Chen Hao (Orcid ID: 0000-0002-8400-3780)

Potentials of AI in Medical Image Analysis in Gastroenterology and Hepatology

Hao Chen PhD[1]

Joseph JY Sung MD, PhD[2]

[1]Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong
[2] Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Shatin, Hong Kong

Correspondence:
Professor Joseph JY Sung, Department of Medicine and Therapeutics, The Chinese University of Hong Kong.
Address: 9/F Lui Che Woo Clinical Sciences Building, Prince of Wales Hospital, Shatin, Hong Kong.
Email: jjysung@cuhk.edu.hk

Key words:
Artificial intelligence (AI), machine learning (ML), deep learning (DL), gastroenterology, hepatology, endoscopy, pathology, radiology.

Declaration of conflict of interest:
The authors declare no conflict of interest.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jgh.15327

Abstract

With the advancement of Artificial intelligence (AI) technology, it comes in a big wave carrying possibly huge impact in the field of Medicine. Gastroenterology and Hepatology, being a specialty relying much on diagnostic imaging, endoscopy and histopathology, AI technology has promised improving the quality and consistency of care to the patients. In this review, we will elucidate the development of machine learning methods, especially the visual representation mechanism in deep learning (DL) on recognition tasks. Various AI image analysis applications in endoscopy, radiology and pathology are covered in Gastroenterology and Hepatology and reveal the enormous potentials for AI in assisting diagnosis, prognosis and treatment. We also discuss the promises as well as pitfalls for AI in medical image analysis and pointing out future research directions.

Introduction

During the last decade, artificial intelligence (AI) has achieved groundbreaking success in many pattern recognition scenarios such as video surveillance[1], autonomous driving[2], natural language processing[3] and medical image analysis (MIA) [4,5]. Medical AI refers to emulating human cognition in the analysis, interpretation and comprehension of healthcare data. MIA falls into the category of medical AI and focuses on analyzing biomedical images by computer techniques such as object detection, segmentation and registration with various applications to diagnosis, prognosis and computer-assisted interventions.

With the increasing maturity of digitalization and low-cost storage, large amounts of heterogeneous medical data, such as human structural images, electronic health records, genetic information, have become widely available. These data could be an enabling resource for deriving medical AI solutions to improve care delivery. In particular, the high-volume and well-structured medical images serve as excellent basis for AI assisted MIA algorithm development, such as X-ray/computed tomography (CT)/ Magnetic Resonance Imaging (MRI)/ultrasound for disease diagnosis in radiology, digitalized whole slide image classification in pathology and lesion recognition (e.g., polyps, bleeding pathology) in endoscopy, etc.

Nowadays, hospitals are looking into AI solutions to support operational initiatives that increase cost saving, improve diagnosis accuracy and enhance patient satisfaction. AI, especially deep learning, has achieved remarkable progress in medical image-based interpretation, assisted-diagnosis and treatment. In this review, we will overview the advancement of medical image-derived AI solutions with emphasis on its applications in the practice of gastroenterology and hepatology.

Machine learning in medical image analysis

Machine learning is a subset of AI that empowers computer systems the ability to automatically learn from experience without being explicitly programmed[6]. It focuses on the development of computer algorithms that can be learned from the training data and make predictions or decisions on the unseen data. The most common machine learning approaches in medical image analysis can be divided into supervised learning, unsupervised learning and semi-supervised learning (Table 1), hinging on how much annotation information is provided to the computer learning system

Supervised learning: The learning algorithm is presented with clinical inputs and their desired outputs as paired training samples. The goal is to learn an inferenced function that maps inputs to correct outputs. The typical techniques in supervised learning for image recognition include regression and classification[7].

Unsupervised learning: There are no annotation labels being provided to the learning algorithm. Therefore, the computer algorithm needs to find underlying pattern by itself[8]. Two main methods in unsupervised learning for image analysis are principal component analysis and clustering.

Semi-supervised learning: There are both labeled and unlabeled data available for learning algorithms, which falls between supervised and unsupervised learning. Typically, a small amount of labeled data can only be acquired due to expensive cost while a large amount of unlabeled data can be easily acquired[9]. This is especially appealing in MIA scenarios where labels are usually expensive to be collected or annotated while abundant of unlabeled or weakly labeled image-structured data are available.

