Journal of Oral Biosciences 64 (2022) 312–320

Contents lists available at ScienceDirect

Journal of Oral Biosciences

journal homepage: www.elsevier.com/locate/job

Review

Deep learning models in medical image analysis

Masayuki Tsuneki [a], [b], *, [1]

[a] Medmain Research, Medmain Inc., Fukuoka, Japan
[b] Division of Anatomy and Cell Biology of the Hard Tissue, Department of Tissue Regeneration and Reconstruction, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan

A R T I C L E I N F O

Article history:
Received 14 February 2022
Received in revised form
8 March 2022
Accepted 9 March 2022
Available online 17 March 2022

Keywords:
Medical image analysis
Computer-aided diagnosis
Computer vision
Deep learning
Artificial intelligence

A B S T R A C T

Background: Deep learning is a state-of-the-art technology that has rapidly become the method of choice for medical image analysis. Its fast and robust object detection, segmentation, tracking, and classification of pathophysiological anatomical structures can support medical practitioners during routine clinical workflow. Thus, deep learning-based applications for diseases diagnosis will empower physicians and allow fast decision-making in clinical practice.
Highlight: Deep learning can be more robust with various features for differentiating classes, provided the training set is large and diverse for analysis. However, sufficient medical images for training sets are not always available from medical institutions, which is one of the major limitations of deep learning in medical image analysis. This review article presents some solutions for this issue and discusses efforts needed to develop robust deep learning-based computer-aided diagnosis applications for better clinical workflow in endoscopy, radiology, pathology, and dentistry.
Conclusion: The introduction of deep learning-based applications will enhance the traditional role of medical practitioners in ensuring accurate diagnoses and treatment in terms of precision, reproducibility, and scalability.

Contents

Abbreviations: CAD, computer-aided diagnosis; AI, artificial intelligence; US, ultrasound; CT, computed tomography; MRI, magnetic-resonance imaging; PET, positron emission tomography; WSI, whole-slide image; TCGA, the Cancer Genome Atlas; CNN, convolutional neural network; GPU, graphic processing unit; ReLU, rectified linear unit; FDA, Food and Drug Administration; ROC, receiver operator curve; AUC, area under the curve; NLP, natural language processing; GAN, generative adversarial network; EUS-FNB, endoscopic ultrasound-guided fine-needle biopsy; RNN, recurrent neural network.

* Board Director & VP of Medical Research Medmain Research, Medmain Inc. 2-4-5-104, Akasaka, Chuo-ku, Fukuoka, 810-0042 Japan.
E-mail address: tsuneki@medmain.com.
[1] **Present/permanent address**: 2-4-5-104, Akasaka, Chuo-ku, Fukuoka, 810-0042 Japan

https://doi.org/10.1016/j.job.2022.03.003

1. Introduction

In the healthcare field, there is a demand to create an environment where patients can receive optimal medical and nursing care in a sustainable and efficient manner. Artificial intelligence (AI) is expected to be used in preventive medicine, diagnostic support, personalized medicine, discovery of new treatments, extension of healthy life expectancy, and nursing care services, which reduce the burden on caregivers. Medical areas where AI is expected to be applied to practical use quickly include genomic medicine, diagnostic imaging support (medical image analysis), diagnosis and treatment support, and drug discovery.

Healthcare is a huge and analog business field, which makes it a target for the world's leading information technology (IT) companies. In 2018, investors in the United States invested more than 11 billion US dollars in digital healthcare startups, a 16% increase over the previous year. Clinical medicine has always required doctors to handle enormous amounts of data (e.g., clinical records, clinical examination, and medical images). In the routine clinical practices, medical image interpretation has mostly been performed by human medical practitioners; however, they have recently begun to benefit from computer-assisted interventions because of various data provided by a diverse range of clinical examinations. The rapid development and use of big data and AI technologies has led to the widespread use of data-based problem-solving processes that enable precise, real-time prediction of various diseases, exhaustive examination of various treatment options, and automatic execution of large–scale complex tasks.

Deep learning methods are highly effective when the number of available data is large during a training stage. Particularly, great improvements in computer vision inspired the use of deep learning in medical image analysis (e.g., segmentation, object detection, classification, prognosis prediction, and microscopic imaging analysis). Among computerized tools, deep learning-based applications are proving to be the state-of-the-art foundation, leading to improved accuracy, which allowed new frontiers in medical image analysis.

This article explains the fundamentals and theoretical approaches of deep learning-based medical image analysis and its applications in clinical practices (e.g., endoscopy, radiology, pathology and dentistry).

2. Application of deep learning in medical image analysis

2.1. Digital image analysis and computer vision

Recently, the combination of powerful hardware and software packages and the fact that nearly everyone has some type of device for digital image acquisition (e.g., smartphone camera, digital camera, and scanner) have increased the number of digital images. Therefore, digital image processing has become a common task. IT professionals especially software engineers and computer scientists are increasingly confronted with developing programs, databases, and associated systems that must deal with digital images.

Computer vision tackles the problem of artificial visual systems with the ability to comprehend and interpret the real world. Major topics in computer vision include object recognition, tracking (motion interpretation), autonomous navigation, scene understanding, and robotic manipulation. Since computer vision has its roots in AI, several AI methods were originally developed to either tackle or represent a problem in computer vision. When computer vision started in the early 1970s, it was viewed as the visual perception component of an ambitious agenda to mimic human intelligence. Today, computer vision is being used in various real-world applications, including machine inspection, optical character recognition, object recognition for automated checkout lanes in retail, 3D model building, medical imaging, automotive safety, match move, motion capture, surveillance (e.g., analyzing traffic), and fingerprint recognition and biometrics (e.g., automatic access authentication). Computer vision attempts to imitate human vision capabilities by providing methods for image formation and machine perception. The modern approach to imitating human cognitive systems consists of machine learning methods that extract information from images. Machine learning is a subfield of AI that teaches computers by showing them a large amount of data and instructing them to learn from these data. The latest trend is the application of sophisticated machine learning techniques to computer vision problems, which coincides with the increased availability of much partially labeled data on the Internet that makes it more feasible to learn object categories without careful human supervision.

