Journal of the National Cancer Center 5 (2025) 113–131

Contents lists available at ScienceDirect

Journal of the National Cancer Center

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

Review

The application of artificial intelligence in upper gastrointestinal cancers

Xiaoying Huang[1,2,†], Minghao Qin[1,2,3,†], Mengjie Fang[4,5,†], Zipei Wang[1,2], Chaoen Hu[1,2], Tongyu Zhao[1,2,6], Zhuyuan Qin[7,8], Haishan Zhu[9], Ling Wu[9], Guowei Yu[9], Francesco De Cobelli[10], Xuebin Xie[9,*], Diego Palumbo[10,*], Jie Tian[1,4,5,*], Di Dong[1,2,*]

[1] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[2] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China
[3] University of Science and Technology Beijing, Beijing, China
[4] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing, China
[5] Key Laboratory of Big Data-Based Precision Medicine, Beihang University, Ministry of Industry and Information Technology, Beijing, China
[6] University of Science and Technology of China, Hefei, China
[7] Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China
[8] Beijing University of Chinese Medicine, Beijing, China
[9] KiangWu Hospital, Macau, China
[10] Department of Radiology, IRCCS San Raffaele Scientific Institute, Milan, Italy

ARTICLE INFO

Keywords:
Upper gastrointestinal cancers
Artificial intelligence
Radiomics
Esophageal cancer
Gastric cancer

ABSTRACT

Upper gastrointestinal cancers, mainly comprising esophageal and gastric cancers, are among the most prevalent cancers worldwide. There are many new cases of upper gastrointestinal cancers annually, and the survival rate tends to be low. Therefore, timely screening, precise diagnosis, appropriate treatment strategies, and effective prognosis are crucial for patients with upper gastrointestinal cancers. In recent years, an increasing number of studies suggest that artificial intelligence (AI) technology can effectively address clinical tasks related to upper gastrointestinal cancers. These studies mainly focus on four aspects: screening, diagnosis, treatment, and prognosis. In this review, we focus on the application of AI technology in clinical tasks related to upper gastrointestinal cancers. Firstly, the basic application pipelines of radiomics and deep learning in medical image analysis were introduced. Furthermore, we separately reviewed the application of AI technology in the aforementioned aspects for both esophageal and gastric cancers. Finally, the current limitations and challenges faced in the field of upper gastrointestinal cancers were summarized, and explorations were conducted on the selection of AI algorithms in various scenarios, the popularization of early screening, the clinical applications of AI, and large multimodal models.

1. Introduction

Upper gastrointestinal (UGI) cancers are among the most common cancers, primarily including esophageal cancer (EC) and gastric cancer (GC). According to the latest data, there were 510,716 new cases of EC and 968,350 new cases of GC worldwide in 2022, ranking 11th and 5th respectively in global cancer incidence. The geographical distribution of these cancers is uneven, with the highest incidence rates of EC in East Asia and East Africa, and the highest incidence rates of GC in East Asia, Eastern Europe, and South America.[1] In the same year, there were 224,000 new cases of EC and 358,700 new cases of GC in China,[2] accounting for approximately 43.86 % and 37.04 % of the global cases of these cancers, respectively. Alongside the high incidence rates, UGI cancers have high mortality rates and poor prognosis. In 2022, the number of deaths worldwide due to UGI cancers was >1.1 million, with 445,129 deaths from EC and 659,853 deaths from GC.[1] In China, >187,000 people died from EC and over 260,000 people died from GC[2], accounting for 42.12 % and 39.46 % of the global deaths from these cancers, respectively. The 5-year survival rate for UGI cancers was approximately 20 %.[3,4] Specifically, age-standardized 5-year net survival rates for EC and GC were generally in the range of 10–30 %[5] and 20–40 %,[6] respectively, with significant variations in prognosis across different countries.[6–8] The survival rate of UGI cancers is closely related to the time of discovery. Utilizing cancer statistics from the United States between 2013 and 2019 as an example. The 5-year survival rates for EC and GC were estimated to be approximately 49 % and 75 %, respectively, when detected in the early stages. Conversely, when detected in advanced stages, particularly when the cancer had metastasized, the 5-

* Corresponding authors.
E-mail addresses: xiexuebin@kwh.org.mo (X. Xie), palumbo.diego@hsr.it (D. Palumbo), jie.tian@ia.ac.cn (J. Tian), di.dong@ia.ac.cn (D. Dong).
† These authors contributed equally to this work.

https://doi.org/10.1016/j.jncc.2024.12.006
Received 13 June 2024; Received in revised form 17 September 2024; Accepted 20 December 2024

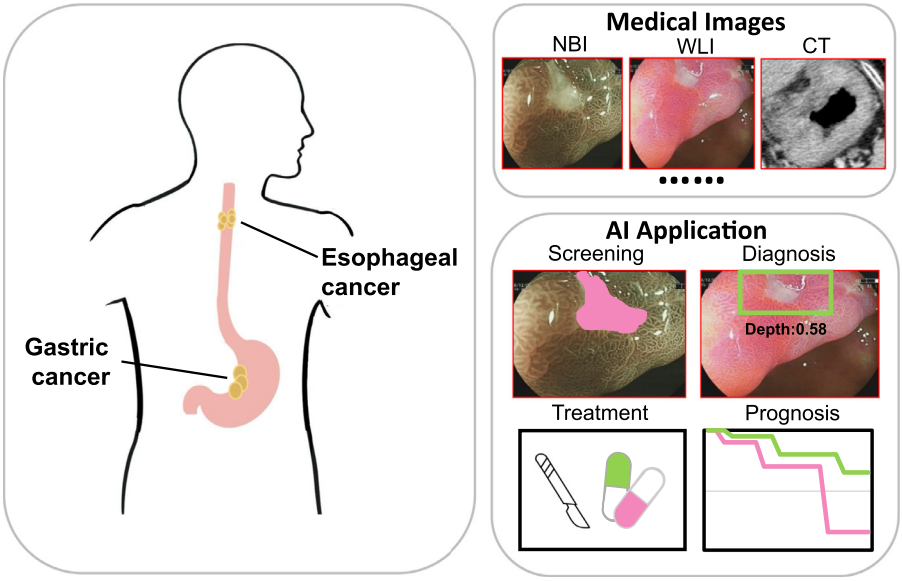

Fig. 1. The application of artificial intelligence in upper gastrointestinal cancers research primarily focuses on four aspects: screening, diagnosis, treatment, and prognosis. CT, computed tomography; NBI, narrow-band imaging; WLI, white-light imaging.

year survival rates for both GC and EC fell to <10 %.[9] Moreover, precise diagnosis and suitable treatment plans also have great significance for patients. Determining the N stage is crucial in the diagnosis of GC, the location and number of metastatic lymph nodes are one of the most important prognostic factors in GC surgery.[10] For EC patients, neoadjuvant chemotherapy combined with surgery improves survival more than surgery alone.[11] In general, there are numerous challenges in various clinical issues of UGI cancers. Thus, addressing or improving these challenges is critical and necessary.

In recent decades, rapid advancements have been witnessed in medical imaging technologies, significantly enhancing the speed and resolution of imaging devices. At present, various medical imaging technologies play a pivotal role in the diagnosis and prognosis of UGI cancers. For example, endoscopic ultrasound can better discriminate the T1 and T2 stages of GC; PET-CT can help detect occult metastasis of GC[12]; chromoendoscopy can detect pre-cancerous squamous dysplasia of EC[13]; and MRI is often used to detect a complete pathological response after neoadjuvant therapy for EC.[14,15] However, the volume of radiological imaging data is expanding significantly faster than the number of available trained doctors. This leads to a significant increase in the workload of radiologists.[16] At the same time, artificial intelligence (AI) technology has developed to be capable of extracting effective image information that the human eye cannot recognize for clinical doctors, as well as modeling complex associations between clinical indices and image features, which significantly alleviates the burden on doctors and improves the efficiency of medical image analysis.

In recent years, an increasing number of researchers have applied AI to clinical issues related to UGI cancers, leading to a series of significant advancements in the field.[17] In this review, our main emphasis is on the latest AI technologies employed in the clinical analysis of EC and GC (Fig. 1), along with an exploration of current constraints and potential avenues for future research.

2. AI in clinical medical analysis

The abundance of medical images provides a strong data foundation for AI applications, making AI well-suited for the clinical domain. AI can rapidly extract key features from medical image data, utilizing acquired knowledge to solve specific clinical problems. It also possesses the ability for continual learning and self-correction, enhancing accuracy through feedback.[18] Currently, numerous studies focus on AI in medical image analysis, with radiomics and deep learning (DL) being the primary technologies used.

