can be repetitive after reviewing several images.[25] Implementing DL algorithms could help to relieve experts of this repetitive task, and also allow real-time analysis of intravascular images.[25] Going further, researchers have developed neural networks for this task, such as the one developed by Chu et al., where a neural network could automatically segment a single OCT frame in a remarkable 0.07 seconds.[25,28] Furthermore, pixel-based DL algorithms could allow for the incorporation of three-dimensional spatial data and also the segmentation of individual plaque components.[25] This could allow for more detailed and accurate identification of different plaques, aiding in clinical decision-making.[25,27]

However, despite the potential benefits of using AI models for intravascular images, more research is required before integration into mainstream clinical practice. As for AI use in ECG and echocardiography interpretation, AI for IVUS and OCT still requires large annotated datasets to test external validation on other datasets.[25] Further, models must be tested on datasets encompassing the wider population to ensure that they have similar performance in real-world clinical practice.[25]

AI use in Gastroenterology: Augmenting disease detection from imaging-based investigations

With high-definition photographic visuals being a crucial component for procedures in gastroenterology, the integration of CNNs into picture-based investigations could prove highly useful, as shown in ***Figure 1***. Currently, colonoscopy and small-bowel capsule endoscopy (SB-CE) are some of the most popular investigations to integrate AI models.[29-30]

Colonoscopy

The integration of AI models into colonoscopies could be useful in identifying polyps. With current estimates indicating that physicians can miss colorectal polyps in colonoscopy up to 28% of the time, this type of AI could help to reduce the chances of missing such lesions.[31]

One way that AI is being integrated into colonoscopy is through devices such as 'GI Genius', a medical device that is built on deep learning for computer aided diagnosis (CAD) and is approved for use in the United States and the European Union.[29] The usefulness of this device was tested in a large, randomized multi-center trial where adenoma miss rate (AMR) was calculated for colonoscopies done with or without AI to identify colorectal neoplasia, a risk factor for the development of colorectal cancer.[29] Wallace et al. conducted this study in 2 groups, with 2 different arms, where in 1 group, colonoscopy was done with AI-enabled (GI-Genius enabled), followed by colonoscopy without AI-enabled, and vice versa in the other group.[29] Using this design, Wallace et al. showed that AMR was significantly lower at 15.5% in the group with AI first, compared to 32.4% AMR in the group with colonoscopy first, which is more than a 2-fold difference.[29] With the necessary further studies, AI could be used in colonoscopy to aid physicians, reducing the risk of missing colorectal polyps.

Capsule endoscopy

With the success of AI in detecting colorectal polyps, there is great potential to incorporate AI into endoscopies, helping to increase accuracy and consistency in detecting gastrointestinal lesions.[30]

The introduction of capsule endoscopy (CE) provided a breakthrough for gastroenterologists to investigate the small bowel in a non-invasive manner for conditions such as blood content, vascular lesions, and inflammatory bowel diseases.32 Although CE is beneficial in diagnosing and managing small bowel diseases, analyzing full-length CE videos with approximately 50,000 images can be a tedious and time-consuming task.[32] This can take between 30-120 minutes per video, leading to physicians reviewing CE videos at a great pace, with a recent study reporting a CE miss rate of 11% for all SB findings and 18.9% for single-mass lesions.[32-33] Using AI to aid specialists in reviewing CE videos could help reduce the time taken and could improve the miss rate of SB lesions.[34]

Ongoing research to implement AI systems into the analysis of SB-CE videos has shown promising potential. Ding et al. developed a CNN model to aid in the detection of multiple SB conditions, including ulcers, polyps, inflammation, vascular lesions, and lymphangiectasia through SB-CE.[34] The CNN model outperformed physicians with a higher sensitivity for per-patient analysis (99.88% vs 74.57%, respectively) and per-lesion analysis (99.90% vs 76.89%, respectively).[34] Additionally, the CNN model achieved a substantially shorter reading time than the physicians (5.9 ± 2.23 minutes vs 96.6 ± 22.53 minutes, respectively).[34] With a CNN reading time that is over 90 minutes less than conventional reading, the use of the model in clinical practice could potentially save significant time when reviewing SB-CE videos.

Further, a study highlighted the use of a CNN model for the detection and classification of SB lesions with hemorrhagic potential using CE images.[32] Similar to the study conducted by Ding et al., the researchers showed a high overall accuracy of 99%, sensitivity of 88%, and specificity of 99%.[32,34] Crucially, in addition to identifying lesions, this CNN model could also classify lesions from CE images, suggesting that such CNNs could soon have the capacity to classify other SB lesions.[32] If introduced into clinical practice, these CNNs could become key players for SB-CE analysis with minimal input from physicians, thereby reducing their workload, and importantly, could address the physician miss rate of SB lesions from CE.[33]

Nevertheless, in both studies, the researchers report several limitations, with the most important one being that still frames were used as opposed to moving images.[32,34] In reality, SB-CE provides moving, full-length videos, and further research is required to evaluate the true performance of these AI models on full-length SB-CE videos, helping to assess generalizability. Additionally, these studies have assessed the performance of CNNs on retrospective, existing data.[32,34] Moving forward, as highlighted by several researchers, prospective studies are necessary and are currently in progress to accurately evaluate the true clinical benefit of these AI-based models in patient care.[34]