2.2. Deep learning for medical image analysis

Medical images are different from other pictures where they depict distributions of various physical features of the human body. These medical images carry different information, and many analytical tasks are related to the individual quantification of entities in the human body. Medical images provide information on the anatomy and physiology of different organs, which is required to precisely diagnose lesions. Recently, the use of big data analytics has substantially increased in healthcare, with medical imaging playing a key role in it. Big data analytics platforms provide great benefits in handling diverse medical images, which include ultrasound (US), X-ray, computed tomography (CT) scans, magnetic-resonance imaging (MRI) scans, positron emission tomography (PET) scans, retinal photography, histopathological and cytopathological whole-slide images (WSIs), and endoscopy and dermoscopy images. Using big data analytics, disease surveillance should be effectively improved. Unstructured medical image data sets can be efficiently evaluated to make better discernments about diseases and necessary preventive and therapeutic methodologies, resulting in much better critical medical decision making. For example, The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/) collected over 20,000 primary cancer and matched normal specimens in 33 cancer types, which generated over 2.5 petabytes of genomic, epigenomic, transcriptomic, and proteomic data since 2006. To handle such a huge amount of image data, dedicated analytics platforms for computations and predictive analysis are necessary to analyze these big data in a distributed environment. Using big data analytic platforms, advanced analytics of medical images allowed much faster diagnosis and prediction of treatment outcomes. To achieve this success, parallel programming and cloud-based computation resources have played a pivotal role in overcoming the challenges of big data computation. To extract features from medical images and recognize patterns in the extracted data for medical image analysis, machine and deep learning techniques provide great benefits in performed the required analyses.

Deep learning is an improvement of artificial neural networks, consisting of multiple processing layers to learn representations of data with multiple levels of abstraction, which have dramatically improved the state-of-the-art in medical image analysis [1]. Deep learning discovers complex structures in large datasets using an algorithm called backpropagation to show how a machine should change its internal parameters used to compute representations in each layer from representations in the previous layers [1]. Particularly, deep convolutional neural networks (CNNs) automatically learn mid- and high-level abstractions from raw data, which are powerful tools for computer vision tasks, including medical image analysis. Moreover, the demand for reducing manual costs in various industries has promoted the growth of automation and computer-aided technologies. Additionally, the availability of low-cost graphical processing units (GPUs) and memory from the video-game industry has made it possible to introduce CNNs with several layers and kernels [2].

Deep learning methods are highly effective when the amount of data is large for training [3]. However, as for medical image analysis, usually, the amount of available data from medical institutions is limited. Additionally, because medical image labeling can be performed by professional medical practitioners, correctly labeled data are even more scarce. These factors prohibited the expansion of CNNs in medical image analysis. Therefore, one of the main challenges in applying deep learning to medical image analysis is the limited number of available training and test datasets to develop models with high accuracy without suffering from overfitting [3]. To reduce overfitting, there are good algorithmic techniques to better train models (e.g., initialization and momentum [4,5], rectified linear unit (ReLU) [6], denoising [7], dropout [8] and dropconnect [9], and batch normalization [10,11]). Moreover, to overcome the aforementioned obstacles, transfer learning is providing great benefits for deep learning models in medical image analysis [12,13]. Transfer learning can help promote deep learning models in less-developed application areas, as well as less technically developed geographical areas, even when not much labeled data are available in medical image analysis. Traditional machine learning trains each model in isolation based on a specific domain (Fig. 1a). As for transfer learning, pretrained supervised CNN models using natural image datasets or different medical domains can be used for developing a new deep learning model without training from scratch, which greatly reduces the training time (Fig. 1b). Because only a small amount of labeled data is required to fine-tune a pre-trained CNN structure, transfer learning is more suitable for developing deep learning models in medical image analysis [11,14,15]. Fine-tuning and feature extractor are the two most common strategies for applying transfer learning in medical image analysis. For fine-tuning, when a small-to-medium dataset does exist for the task at hand, using a pretrained CNN would be possible as the initialization of the network, following which further supervised training of several network layers is conducted, using the new data for the task at hand [14]. In contrast, the approach of feature extractor is freezing all other layers in pretrained CNNs, except for the last few layers, and then splicing its own classifier to reconstruct a new CNN model. If the labeled dataset is scarce, in the feature extractor methods, overfitting would be easier to avoid because this method enables the newly constructed networks to earn the powerful feature extraction performance from the pre-trained network with only low training resources [11]. Based on the transfer learning methodologies, once we establish a well-trained model in one domain, we can bring this model to benefit other similar domains. Therefore, having an accurate distance measure between any task domains is necessary for developing an appropriate transfer learning methodology. If the distance between two domains is large, applying transfer learning is impossible as learning might turn out to produce a negative effect. However, if two domains are near, transfer learning can be effectively applied.

Although CNNs and derivatives are clearly the top performers among most medical image analysis platforms, the exact architecture is not the most important determinant in obtaining a good solution. A key aspect is that professional knowledge on the task to be solved can provide advantages that go beyond adding more layers to a CNN [16]. As deep learning methods have achieved the state-of-the-art performance over different medical applications, its use for further improvement can be a major step in medical image analysis. Deep learning models for medical image analysis have great impacts on both clinical applications and scientific studies.

2.3. Deep learning for computer-aided diagnosis (CAD)

Deep learning is the state-of-the-art approach, which can bring evolutionary changes in healthcare. Medical image analysis plays a key role in clinical decision making at many stages in the patient care process. The potential application for deep learning based medical image analysis is computer-aided diagnosis (CAD), which provides decision-making support to medical practitioners. The attempt of using computers in automated medical image analysis emerged in the 1960s [17–19]. After 30 years, in 1998, the first CAD commercial system was approved by the United States Food and Drug Administration (FDA) for use as a second reader to assist in detecting breast cancer in screening mammography. In the past few decades, CAD has been a major part of research and development in medical image analysis. However, conventional machine learning approaches have limitations where human developers may not be able to translate the complex disease patterns into a finite number of feature descriptors even if they have seen several cases from the patient population; thus, given the limitations of machine learning technology in the early days of CAD, the performance of conventional machine learning-based CAD systems can achieve a sensitivity comparable to that of radiologists but at the expense of relatively high false-positive rates [2,20]. Importantly, deep learning-based medical image analysis brings breakthroughs in CAD performance and allows the widespread use of deep learning-based CAD to various tasks in routine clinical workflow. Since deep learning processes are automated, deep learning models can easily analyze millions of cases without interval. In the last several years, deep learning has been applied to medical image analysis tasks for CAD [16,21,22]. The major areas of deep learning-based CAD application include classification of diseases and normal patterns, classification of malignant and benign lesions, and prediction of high- and low-risk patterns of developing cancers [20]. Although CAD systems using deep learning approaches have not been tested in large-scale clinical trials, the experiences of using CAD in mammography screening may provide some insights into what to expect from CAD tools in clinical practice [2,20]. In the recent challenges of developing CAD methods for various classification tasks in medical images (e.g., Kaggle competitions and DREAM Challenges), all winning teams used deep learning approaches [2]. Although deep learning models can be more accurate and robust than conventional machine learning models in many CAD applications, deep learning algorithms have not been extensively tested in routine clinical settings, where many seemingly ideal hardware and software applications could fail. Unraveling black-box predictions from deep learning and discovering correlations and causal relationships of machine findings to other clinical data from patients will be a key area of investigation [2]. For CAD to be more widely accepted as a clinical decision-making support tool, it