Radiomics, initially proposed by Philippe Lambin and others in 2012, entails extracting high-dimensional quantitative image features from medical images and modeling to identify correlations between these features and specific clinical tasks.[19] Upon its proposal, it quickly attracted significant attention in the academic community and has experienced rapid development in recent years. Radiomics has already yielded a series of research outcomes in the field of cancer.[20]

DL is a subset of machine learning. Its basic idea is to use neural networks to automatically extract and filter feature representations from data.[21] Features extracted through this method more effectively characterize the inherent patterns of data and can be transferred to different clinical tasks. Currently, a multitude of DL models, including the convolutional neural network (CNN) and generative adversarial network (GAN), demonstrate immense potential in the field of medical image analysis.[22,23]

2.1. The pipeline of radiomics

The pipeline of radiomics mainly includes the following four steps: image segmentation, feature extraction, feature selection, and modeling (Fig. 2).

Image segmentation: This process refers to separating the region of interest (ROI) from the surrounding tissues by manual, semi-automatic, or automatic segmentation methods. This step determines which region of the image will be used for subsequent analysis, so segmentation accuracy is critically important. However, accurate segmentation is challenging due to the lack of an objective golden standard for tumor segmentation,[24] the partial volume effect on tumor edges, etc. Presently, extensive automated segmentation methods have been proposed,[25–27] but for most radiomics studies, manual ROI segmentation is still employed for research.[20]

Feature extraction: Feature extraction refers to extracting high-throughput features from images, which can be manually extracted or automatically extracted using DL models such as CNN. Manually-defined features often encompass shape features describing the tumor, first-order and second-order features constructed based on ROI voxel intensities, texture features, wavelet features, etc.[28] Features automatically extracted by DL models often lack interpretability, but their advantage lies in being more specific to clinical outcomes[29] and containing more comprehensive information. They are capable of extracting multi-level tumor information from low-level visual features to high-level abstract features.

Feature selection: The fundamental concept of feature selection is to eliminate redundant features through algorithms and identify true

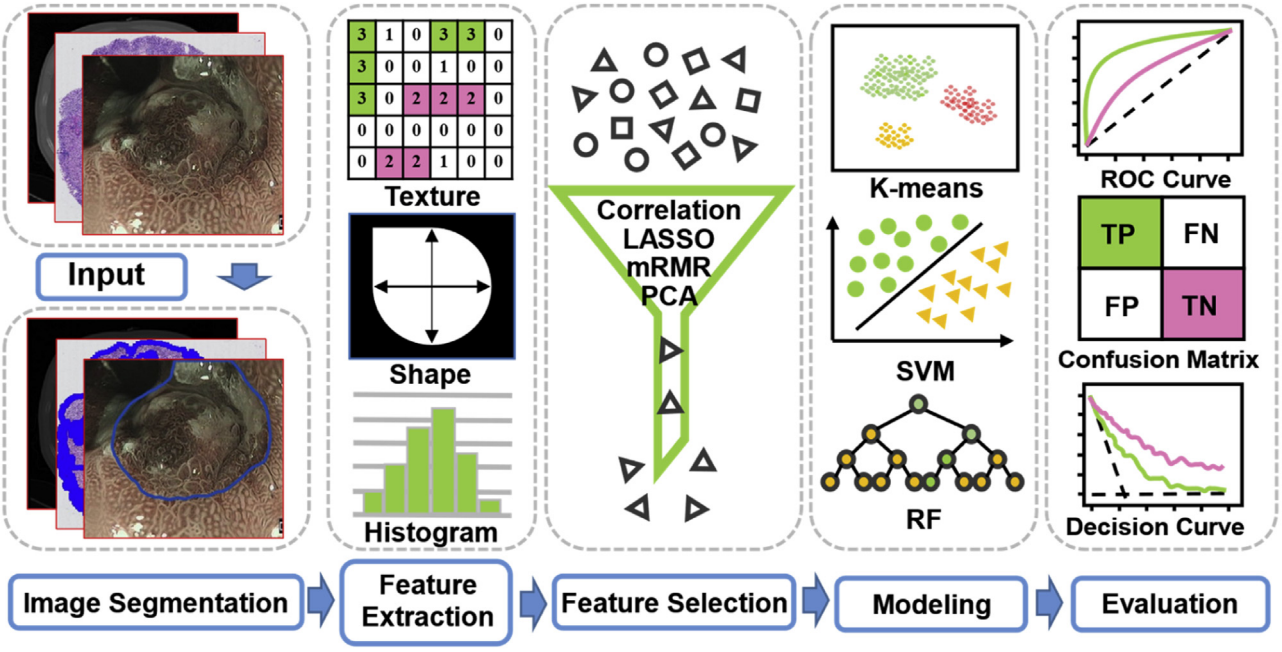

Fig. 2. The pipeline of radiomics. The input images undergo image segmentation initially to acquire the region of interest. Subsequently, features such as texture, shape, and histogram are extracted from the segmented regions. LASSO or other methods are then employed for feature selection. Following this, machine learning methods such as K-means, SVM, and RF are utilized for modeling based on the selected features. Finally, the performance of the model is evaluated using metrics like the ROC curve, confusion matrix, and decision curve. FN, false negative; FP, false positive; LASSO, least absolute shrinkage and selection operator; mRMR, max-relevance and min-redundancy; PCA, principal component analysis; RF, random forest; ROC, receiver operating characteristic; SVM, support vector machine; TN, true negative; TP, true positive.

features that are beneficial for clinical task prediction.[30] Therefore, feature selection is a very critical step in radiomics, which can improve the efficiency and accuracy of the model. Common feature selection methods in radiomics include least absolute shrinkage and selection operator (Lasso) regression, minimal redundancy maximal relevance (mRMR), chi-square test, analysis of variance, etc.

Modeling: Modeling in radiomics plays a critical role in developing predictive models that can perform various clinical tasks based on selected imaging features. These models usually rely on machine learning to extract meaningful patterns from the data. For example, the Cox proportional-hazards model is a widely used and effective tool for prognosis tasks such as predicting overall survival and progression-free survival.[31–33] Logistic regression is another powerful and efficient classification algorithm commonly applied in radiomics. It can be utilized either directly or indirectly in various clinical tasks, such as predicting distant metastasis,[34] differentiating between different types of tissue or disease states,[35] and, when combined with nomograms, enhancing prognostic predictions.[36,37] Support vector machines (SVMs) also play a significant role in radiomics modeling, particularly in tasks that involve high-dimensional data and require robust classification. SVMs can be employed for tasks such as tumor classification, treatment response prediction,[38] and survival prediction,[39] making them valuable tools in the clinical decision-making process. Moreover, other machine learning algorithms such as random forest (RF)[40] and XGBoost[38,41] are also widely applied in these clinical tasks, further enhancing the accuracy and reliability of predictive models in radiomics.

2.2. The pipeline of deep learning

DL methods offer greater flexibility compared to radiomics. They can select various DL models and design customized pipelines based on specific task requirements (Fig. 3). The main model used in medical image analysis is commonly based on these three architectures, CNN, GAN, and Transformer. The following will briefly introduce some classic models for medical image analysis.

CNN is a classic DL model primarily composed of convolutional layers, pooling layers, and fully connected layers. The role of convolutional layers and pooling layers is to extract features, while fully connected layers map features to outputs.[42] Since CNN was proposed, it has demonstrated outstanding performance in various fields, including medical image analysis. Representative variants based on CNN architecture include U-net,[43] VGG,[44] and SegNet.[45] U-net is a significant variant of CNN in medical image analysis. It is modified and extended from fully convolutional networks, characterized by a symmetric contracting path (Contracting Path) and an expanding path (Expansive Path) to transmit and fuse information at different network levels. This architecture allows it to produce more precise segmentation with few training images.[43] A multitude of studies have showcased the strong performance of U-net in medical image segmentation tasks of various cancers and other structures.[46–48]

GAN was introduced by Ian J. Goodfellow et al. in 2014. It serves as a framework for estimating generative models through an adversarial process. The core idea is to simultaneously train two models: a generative model G and a discriminative model D, and have them constantly oppose each other during training. G will try to generate more realistic data to deceive D, while D will try to improve the ability to distinguish real from fake data.[49] CycleGAN,[50] Pix2Pix,[51] and other similar models are classic works based on GANs. In medical image analysis, GAN is extensively used in various medical image synthesis tasks. Whether for unconditional synthesis or cross-modality synthesis, GAN performs excellently.[52] In addition, GAN also performs well in medical image segmentation, detection, etc.[52–54]