AI use in Dermatology: Improving skin cancer detection

Using CNN models for image interpretation and pattern analysis can aid in recognizing skin conditions, especially skin cancers, as outlined in ***Figure 1.***[35]

Researchers trained a CNN model using over 129,000 clinical images comprising over 2,000 skin diseases to distinguish between melanocytic and keratinocytic lesions.[35-36] The CNN model trained to classify epidermal and keratinocytic lesions achieved an accuracy of over 91%, and performed on par or even outperformed 21 board-certified dermatologists using clinical images.[35] Furthermore, the ability of the CNN model to classify melanomas using dermoscopic images, as opposed to clinical images, was also matched to the accuracy levels of dermatologists.[35] In another similar study from China, researchers also demonstrated high diagnostic values for a novel CNN model to recognize certain skin diseases from a dataset comprising 14 different categories of common cutaneous diseases.[37] Similar to research conducted by Esteva et al., this CNN also showed a high overall accuracy of 94.8%, with a sensitivity of 93.4% and a specificity of 95.0%.37 Although in a separate test against 280 board-certified dermatologists with 200 different images, the CNN and the dermatologists both had like-for-like figures for average accuracy (92.75% vs 92.13%) and specificity (94.07% vs 95.50%), the sensitivity was significantly higher compared to dermatologists (83.50% vs 68.51%).[37] These figures again emphasize the potential for the CNN model to perform at the same competency as dermatologists, and at times, at higher levels.[37]

The positive results of these trials suggest that AI could be used to classify skin cancers and even aid dermatologists in reducing workload and providing diagnoses. Moreover, Esteva et al. denote that the ability of the CNN model to classify lesions using clinical images as opposed to widely used dermoscopic images could be highly useful in introducing the technology to a smartphone app.[35] With the increasing use of smartphones across the world, such AI could be integrated into apps, providing skin lesion classification from just a smartphone camera.[35] This, in turn, could allow skin conditions, such as cancers, to be detected and classified in primary care, thereby allowing general practitioners to prioritize urgent or non-urgent referrals to dermatologists in secondary care.[35]

Down the line, this could allow for improved management of these conditions, reducing the risk of cancer development, and thereby providing a great public health benefit.[38] In addition, prioritizing appointments in this way would free up time for dermatologists, allowing them to use that time to complete other tasks or see more patients, as seen in a recent pilot project where the Skin Analytics AI-Powered Teledermatology was reviewed by the University Hospitals of Leicester NHS Trust in the UK.[32] This AI tool, known as DERM, which recently received conditional recommendation for use by the National Institute for Health and Care Excellence (NICE), reviewed skin lesion images and classified them as either "diagnosis of concern" or "benign".[39-40] The project showed there was a reduction in the NHS two-week wait referrals for cancers while freeing up 1,450 outpatient appointments, and achieving a clinical time saving of 263 minutes per 100 patients.[41] This pilot project showed promising results, suggesting there could be several benefits from such AI-based technologies, provided that more encouraging studies are carried out. Further real-world prospective studies are still required to evaluate the true benefit in the clinic and address any potential challenges before this AI technology can be introduced into clinical practice.[35]

AI use in Ophthalmology: Enhancing pathology detection from retinal imaging

Retinal imaging is used widely to diagnose several retinal pathologies, including diabetic retinopathy (DR), glaucoma, and age-related macular degeneration (AMD).[42] To make a diagnosis, an ophthalmologist is required to manually analyze and evaluate retinal fundus images, which is a time-consuming process, as with other imaging-based investigations in other specialties.[43] CNN models that can automatically analyze fundus images and categorize the pathology can assist ophthalmologists in making a diagnosis and point towards a management plan, as shown in ***Figure 1.***

In a recent study, Pandey et al. trained an ensemble of 5 CNNs to recognize and classify retinal pathologies into the following 4 categories: DR, glaucoma, AMD, and normal eyes from 100 unseen fundus images.[42] With the performance of the CNN directly compared to 7 board-certified ophthalmologists, the CNN had a higher overall accuracy over all 4 categories, with a score of 79.2%, while the doctors scored 72.7%.[42] Furthermore, the CNN ensemble had a higher overall score for correctly classifying images as DR, with a mean score of 76.8%, while the doctors had a mean score of 57.5%.[42] The remarkable difference of over 19% highlights the impressive accuracy of the CNNs, which outperformed ophthalmologists.[42] The CNN ensemble and ophthalmologists had similar classification scores for the other 3 categories, which were not statistically significant.[42] These favorable results emphasize the great potential for using CNNs and other deep-learning algorithms to aid in the detection of retinal pathologies.

Nevertheless, although Pandey et al. used unseen clinical images to train the CNN, using only 100 images is not representative of all the pathologies seen in the real world.[42] In clinical practice, ophthalmologists often see a range of retinal diseases, more than the 4 conditions used in the study.[42] Furthermore, the images seen may not always be of high quality that CNNs can read with ease. Hence, CNNs trained on a large variable dataset are needed to target generalizability and avoid overfitting to one particular dataset, allowing them to be integrated for mainstream clinical use.