There is another type of machine learning method called **reinforcement learning**, which involves a computer program interacting with a dynamic environment to achieve a certain goal, such as game playing. A few methods have been explored in MIA, such as 3D landmark detection from CT images[10].

Deep learning: Deep learning (DL) is a branch of machine learning family with multi-layer artificial neural networks (ANN) for hierarchical representation learning. Different from traditional machine learning algorithms such as random forest and support vector machine, DL learns features automatically from input data without relying heavily on engineering-designed domain knowledge. Take visual image recognition[11] as example (Fig. 1), the low-level features such as edges or blocks are first extracted from the raw input images in the hierarchical feature learning. Then, the subsets of these low-level features are composed together to form higher-level invariant features, exemplified by the pooling and subsampling layers in the most successful DL architecture of convolutional neural networks (CNN). The high-level invariant representation offers more discriminative capability towards final specific task. With deeper layers of visual representation, the computer system usually can achieve better performance in image recognition[12].

The theory of DL can date back to 1950s. Nowadays DL methods have been the primary choice for MIA research due to the availability of optimization algorithms, large-scale annotated dataset and powerful computation infrastructure. Since 2012, most of studies focused on supervised DL for improving classification accuracy, such as the most successful AlexNet on ImageNet classification[13]. It achieved substantial improvement in comparison with traditional methods by learning hierarchical feature representation. In addition, **transfer learning** has demonstrated the effectiveness by pre-training on large image dataset first and then follow fine-tuning on the specific visual recognition task[5,14]. However, the annotation supervisions are not always available in the medical domain. Therefore, there is a trend that unsupervised DL algorithms for effective feature representation learning from large-scale unlabeled image dataset gain popularity, such as momentum contrast for visual representation learning[15]. Recently, a variety of deep semi-supervised learning methods have been developed. Technologies such as consistency regularization[16] and virtual adversarial training[17] are promising for MIA with a handful of fine-labeled data while large-scale unlabeled data can be exploited. To overcome the issue of overfitting in MIA, various data augmentation methods are extensively used in the training phase[4].

AI image analysis in radiology

AI technology has dramatically advanced the radiology image analysis in gastroenterology and hepatology applications. During last few years, DL made significant contributions for gastroenterology and hepatology disease classification and characterization from diagnostic images. AI technologies in MIA, consisting of classification, detection, segmentation and registration, have been widely applied in the radiology image analysis, including ultrasound, CT and MRI. Wang K et al.[18] proposed a DL derived radiomics approach for assessing liver fibrosis stages in elastography. A prospective multicenter study was conducted to evaluate the accuracy in patients with chronic hepatitis B (CHB) and showed overall higher performance in classifying liver fibrosis stages than existing liver stiffness measure from two-dimensional shear wave elastography (2D-SWE). Togo et al.[19] designed a deep CNN for detection of gastritis from double-contrast upper gastrointestinal barium X-ray radiography. Analyzing results on 6520 images with sensitivity 0.962 and specificity 0.983, they demonstrated the profound efficacy of DL techniques in gastritis detection. Yasaka et al.[20] investigated the differentiation of liver masses at dynamic contrast-enhanced CT images with deep CNN. Their method showed a high degree of AUC score 0.92 on differential diagnosis of liver masses.

Furthermore, AI has been extensively applied into the lesion detection and segmentation from radiology images. Näppi et al.[21] employed a deep CNN for computer-aided detection of serrated polyps in CT colonography. They achieved substantially high performance for serrated polyp ≥6 mm in size. In collaboration of multiple hospitals, a liver and liver tumor segmentation benchmark dataset[22] was released with a total of 201 CT volumes (varying amounts of lesions). Li et al.[23] proposed a hybrid densely connected network (H-DenseUNet) for aggregating volumetric contexts and achieved top performance on the challenge dataset for liver and tumor segmentation (Fig. 2). This demonstrated the feasibility of AI in assisting doctors in diagnosis and treatment planning of liver cancer. In addition, AI has been applied to the pancreas segmentation from both CT and MRI images. Roth et al.[24] proposed a spatial aggregation of holistically-nested CNN for pancreas localization and segmentation from CT images. They validated their method on 82 patients CT images and achieved very good performance. Zheng et al.[25] studied a DL approach for pancreas segmentation from MRI images. Promising performance has been achieved on two challenging cancer MRI datasets. These AI radiology studies in gastroenterology and hepatology corroborated the promising potential of AI in assisting clinical practice, which can potentially improve both the efficiency and accuracy in diagnostic reporting.