Transformer was originally a DL framework applied to natural language processing, demonstrating excellent performance in machine translation, text understanding, etc. The proposal of vision-transformer (ViT) shows that this framework is also effective in computer vision.[55] Subsequent studies have revealed that transformer architecture has huge potential in the visual domain, such as swin-transformers[56] and TransUNet,[57] surpassing CNN in many tasks.[58] Specifically in medical image analysis, transformer architecture also has significant effects on medi-

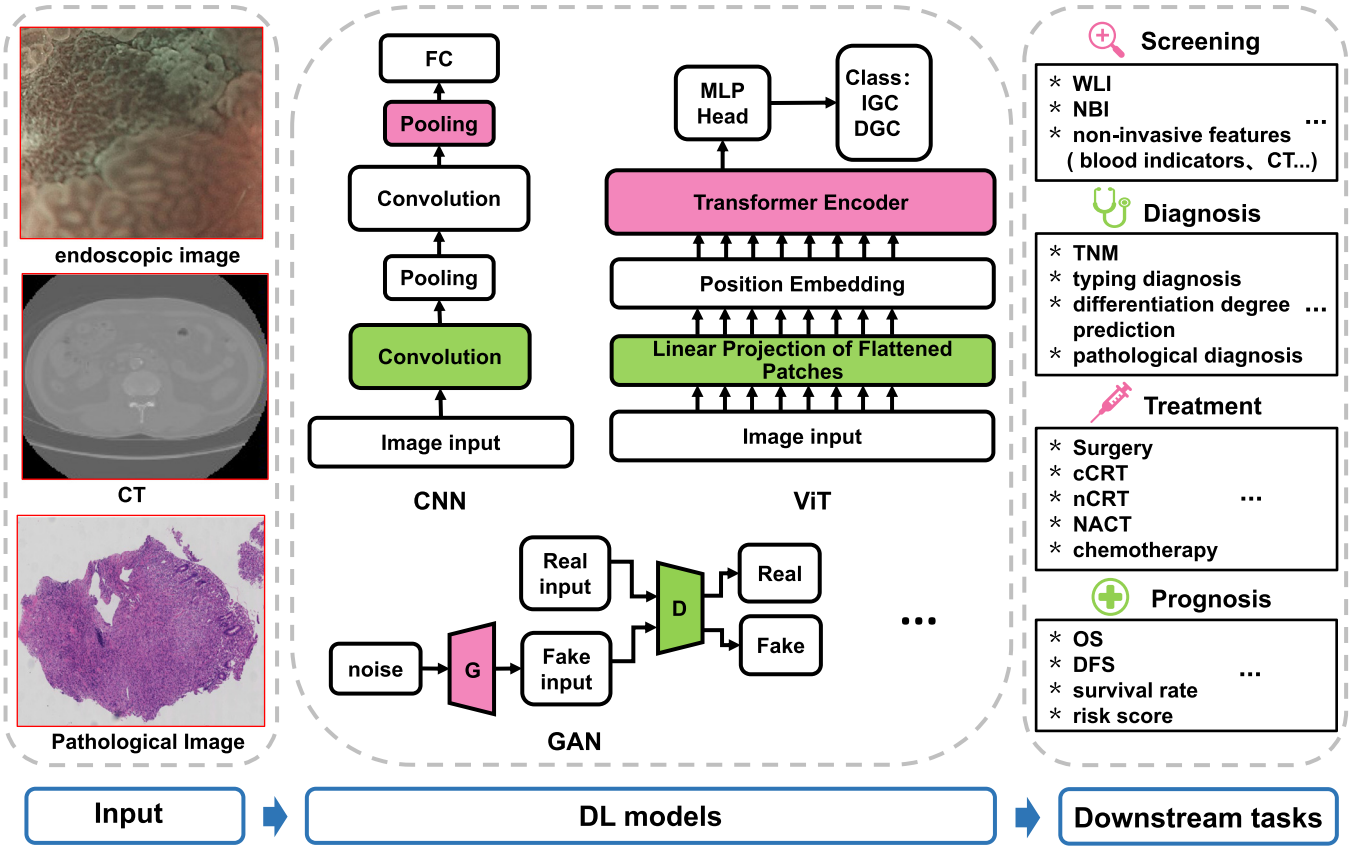

Fig. 3. The pipeline and classic model of deep learning in medical imaging. CNN, ViT, and GAN are commonly used in deep learning models in medical image analysis. Upon training these models, they can be applied to specific downstream tasks. The pipeline of deep learning is highly flexible, allowing for modification of model architectures and autonomous design of training tasks according to specific requirements. cCRT, concurrent chemoradiotherapy; CNN, convolutional neural network; D, discriminator; DFS, disease-free survival; DGC, diffuse gastric cancer; DL, deep learning; FC, fully connected layer; G, generator; GAN, generative adversarial networks; IGC, intestinal gastric cancer; MLP, multilayer perceptron; NACT, neoadjuvant chemotherapy; NBI, narrow-band imaging; nCRT, neoadjuvant chemoradiotherapy; OS, overall survival; ViT, vision transformer; WLI, white-light imaging.

cal image segmentation, medical image classification, reconstruction, etc.[56,59,60]

3. AI in esophageal cancer

EC refers to cancers occurring in the esophagus, which can mainly be categorized into two types: esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC).[61] EC ranks as the eighth most common cancer in the world, with approximately 500,000 new cases each year.[62] The effectiveness of its treatment is associated with the local medical level, and the 5-year survival rate is often low.[63] To improve the survival rate of patients, in recent years, many scholars and researchers have been dedicated to applying AI technology to EC screening, diagnosis, treatment, and prognosis. These studies have made important contributions to the medical treatment of EC patients.

3.1. Screening

The early symptoms of EC are obscure, lacking specific clinical features, with approximately 90 % of EC patients already being diagnosed with advanced-stage disease at the time of diagnosis. Effective early screening can greatly improve the survival rate of patients with EC. Currently, endoscopic examination is the mainstream screening method, but due to its high cost and low efficiency, it is not suitable for large-scale population screening. Therefore, it is necessary to utilize AI to improve the efficiency and accuracy of early endoscopic screening (Table 1).

White-light imaging (WLI) endoscopy is the most commonly used screening technique for early detection of EC, but inexperienced doctors often find it challenging to detect early EC with WLI endoscopy.[64] AI can enhance the detection efficiency and accuracy of WLI endoscopy. Currently, numerous studies have been conducted to explore this domain. Hou et al.[65] added an attentive hierarchical aggregation module and an efficient self-distillation mechanism based on the teacher-student architecture, with SEResNet50 as the backbone, to ensure that the model can selectively integrate the logic of each feature. The general framework of this model is shown in Fig. 4. The model demonstrated good prediction and generalization capabilities in the task of identifying early EAC. Liu et al.[66] used a CNN for the early detection of ESCC, and the best-trained model outlined the edges of EC more accurately than expert endoscopists and senior endoscopists.

Narrow-band imaging (NBI) endoscopy is another frequently used screening technique. It relies on spectral technology to display the range of lesions and the morphology of intrapapillary capillary loops (IPCL).[67,68] Compared to WLI, NBI enhances the visibility of lesions, thereby improving the detection rate of endoscopists for lesions. It exhibits better screening performance than WLI but still demands high experience from doctors. Several studies have achieved notable results by using a combination of NBI images and WLI images as training data. Horie et al.[69] trained CNN on 8428 EC images from 384 patients, encompassing both WLI and NBI images. The trained model achieved a sensitivity of 77 % for diagnosis on the test set including WLI and NBI, correctly detecting 125 cancer lesions out of 162 cancer images. Wang et al.[70] built a single-stage multi-box detector based on CNN and trained it on 498 WLI images and 438 NBI images, achieving fast detection speed and excellent model performance. Additionally, they found that this model performed better in diagnosing NBI (95 %) images than WLI

Table 1
Radiomics and DL studies of screening for EC.

Authors (year)	Type of task	Modeling	Evaluation
Horie Y. (2019)	Detect ESCC and EAC	CNN	Accuracy 98 %,
Wang Y. (2021)	Detect ESCC	Single-shot multi-box detector-based CNN	Accuracy 96.2 %, Sensitivity 70.4 %, Specificity 90.9 %
Chou C. (2023)	Detect ESCC	A hybrid model integrating EfficientNet and ViT	Accuracy 96.3 %, F1-score 96.0 %
Hou W. (2021)	Detect EAC	Teacher-student architecture with SEResNet50 as the backbone	Accuracy 92.0 %, Sensitivity 89.0 %, Specificity 87.0 %, F1-score of 89.0 %
Liu W. (2022)	Detect ESCC	CNN	Accuracy 84.5 %
Feng Y. (2022)	Detect ESCC	AI-assisted sponge cytology	-
Zhang P. (2022)	Detect EC	Two-stage DL-based R-CNN	Accuracy 90.3 %, Sensitivity 92.5 %, Specificity 88.7 %
Takeuchi M. (2021)	Detect EC	VGG16	Accuracy 84.2 %, Sensitivity 71.7 %, Specificity 90.0 %

Abbreviations: AI, artificial intelligence; CNN, convolutional neural network; DL, deep learning; EAC, esophageal adenocarcinoma; EC, esophageal cancer; ESCC, esophageal squamous cell carcinoma; R-CNN, region-CNN; VGG16, visual geometry group 16; ViT, Vision Transformer.