AI use in Radiology: Improving image interpretation and disease detection

The use of AI through CNNs has grown massively in radiology, with extensive research highlighting recent advances in image interpretation. Radiological imaging is an essential aspect of medical care, assisting clinicians in making crucial decisions in the management of a patient. Medical imaging consists of various modalities, including X-ray, ultrasound, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET), which are central to various clinical specialties.[44] Traditionally, radiologists manually analyze these images and provide a report for clinicians to use in their decision-making process.[45] Extensive research has been carried out to evaluate the use of AI models to automate the analysis of radiological images to an extent, saving time for radiologists and potentially increasing accuracy.[46]

Research using CNNs for disease detection and classification has been promising, with prominent advances highlighted in ***Figure 1.*** Several studies have shown that CNNs can be trained to recognize certain patterns, shapes and contours in imaging to identify certain pathologies.[45] One meta-analysis comprising 20 studies highlighted that DL models could identify intracranial aneurysms, with excellent accuracy in detecting aneurysms more than 3 mm in size.[47] More importantly, as highlighted by Abdollahifard et al., using AI to assist clinicians with detecting intracranial aneurysms increased the clinicians' sensitivity by 12.8%.[47] Carefully using such AI systems to support radiologists could lead to better interpretation of medical imaging and increased accuracy. Going further, other studies detail the use of AI models to detect and classify disease on imaging. AI models such as Brainiomix and iSchemaView automatically complete the Alberta Stroke Program Early CT Score (ASPECTS) from non-contrast CT scans (NCCT) for patients with an acute stroke.48 Grading using the ASPECTS system can be difficult due to the subtlety of ischemic changes on NCCTs, with a variable interobserver agreement, indicating that the grading of certain strokes may not always be unanimous.[48] Implementing such AI-based software into clinical practice can help to enhance consistency and reduce diagnostic uncertainty in medical imaging.

Additionally, CNNs could be widely utilized for image segmentation. Image segmentation is essentially the division of a medical image into distinct regions that correspond to specific anatomical or pathological regions, helping to inform clinicians on potential diagnosis and treatment planning.[49] Traditional segmentation algorithms have been in place for several years, allowing radiologists to see abnormalities in imaging. However, the performance of these segmentation methods is complicated by complex medical images with unclear contours, ambiguous boundaries and variations in intensity, leading to poor images.49 Implementing CNNs to automatically extract features from highly complex, three-dimensional images to identify and outline pathological regions, and deliver superior performance can go a long way in detecting lesions.[49] One recent meta-analysis comprising nine studies highlights the strong performance of CNNs in meningioma segmentation from MRI scans, with a pooled Dice score of 89%.50 Notably, CNN models trained on multiple MRI sequences performed better than those trained on single MRI sequences, emphasizing the need for high-quality datasets to develop robust and clinically viable CNN models.50 Interestingly, the authors noted that the dataset size did not significantly impact the accuracy of the CNN models, underlining that the quality of the dataset outweighs the quantity of the data available.[50] However, the lack of performance differences could suggest possible overfitting to the specific datasets used, which could ultimately limit generalizability.[50]

Concerns around generalizability are significant, limiting the widespread integration of such AI models into clinical practice.[51] Several CNN models have been developed to assist with diagnostic tasks, and many excel with remarkable accuracy and performance. However, when implementing the same model on a different dataset or another task with minor differences, effective generalization and similar performance are not always seen.[51] Similarly, the lack of numerous well-annotated datasets presents a challenge in effectively training CNNs.51 Using large-sized, well-labeled datasets helps to train CNNs to recognize complex patterns and features and also aids in reducing overfitting.[52] However, these annotations are only completed by radiologists, and being a significantly lengthy and tedious process has contributed to the scarcity of well-labeled datasets.[52] Using data augmentation techniques by applying transformations can help not only to expand the dataset, but also to increase the diversity of the existing data.51 This can help to train CNNs on a more robust dataset and avoid learning patterns and features only from the original dataset, addressing generalizability.[53]

AI use in Surgery: Providing intraoperative assistance in robot-assisted surgery

Using the principles of image recognition and classification, CNN models could also be incorporated into surgical techniques, allowing for improved patient outcomes, as shown in ***Figure 1.*** Several studies described in this review have highlighted that AI can be useful in the detection and classification of pathologies into different categories.[32,35,42] This capability of CNN models can be useful for the pre-operative assessment of a patient's clinical condition, before proceeding with surgery.

There is great potential for integrating AI intraoperatively, though research is still emerging on this aspect of AI integration into surgery. One example includes an ML-based model that can predict the risk of developing hypoxemia, assisting anesthesiologists in anticipating such an event and proactively intervening.[54] Such AI tools can augment clinical decision-making during surgery, helping to improve patient safety.[54] Meanwhile, further research into intraoperative AI use has shown significant advancements in robot-assisted surgery, with autonomous and semi-autonomous surgeries coming to light due to the integration of algorithms and computer vision.[55] In recent years, there have been advances in orthopedics where specialists use robots such as the MAKO system for semi-autonomous robotic-arm assisted total knee arthroplasty (RATKA).[56] This system runs on complex pre-operative planning through a computer program to map the joint in a three-dimensional view.[56] From this, the surgeon guides the robotic arm to operate within the pre-defined areas, reducing the chances of an accident.[56-57] Moreover, a study highlights that RATKA can also lead to superior surgical precision and better positioning of implants than manual surgical methods.56 These outcomes suggest that using such robotic surgical systems could lead to better outcomes and greater quality of care.