AI image analysis in pathology

Pathology analysis is the gold standard in cancer diagnosis. Previous research in image patch analysis has been extensively studied[26]. With the increasing availability of whole slide imaging (WSI) scanner, digital pathology integrating with AI can provide a unique opportunity to increase diagnostic accuracy, efficiency and reproducibility of histopathology. Several WSI benchmark datasets[4] were released and dramatically advanced the computational studies.

One of the main challenges is accurately recognizing interested regions from the histology slides. A series of robust and scalable AI methods have been developed in this direction during the last few years. Wang et al.[27] proposed a recalibrated multi-instance DL network (RMDL) for gastric cancer WSI classification (Fig. 3). A large gastric histopathology dataset with pixel-level annotations was constructed and they achieved an accuracy outperforming conventional methods. Heinemann et al.[28] automated histological feature (ballooning, inflammation, steatosis and fibrosis) scoring for non-alcoholic steatohepatitis (NASH) with DL methods. The DL algorithm output continuous scores for quantifying the extent of each feature and the quantitative comparison with human pathologists achieved very good agreement, such as steatosis. Zhou et al.[29] employed DL approach for Helicobacter pylori detection from histopathology images. They found assisted diagnosis with DL was faster and much more accurate than that without on positive cases. But the real-world clinical practice integration requires further optimization on mitigating diagnostic uncertainty which might be attributed to AI assistance, especially on negative cases.

Apart from extensive AI studies in diagnosis, there is an increasing interest in deriving pathology image biomarkers for overall survival and molecular alternation prediction. Kather et al.[30] adopted DL for predicting molecular features directly from histology in gastrointestinal cancer. The DL method can predict microsatellite instability (MSI) with reasonably good performance. Because MSI can determine the quality of gastrointestinal cancer patients responding to immunotherapy, this study highlighted the potential of AI in assisting immunotherapy. Furthermore, they validated the efficacy of DL in detecting colorectal cancers specimens with MSI and mismatch-repair deficiency (dMMR) on a multi-ethnical dataset with 8836 colorectal tumors[31]. In another study[32], DL has validated its feasibility in predicting overall survival from colorectal cancer histology slides with a hazard ratio of 1.63 in a multivariable Cox proportional hazard model, which can possibly be derived as image-derived prognostic biomarkers. In addition, Saillard et al.[33] investigated the DL based algorithm for patient survival prediction from histology slides after resection of liver cancer. They achieved c-indexes for survival prediction above 0.75 and further validated on The Cancer Genome Atlas (TCGA) dataset, corroborating the capability of AI in prognosis prediction on HCC patients. With the readily available of WSIs, AI studies in pathology applications have reached an unprecedented scale. Extensive on-going research and multi-ethnical validation will lay the foundation for the deployment of computational pathology analysis systems in future practice.

AI image analysis in endoscopy

Since the advent of fiberoptics, endoscopy has played an instrumental role in diagnosis and therapy of gastrointestinal diseases. Early progress has been witnessed in a variety of applications of MIA, such as polyp detection (Fig. 4) and characterization from colonoscopy, gastrointestinal bleeding and lesion detection in capsule endoscopy, identification of esophageal neoplasia and gastric cancer, etc.

Colonoscopy screening and polypectomy is the gold standard in colorectal cancer prevention. However, studies have reported the miss detection rate of neoplasia can be as high as 25%[34]. Urban et al.[35] utilized DL to detect polyps from colonoscopy with a cross-validation accuracy of 96.4%, outperforming conventional colonoscopy. Bernal J et al.[36] constructed open-access annotated benchmark and reported the results of comparative evaluation of polyp detection methods, which indicated CNN was the state-of-the-art for polyp detection. Yu et al.[37] integrated online and offline 3d DL for automated polyp detection from colonoscopy videos by leveraging fully convolutional networks. Their method can effectively suppress the false positives while retaining the high sensitivity.