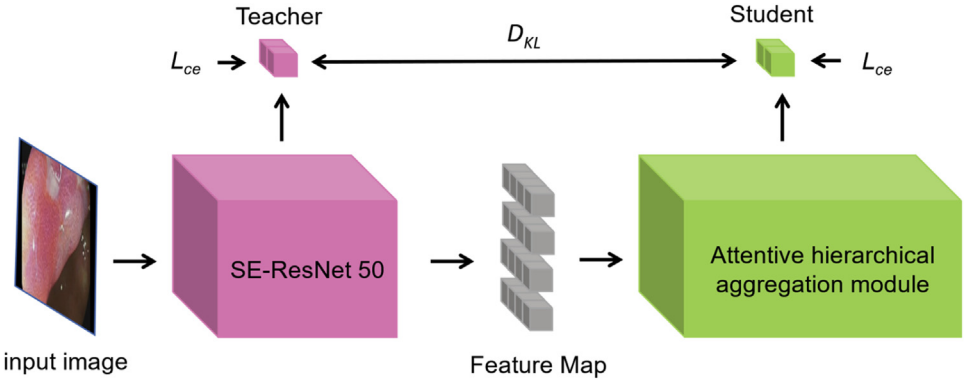

Fig. 4. The general framework of the model proposed by Hou,[65] consisting mainly of two modules: SE-ResNet 50 and the attentive hierarchical aggregation module. These serve as the student and teacher, respectively. The entire model is trained in a supervised learning setting with self-distillation. L_{ce} represents the cross-entropy loss, while D_{KL} denotes the KL divergence.

(89 %) images in terms of accuracy. Some other studies explored the impact of NBI images on models. Chou et al.[71] proposed a hybrid model combining EfficientNet and ViT. They conducted comparative experiments using different training data (only WLI images, only NBI images, mixed WLI and NBI images). The experiments found that the model trained only on NBI data performed worst, and mixing WLI and NBI images as training data could improve model performance.

In addition to applying AI to endoscopic examination, certain studies have also explored the feasibility of integrating other screening methods with AI. Feng et al.[72] considered the high cost of endoscopy screening and investigated the possibility of AI-assisted sponge cytology as a pre-endoscopy detection tool. The results prove that it significantly reduces the average cost of endoscopy, while also improving the accuracy of large-scale screening. Zhang et al.[73] established a DL model based on Faster R-CNN to detect EC through an esophageal barium meal. With the assistance of the model, the interpretation time of radiologists was considerably shortened, and the diagnostic accuracy increased from 89.3 % to 96.8 %. Takeuchi et al.[74] employed AI to screen EC from CT images. They trained a CNN-based model on thousands of CT images with an average accuracy of 86.4 % ± 5.6 % in the validation set and an accuracy of 84.2 % in the test set.

3.2. Diagnosis

Typical diagnoses include histological diagnosis, differentiation degree diagnosis, and tumor-node-metastasis (TNM) staging diagnosis. Among these, TNM staging diagnosis holds paramount importance. In T staging diagnosis, the depth of invasion of the primary tumor is important for determining the necessity of chemotherapy or radiotherapy and the choice of surgical approach; in N staging diagnosis, assessment of lymph node metastasis (LNM) guides the strategy for lymph node dissection during surgery; and in M staging diagnosis, the presence of distant metastasis plays a decisive role in determining the suitability for surgery. In the following, we will introduce relevant research on applying AI technology to EC diagnosis (Table 2).

To assess the depth of EC invasion, Tokai et al.[75] trained a CNN model based on 1751 ESCC endoscopic images, which could detect ESCC in 95.5 % (279/291) of test images within 10 s, analyze 279 images within 6 s and correctly estimate the depth of invasion of ESCC with a sensitivity of 84.1 % and accuracy of 80.9 %. Shimamoto et al.[76] utilized richer training data, including 6857 WLI images and 17,120 NBI/blue laser imaging (BLI) images. They used the visual geometry group network as the backbone and the single shot multibox detector as the detection algorithm. The trained model achieved an accuracy of 87.3 % and 89.2 % in detecting invasion depth in magnified endoscopy (ME) and non-ME images, respectively. These results were comparable to or even better than those of expert endoscopists with extensive experience.

In the domain of predicting LNM, Peng et al.[77] extracted radiomic features from different phases (non-enhanced images and enhanced images) and used LASSO to eliminate redundancy in the feature group. The radiomics model built on these combined features achieved an area under the curve (AUC) of 0.811; Wang et al.[78] developed a multi-instance DL model to predict lymph node station metastasis, demonstrating a clinical diagnostic recall rate of 63.21 %, which is higher than that of traditional predictive models; and Zhang et al.[79] developed an AI-based computer-aided diagnosis system that emulates the diagnostic logic of

Table 2
Radiomics and deep learning for the diagnosis of esophageal cancer.

Authors (year)	Type of task	Modeling	Evaluation
Tokai Y. (2020)	Predict invasion	CNN	Accuracy 80.9 %, Sensitivity 84.1 %
Shimamoto Y. (2020)	Predict invasion	CNN	Accuracy 87 %, Sensitivity 50 %, Specificity 99 %
Peng G. (2022)	Predict LNM	CT radiomics, Multivariate logistic regression analysis	AUC 0.811
Wang Y. (2023)	Predict LNM	multi-instance deep learning model	AUC 0.7539
Zhang S. (2023)	Predict LNM	artificial intelligence-based computer-aided diagnosis system	Accuracy 0.744
Zhu C. (2022)	Predict distant metastasis	CT radiomics, LASSO logistic regression, Nomogram	AUC 0.827
Akashi T. (2023)	Predict distant metastasis	Classifier based CNN	Accuracy 99.8
Zhou Z. (2022)	Predict typing	Multi-modal CNN	AUC 0.92
Kawahara D. (2021)	Predict differentiation degree	CT radiomics, LASSO logistic regression	Accuracy 85.4 %, Specificity 88.6 %, Sensitivity 80.0 %, AUC 0.92

Abbreviations: AUC, area under the curve; CNN, convolutional neural network; CT, computed tomography.

radiologists. This system first utilizes U-NET to obtain anatomical information from CT images, then uses gradient-based methods to extract metabolic information from PET images, and subsequently predicts LNM, achieving good consistency with human experts in external validation.

To predict distant metastasis of EC, Zhu et al.[34] used logistic univariate and multivariate regression analyses to identify independent clinical predictive factors and create clinical flowcharts for predicting distant metastasis of EC. The AUC of this model is 0.827, exhibiting good prediction performance. Circulating tumor cells (CTCs) are considered predictive factors for distant metastasis and cancer recurrence.[80] Efficient and accurate identification of CTCs aids in assessing distant metastasis of EC. Akashi et al.[81] first established a DL model based on CNN, then trained it to distinguish ESCC cell lines from peripheral blood mononuclear cells (PBMCs), and subsequently applied it to detect CTCs in patients with ESCC. Comparative experiments found that this model highlighted high accuracy and shorter analysis time, making it suitable for clinical application in ESCC patients. In addition, AI can distinguish CTCs based on unknown features, regardless of known marker expression.

Apart from TNM staging diagnosis, many studies also concentrate on various sub-directions of diagnosis, such as typing diagnosis, differentiation degree prediction, etc. EC is typically categorized as ESCC and EAC based on histological type. Predicting the type of EC aids in formulating personalized treatment plans. Zhou et al.[82] proposed a multi-modal convolutional neural network (CNN) architecture employing dynamic CT and preoperative WSI for EC classification. The model achieved an AUC of 0.92, proving that the fusion of the two modalities helps the model's typing diagnosis of EC. The differentiation degree is a crucial indicator for measuring the invasiveness of EC. Preoperative prediction of the differentiation degree through AI technology can help reduce the need for biopsy pathology examinations and predict local control. Kawahara et al.[35] used LASSO logistic regression to construct a model based on CT radiomic features for differentiation prediction in locally advanced EC patients. The best predictive model has an AUC of 0.92, indicating the approach has the potential to diminish the need for pathological examination via biopsy and enhance predictions for local control.