AI models can also be used to improve visuals of the surgical field.[58] During surgery, electrocautery devices are used for dissection and ligation of tissues, which subsequently creates smoke that can obscure the surgical field temporarily.[58] Wang et al. proposed a CNN model, linked to a Swin transformer that can remove surgical smoke from the surgical footage in robotic surgery, improving image quality and producing a smoke-free surgical view.[59] Augmented reality (AR) surgery is another surgical discipline that has the potential to enhance a surgeon's capabilities by delivering real-time information in the surgeon's field of view to improve accuracy and safety.60 Differentiating between native tissue and non-native surgical tools is a crucial challenge in AR for robotic surgery, which could be addressed by CNN models as shown by De Backer et al.[60] This DL model developed for AR-guided robot-assisted kidney transplantation, achieved an impressive Dice score of 97.1% in correctly identifying surgical instruments, suggesting that such models can be integrated into AR surgery.[58] Furthermore, studies highlight using CNN models to complete real-time robotic suturing.58 Saeidi et al. developed a CNN model integrated into a robotic system, which could complete fully automated laparoscopic bowel anastomoses.[58,61] Compared to manual laparoscopic surgery and traditional robot-assisted surgery, the model showed superior consistency and accuracy when considering metrics such as needle placement, suture placement and completion time.[61]

Despite promising studies highlighting advances in using AI models in surgery, there has not been enough research carried out to integrate such technologies into mainstream surgery confidently. Applying AI to surgical fields can present a substantial challenge as surgical interventions rely heavily on a surgeon's practical skills, often in a high-risk, high-pressure, and highly dynamic operating theatre.[62] AI systems have not reached this capability yet, and further studies exploring this critical issue are pertinent. Moreover, training AI algorithms requires a large range of annotated surgical images taken from real-time surgeries. The scarcity and difficulty in obtaining such data from surgical environments present a further challenge, hindering the training of AI-based models.[62] Accountability remains a crucial issue that needs to be addressed to integrate AI-based surgical technologies into real-world clinical practice.[58] With the use of AI spanning from diagnosis, treatment planning, to robot-assisted procedures, it can be difficult to determine who would be responsible if a negative patient outcome occurs.[58] Further clinical research is necessary to address these limitations and integrate AI into surgical fields safely and effectively.

Current challenges, emerging AI models and pathways to clinical integration

The existing literature has shown that there have been significant advancements in AI systems to assist doctors, perhaps allowing for the provision of better care. However, one major limitation preventing the integration of AI models into mainstream clinical practice is the "AI blackbox theory".[63] The theory stems from the issue that although many AI models, such as CNNs, are shown to be highly effective in various specialties, they perform tasks through highly complex computational layers, making interpretability difficult.[64] Essentially, this leads to situations where the rationale behind decisions and recommendations from AI models cannot be explained, which is crucial for patient safety and clinician trust.[63] This also raises concerns about allocating responsibility if an AI system contributes to an adverse patient outcome, especially if clinicians cannot explain AI recommendations.[65] Furthermore, the lack of explainability of "blackbox" models can lead to issues with transparency requirements set by regulatory authorities when looking for integration into clinical practice.[63,66]

These issues have prompted research into Explainable AI (XAI), where AI models use tools such as saliency maps or heatmaps to improve the interpretability of algorithms and how certain decisions are reached.[67] These tools can improve AI models' interpretability and help reduce the "blackbox" nature.[68] However, as highlighted by Ghassemi et al., not all XAI models provide clinically relevant explanations.[69] Some XAI methods, such as LIME and SHAP, generate rationales including heat maps to explain decisions, but these lack medical causality and instead provide technical reasons.[69] Moreover, significant XAI research has been done using retrospective studies, which do not test real-world utility.[69] Further research is required to develop transparent models that can give explanations grounded in medical science and trained on real-world data to improve confidence among clinicians and patients. Real-world, large-dataset prospective studies with such models are critical to validate their clinical benefits and facilitate their integration into mainstream clinical practice.[69]

Another consideration is the recent emergence of open-source large language models (LLM) such as DeepSeek, which has altered the clinical AI landscape.[70] Many other LLMs rely on expensive application programming interfaces (API) or external cloud infrastructure, making it difficult for resource-limited healthcare institutions to access AI technologies, potentially widening global health disparities.[70-71] DeepSeek, unlike other LLMs, enables local deployment, allowing institutions to run such LLMs on their own network and has capabilities for continuous learning from publicly available open-source datasets.[70] This can allow adoption by healthcare institutions without the financial burden of costly APIs or cloud subscriptions.[70] Moreover, DeepSeek supports offline deployment, avoiding the need to transmit sensitive patient information through third-party servers, strengthening data security.[70] The cost-saving and data privacy benefits have already led to over 90 Chinese tertiary hospitals adopting DeepSeek for diagnostic image analysis, administrative tasks, and clinical decision support.[72] Nevertheless, despite these advantages with models like DeepSeek, these LLMs must be thoroughly investigated to ensure data privacy is intact and AI hallucinations do not lead to incorrect outcomes.[73]

Despite several studies showing great potential for AI to assist clinicians in providing enhanced healthcare, it is also paramount that other ethical considerations of using AI technology are not overlooked.[74] While using AI to process sensitive patient data can be beneficial, it is important to ensure that there are robust data protection measures to protect patient information.[74] Additionally, using AI to aid in the decision-making process should also be balanced with important input from clinicians.[75] Crucially, AI must explicitly serve as a tool to enhance the decision-making process, rather than replace human judgment.[71] Using clinician expertise to identify errors made by AI and considering patient preferences should be of the utmost priority to provide the best patient-centered care.[74] Moving forward, further ethical considerations should be taken into account to make the most of AI in clinical practice, while maintaining the highest ethical standards.