Beside detection, AI-assisted colonoscopy can also achieve polyp characterization. With an AI prediction of histopathological analysis, the endoscopist can adopt the diagnosis-and-leave strategy and reduce unnecessary polypectomy during colonoscopy. Chen et al.[38] used DL with Narrow-band imaging (NBI) for histologic classification of diminutive neoplastic or hyperplastic polyps with 96.3% sensitivity and 78.1% specificity. Byrne et al.[39] also showed AI model trained on endoscopic video can differentiate diminutive adenomas from hyperplastic polyps with high accuracy of 94%. Mori et al.[40] investigated the use of AI in real-time identification of diminutive polyps (≤5 mm) during colonoscopy. This study validated the capability of AI assistance in meeting the requirements of promoting diagnosis-and-leave strategy for diminutive polyps.

In esophagogastroduodenoscopy (EGD), AI studies have achieved high accuracies in diagnosing a wide range of gastrointestinal conditions including gastric cancer, neoplasia in Barrett’s esophagus and Helicobacter pylori gastritis. DL system detected neoplasia in patients with Barrett’s esophagus (BE) with higher accuracy than endoscopists[41].

On the other hand, wireless capsule endoscopy (WCE) enables endoscopists to examine the digestive tract without surgical intervention. AI can be a perfect assistant in analyzing this high-volume data and highlight the suspicious regions. Soffer et al[42] investigated nineteen retrospective studies and showed DL has been widely adopted for WCE applications including detection of ulcers, polyps, celiac disease, bleeding and hookworm with detection accuracy above 90% for most studies.

Promises in AI medical image analysis

Besides the MIA assisted diagnosis and prognostication, the expanding use of AI into image acquisition such as fast imaging and quality assessment is an active research field. In

WSI of gastric cancer pathology slides, the identification of out-of-focus image regions could allow scanners focusing on blurred regions and take clearer views, which speed up the advent of full digital pathology era. In the MRI imaging of liver cancer patients, AI model can accelerate the image reconstruction without undermining perceptual image quality details. In the future, personalized imaging protocols can be developed with AI techniques for improving patients' diagnostic efficiency and accuracy.

The development of AI image-derived biomarkers holds promise for interdisciplinary clinical study and potentially provide readily available, economic and efficient solutions at point of care. New dimensions such as histological biomarkers from WSIs for genetic defects or molecular alternation, optical biopsy from AI-assisted endoscopic diagnosis, radiomics biomarker from diagnostic images for prediction of prognosis are promising tools on the horizon. The publicly available TCGA datasets have demonstrated substantial value for extracting image-based prognostic biomarkers in preliminary studies of gastrointestinal and liver cancers[43]. Such MIA tools may directly infer treatment response and disease recurrence from routine medical images. Furthermore, the combination of multi-modality images and omics data offers an unparalleled opportunity to uncover the mechanism of disease progression and treatment. AI excels at mining high-dimensional data including various imaging and omics features, which holds great promise for developing more accurate clinical outcome predictive tools. This will also provide exceptional insights into precision medicine by harnessing full potential of integrative data at the individual patient level.

Challenges in implementation of AI image analysis in Clinical Practice

Despite great advancements in MIA over the last few years, implementation of MIA is still slow in clinical practice. There are several that needs to be resolved. First, most of the studies testing and verifying AI-image analysis tools are retrospective based on single-center study with potential data selection bias. Large-scale prospective studies from multi-center recruitment will be needed to ensure validity and reproducibility of the investigational tool. In addition, DL is a data-hungry approach excelling at visual image recognition, thus large-scale datasets with international collaborative efforts are required to be established. Yet, data privacy should be respected and protection of health identification information must be secured. Federated learning[44] offers a possible solution to allow sharing knowledge learned from AI models without compromising patient privacy. Meanwhile, rigorous criteria for evaluation of AI models with widely consented standards will be needed. Although AI has been very effective for medical image analysis tasks, the black-box nature of DL algorithms lacking clinical interpretability and transparency has restricted its clinical adoption. It will be difficult for clinician, as well as patients, to base their clinical decision on something that they don't understand. Explainable AI (XAI) for interpretable and reliable healthcare is worth much investigation.

One of major criticisms of AI image analysis is the robustness and generalization capability when applied to unseen data. The performance may degrade significantly when generalizing to unseen data samples or suffering vulnerabilities from even small alternations[45]. An emerging field of adversarial learning in AI research has shown exciting results in improving the generalization capability of AI systems in image recognition tasks[46]. Uncertainty estimation is also certainly one of key elements in designing trustable AI models for healthcare tasks since doctors should exercise discretion when needed.