3.3. Treatment response prediction

Currently, clinical treatment methods for EC patients include concurrent chemoradiotherapy (cCRT), neoadjuvant chemoradiotherapy (nCRT), definite chemoradiotherapy, esophagectomy, etc. These treatment methods can, to some extent, reduce tumor burden, lower recurrence and metastasis risks, and prolong patient survival. However, not all EC patients can benefit from these treatments. Therefore, to formulate more personalized treatment plans to improve patient outcomes, reduce postoperative risks, and reduce the possibility of complications, it is necessary to employ AI technology to predict the impact of treatment modalities on patients (Table 3). cCRT is a combined modality treatment method using chemotherapy and radiotherapy. Research has demonstrated the superiority of this approach compared to standalone radiotherapy or chemotherapy. However, this approach was also associated with severe or life-threatening toxicity in 44 % and 20 % of patients, respectively, with an overall compliance rate of 54 %.[83] Currently, there have been research efforts utilizing AI technology to predict the therapeutic response of cCRT, thereby optimizing treatment regimens and avoiding adverse events. Li et al.[84] developed and validated a 3D DL radiomics model based on pre-treatment CT images to predict the response to cCRT. The model performed well in all three prediction categories, with an AUC > 0.8 for each category and a positive predictive value of approximately 100 %. An et al.[85] conducted a univariate analysis to explore the associations between clinical features, apparent diffusion coefficient (ADC) values, radiomic features, and treatment response to cCRT. Given the large number of features, they first applied the minimum redundancy maximum relevance (mRMR) algorithm to identify and rank the top 30 radiomic features for training and testing. Ultimately, they found that the primary tumor site was associated with treatment response, while ADC and delta-ADC values showed no significant correlation with treatment response. Jin et al.[38] separately used SVM and XGBoost to build models based on radiomic features and dosimetric parameters, and the experiments demonstrated that the combination of radiomic features and dosimetric parameters helped their models predict the treatment response to cCRT. nCRT is the administration of chemotherapy and radiation therapy before surgery. This approach aims to shrink the tumor, making it easier to remove during surgery, and also to potentially improve long-term outcomes by targeting any remaining cancer cells in the area. Approximately 29 % of patients achieve pCR following nCRT, indicating that patients undergoing nCRT may achieve outcomes similar to esophagectomy while preserving organs.[86] Predicting pCR following nCRT through AI can determine the necessity of subsequent surgical intervention, facilitating the optimization of treatment strategies. Some studies predict treatment response based on PET image radiomics. For example, Murakami et al.[87] extracted radiomic features from pretreatment fluorodeoxyglucose positron emission tomography (FDG PET) images and established a prediction model based on them. After 5-fold validation, it was demonstrated that the use of FDG PET radiomics to predict local response in locally advanced ESCC treated by nCRT is feasible. Beukinga et al.[88] created a radiomics model based on

Table 3
Radiomics and DL studies of treatment response prediction for esophageal cancer.

Authors (year)	Type of task	Modeling	Evaluation
Li X. (2021)	Predict treatment response of cCRT	3D DL radiomics model	AUC 0.833
An D. (2020)	Predict treatment response of cCRT	Radiomics, Multivariate logistic regression	AUC 0.823
Jin X. (2019)	Predict treatment response of cCRT	Radiomics, SVM, XGBoost	Accuracy 0.708, AUC 0.541
Murakami Y. (2021)	Predict treatment response of nCRT	FDG PET radiomics	AUC 0.927
Beukinga R. (2021)	Predict treatment response of nCRT	FDG PET radiomics	AUC 0.857
Liu Y. (2023)	Predict treatment response of nCRT	MRI radiomics	AUC 0.800, Sensitivity 79.2 %, Specificity 83.7 %
Lu S. (2023)	Predict treatment response of nCRT	MRI radiomics	AUC 0.831
Chufal KS. (2021)	Predict treatment response of nCRT	CNN	Accuracy 0.74, AUC 0.697
Wang Q. (2023)	Predict treatment response of nCRT	Radiomics, Disentangled representation network	AUC 0.8429
Puttanawarut. C. (2021)	Predict RP	Dosiomics, Multivariate logistic regression	AUC 0.71
Sheng. L. (2023)	Predict RP	Hybrid DL model based on Resnet18	Accuracy 0.85, AUC 0.91
Li. Z. (2023)	Predict EF	Radiomics, Univariate and multivariate stepwise logistic regression analysis	AUC 0.896
Xu. Y. (2021)	Predict EF	Attentional multi-view multi-scale CNN	C-index 0.901

Abbreviations: AUC, area under the curve; CNN, convolutional neural network; cCRT, concurrent chemoradiotherapy; nCRT, neoadjuvant chemoradiotherapy; DL, deep learning; EF, esophageal fistula; FDG PET, fluorodeoxyglucose positron emission tomography; RP, radiation pneumonia; SVM, support vector machine.

^{18}F-FDG PET and attempted to introduce tumor markers, demonstrating that incorporating both HER2 and CD44 tumor markers improved the overall performance of the model. There are also radiomic studies based on other modalities. Liu et al.[89] validated the feasibility of this method using clinical blood markers and MRI radiomics. Lu et al.[90] first explored the value of MRI radiomic features based on T2-TSE-BLADE images in predicting the response of ESCC patients to neoadjuvant chemotherapy (NACT). Furthermore, certain investigations utilize CT imaging as the foundation for modeling and predictive analysis. Chufal et al.[91] trained a CNN model using pre-nCRT CT images, achieving an accuracy of 74 % on the test set, and found that the region within the gross tumor volume (GTV), excluding the esophageal lumen, was most predictive of pathological response. Wang et al.[92] combined DL with radiomics, inputting radiomic features into a disentangled representation network to explore the relationship between dynamic changes in tumors and pathological response before and after nCRT.

EC may experience some complications or side effects during or after treatment, such as radiation pneumonia (RP) and esophageal fistula (EF). These complications can disrupt subsequent treatment processes while seriously affecting the patient's quality of life. RP is one of the major complications of thoracic radiation therapy.[93] In terms of predicting RP, Puttanawarut et al.[94] established a model based on dose-volume histograms (DVH) and dosimetric features using multivariate logistic regression with L1 norm regularization, which demonstrated good predictive ability for RP occurrence. This study also proved that dosiomic features extracted from dose distributions can improve the performance of RP prediction models. Sheng et al.[95] designed a novel hybrid DL network based on ResNet18, integrating clinical features and dose distribution matrics to predict RP and achieving an accuracy of 0.85 and an AUC of 0.91. EF is another serious complication, with about 4–24 % of patients at risk after radiation therapy for EC.[96–99] In the realm of predicting EF, Li et al.[100] recorded and extracted clinical and radiomic features of included patients, determined risk factors related to EF by univariate and multivariate stepwise logistic regression analysis, proving tumor length, tumor volume, T stage, lymphocyte rate, and grade 4 esophageal stenosis were related to EF, and used a nomogram to predict EF with an AUC of 0.896. Xu et al.[101] developed an attentional multi-view multi-scale CNN model to extract features from CT, then used fully connected layers to combine radiomic and clinical features, and finally used SoftMax classifier to predict EF. The model achieved a C-index of 0.901 on the validation set. By comparing with models utilizing CT imaging alone or clinical data alone, it was demonstrated that the combination of clinical data and radiomic features improved the model's predictive performance.

3.4. Prognosis

Despite continuous advances in surgical techniques and medical methods, most EC patients have a poor prognosis, with an overall 5-year survival rate ranging from only 9 % to 22 %.[102] Improved accuracy and effectiveness in prognosis prediction are conducive to offering more reasonable guidance for individualized treatment decisions. In recent years, several studies have applied AI to tumor prognosis prediction and achieved good results. The following will introduce present pertinent research on EC prognosis prediction (Table 4).

Regarding EC, research on prognosis prediction can be broadly divided into two categories according to the treatment methods of the study objects: one is the prognosis prediction of patients who underwent esophageal resection, and the other is patients who received chemoradiotherapy. We first introduce studies on prognosis prediction related to esophagectomy. Gujjuri et al.[31] established overall survival (OS) and disease-free survival (DFS) prediction models using Cox proportional hazard (CPH) and random survival forest (RSF), effectively predicting long-term survival rates and recurrence times after esophagectomy; Rahman et al.[32] developed an OS prediction model using the RSF method and Cox regression based on 41 patients and disease characteristics with an AUC of 83.9 %; and Jung et al.[103] also used RSF to predict prognosis for patients who underwent esophagectomy. In addition, by comparing various machine learning algorithms, their demonstration highlighted that RSF is particularly suitable for long-term prognostic prediction. Besides work focusing solely on esophagectomy, there are also studies on prognostic prediction of neoadjuvant treatment. DeFreitas et al.[104] validated through variable analysis that CT-derived body composition measurements can serve as predictive factors for prognosis in EAC patients after neoadjuvant treatment. Based on this, they established a Cox model that effectively predicted the risk scores for these patients.