Limitations of this Narrative Review

The fact that this paper is a narrative review poses a limitation, as there was no quantitative question being addressed. Nevertheless, the literature was reviewed and summarized using the SANRA guidelines to effectively address the qualitative research question outlined in the introduction. Additionally, the uses of AI were not covered in every medical and surgical specialty, as there would be too much to include in one narrative review. The specialties that had major developments were selected, and key AI advances were covered in this review

Conclusion

Integrating AI into the clinical setting can revolutionize healthcare in countless ways with recent studies showing encouraging results. The reviewed literature for this article has demonstrated a massive potential for the use of AI in several specialties that could lead to better patient care. There has been extensive research into the applications of AI in medicine and surgery, namely through the integration of deep learning algorithms with some systems using computer vision. Medically, researchers have achieved encouraging results in developing CNN models to detect, recognize, and classify clinical images to aid physicians in various disciplines and specialties. Surgical studies have highlighted using AI models in robot-assisted surgery to guide surgeons, helping to enhance accuracy and reduce the risk of complications, leading to better outcomes. However, despite these studies, more research is required to move ahead and implement AI into everyday clinical care, as many developed CNN models are still being tested on existing data from retrospective studies. Other prominent limitations in existing research include the development of "blackbox" models that lack interpretability, and limited models assessed on generalizability and external validation. Long-term, real-world prospective studies comprising diverse datasets are imperative to assess the true clinical benefits and address any potential drawbacks and limitations before AI can be introduced into mainstream clinical practice.

Summary – Accelerating Translation

Artificial intelligence (AI) is currently being used in several sectors around the world to automate tasks. Clinical medicine is one field where AI can be beneficial to automate tasks, with several studies demonstrating that AI models can carry out tasks with impressive accuracy and efficiency. The main aim of this article is to evaluate the recent advances and applications of various types of AI models in different clinical specialties. Research was carried out through analysis of numerous peer-reviewed articles taken from online medical research databases, such as PubMed.

The research showed increasing evidence of AI models, such as convolutional neural networks (CNN), having the capacity to carry out complex tasks to aid physicians in their clinical decision-making across several specialties. CNNs, a type of deep learning (DL) model, can inherently recognize and classify patterns, making them great candidates for use in diagnostic investigations. In cardiology, several researchers showed the potential for CNNs to aid clinicians through automated analysis of electrocardiograms (ECGs) and echocardiography, allowing for recognition of various pathologies. Furthermore, DL algorithms can be applied to wearable devices, such as smartwatches, allowing them to passively monitor for arrhythmias with precision, which could be useful for clinicians to review if widely adopted. Other research also shows the potential for AI models to assist in complex intravascular imaging by recognizing image components with more accuracy, allowing for greater detection of cardiac arterial pathologies. In gastroenterology, greater detection rates of abnormal growths and other pathologies through CNN-based colonoscopy could lead to earlier identification of colorectal cancers. Similarly, CNNs can be useful in analyzing captured footage from capsule endoscopy, where a small camera attached to a pill is swallowed to take pictures and videos of the small intestine. This analysis is a typically time-consuming task for physicians, and CNNs could help by detecting abnormalities at a much greater pace and accuracy. Comparably, the same principle applies in ophthalmology, where CNNs could be used to evaluate retinal images, aiding ophthalmologists in detecting more pathologies of the eye.

In dermatology, research shows that CNNs can aid in recognizing skin cancers from clinical images with ease and accuracy, allowing for earlier detection rates. Moreover, recent studies show the integration of CNNs into software to detect and analyze images taken from a smartphone camera, allowing the expansion and revolution of teledermatology. Radiology is one field where the use of AI models is substantial and could be greatly beneficial. As with intravascular imaging, CNNs and other AI models could help radiologists in analyzing shapes and contours with more precision, allowing for better detection and classification of pathologies from imaging. These same principles could aid in surgical techniques and robot-assisted surgery, where AI models view and analyze the surgical field before the surgery is commenced, assisting with the surgeon's precision. Furthermore, AI models could help to provide a clear view for the surgeon in augmented reality (AR) surgery, allowing for improved accuracy and safety. Remarkable studies have also shown that robotic surgery equipped with AI models can complete automated suturing, providing higher accuracy and consistency than traditional surgical techniques.

Recently, there has also been the introduction of large language models (LLMs) which can provide quick support and analysis for administrative tasks, diagnostic imaging and clinical decision support. Models such as DeepSeek can also run offline or on institutions' own networks, which provides a substantial financial benefit for resource-limited institutions and hospitals where expensive online services cannot be afforded. Nevertheless, despite the added advantage of these LLMs, further extensive research must be done to ensure that there are robust data protection measures.