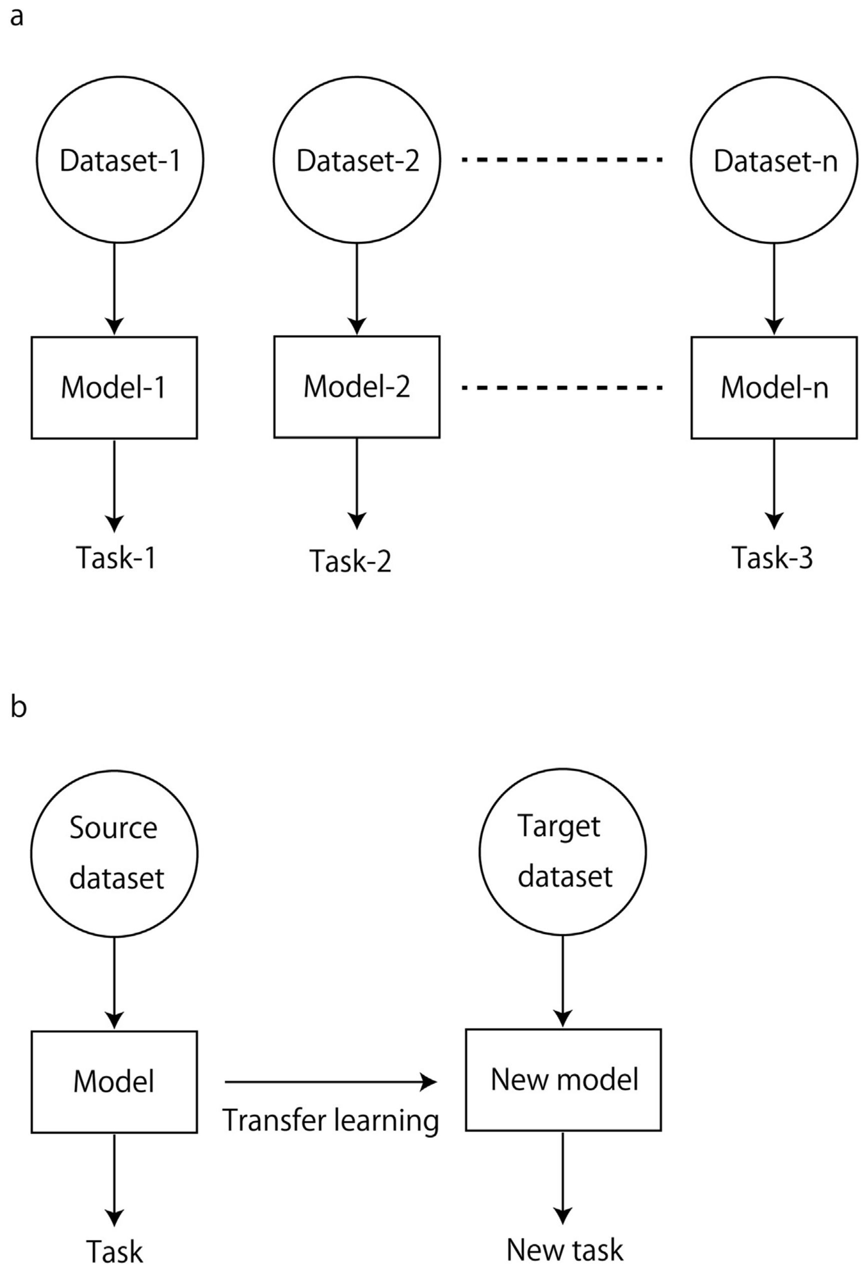

Fig. 1. Schematic diagrams for conventional machine learning approach and transfer learning process. (a): An illustration of conventional machine learning. (b): An illustration of transfer learning.

should ideally be able to more intelligently present recommendations based on inferences to clinicians, correlate the findings with the patient's condition and data, and provide further explanation if the clinician has questions about the CAD recommendations. Uncovering the relationship between machine-generated findings and a patient's medical condition and using deep learning and big data analytics to discover new links between diseases and clinical data are important areas of research that will enable CAD to provide interpretable diagnoses to clinicians and advance CAD toward AI in medical practices [20].

Importantly, on September 21, 2021, the United States FDA authorized the marketing of software to assist pathologists in detecting areas suspicious for cancer as an adjunct to reviewing WSIs from prostate biopsy specimens. The software, called Paige Prostate, is based on a weakly-supervised deep learning algorithm [23] and is the first deep learning-based software designed to identify an area of interest on prostate biopsy images with the highest likelihood of harboring cancer. It has been reported that without Paige Prostate, pathologists had an average sensitivity of 74% and an average specificity of 97%; however, with Paige Prostate, the average sensitivity of pathologists significantly increased to 90% without a statistically significant change in specificity. Moreover, with Paige Prostate, pathologists more often correctly classified smaller foci and lower-grade tumors and spent less time analyzing each WSI [24]. According to the FDA comments, potential risks include false-negative and false-positive results, which are mitigated by the device's use as an adjunct and by the professional evaluation by a qualified pathologist who considers patient history among other relevant clinical data and who may perform additional laboratory studies on the samples before rendering a final diagnosis. It was good news that the United States FDA granted marketing authorization of Paige Prostate to Paige.AI., suggesting that deep learning-based CAD applications will be rapidly introduced to the worldwide healthcare market.