Studies on prognostic prediction for chemoradiotherapy are more diverse. Kawahara et al.[36] established a nomogram based on CT, PET,

Table 4
Radiomics and DL studies of prognosis for EC.

Authors (year)	Type of task	Modeling	Evaluation
Gujjuri RM. (2023)	Predict survival rate after esophageal resection	CPH, RSF	AUC 0.771(OS, RSF), AUC 0.782 (OS, CPH), AUC 0.786 (DFS, RSF), AUC 0.794 (DFS, CPH)
Rahman SA. (2023)	Predict survival rate after esophageal resection	RSF, Cox	AUC 0.839 (RSF), AUC 0.823 (Cox)
Jung JO. (2023)	Predict risk score	RSF	AUC 0.814
DeFreitas MR. (2023)	Validate predictive factor	U-net, Univariate and multivariate analyses	P values 0.01 (BMI), P values 0.004 (SMI)
Kawahara D. (2023)	Predict OS and PFS after radiotherapy	CT radiomics, PET radiomics, Nomogram	C-index 0.92
Lin Z. (2022)	Predict survival rate after cCRT	Coordinate Attention Convolutional AutoEncoder, uncertainty-based jointly Optimizing Cox Model.	C-index 0.72
Gong J. (2022)	Predict LRFS after cCRT	3D Densenet	C-index 0.7167
Yu N. (2022)	Predict survival rate after definite chemoradiotherapy	MRI radiomics, CT radiomics, LASSO	C-index 0.689
Cui J. (2023)	Predict OS and PFS after definite chemoradiotherapy	Radiomics, Genomics, Univariate and multivariate Cox analyses	C-index 0.659

Abbreviations: AUC, area under the curve; BMI, body mass index; cCRT, concurrent chemoradiotherapy; CPH, Cox proportional hazard; CT, computed tomography; DFS, disease-free survival; DL, deep learning; EC, esophageal cancer; OS, overall survival; PET, positron emission tomography; PFS, progression-free survival; RSF, random survival forest; SMI, skeletal muscle index.

and dosiomic radiomic features selected by LASSO combined with logistic regression, to predict OS for patients receiving radiotherapy. Lin et al.[105] proposed a survival prediction framework combining 3D coordinate attention convolutional AutoEncoder and uncertainty-based jointly optimizing Cox model for survival prediction of cCRT. Gong et al.[106] developed a local recurrence-free survival (LRFS) prediction model based on a 3D DenseNet, which effectively captures and models the relationships between features relevant to the target task. The experimental results, achieving a C-index of 0.7167 on an external dataset, strongly demonstrated the capability of this method. Yu et al.[107] combined MRI and CT radiomic features to predict the 2-year OS of locally advanced EC patients treated with definitive chemoradiotherapy. The test results demonstrated that the combination of these two features significantly improved prognostic prediction; Cui et al.[33] established Cox proportional hazard regression models combining radiomics, genomics, and clinical factors, achieving a C-index of 0.649 on the validation cohort. This performance surpassed that of the radiomic-only model (C-index: 0.625) and the genomic-only model (C-index: 0.586).

4. AI in gastric cancer

GC is the fifth most common malignant tumor worldwide and the fourth leading cause of cancer-related deaths globally.[108] China is a country with a high incidence of gastric cancer, second only to lung cancer in terms of its occurrence rate.[109] Although there has been a downward trend in the incidence and mortality rates of GC in recent years, the cancer continues to impose a significant social burden due to its large population base, increasing aging population, and a high proportion of patients being diagnosed at advanced stages.[110]

Currently, the clinical diagnosis and treatment of gastric cancer heavily rely on gastroscopy, CT, and qualitative judgments made by physicians. However, these methods are highly dependent on the personal experience of doctors, leading to diagnostic uncertainties.[111,112] With the continuous application of AI in the field of medicine, significant developments and promising clinical applications have emerged in the exploration of GC screening, diagnosis, treatment response prediction, and prognosis.

4.1. Screening

Esophagogastroduodenoscopy (EGD) is a fundamental procedure for detecting early gastric cancer (EGC). However, the diagnostic accuracy of EGC is subject to the subjectivity of endoscopists, leading to a missed diagnosis rate of 20 % to 40 %.[113] In recent years, significant advancements have been made in the field of AI in medical image recognition.[114] It has been demonstrated that AI can effectively detect blind spots during the screening of EGD. Some representative works are shown in Table 5.

WLI endoscopy is a standard protocol for examining gastric lesions. Niikura et al.[115] confirmed the non-inferiority of AI in diagnosing GC. About 23,892 white-light EGD images from 500 patients were utilized to compare the diagnostic rates of GC between an AI algorithm based on the single shot MultiBox detector and expert endoscopists. Ultimately, the AI algorithm achieved a higher diagnostic rate for GC (99.87 %) compared to that of the expert endoscopists (88.17 %). Furthermore, Oura et al.[116] developed a double check support system with 12,977 still white-light endoscopic images from 855 cancer patients, which incorporated the function of lesion detection based on cascade region-CNN and low-quality images detection based on DenseNet121. Validated results showed a detection rate of 93.2 % for EGC and 92.5 % for malignant lymphoma, as well as excellent performance in identifying low-quality images. However, the study conducted by Tang et al.[117] not only demonstrated superior diagnostic capabilities of AI compared to endoscopists (accuracy, AI vs. experts vs. trainees: 95.3% vs. 87.3% vs. 73.6 %) but also highlighted its potential to improve the performance of endoscopists (accuracy: AI+experts, 94.3 %; AI+trainees, 96.2 %). They based this on a multicenter retrospective study where they developed and validated a real-time CNN system for detecting EGC. The system was trained and validated on 35,823 and 10,931 endoscopic images, respectively, and achieved an accuracy ranging from 85.1 % to 91.2 % in validation datasets. He et al.[118] extended the EGC screening model from image-based applications to video data and prospectively validated it. Based on ResNet-50 and 10,693 EGD images, they developed an EGC diagnostic system and evaluated it with 187 videos. In 93 internal videos, the system achieved a higher accuracy (90.32 %) compared to that of senior endoscopists (70.16 % ± 8.78 %). In 94 external videos, endoscopists demonstrated improved diagnostic performance with the assistance of the system. Furthermore, the system exhibited excellent clinical

Table 5
Radiomics and DL studies of screening for GC.

Authors (year)	Type of task	Modeling	Evaluation
Niikura R. (2022)	Detect EGC	The Single Shot MultiBox Detector	Accuracy 99.87%
Oura H. (2022)	Detect EGC lesion and malignant lymphoma	Cascade Region-CNN, DenseNet121	Accuracy 93.2 % (EGC), Accuracy 92.5 % (malignant lymphoma)
Tang D. (2020)	Detect EGC	CNN	Accuracy 85.1 %-91.2 %, Sensitivity 85.9 %-95.5 %, Specificity 81.7 %-90.3 %, AUC 0.887-0.940
He X. (2022)	Detect EGC	ResNet-50	Accuracy 83.67 %, Sensitivity 92.59 %
Hu H. (2021)	Detect EGC	VGG19	Accuracy 0.775, Sensitivity 0.767, Specificity 0.742, AUC 0.813
Ling T. (2022)	Delineate EGC margins	U-NET++	Accuracy 82.7 % (differentiated EGC), Accuracy 88.1 % (differentiated EGC)
Gong L. (2022)	Describe the visual features of ME-NBI images for endoscopists	ResNet101 network, Meshed-memory transformer	B1 52.434, C 36.734, M 27.823, R 49.949, S 35.548 (internal test cohort) B1 51.028, C 39.240, M 25.745, R 46.634, S 28.825 (external test cohort)*
Zhu SL. (2020)	Detect GC	Gradient Boosting Decision Tree	Accuracy 83.0 %, Sensitivity 87.0 %, Specificity 84.1 %, AUC 91 %
Taninaga J. (2019)	Detect GC	XGBoost	AUC 0.899
Zhang B. (2022)	Detect GC	H2O framework	Accuracy 84.54 %, Sensitivity 85.44 %, Specificity 83.82 %, AUC 0.9165
Tang JW. (2024)	Detect *Helicobacter pylori* infection	Light Gradient Boosting Machine	Accuracy 99.54 %, Precision 99.54 %, Recall 99.54 %, F1 99.54 %, Accuracy (5-fold CV) 99.21 %
Si YT. (2024)	Detect different stages of gastric disease	XGBoost	Accuracy 96.88 %, Precision 96.88 %, Recall 95.47 %, F1 96.86 %, Accuracy (5-fold CV) 96.17 % (Identify chronic non-atrophic gastritis and intestinal metaplasia) Accuracy 91.67 %, Precision 91.67 %, Recall 92.17 %, F1 91.61 %, Accuracy (5-fold CV) 90.46 % (Identify chronic non-atrophic gastritis with different severities)

* B1, C, M, R, S represent BLEU1, CIDEr, METEOR, ROUGE, and SPICE scores, respectively. These scores are commonly used metrics to evaluate the quantitative results of the captioning model.
Abbreviations: AUC, area under curve; CNN, convolutional neural network; DL, deep learning; EGC, early gastric cancer; GC, gastric cancer; VGG19, visual geometry group-19; XGBoost, eXtreme gradient boosting.

performance in a prospective patient study, with a sensitivity of 92.59 % and an accuracy of 83.67 %.