Despite researchers showing the great potential for AI to be used in clinical practice, several challenges and concerns still exist that are preventing widespread clinical integration. The most significant limitation is the "AI blackbox theory", where the reasoning behind certain AI responses and decisions cannot be explained. This raises concerns around patient safety and clinician trust, and allocating responsibility if AI contributes to an adverse patient outcome. Although efforts have been made to introduce explainable AI (XAI) models, these are still not enough, as XAI models do not always provide medical rationale behind clinical recommendations. Furthermore, concerns remain around generalizability and external validation, where AI models do not have similar performance with different datasets. Additionally, many studies assessing AI models have been carried out on existing retrospective datasets. More research needs to be focused on real-time, large-dataset, diverse prospective studies with transparent AI models, where patient data and outcomes need to be followed to validate their true clinical benefit and facilitate mainstream clinical adoption.

References

1. Reddy S, Fox J, Purohit MP. Artificial intelligence-enabled healthcare delivery. J R Soc Med. 2019;112(1):22–8.
2. Arooj S, Rehman S ur, Imran A, Almuhaimeed A, Alzahrani AK, Alzahrani A. A deep convolutional neural network for the early detection of heart disease. Biomedicines. 2022;10(11):2796.
3. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Gastrointest Endosc. 2020;92(4):807–12.
4. Yu AC, Mohajer B, Eng J. External validation of deep learning algorithms for radiologic diagnosis: a systematic review. Radiol Artif Intell. 2022;4(3):e210064.
5. Park SH, Choi J, Byeon JS. Key principles of clinical validation, device approval, and insurance coverage decisions of artificial intelligence. Korean J Radiol. 2021;22(3):442–50.
6. Buschman H. Artificial intelligence enables rapid COVID-19 lung imaging analysis at UC San Diego Health. 2020.
7. Henry KE, Adams R, Parent C, Soleimani H, Sridharan A, Johnson L, et al. Factors driving provider adoption of the TREWS machine learning-based early warning system and its effects on sepsis treatment timing. Nat Med. 2022;28(7):1447–54.
8. Pham TQ, Matsui T, Chikazoe J. Evaluation of the hierarchical correspondence between the human brain and artificial neural networks: a review. Biology. 2023;12(10):1330.
9. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intell. 2022;15(1):1–22.
10. Kulkarni S, Seneviratne N, Baig MS, Khan AHA. Artificial intelligence in medicine: where are we now? Acad Radiol. 2020;27(1):62–70.
11. Krittanawong C, Johnson KW, Rosenson RS, Wang Z, Aydar M, Baber U, et al. Deep learning for cardiovascular medicine: a practical primer. Eur Heart J. 2019;40(25):2058–73.
12. Cook DA, Oh SY, Pusic MV. Accuracy of physicians' electrocardiogram interpretations: a systematic review and meta-analysis. JAMA Intern Med. 2020;180(11):1461–71.
13. Hughes JW, Olgin JE, Avram R, Abreau SA, Sittler T, Radia K, et al. Performance of a convolutional neural network and explainability technique for 12-lead electrocardiogram interpretation. JAMA Cardiol. 2021;6(11):1285–95.
14. Yoon T, Kang D. Bimodal CNN for cardiovascular disease classification by co-training ECG grayscale images and scalograms. Sci Rep. 2023;13(1):2937.
15. Makimoto H, Höckmann M, Lin T, Glöckner D, Gerguri S, Clasen L, et al. Performance of a convolutional neural network derived from an ECG database in recognizing myocardial infarction. Sci Rep. 2020;10(1):8445.
16. Attia IZ, Tseng AS, Benavente ED, Medina-Inojosa JR, Clark TG, Malyutina S, et al. External validation of a deep learning electrocardiogram algorithm to detect ventricular dysfunction. Int J Cardiol. 2021;329:130–5.