3. Development of deep learning models for clinical applications

3.1. Endoscopy

In gastroenterology, clinical decision-making and treatment are driven by complex endoscopic procedures and visual identification and interpretation. Clinicians come across huge amounts of data of endoscopic images, where deep learning models can be used to aid gastroenterologists in making endoscopic diagnoses, analyzing images, and screening malignancies [25]. For endoscopic image diagnoses of clinical colorectal lesions, high-resolution micro-endoscopy [26], fluorescence imaging [27], and enhanced endoscopy [28] are available to improve the detection rate of tumors using endoscopy. Importantly, automated image analysis techniques can accurately predict histopathologies based on images captured using colonoscopy [29]. The application of deep learning models in CAD can improve the detection rate of early colorectal carcinomas and polyps [30]. A deep learning model could locate and identify colorectal polyps in real time with a cross-validation accuracy of 96.4% and an area under the receiver operating characteristic curve (AUC) of 0.991 in a set of 8641 colonoscopy images containing 4088 unique polyps [31]. In Japan, Cybernet (Cybernet Systems Co., Ltd., Tokyo, Japan) developed an AI system called EndoBRAIN, which can analyze blood vessels and the cellular structure of a lesion and show the tumor probability in an instant that would help identify tumors or nonneoplastic polyps [32]. EndoBRAIN was a CAD system designed to help endoscopists identify pathological changes in colorectal polyps during colonoscopy. EndoBRAIN is directly connected to the endoscopy unit and predicts polyps as either neoplastic or nonneoplastic based on *in vivo* microscopic imaging obtained using an endocytoscope (Olympus Corp., Tokyo, Japan). To detect esophageal cancers, including squamous cell carcinoma and adenocarcinoma, a deep learning model with a sensitivity of 98% and an accuracy of 98% was developed using 8428 training images of esophageal cancer from 384 patients [33]. Interestingly, a deep learning-based CAD system could determine invasion depth of gastric cancer based on conventional endoscopy, with a sensitivity of 76.47%, a specificity of 88.97%, and an overall accuracy of 89.16%; this CAD system distinguished early gastric cancer from deeper submucosal invasion and minimized overestimation of invasion depth, which could reduce unnecessary gastrectomy [34]. Even in *Helicobacter pylori* (HP)-associated chronic gastritis, the deep learning-based CAD system for detecting HP infection achieved a sensitivity of 86.7%, a specificity of 86.7%, and an AUC of 0.956, which seems feasible and is expected to facilitate and improve diagnosis during health checkups [35].

In the near future, endoscopic surgery will be presented in the same way as a surgical Da Vinci robot. With the development of new materials and software technology along with visual processing of computer images and videos, the detection rate of early colorectal carcinomas should be improved by printing gastrointestinal tract models using 3D materials [36] and improving the efficiency of diagnosis and follow-up using 5G remote online technology [37].

3.2. Radiology

Chest radiography has been the cornerstone of radiological imaging for many decades and remains the most commonly performed radiological examination worldwide. Interpreting chest radiographs can be challenging due to the superimposition of anatomical structures along the projection direction, which can make determining abnormalities and accurately distinguishing between different pathological patterns difficult [38]. For chest X-ray public datasets, the established digitization of radiological workflows enables medical institutions to collate and classify large sets of digital radiographs. Moreover, radiological reports can now be automatically analyzed to extract labels of interest for each image using natural language processing algorithms, which have allowed the construction of multiple large labeled chest X-ray public datasets for research (e.g., ChestX-ray14, CheXpert, MIMIC-CXR, PadChest (P), PLCO, Open-i, Ped-pneumonia, JSRT + SCR, RSNA-Pneumonia, Shenzhen, Montgomery, BIMCV, COVIDDSL, COVIDGR, SIIM-ACR, CXR14-Rad-Labels, COVID-CXR, NLST, Object-CXR, and Belarus) [38]. The recent rapid increase in the number of chest X-ray public datasets has positively impacted the promotion of deep learning studies.

Recently, the diagnosis and evaluation of coronavirus disease 2019 infection using chest X-rays is a topic that has attracted much interest from scientists. For example, the disease severity and progression could be predicted by comparing previous examinations of the patient [39]. Intriguingly, image generation techniques have been harnessed for various purposes (e.g., data augmentation, visualization, anomaly detection, and domain adaptation). For anomaly detection, a generative adversarial network (GAN) model trained to reconstruct healthy images would have a high reconstruction error if abnormal images are input at test time, allowing them to be identified [38]. For example, an autoencoder for anomaly detection trained only with healthy chest X-ray images was tailored to not only reconstruct healthy images but also produce uncertainty predictions [40]. Presently, it is expected that AI software will act as assistants to radiologists. However, there should be the demand for AI systems to detect infectious diseases (e.g., tuberculosis) in low-resource settings where radiologists are unavailable. In the future, innovative AI systems for visualizing and quantifying interval changes using one or more previous chest X-ray images from the same patient could substantially improve the efficiency of radiologists.

Not only chest X-ray but also deep learning models have demonstrated increased efficiency and image quality for PET reconstruction from sinogram data. As mentioned above, GANs have been applied in modality transformation, artifact reduction, and synthetic PET image generation, which could generate images directly from sinograms [41]. With the advent of CNN-based deep learning approaches, the feasibility of generating a differential diagnosis directly from images was demonstrated, which outperformed radiologists in some cases [42]. Moreover, CNNs have been applied to predict outcomes; a CNN applied to ^{18}F-FDG PET images could predict the onset of Alzheimer's disease years in advance, exceeding the accuracy of experts [43].

Importantly, the development of automated and reproducible quantification analysis of features alongside with the combination of conventional anatomical and functional characteristics could characterize neoplastic profiles, such as aggressiveness and potential of response to chemotherapy, which are essential for clinical decision-making [44,45]. Radiomics, which is the high-throughput extraction of several medical image features from radiographic images, can address this issue and is an appropriate approach that holds great promises [45]. However, translating and effectively using these multiparametric models combining advanced mathematical models with numerous variables of clinically derived biomarkers are a big challenge. Deep learning-based radiomic features are obtained by normalizing information obtained from deep neural networks. The effectiveness of deep learning radiomic features is highly related to the quality of segmentation and volume of training sets, which requires large datasets to identify a relevant and robust feature subset [46].

In recent radiation oncology, deep learning approaches have found several software applications in many research domains. The exploitation of radiomic features provided a new insight for quantitative image analysis, aiming to support clinical decision-making in characterization and treatment planning in several pathological conditions. Despite the rapid and increasing development of state-of-the-art deep learning approaches in radiomics, the concept of understanding and explaining the way that the predictions are performed, which precludes widespread adoption, remains an issue [47]. A major issue of radiomics is interpretability in clinical routine practices because most radiomic extraction and imaging biomarkers are used as black-boxes, making them impossible to clinically translate [46].