When applying AI in the healthcare-oriented applications, an additional hurdle is the insufficiency of medical supervision annotations due to the expensive cost of detailed labelling by doctors, even in the image-based scenarios. Annotation-efficient learning using weakly supervised, semi-supervised, or unsupervised learning against limited medical labels are attractive and promising to be explored within the next few years.

Last but not least, the boundaries and shared responsibility between doctors and AI systems needs to be handled properly. Instead of considering AI models as a substitute for human doctors, AI systems should serve as a helpful adjunct to clinical practice and empower doctors in decision-making with improved efficiency and accuracy. Issues such as legal liability, ethical clearance and legislation amendment, etc., should be delineated in the early phase of development to maintain well balanced between maximal patient benefits and minimal cost of technology[47,48].

Conclusion

AI can produce transformative impact in healthcare, especially in various image-based applications covering radiology, pathology and endoscopy in Gastroenterology and Hepatology. AI performance carries tremendous potential in assisting doctors improving efficiency and accuracy in clinical practice. Yet, there are lots of challenges existing before AI can be successfully deployed in the daily practice of Gastroenterology and Hepatology. Medical, technological, legal and ethical experts should work hand-in-hand in pursuit of the patient safety, trust between man and machine, and generalizability of AI-assisted medical image analysis.

References

1. LeCun Y, Bengio Y, Hinton G. Deep learning. nature. 2015 May;521(7553):436-44.
2. Daily M, Medasani S, Behringer R, Trivedi M. Self-driving cars. Computer. 2017 Dec 18;50(12):18-23.
3. Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing. IEEE Computational Intelligence Magazine. 2018 Jul 20;13(3):55-75.
4. Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Medical image analysis. 2017 Dec 1;42:60-88.
5. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annual review of biomedical engineering. 2017 Jun 21;19:221-48.
6. Vapnik V. The nature of statistical learning theory. Springer science & business media; 2013 Jun 29.
7. Mohri M, Rostamizadeh A, Talwalkar A. Foundations of machine learning. MIT press; 2018 Nov 30.
8. Hinton GE, Sejnowski TJ, Poggio TA, editors. Unsupervised learning: foundations of neural computation. MIT press; 1999.
9. Oliver A, Odena A, Raffel CA, Cubuk ED, Goodfellow I. Realistic evaluation of deep semi-supervised learning algorithms. In Advances in neural information processing systems 2018 (pp. 3235-3246).
10. Ghesu FC, Georgescu B, Zheng Y, Grbic S, Maier A, Hornegger J, Comaniciu D. Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. IEEE transactions on pattern analysis and machine intelligence. 2017 Dec 12;41(1):176-89.
11. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In European conference on computer vision 2014 Sep 6 (pp. 818-833). Springer, Cham.
12. Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence. 2013 Mar 7;35(8):1798-828.
13. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems 2012 (pp. 1097-1105).
14. Chen H, Dou Q, Ni D, Cheng JZ, Qin J, Li S, Heng PA. Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2015 Oct 5 (pp. 507-514). Springer, Cham.
15. He K, Fan H, Wu Y, Xie S, Girshick R. Momentum contrast for unsupervised visual representation learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2020 (pp. 9729-9738).