17. Chan JSK, Tse G, Zhao H, Luo XX, Jin CN, Kam K, et al. Echocardiography update for primary care physicians: a review. Hong Kong Med J. 2020;26(1):44–55.
18. Madani A, Ong JR, Tibrewal A, Mofrad MRK. Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. NPJ Digit Med. 2018;1(1):1–11.
19. Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view classification of echocardiograms using deep learning. NPJ Digit Med. 2018;1(1):1–8.
20. Naser JA, Lee E, Pislaru SV, Tsaban G, Malins JG, Jackson JI, et al. Artificial intelligence-based classification of echocardiographic views. Eur Heart J Digit Health. 2024;5(3):260–9.
21. Biondi-Zoccai G, D'Ascenzo F, Giordano S, Mirzoyev U, Erol Ç, Cenciarelli S, et al. Artificial intelligence in cardiology: general perspectives and focus on interventional cardiology. Anatol J Cardiol. 2025;29(4):152–63.
22. Bayoumy K, Gaber M, Elshafeey A, Mhaimeed O, Dineen EH, Marvel FA, et al. Smart wearable devices in cardiovascular care: where we are and how to move forward. Nat Rev Cardiol. 2021;18(8):581–99.
23. Papalamprakopoulou Z, Stavropoulos D, Moustakidis S, Avgerinos D, Efremidis M, Kampaktsis PN. Artificial intelligence-enabled atrial fibrillation detection using smartwatches: current status and future perspectives. Front Cardiovasc Med. 2024;11:1432876.
24. Smith S, Maisrikrod S. Wearable electrocardiogram technology: help or hindrance to the modern doctor? JMIR Cardio. 2025;9:e62719.
25. Föllmer B, Williams MC, Dey D, Arbab-Zadeh A, Maurovich-Horvat P, Volleberg RHJA, et al. Roadmap on the use of artificial intelligence for imaging of vulnerable atherosclerotic plaque in coronary arteries. Nat Rev Cardiol. 2024;21(1):51–64.
26. Ono M, Kawashima H, Hara H, Gao C, Wang R, Kogame N, et al. Advances in IVUS/OCT and future clinical perspective of novel hybrid catheter system in coronary imaging. Front Cardiovasc Med. 2020;7:119.
27. Subhan S, Malik J, Haq A ul, Qadeer MS, Zaidi SMJ, Orooj F, et al. Role of artificial intelligence and machine learning in interventional cardiology. Curr Probl Cardiol. 2023;48(7):101698.
28. Chu M, Jia H, Gutiérrez-Chico JL, Maehara A, Ali ZA, Zeng X, et al. Artificial intelligence and optical coherence tomography for the automatic characterisation of human atherosclerotic plaques. EuroIntervention. 2021;17(1):41–50.
29. Wallace MB, Sharma P, Bhandari P, East J, Antonelli G, Lorenzetti R, et al. Impact of artificial intelligence on miss rate of colorectal neoplasia. Gastroenterology. 2022;163(1):295–304.e5.
30. El Hajjar A, Rey JF. Artificial intelligence in gastrointestinal endoscopy: general overview. Chin Med J (Engl). 2020;133(3):326–34.
31. Kim NH, Jung YS, Jeong WS, Yang HJ, Park SK, Choi K, et al. Miss rate of colorectal neoplastic polyps and risk factors for missed polyps in consecutive colonoscopies. Intest Res. 2017;15(3):411–8.
32. Mascarenhas Saraiva MJ, Afonso J, Ribeiro T, Ferreira J, Cardoso H, Andrade AP, et al. Deep learning and capsule endoscopy: automatic identification and differentiation of small bowel lesions. BMJ Open Gastroenterol. 2021;8(1):e000753.
33. Shim KN, Jeon SR, Jang HJ, Kim J, Lim YJ, Kim KO, et al. Quality indicators for small bowel capsule endoscopy. Clin Endosc. 2017;50(2):148–60.
34. Ding Z, Shi H, Zhang H, Meng L, Fan M, Han C, et al. Gastroenterologist-level identification of small-bowel diseases and normal variants by capsule endoscopy using a deep-learning model. Gastroenterology. 2019;157(4):1044–54.e5.
35. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–8.
36. Leachman SA, Merlino G. The final frontier in cancer diagnosis. Nature. 2017;542(7639):36–8.
37. Zhu CY, Wang YK, Chen HP, Gao KL, Shu C, Wang JC, et al. A deep learning based framework for diagnosing multiple skin diseases in a clinical environment. Front Med. 2021;8:–.
38. Koka SSA, Burkhart CG. Artificial intelligence in dermatology: current uses, shortfalls, and opportunities for further implementation in diagnostics and care. Open Dermatol J. 2023;17:–.
39. Health Innovation East Midlands. Skin Analytics. AI powered teledermatology shows considerable potential to reduce waiting lists. 2023 [cited 2025 Mar 9]. Available from: URL.
40. The National Institute for Health and Care Excellence. AI skin cancer detection system gets green light for conditional NHS use. 2025 [cited 2025 May 7]. Available from: URL.
41. An evaluation of AI Powered Tele Dermatology for Skin Cancer 2WW Pathway. 2023 [cited 2024 Jul 2]. Available from: URL.
42. Pandey PU, Ballios BG, Christakis PG, Kaplan AJ, Mathew DJ, Tone SO, et al. Ensemble of deep convolutional neural networks is more accurate and reliable than board-certified ophthalmologists at detecting multiple diseases in retinal fundus photographs. Br J Ophthalmol. 2024;108(3):417–23.
43. Bhulakshmi D, Rajput DS. A systematic review on diabetic retinopathy detection and classification based on deep learning techniques using fundus images. PeerJ Comput Sci. 2024;10:e1947.
44. Chen C, Mat Isa NA, Liu X. A review of convolutional neural network based methods for medical image classification. Comput Biol Med. 2025;185:109507.
45. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500–10.
46. van Leeuwen KG, de Rooij M, Schalekamp S, van Ginneken B, Rutten MJCM. How does artificial intelligence in radiology improve efficiency and health outcomes? Pediatr Radiol. 2022;52(11):2087–93.
47. Abdollahifard S, Farrokhi A, Kheshti F, Jalali M, Mowla A. Application of convolutional network models in detection of intracranial aneurysms: a systematic review and meta-analysis. Interv Neuroradiol. 2023;29(6):738–47.
48. Soun JE, Chow DS, Nagamine M, Takhtawala RS, Filippi CG, Yu W, et al. Artificial intelligence and acute stroke imaging. AJNR Am J Neuroradiol. 2021;42(1):2–11.
49. Xu Y, Quan R, Xu W, Huang Y, Chen X, Liu F. Advances in medical image segmentation: a comprehensive review of traditional, deep learning and hybrid approaches. Bioengineering. 2024;11(10):1034.
50. Wang TW, Hong JS, Lee WK, Lin YH, Yang HC, Lee CC, et al. Performance of convolutional neural network models in meningioma segmentation in magnetic resonance imaging: a systematic review and meta-analysis. Neuroinformatics. 2025;23(1):14.
51. Quanyang W, Yao H, Sicong W, Linlin Q, Zewei Z, Donghui H, et al. Artificial intelligence in lung cancer screening: detection, classification, prediction, and prognosis. Cancer Med. 2024;13(7):e7140.
52. Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K, et al. Recent advances and clinical applications of deep learning in medical image analysis. Med Image Anal. 2022;79:102444.
53. Chlap P, Min H, Vandenberg N, Dowling J, Holloway L, Haworth A. A review of medical image data augmentation techniques for deep learning applications. J Med Imaging Radiat Oncol. 2021;65(5):545–63.
54. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. Nat Biomed Eng. 2018;2(10):749–60.
55. Feizi N, Tavakoli M, Patel RV, Atashzar SF. Robotics and AI for teleoperation, tele-assessment, and tele-training for surgery in the era of COVID-19: existing challenges and future vision. Front Robot AI. 2021;8:–.
56. Hampp EL, Chughtai M, Scholl LY, Sodhi N, Bhowmik-Stoker M, Jacofsky DJ, et al. Robotic-arm assisted total knee arthroplasty demonstrated greater accuracy and precision to plan compared with manual techniques. J Knee Surg. 2019;32(3):239–50.
57. Wall J, Krummel T. The digital surgeon: how big data, automation, and artificial intelligence will change surgical practice. J Pediatr Surg. 2020;55S:47–50.