3.3. Pathology

Deep learning is taking the scope of digital pathology beyond mere digitization and telepathology. The incorporation of deep learning-based computer vision algorithms in routine workflow is on the horizon and must be a disruptive technology, reducing processing time, and increasing the detection rate of anomalies. Recent developments in hardware and software have expanded the possibilities for modeling and analyzing WSIs in pathology [48]. Particularly, WSIs may play an important role in reviewing central pathologies and substantially reduce the overall turnaround time required for slide review at the central location. WSI-based approaches may benefit the large clinical trials organized by oncology cooperative groups since most trials involve complicated logistics owing to the enrollment of several patients at several remotely located participating institutions [49]. In radiology, digital images (e.g., X-ray and CT) are obtained directly from medical laboratory devices, whereas, in pathology, two steps (i.e., stained slide glass preparation and scanning) are required to obtain WSIs, which preclude hospitals to introduce digital pathology due to the extra cost of purchasing a digital slide scanner and human resources. However, digital pathology is not just a tool for remote pathological diagnosis (telepathology); it can be an innovative technology in various medical studies because the possibility of training deep neural networks combined with the generation of WSIs has generated new opportunities for medical image analysis of pathological specimens. Moreover, pathology-based assessments have been performed to classify diseases and determine efficacy in drug development across various disease areas. Deep learning models in pathology improve quantitative accuracy and allow geographical contextualization of data using spatial algorithms. The development and integration of digital pathology and deep learning-based approaches provide substantive advantages over traditional methods [50]. The increased speed and efficiency gained in image acquisition can enhance the usage options of glass slides (e.g., hematoxylin and eosin [H&E] staining, immunohistochemistry, and in-situ hybridization), which can be converted into a remotely available WSI within minutes and centrally reviewed by multiple pathologists worldwide, with applications including education, research, consultation, and diagnostics [49].

The advent of WSIs led to the application of deep learning techniques for aiding pathologists in inspecting WSIs and diagnosing cancers. Promising and successful computational pathology applications include tumor classification and segmentation, mutation classification, and outcome prediction [23,51–58]. One of the main challenges in computational pathology is the sheer size of a WSI. A single WSI obtained at 20 × magnification can contain several billions of pixels, while the area of interest can be as small as a few thousand pixels. To apply a deep learning classifier, a WSI should be divided into several thousand tiles, with the classifier then applied independently on each tile. Then, the output from all tiles must be aggregated to obtain a final WSI classification (e.g., adenocarcinoma) [59]. Deep learning model training and test process are summarized in Fig. 2. For gastric and colonic epithelial tumor classification into adenocarcinomas, adenomas, and non-neoplastic lesions, supervised deep learning classifiers were trained using manually annotated WSI training sets and achieved AUCs of up to 0.97 and 0.99 for gastric adenocarcinomas and adenomas, respectively, and 0.96 and 0.99 for colonic adenocarcinomas and adenomas, respectively [59]. The fully supervised learning approach using detailed manual annotation (Fig. 2) performed by expert pathologists achieved high AUCs of 0.984 for detecting pancreatic ductal adenocarcinomas on endoscopic US-guided fine-needle biopsy specimens [60], 0.95 for classifying diffuse-type adenocarcinomas (poorly differentiated adenocarcinomas) of the stomach [61], and 0.99 for classifying signet ring cell carcinomas of the stomach [62]; however, preparing a large fully annotated training dataset for WSI cancer classification is a tedious, time-consuming task. In contrast, weakly supervised learning (Fig. 2) is an alternative approach and requires only weakly labeled data. Given that diagnoses from WSIs are readily available from pathological reports, additional manual annotations by pathologists are not required. Weakly supervised learning methods, such as multiple instance learning, can operate directly on WSIs using the diagnoses as slide-level labels. In fact, weakly supervised learning achieved robust AUCs higher than 0.97 for differentiating between lung carcinomas and nonneoplastic lesions [63]. Interestingly, using recurrent neural networks (Fig. 2) to obtain the WSI classification as opposed to using a simple max pooling was highly beneficial for classification that allows uncertainty (e.g., adenocarcinoma and squamous cell carcinoma classification in the presence of poorly differentiated cancer cell components; adenocarcinoma and adenoma classification; and breast invasive ductal carcinoma and ductal carcinoma in-situ classification) [59,64,65]. Moreover, the combination of transfer learning from existing models and weakly supervised learning approaches achieved high AUCs of 0.98 for classifying breast invasive ductal carcinomas [66] and 0.95 for classifying poorly differentiated colorectal adenocarcinomas (a rare subtype of colorectal adenocarcinoma) [67]. Deep learning models for classifying histopathological WSIs could be used as part of an integrated workflow where, as soon as the glass slides are scanned, a diagnosis is predicted automatically and used to rank cases based on order of priority for review by pathologists. This would allow pathologists to inspect first the cases that potentially require the most attention, allowing faster turnaround. Deep learning-based applications could be applied on a large database of WSIs to organize and retrieve slides based on a given predicted diagnosis and could serve as a second reader to confirm the diagnosis or alert the pathologists in cases of disagreement on the primary diagnosis, requesting further inspection. Hence, integrating deep learning models into the computational pathology (digital pathology) workflow would be beneficial for easing the ever-increasing workloads on pathologists, especially in regions that have shortages in access to pathological diagnosis services.

Furthermore, deep learning-based approaches may be applied in translational medicine and clinical practice by predicting genetic mutations from routine histopathological H&E slides. Deep learning models may be particularly useful for evaluating genomic instability and the mutational landscape, with the possibility of assessing pathological and genetic features [50]. For example, a deep learning model trained using WSIs of H&E slides of hepatocellular carcinoma could predict CTNNB1, FMN2, TP53, and ZFX4 genetic mutations with AUCs ranging from 0.71 to 0.89 [68].

Despite the advantages of incorporating deep learning models and digital pathology into clinical settings, technical concerns