16. Tarvainen A, Valpola H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In Advances in neural information processing systems 2017 (pp. 1195-1204).
17. Miyato T, Maeda SI, Koyama M, Ishii S. Virtual adversarial training: a regularization method for supervised and semi-supervised learning. IEEE transactions on pattern analysis and machine intelligence. 2018 Jul 23;41(8):1979-93.
18. Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, Wu C, Liu C, Huang L, Jiang TA, Meng F. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. Gut. 2019 Apr 1;68(4):729-41.
19. Togo R, Yamamichi N, Mabe K, Takahashi Y, Takeuchi C, Kato M, Sakamoto N, Ishihara K, Ogawa T, Haseyama M. Detection of gastritis by a deep convolutional neural network from double-contrast upper gastrointestinal barium X-ray radiography. Journal of gastroenterology. 2019 Apr 1;54(4):321-9.
20. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. Radiology. 2018 Mar;286(3):887-96.
21. Näppi JJ, Pickhardt P, Kim DH, Hironaka T, Yoshida H. Deep learning of contrast-coated serrated polyps for computer-aided detection in CT colonography. In Medical Imaging 2017: Computer-Aided Diagnosis 2017 Mar 3 (Vol. 10134, p. 101340H). International Society for Optics and Photonics.
22. Bilic P, Christ PF, Vorontsov E, Chlebus G, Chen H, Dou Q, Fu CW, Han X, Heng PA, Hesser J, Kadoury S. The liver tumor segmentation benchmark (lits). arXiv preprint arXiv:1901.04056. 2019 Jan 13.
23. Li X, Chen H, Qi X, Dou Q, Fu CW, Heng PA. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE transactions on medical imaging. 2018 Jun 11;37(12):2663-74.
24. Roth HR, Lu L, Lay N, Harrison AP, Farag A, Sohn A, Summers RM. Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation. Medical image analysis. 2018 Apr 1;45:94-107.
25. Zheng H, Chen Y, Yue X, Ma C, Liu X, Yang P, Lu J. Deep pancreas segmentation with uncertain regions of shadowed sets. Magnetic Resonance Imaging. 2020 May 1;68:45-52.
26. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: A review. IEEE reviews in biomedical engineering. 2009 Oct 30;2:147-71.
27. Wang S, Zhu Y, Yu L, Chen H, Lin H, Wan X, Fan X, Heng PA. RMDL: Recalibrated multi-instance deep learning for whole slide gastric image classification. Medical image analysis. 2019 Dec 1;58:101549.
28. Heinemann F, Birk G, Stierstorfer B. Deep learning enables pathologist-like scoring of NASH models. Scientific reports. 2019 Dec 5;9(1):1-0.

29. Zhou S, Marklund H, Blaha O, Desai M, Martin B, Bingham D, Berry G, Gomulia E, Ng AY, Shen J. Deep learning-based Helicobacter pylori detection: A diagnostic pathology study. medRxiv. 2020 Jan 1.
30. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, Marx A, Boor P, Tacke F, Neumann UP, Grabsch HI. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nature medicine. 2019 Jul;25(7):1054-6.
31. Echle A, Grabsch HI, Quirke P, van den Brandt PA, West NP, Hutchins GG, Heij LR, Tan X, Richman SD, Krause J, Alwers E. Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning. Gastroenterology. 2020 Jun 17.
32. Kather JN, Krisam J, Charoentong P, Luedde T, Herpel E, Weis CA, Gaiser T, Marx A, Valous NA, Ferber D, Jansen L. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. PLoS medicine. 2019 Jan 24;16(1):e1002730.
33. Saillard C, Schmauch B, Laifa O, Moarii M, Toldo S, Zaslavskiy M, Pronier E, Laurent A, Amaddeo G, Regnault H, Sommacale D. Predicting survival after hepatocellular carcinoma resection using deep learning on histological slides. Hepatology. 2020 Feb 28.
34. Leufkens AM, Van Oijen MG, Vleggaar FP, Siersema PD. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. Endoscopy. 2012 May;44(05):470-5.
35. Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology. 2018 Oct 1;155(4):1069-78.
36. Bernal J, Tajkbaksh N, Sánchez FJ, Matuszewski BJ, Chen H, Yu L, Angermann Q, Romain O, Rustad B, Balasingham I, Pogorelov K. Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge. IEEE transactions on medical imaging. 2017 Feb 2;36(6):1231-49.
37. Yu L, Chen H, Dou Q, Qin J, Heng PA. Integrating online and offline three-dimensional deep learning for automated polyp detection in colonoscopy videos. IEEE journal of biomedical and health informatics. 2016 Dec 7;21(1):65-75.
38. Chen PJ, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate classification of diminutive colorectal polyps using computer-aided analysis. Gastroenterology. 2018 Feb 1;154(3):568-75.
39. Byrne MF, Chapados N, Soudan F, Oertel C, Pérez ML, Kelly R, Iqbal N, Chandelier F, Rex DK. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. Gut. 2019 Jan 1;68(1):94-100.
40. Mori Y, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S, Ogawa Y, Maeda Y. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. Annals of internal medicine. 2018 Sep 18;169(6):357-66.