Similar to NBI endoscopy for EC, NBI endoscopy, compared to WLI endoscopy, can enhance the visualization of abnormal mucosal lesions, thereby improving the detection rate of EGC. It also requires endoscopists to possess substantial expertise and extensive clinical experience, and the application of AI in NBI can effectively improve the diagnostic rate of EGC. Hu et al.[119] developed an EGC diagnostic model based on the visual geometry group 19 (VGG19) architecture. The model was trained and tested using 1777 magnifying endoscopy NBI (ME-NBI) images from 295 cases collected from three centers. The performance of the model was compared to that of eight endoscopists with varying levels of experience. The results demonstrated that the model achieved comparable diagnostic accuracy to highly skilled endoscopists in diagnosing EGC. After referring to the model's results, the average diagnostic ability of the endoscopists significantly improved in terms of accuracy, sensitivity, and other relevant aspects. In addition to improving the rate of EGC diagnosis, accurately identifying the differentiation status and margins of EGC is also crucial for determining the surgical strategy for patients. Ling et al.[120] developed a real-time system that included a CNN2 model (based on 1670 ME-NBI images) for delineating the margins of EGC. Upon validation, the system demonstrated excellent performance compared to experts and was capable of margin delineation of EGC in actual videos. Furthermore, Gong et al.[121] developed an EGC captioning model utilizing ME-NBI images to automatically describe the visual features of ME-NBI images and assist endoscopists in enhancing their diagnostic capabilities for EGC. In this study, the EGC captioning model can generate descriptions for ME-NBI images to assist endoscopists in quickly forming diagnostic impressions of suspicious lesions, thereby enhancing EGC diagnosis. It can help less experienced junior endoscopists reduce their reliance on experience, bringing their diagnostic capabilities closer to expert levels. The model also ensures

the interpretability of decision-making, enhancing physicians' trust in AI-assisted diagnosis. Given the various advantages of generating crucial diagnostic captions for medical images, it is particularly suitable for various applications. For instance, it can be used in clinical practice to generate descriptive captions on X-rays, CT scans, MRI, and pathology images, aiding doctors in identifying abnormal structures and providing timely, robust support for clinical decisions. It can be utilized for educational training, assisting novice doctors in understanding complex pathological features, rapidly acquiring lesion diagnostic skills, and enhancing clinical judgment. Moreover, integrating the captioning system with diagnostic systems can automatically generate comprehensive examination reports, improving diagnostic accuracy and efficiency, reducing the workload on physicians, and ensuring standardization and traceability of diagnostic results.

Just as in EC research, in GC studies, researchers not only utilized AI in endoscopic examinations but also investigated the feasibility of integrating AI with other screening methods, such as combining AI with some non-invasive indicators like blood markers, biological features, *Helicobacter pylori* (*H. pylori*) infection status, etc. Considering the invasiveness and high cost of endoscopic examinations and aiming to reduce patients' medical expenses, Zhu et al.[122] developed a high-accuracy prediction model for GC diagnosis based on non-invasive features, such as gender, age, neutrophil-to-lymphocyte ratio, hemoglobin, albumin, carcinoembryonic antigen, carbohydrate antigen 125, and carbohydrate antigen 199, combined with the algorithm of gradient boosting decision trees. Taninaga et al.[41] utilized machine learning and comprehensive examination parameters including biological features, *H. pylori* infection status, endoscopic findings, and blood test results to classify patients with GC into high-risk or low-risk categories. Among the ten trained classification models, the XGBoost algorithm achieved the highest AUC (0.899). This approach improved the classification accuracy and helped identify patients at a high risk of GC. Additionally, liquid biopsy can serve as a non-invasive and rapid method in GC screening. Zhang et al.[123] constructed a GC identification system based on the H2O framework[124] and 33 blood biochemical indicators. The system achieved sensitivity, specificity, accuracy, and AUC on the cross-validation set of 85.44 %, 83.82 %, 84.54 %, and 0.9165, respectively, offering efficient, non-invasive, and cost-effective advantages for GC detection.

There are also some studies combining surface-enhanced Raman spectroscopy (SERS) technology with machine learning algorithms, showing great potential.[125] *H. pylori* is a major risk factor for gastric cancer. Tang et al.[126] and his team used the light gradient boosting machine (LightGBM) framework to build a machine learning prediction model based on 12,000 SERS spectra data from 100 participants to predict *H. pylori* infection. This model not only outperformed traditional methods in terms of accuracy but also demonstrated significant advantages in time efficiency. On the other hand, Si et al.[127] utilized SERS spectra data to classify different stages of gastric diseases. They first applied orthogonal partial least squares discriminant analysis (OPLS-DA) to deeply analyze the preprocessed spectral data, distinguishing between different stages of gastric diseases. Based on this, they used XGBoost to build a diagnostic model, ultimately achieving the highest accuracy of 96.88 % and 91.67 %, respectively, in distinguishing chronic non-atrophic gastritis from intestinal metaplasia and different subtypes of gastritis (mild, moderate, and severe).

4.2. Diagnosis

The diagnosis of GC is an essential step that significantly affects patients' treatment and prognosis. Accurate diagnosis of GC helps physicians understand its molecular characteristics and genetic variations and select the most appropriate treatment strategies for patients, improving treatment effectiveness while reducing unnecessary side effects. Many studies have applied AI for the diagnosis of GC. Some representative research works are summarized in Table 6.

The TNM staging diagnosis system for GC, developed by the Union for International Cancer Control (UICC) and the American Joint Committee on Cancer (AJCC),[128,129] holds substantial value in clinical practice for treatment planning and patient prognosis evaluation,[130,131] similarly to its application in EC.

For predicting the depth of GC invasion, Sun et al.[132] built a nomogram based on hand-crafted features and DL features from three-phase CT images to predict the serosa invasion depth of advanced gastric cancer (AGC) patients. Its AUC is higher than 0.85 in both the phase I retrospective study and phase II validation study. Unlike Sun et al.'s work, Goto et al.[133] focused on predicting the invasion depth of EGC. They fine-tuned the EfficientnetB1 model to distinguish between intramucosal and submucosal GC. The fine-tuned model worked with doctors to jointly determine the infiltration degree, and finally performed well in the test images. Nagao et al.[134] conducted work that was more comprehensive. They fine-tuned ResNet50 separately on three image datasets: WLI, NBI, and indigo-carmine dye contrast imaging (indigo). The trained models were then evaluated for invasion depth prediction. The experiments found that the prediction accuracies of the three models were 94.5 %, 94.3 %, and 95.5 %, respectively, with no significant differences.

For predicting LNM, Wang et al.[40] constructed a radiomics model using preoperative CT images from 247 confirmed GC patients, employing the RF algorithm. The CT images were segmented semi-automatically using a dedicated semi-processing prototype software, "Radiomics." They combined radiomic scores with selected clinical predictors to establish a radiomic nomogram. Both the radiomics model and nomogram demonstrated good performance in discriminating LNM. Dong et al.[37] utilized CT images from 730 patients with locally advanced gastric cancer (LAGC) from six centers. They established a deep learning radiomic nomogram (DLRN) based on CNN using DenseNet-201 architecture, which involved extracting 112 DL features and 289 handcrafted features from each ROI in the images. The radiomic nomogram demonstrated good discrimination for the number of LNM (overall C-index: primary cohort, 0.821, all external validation cohorts, 0.797, international validation cohort, 0.822), outperforming conventional clinical N stage, tumor size, and clinical model. Furthermore, Zhao et al.[135] compared four different methods, namely the radiomic nomogram, decision trees, naive Bayes, and DL, to establish an LNM prediction model. Ultimately, the DL model based on a fully convolutional network showed the most accurate LNM prediction with an AUC of 0.79.