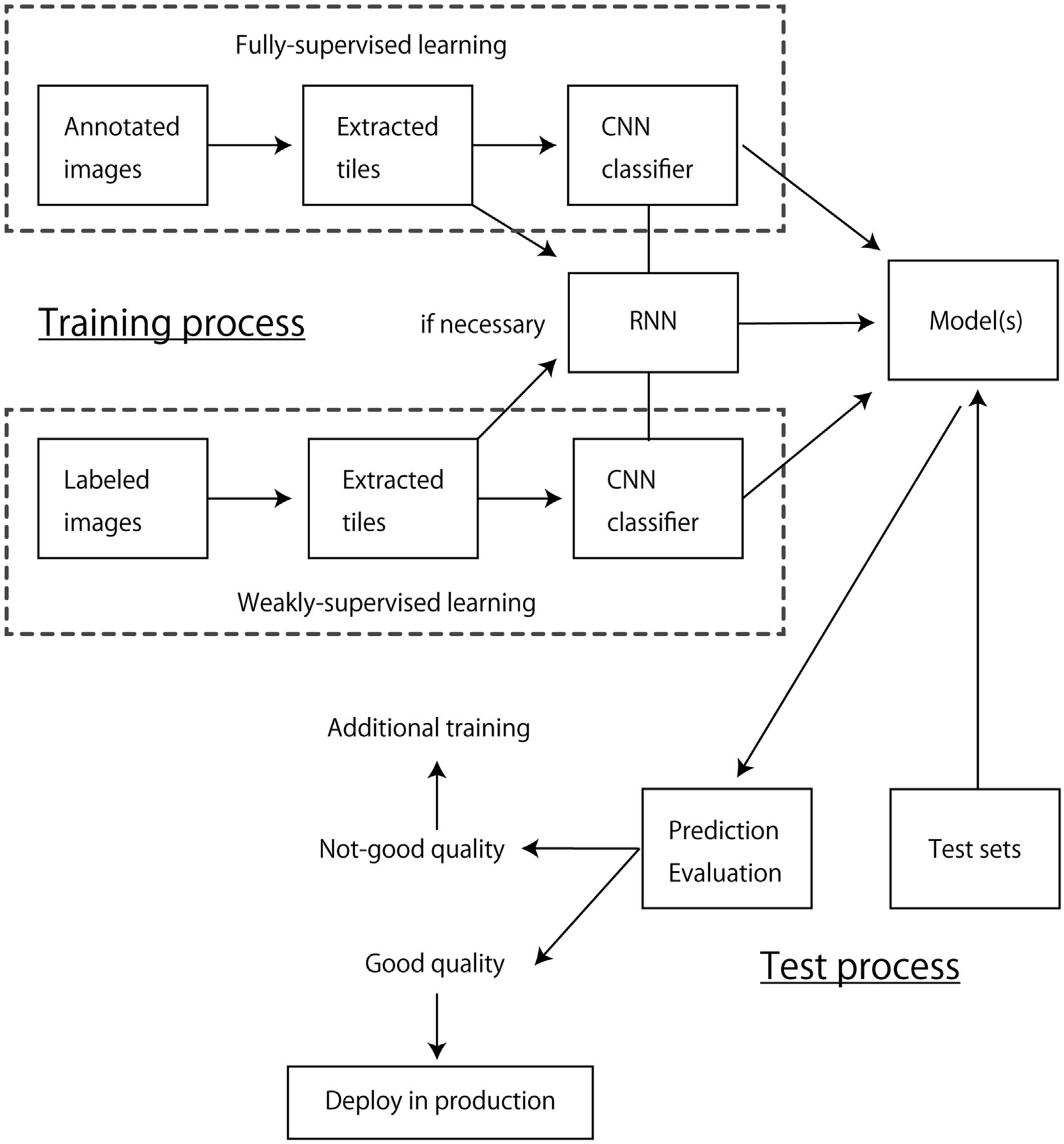

Fig. 2. Schematic diagrams for deep learning model training and test processes.

remain related to reproducibility, interpretability, the accuracy of competing devices, financial costs, and regulatory approvals [50]. A strategy toward the implementation of digital pathology may involve several phases culminating in the adoption of WSIs and deep learning-based technology in clinical practices. The first step involves demonstrating the reliability of digital pathology using a biomarker that has shown clinical impact with manual pathology (e.g., complementary diagnostics). The next steps would be introducing digital pathology as a diagnostic aid with novel biomarkers, with the aim of demonstrating the clinical use of the biomarker by digital quantification, which requires the development of deep learning-based software for use in prospective clinical trials to evaluate the selected biomarker for patient stratification. In the final step, deep learning models trained using large quantities of data (big data) can predict patient responses (clinical outcomes) based on WSIs [50].

3.4. Dentistry

In dentistry, deep learning approaches have produced interesting results in diagnosis and prediction. Deep learning models demonstrated successful detection of dental caries from periapical radiographic images [69–71], near-infrared light transillumination images [72,73], and clinical features [74]. On periapical radiographs, deep learning algorithms achieved AUCs of 0.917 (95% confidence interval [CI], 0.860–0.975) on premolar models, 0.890 (95% CI, 0.819–0.961) on molar models, and 0.845 (95% CI, 0.790–0.901) on both premolar and molar models [69]. Interestingly, to detect caries lesions on bitewing radiographs, deep learning models were found to be significantly more accurate than dentists, indicating that deep learning models can assist dentists in detecting especially initial caries on bitewings [70]. Dental caries are a challenge because of their high prevalence, besides being a chronic but preventable disease, which can occur depending on the consumption of certain nutritional elements interacting simultaneously with different factors, such as socioeconomic factors. Deep learning models could classify subjects without caries from those with caries or restorations with moderate accuracy according to their demographic and dietary factors [74]. In the field of prosthodontics, deep learning models can be applied to CAD/computer-aided manufacturing (CAM) systems, implant prosthetics, tooth preservation, and orofacial anatomy because deep learning-based technologies are particularly suited to dealing with complex situations with multiple possible factors [75]. For example, a deep learning model for predicting the debonding probability of CAD/CAM CR crowns from 2D images captured from 3D stereolithography models achieved a high AUC of 0.998 [76]. Deep learning model-based treatment planning in CAD/CAM implant dentistry could be crucial to simplify virtual 3D treatment planning and robotic insertion of dental implants using deep learning-based applications [77]. Moreover, there was a dental CT image denoizing method based on transfer learning of a GAN from the public-domain CT images. A GAN with the Wasserstein loss function was trained using high- and low-dose medical CT image pairs of human chest and abdomen; then, the network was fine-tuned using dental CT image pairs of two human skull phantoms. The fine-tuning procedure in the transfer learning scheme enhanced the network performance in terms of the quantitative metrics and improved the visual appearance of the processed images [78]. Moreover, a hybrid deep learning model consisting of Cycle-GAN and U-Net has been developed to measure bone mineral density quantitatively and directly from Cone-beam CT images. The Cycle-GAN improved the contrast of the bone images by reflecting the original bone mineral density distribution of the quantitative CT images locally, whereas the two-channel U-Net improved the spatial uniformity of the bone images by globally suppressing the image artifacts and noise. As a result, the Cycle-GAN and two-channel U-Net worked to provide complementary benefits in improving the contrast and uniformity of the bone image locally and globally; therefore, the hybrid deep learning model could substantially enhance the linearity, uniformity, and contrast as well as the anatomical and quantitative accuracy of the bone images to quantitatively measure bone mineral density from Cone-beam CT images [79].

AI technology in dentistry plays a key role in patient data management, medical applications, and services and can facilitate the future development of patient-centered, personalized treatment. Deep learning-based applications are particularly beneficial for processing and analyzing large amounts of data to classify outcomes and for processing repetitive workflows, which would likely provide support in evidence-based dental decision-making, particularly for less experienced practitioners, and facilitate the analysis of individual patient cases [79].