41. de Groof AJ, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O. Deep-learning system detects neoplasia in patients with barrett's esophagus with higher accuracy than endoscopists in a multistep training and validation study with benchmarking. Gastroenterology. 2020 Mar 1;158(4):915-29.
42. Soffer S, Klang E, Shimon O, Nachmias N, Eliakim R, Ben-Horin S, Kopylov U, Barash Y. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. Gastrointestinal Endoscopy. 2020 Apr 22.
43. Kather JN, Calderaro J. Development of AI-based pathology biomarkers in gastrointestinal and liver cancer. Nature Reviews Gastroenterology & Hepatology. 2020 Jul 3:1-2.
44. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST). 2019 Jan 28;10(2):1-9.
45. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. Science. 2019 Mar 22;363(6433):1287-9.
46. Li H, Jialin Pan S, Wang S, Kot AC. Domain generalization with adversarial feature learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018 (pp. 5400-5409).
47. Sung JJ, Poon NC. Artificial intelligence in gastroenterology: where are we heading? Frontiers of Medicine. 2020 May 26.
48. Poon NC, Sung JJ. Self−driving cars and AI−assisted endoscopy: Who should take the responsibility when things go wrong? Journal of gastroenterology and hepatology. 2019 Apr;34(4):625-6.

Table 1. Summary of machine learning methods in medical image analysis

Type	Annotation	Examples of methods	Applications
Supervised learning	Yes	Logistic regression, linear discriminant analysis, k-nearest neighbors, decision trees, naïve Bayes, artificial neural networks (e.g., CNN) and support vector machines	Disease classification, organ/lesion segmentation, image retrieval, prognosis, etc.
Unsupervised learning	No	K-means, hierarchical clustering, mixture models, neural networks (e.g., auto-encoders, generative adversarial network), expectation maximization, principal component analysis, independent component analysis	Anomaly detection, density estimation, cluster analysis, image retrieval, segmentation, dimensionality reduction, etc.
Semi-supervised learning	Typical small labeled dataset and large unlabeled dataset	A variant making use of both supervised and unsupervised techniques	Both supervised and unsupervised learning applications

Figure 1. Visual hierarchical representation learning in deep neural networks. The CNN is first trained on large-scale visual images and then the filters (trained convolutional parameters) in CNN from low layer to high layers are visualized. The corresponding complexity of feature composition increases with deeper layers and provides stronger discrimination capability.

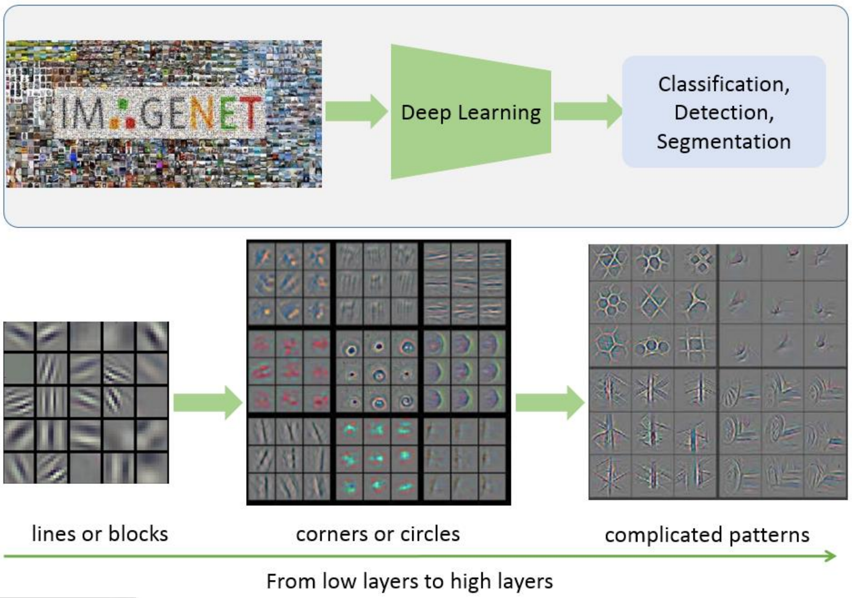

Figure 2. Radiology application of AI image analysis in liver tumor diagnosis and prognosis, including malignancy classification, liver/tumor localization and segmentation using DL method H-DenseUNet[23] for quantitative evaluation, and follow-up registration for change monitoring.