58. Knudsen JE, Ghaffar U, Ma R, Hung AJ. Clinical applications of artificial intelligence in robotic surgery. J Robot Surg. 2024;18(1):102.
59. Wang F, Sun X, Li J. Surgical smoke removal via residual Swin transformer network. Int J Comput Assist Radiol Surg. 2023;18(8):1417–27.
60. De Backer P, Van Praet C, Simoens J, Peraire Lores M, Creemers H, Mestdagh K, et al. Improving augmented reality through deep learning: real-time instrument delineation in robotic renal surgery. Eur Urol. 2023;84(1):86–91.
61. Saeidi H, Opfermann JD, Kam M, Wei S, Leonard S, Hsieh MH, et al. Autonomous robotic laparoscopic surgery for intestinal anastomosis. Sci Robot. 2022;7(62):eabj2908.
62. Bodenstedt S, Wagner M, Müller-Stich BP, Weitz J, Speidel S. Artificial intelligence-assisted surgery: potential and challenges. Visc Med. 2020;36(6):450–5.
63. Abgrall G, Holder AL, Chelly Dagdia Z, Zeitouni K, Monnet X. Should AI models be explainable to clinicians? Crit Care. 2024;28:301.
64. Yang G, Ye Q, Xia J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond. Int J Inf Fusion. 2022;77:29–52.
65. Naik N, Hameed BMZ, Shetty DK, Swain D, Shah M, Paul R, et al. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? Front Surg. 2022;9:862322.
66. Palaniappan K, Lin EYT, Vogel S. Global regulatory frameworks for the use of artificial intelligence (AI) in the healthcare services sector. Healthcare. 2024;12(5):562.
67. Borys K, Schmitt YA, Nauta M, Seifert C, Krämer N, Friedrich CM, et al. Explainable AI in medical imaging: an overview for clinical practitioners – saliency-based XAI approaches. Eur J Radiol. 2023;162:110787.
68. Chaddad A, Peng J, Xu J, Bouridane A. Survey of explainable AI techniques in healthcare. Sensors. 2023;23(2):634.
69. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3(11):e745–50.
70. Temsah A, Alhasan K, Altamimi I, Jamal A, Al-Eyadhy A, Malki KH, et al. DeepSeek in healthcare: revealing opportunities and steering challenges of a new open-source artificial intelligence frontier. Cureus. 2025;17(2):e79221.
71. Jeyaraman M, Balaji S, Jeyaraman N, Yadav S. Unraveling the ethical enigma: artificial intelligence in healthcare. Cureus. 2023;15(8):e43262.
72. Chen J, Miao C. DeepSeek deployed in 90 Chinese tertiary hospitals: how artificial intelligence is transforming clinical practice. J Med Syst. 2025;49(1):53.
73. Zeng D, Qin Y, Sheng B, Wong TY. DeepSeek's "low-cost" adoption across China's hospital systems: too fast, too soon? JAMA. 2025;333(21):1866–9.
74. Elendu C, Amaechi DC, Elendu TC, Jingwa KA, Okoye OK, John Okah M, et al. Ethical implications of AI and robotics in healthcare: a review. Medicine (Baltimore). 2023;102(50):e36671.
75. Scott IA, Carter SM, Coiera E. Exploring stakeholder attitudes towards AI in clinical practice. BMJ Health Care Inform. 2021;28(1):e100450.

Acknowledgments

None.

Conflict of Interest Statement & Funding

The Authors have no funding, financial relationships or conflicts of interest to disclose.

Author Contributions

Conceptualization: SSRK. Data Curation: SSRK. Formal Analysis: SSRK. Investigation: SSRK. Methodology: SSRK. Supervision: DS. Writing - Original Draft: SSRK. Writing - Review Editing: SSRK, DS.

Cite as

Rohit S. S. K, Sunnucks D, Navigating the Digital Frontier: A Review on the Clinical Applications of Convolutional Neural Networks and Emerging AI Models in Medicine and Surgery. Int J Med Stud. 2025 Oct-Dec;13(4): 441-451.

ISSN 2076-6327

This journal is published by [**Pitt Open Library Publishing**]