4. Conclusions

Deep learning models for medical image analysis provide great impacts on both clinical and research applications and are expected to revolutionize CAD in medicine. Moreover, the research and development of deep learning models and their implementation are expected to advance dramatically in the medical imaging field (i.e., endoscopy, radiology, pathology, and dentistry). Additionally, AI will be able to make developing diagnoses and making treatment decision easier in the near future. In contrast, medical practitioners will be able to return to their original jobs of healing patients by gaining more time and mental space. Furthermore, big data and AI will transform healthcare.

As Drs. Obermeyer and Emanuel have reported in 2016, AI development will create winners and losers in medicine as in other industries; however, patients will ultimately emerge as the biggest winners as deep learning transforms clinical medicine [80]. I cannot wait to see what the world of medicine will look like 5 or 10 years from now.

Ethical approval

Ethical approval is not required for this review article.

CRediT authorship contribution statement

Masayuki Tsuneki: Literature search, figure preparation, Writing this manuscript, Supervision.

Conflicts of interest

M.T. is an employee of Medmain Inc. The author declares no competing financial interests associated with this publication. This review article did not receive any financial supports from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgement

The author thanks to Dr. Fahdi Kanavati of the Medmain Research, Medmain Inc.

References

[1] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.
[2] Chan HP, Hadjiiski LM, Samala RK. Computer-aided diagnosis in the era of deep learning. Med Phys 2020;47:e218–27.
[3] Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017;19:221–48.
[4] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. Proc Mach Learn Res 2010;9:249–56.
[5] Sutskever I, Martens J, Dahl G, Hinton G. On the importance of initialization and momentum in deep learning. Proc Mach Learn Res 2013;28:1139–47.
[6] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. Proc Mach Learn Res 2011;15:315–23.
[7] Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol PA. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res 2010;11:3371–408.
[8] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15:1929–58.
[9] Wan L, Zeiler M, Zhang S, Cun YL, Fergus R. Regularization of neural networks using DropConnect. Proc Mach Learn Res 2013;28:1058–66.
[10] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. Proc Mach Learn Res 2015;37:448–56.
[11] Kanavati F, Tsuneki M. Partial transfusion: on the expressive influence of trainable batch norm parameters for transfer learning. Proc Mach Learn Res 2021;143:338–53.
[12] Armato 3rd SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 2011;38:915–31.
[13] Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
[14] Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imag 2016;35:1299–312.
[15] Cheng PM, Malhi HS. Transfer learning with convolutional neural networks for classification of abdominal ultrasound images. J Digit Imag 2017;30:234–43.
[16] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.
[17] Winsberg F, Elkin M, Macy J, Bordaz V, Weymouth W. Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. Radiology 1967;89:211–5.
[18] Spiesberger W. Mammogram inspection by computer. IEEE Trans Biomed Eng 1979;26:213–9.
[19] Semmlow JL, Shadagopappan A, Ackerman LV, Hand W, Alcorn FS. A fully automated system for screening xeromammograms. Comput Biomed Res 1980;13:350–62.
[20] Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis. Adv Exp Med Biol 2020;1213:3–21.
[21] Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, et al. Deep learning in medical imaging and radiation therapy. Med Phys 2019;46:e1–36.
[22] Mazurowski MA, Buda M, Saha A, Bashir MR. Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI. J Magn Reson Imag 2019;49:939–54.
[23] Campanella G, Hanna MG, Geneslaw L, Miraflor A, Silva VWK, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med 2019;25:1301–9.