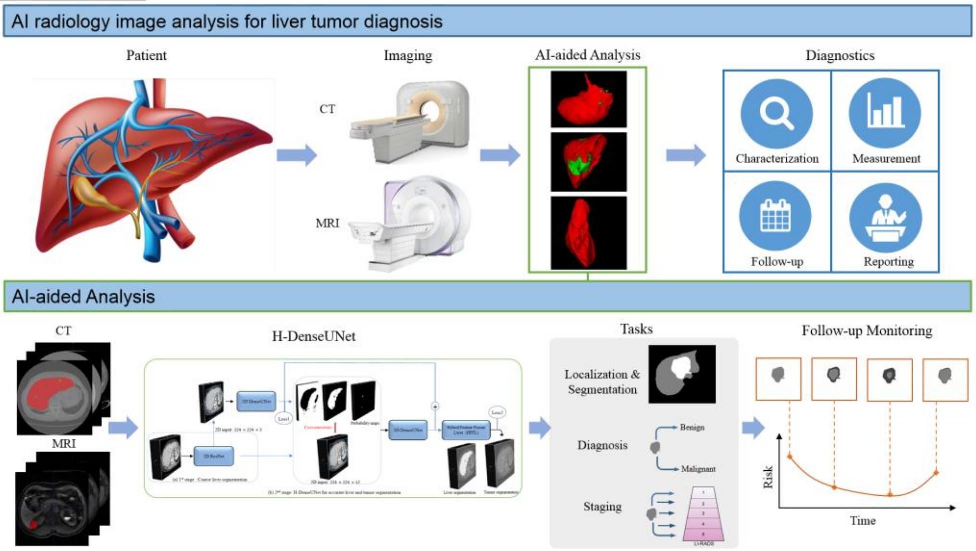

Figure 3. Pathology application of AI image analysis in gastric cancer classification from WSIs, including discriminative patch instance detection (stage1) and whole-slide image-level classification (stage2) from RMDL[27] method. Visualization of DL prediction results and ground-truth annotations shows high concordance.

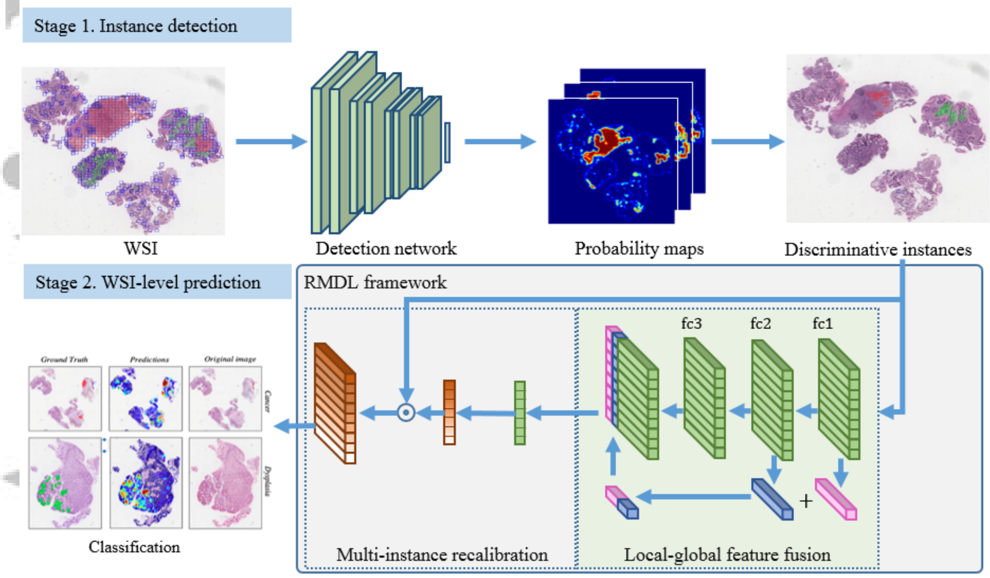

Figure 4. Endoscopy application of AI image analysis for polyp detection and segmentation from endoscopy videos. AI algorithms are trained on massive endoscopic images labeled by experienced endoscopists and can be readily deployed in the clinical scenario for assisting diagnosis.