[24] Raciti P, Sue J, Ceballos R, Godrich R, Kunz JD, Kapur S, et al. Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies. Mod Pathol 2020;33:2058–66.
[25] Anirvan P, Meher D, Singh SP. Artificial intelligence in gastrointestinal endoscopy in a resource-constrained setting: a reality check. Euroasian J Hepato-Gastroenterol 2020;10:92–7.
[26] Tan T, Qu YW, Shu J, Liu ML, Zhang L, Liu HF. Diagnostic value of high-resolution micro-endoscopy for the classification of colon polyps. World J Gastroenterol 2016;22:1869–76.
[27] Mizrahi I, Wexner SD. Clinical role of fluorescence imaging in colorectal surgery - a review. Expet Rev Med Dev 2017;14:75–82.
[28] Ho SH, Uedo N, Aso A, Shimizu S, Saito Y, Yao K, et al. Development of image-enhanced endoscopy of the gastrointestinal tract: a review of history and current evidences. J Clin Gastroenterol 2018;52:295–306.
[29] Tajbakhsh N, Gurudu SR, Liang J. A comprehensive computer-aided polyp detection system for colonoscopy videos. Inf Process Med Imaging 2015;24:327–38.
[30] Cai YW, Dong FF, Shi YH, Lu LY, Chen C, Lin P, et al. Deep learning driven colorectal lesion detection in gastrointestinal endoscopic and pathological imaging. World J Clin Cases 2021;9:9376–85.
[31] Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology 2018;155:1069–78.
[32] Mori Y, Kudo SE, Mori K. Potential of artificial intelligence-assisted colonoscopy using an endocytoscope (with video). Dig Endosc 2018;30:52–3.
[33] Horie Y, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest Endosc 2019;89:25–32.
[34] Zhu Y, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, et al. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. Gastrointest Endosc 2019;89:806–15.
[35] Itoh T, Kawahira H, Nakashima H, Yata N. Deep learning analyzes Helicobacter pylori infection by upper gastrointestinal endoscopy images. Endosc Int Open 2018;6:E139–44.
[36] Mishra S. Application of 3D printing in medicine. Indian Heart J 2016;68:108–9.
[37] Li D. 5G and intelligence medicine-how the next generation of wireless technology will reconstruct healthcare? Precis Clin Med 2019;2:205–8.
[38] Çallı E, Sogancioglu E, van Ginneken B, van Leeuwen KG, Murphy K. Deep learning for chest X-ray analysis: a survey. Med Image Anal 2021;72:102125.
[39] Cohen JP, Dao L, Roth K, Morrison P, Bengio Y, Abbasi AF, et al. Predicting COVID-19 pneumonia severity on chest X-ray with deep learning. Cureus 2020;12:e9448.
[40] Mao Y, Xue FF, Wang R, Zhang J, Zheng WS, Liu H. Abnormality detection in chest X-ray images using uncertainty prediction autoencoders. In: Martel AL, et al., editors. Medical image computing and computer assisted intervention – MICCAI 2020. Lecture notes in computer science, vol. 12266. Cham: Springer; 2020. https://doi.org/10.1007/978-3-030-59725-2_51.
[41] Herskovits EH. Artificial intelligence in molecular imaging. Ann Transl Med 2021;9:824.
[42] Liu M, Cheng D, Yan W, Alzheimer's Disease Neuroimaging Initiative. Classification of alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images. Front Neuroinf 2018;12:35.
[43] Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, et al. A deep learning model to predict a diagnosis of alzheimer disease by using 18F-FDG PET of the brain. Radiology 2019;290:456–64.
[44] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, C de Jong EE, Timmeren JV, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
[45] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Stiphout RGPMV, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.
[46] Avanzo M, Wei L, Stancanello J, Vallières M, RaoA, Morin O, et al. Machine and deep learning methods for radiomics. Med Phys 2020;47:e185–202.
[47] Papadimitroulas P, Brocki L, Christopher Chung N, Marchadour W, Vermet F, Gaubert L, et al. Artificial intelligence: deep learning in oncological radiomics and challenges of interpretability and data harmonization. Phys Med 2021;83:108–21.
[48] Smith B, Hermsen M, Lesser E, Ravichandar D, Kremers W. Developing image analysis pipelines of whole-slide images: pre- and post-processing. J Clin Transl Sci 2020;5:e38.
[49] Mroz P, Parwani AV, Kulesza P. Central pathology review for phase III clinical trials: the enabling effect of virtual microscopy. Arch Pathol Lab Med 2013;137:492–5.
[50] Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. Mod Pathol 2022;35:23–32.
[51] Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep 2016;6:26286.
[52] Kraus OZ, Ba JL, Frey BJ. Classifying and segmenting microscopy images with deep multiple instance learning. Bioinformatics 2016;32:i52–9.
[53] Korbar B, Olofson AM, Miraflor AP, Nicka CM, Suriawinata MA, Torresani L, et al. Deep learning for classification of colorectal polyps on whole-slide images. J Pathol Inf 2017;8:30.
[54] Luo X, Zang X, Yang L, Huang J, Liang F, Rodriguez-Canales J, et al. Comprehensive computational pathological image analysis predicts lung cancer prognosis. J Thorac Oncol 2017;12:501–9.
[55] Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med 2018;24:1559–67.
[56] Wei JW, Tafe LJ, Linnik YA, Vaickus LJ, Tomita N, Hassanpour S. Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. Sci Rep 2019;9:3358.
[57] Gertych A, Swiderska-Chadaj Z, Ma Z, Ing N, Markiewicz T, Cierniak S, et al. Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. Sci Rep 2019;9:1483.
[58] Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. Cell Rep 2018;23:181–93. e7.
[59] Iizuka O, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep learning models for histopathological classification of gastric and colonic epithelial tumours. Sci Rep 2020;10:1504.
[60] Naito Y, Tsuneki M, Fukushima N, Koga Y, Higashi M, Notohara K, et al. A deep learning model to detect pancreatic ductal adenocarcinoma on endoscopic ultrasound-guided fine-needle biopsy. Sci Rep 2021;11:8454.
[61] Kanavati F, Tsuneki M. A deep learning model for gastric diffuse-type adenocarcinoma classification in whole slide images. Sci Rep 2021;11:20486.
[62] Kanavati F, Ichihara S, Rambeau M, Iizuka O, Arihiro K, Tsuneki M. Deep learning models for gastric signet ring cell carcinoma classification in whole slide images. Technol Cancer Res Treat 2021;20. 15330338211027901.
[63] Kanavati F, Toyokawa G, Momosaki S, Rambeau M, Kozuma Y, Shoji F, et al. Weakly-supervised learning for lung carcinoma classification using deep learning. Sci Rep 2020;10:9297.
[64] Kanavati F, Toyokawa G, Momosaki S, Takeoka H, Okamoto M, Yamazaki K, et al. A deep learning model for the classification of indeterminate lung carcinoma in biopsy whole slide images. Sci Rep 2021;11:8110.
[65] Kanavati F, Ichihara S, Tsuneki M. A deep learning model for breast ductal carcinoma in situ classification in whole slide images. Virchows Arch 2022 [in-press].
[66] Kanavati F, Tsuneki M. Breast invasive ductal carcinoma classification on whole slide images with weakly-supervised and transfer learning. Cancers 2021;13:5368.
[67] Tsuneki M, Kanavati F. Deep learning models for poorly differentiated colorectal adenocarcinoma classification in whole slide images using transfer learning. Diagnostics 2021;11:2074.
[68] Chen M, Zhang B, Topatana W, Cao J, Zhu H, Juengpanich S, et al. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. NPJ Precis Oncol 2020;4:14.
[69] Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent 2018;77:106–11.
[70] Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, Gaudin R, et al. Detecting caries lesions of different radiographic extension on bitewings using deep learning. J Dent 2020;100:103425.
[71] Javed S, Zakirulla M, Baig RU, Asif SM, Meer AB. Development of artificial neural network model for prediction of post-streptococcus mutans in dental caries. Comput Methods Progr Biomed 2020;186:105198.
[72] Schwendicke F, Elhennawy K, Paris S, Friebertshäuser P, Krois J. Deep learning for caries lesion detection in near-infrared light transillumination images: a pilot study. J Dent 2020;92:103260.
[73] Holtkamp A, Elhennawy K, Cejudo Grano de Oro JE, Krois J, Paris S, Schwendicke F. Generalizability of deep learning models for caries detection in near-infrared light transillumination images. J Clin Med 2021;10:961.
[74] Zanella-Calzada LA, Galván-Tejada CE, Chávez-Lamas NM, Rivas-Gutierrez J, Magallanes-Quintanar R, Celaya-Padilla JM, et al. Deep artificial neural networks for the diagnostic of caries using socioeconomic and nutritional features as determinants: data from NHANES 2013-2014. Bioengineering 2018;5:47.
[75] Bernauer SA, Zitzmann NU, Joda T. The use and performance of artificial intelligence in prosthodontics: a systematic review. Sensors 2021;21:6628.
[76] Yamaguchi S, Lee C, Karaer O, Ban S, Mine A, Imazato S. Predicting the debonding of CAD/CAM composite resin crowns with AI. J Dent Res 2019;98:1234–8.
[77] Kurt Bayrakdar S, Orhan K, Bayrakdar IS, Bilgir E, Ezhov M, Gusarev M, et al. A deep learning approach for dental implant planning in cone-beam computed tomography images. BMC Med Imag 2021;21:86.
[78] Hegazy MAA, Cho MH, Lee SY. Image denoising by transfer learning of generative adversarial network for dental CT. Biomed Phys Eng Express 2020;6:055024.
[79] Yong TH, Yang S, Lee SJ, Park C, Kim JE, Huh KH, et al. QCBCT-NET for direct measurement of bone mineral density from quantitative cone-beam CT: a human skull phantom study. Sci Rep 2021;11:15083.
[80] Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med 2016;375:1216–9.