For predicting distant metastasis of GC, Mirniaharikandehei et al.[136] addressed the issue of predicting peritoneal metastases (PMs) based on radiomic features of the primary lesions of GC on venous phase CT images. They compared the predictive performance of five feature selection or dimensionality reduction methods combined with a gradient boosting machine. The results indicated that the proposed random projection algorithm can generate feature vectors with strong discriminative power. Furthermore, Dong et al.[137] developed individualized radiomic nomograms for preoperative identification of occult peritoneal metastases in patients based on 266 quantitative image features from 554 AGC patients from four centers. These nomograms incorporated primary tumor (RS1), peritoneum region (RS2), or clinical factors to assess the PM status, preventing the omission of patients with PMs during CT evaluation. The study ultimately demonstrated that RS1, RS2, and Lauren type were important predictive factors for occult PMs. In contrast to Dong et al.'s work, Jiang et al.[138] established a densely connected convolutional network combined with long-short connections to predict occult PMs based on preoperative CT images of 1978 patients. The results showed that the model had significantly higher discriminative performance (AUCs in the external validation cohorts 1 and 2 were 0.946 and 0.920, respectively) compared to conventional clinic pathological factors (AUC range, 0.51–0.63). The model was identified as an independent predictor for occult PMs.

In addition to the TNM staging diagnosis, many studies also focus on other sub-directions of diagnosis, such as histological diagnosis, patho-

Table 6
Radiomics and DL studies for the diagnosis of gastric cancer.

Authors (year)	Type of task	Modeling	Evaluation
Sun RJ. (2020)	Predict invasion	Radiomics nomogram	AUC 0.87 (test set I), AUC 0.90 (test set II)
Goto A. (2023)	Predict invasion	EfficientnetB1	Accuracy 77.0 % (AI alone), Accuracy 78.0 % (AI with endoscopists)
Nagao S. (2020)	Predict invasion	ResNet50	Accuracy 94.5 % (WLI), Accuracy 94.3 % (NBI), Accuracy 95.5 % (Indigo)
Wang Y. (2020)	Predict LNM	RF, Nomogram	AUC 0.837 (RF), AUC 0.881 (Nomogram)
Dong D. (2020)	Predict LNM	DLRN	C-index 0.797 (external validation), C-index 0.822 (international validation)
Zhao L. (2023)	Predict LNM	Nomogram, Decision Tree, Naive Bayes, DL	AUC 0.78 (Nomogram), AUC 0.76 (Decision Tree), AUC 0.77 (Naive Bayes), AUC 0.79 (DL)
Mirniaharikandehei S. (2021)	Predict peritoneal metastasis	Gradient boosting machine model embedded with a random projection algorithm	AUC 0.69±0.019, Accuracy 71.20 %, Precision 65.78 %, Sensitivity 43.10 %, Specificity 87.12 %
Dong D. (2019)	Predict occult peritoneal metastasis	Nomogram	AUC 0.928 (external validation cohort 1), AUC 0.920 (external validation cohort 2)
Jiang Y. (2021)	Predict occult peritoneal metastasis	Densely connected convolutional network	AUC 0.946 (validation cohort 1), AUC 0.920 (validation cohort 2)
Wang XX. (2020)	Predict Lauren classification	Radiomics, Nomogram	AUC 0.758
Wang S. (2021)	Predict specific Borrmann classification	Radiomics, multilayer perceptron network	AUC 0.702 (Borrmann I/II/III vs. IV), AUC 0.731 (Borrmann II vs. III)
Muti HS. (2021)	Detect EBV and MSI status	DL-based classifier	MI AUC 0.597-0.836 (internal validation), EBV AUC 0.819-0.897 (internal validation), MI AUC 0.723-0.863 (external validation), EBV AUC 0.672-0.859 (internal validation)
Lee S H. (2023)	Detect MSI	DL-based classifier	AUC 0.874
Zheng X. (2022)	Detect EBV	EBVNet	AUC 0.941 (external validation)
Ba W. (2022)	evaluate DL assistance with pathologists' diagnosis of GC	CNN with DeepLab v3 architecture	AUC 0.911, Sensitivity 90.63 %, Specificity 78.23 %

Abbreviations: AI, artificial intelligence; AUC, area under curve; CNN, convolutional neural network; DL, deep learning; DLRN, deep learning radiomics nomogram; EBV, Epstein-Barr virus; GC, gastric cancer; LNM, Lymph node metastasis; MSI, microsatellite instability; NBI, narrow-band imaging; RF, random forest; WLI, white-light imaging.

logical diagnosis, etc. The Lauren classification of GC primarily includes two types: diffuse type and intestinal type. These two types of GC exhibit varying sensitivity and effectiveness to chemotherapy and radiotherapy. The Lauren classification assists physicians in selecting appropriate surgical methods, chemotherapy regimens, and radiotherapy plans based on the pathological characteristics of GC, thereby improving treatment outcomes. For patients with intestinal-type GC, the primary approach is surgical resection, supplemented by preoperative or postoperative chemotherapy to reduce tumor size or recurrence risk. In contrast, patients with diffuse-type GC require systemic treatment primarily with chemotherapy.[139,140] The Lauren classification is also potentially associated with patient prognosis. Studies indicated that patients with intestinal-type GC generally had a better prognosis, whereas those with diffuse-type had a poorer prognosis, with a five-year survival rate lower than that of intestinal-type patients.[141,142] However, research by Tang et al.[143] found that the survival rate for intestinal-type patients was worse than for diffuse-type patients. These contradictory results may be due to clinical heterogeneity and varying sample sizes, highlighting the need for further large-sample studies to validate these conclusions.

Therefore, predicting the Lauren classification preoperatively is particularly important for treatment selection and prognosis assessment in GC. Wang et al.[144] conducted the first preoperative radiomics analysis specifically targeting the Lauren classification in GC. In their study, they used tumor regions and the surrounding area from CT images of 539 GC patients as two sets of radiomic features. They developed a radiomic nomogram that combined radiomic and clinical features. This nomogram outperformed other models, with AUCs of 0.745 and 0.758 in the training and validation cohorts, respectively, enabling better identification of diffuse-type GC patients. Similarly, the Borrmann classification of AGC is crucial for surgical strategy and prognostic evaluation. Wang et al.[145] extracted radiomic features from CT images of 889 AGC patients from two centers. They constructed 15 basic classification models and an ensemble multilayer perceptron network. This network could automatically identify Borrmann IV and improve the discriminatory ability between Borrmann II and III.

Epstein-Barr virus (EBV) positivity and microsatellite instability (MSI) are closely associated with the immune response in GC, making them potential biomarkers for immunotherapy. Muti et al.[146] developed

Table 7
Radiomics and DL studies of treatment response prediction for gastric cancer.

Authors (year)	Type of task	Modeling	Evaluation
Wang W. (2024)	Predict complications	Light Gradient Boosting Machine	Accuracy 87.28 %, AUC 0.9232
Aoyama Y.(2024)	Predict pancreatic fistula	HyperSeg nested with a U-Net architecture	Dice of Pancreas 0.70
Takeuchi M.(2023)	Predict surgical complexity	TeCNO	Accuracy 87 %, AUC 0.859
Wang W. (2021)	Predict treatment response of NACT	LASSO	AUC 0.679
Cui Y. (2022)	Predict treatment response of NACT	DLRN	AUC 0.804 (external validation 1), AUC 0.827 (external validation 2)
Zhong H. (2021)	Predict the response of metastatic lymph nodes to NACT	DLDRN	AUC 0.94
Tan J. (2020)	Predict treatment response of chemotherapy	Semi-automatic delta radiomics model	AUC 0.728 (test cohort) AUC 0.828 (validation cohort)

Abbreviations: AUC, area under curve; DL, deep learning; DLDRN, deep learning delta radiomic nomogram; DLRN, deep learning radiomics nomogram; IoU, intersection over union; LASSO, least absolute shrinkage and selection operator; NACT, neoadjuvant chemotherapy.

a deep neural network-based classifier utilizing routine histopathological slide data from 2823 patients with known MSI status and 2685 patients with known EBV status from seven countries. Their classifier effectively distinguished EBV and MSI and could serve as a predictive biomarker for immunotherapy in GC. Lee et al.[147] designed a DL-based classifier to automate the classification of MSI status. The classifier was able to automatically learn and identify optimal features for MSI status. In the external validation cohort, the classifier demonstrated favorable performance with an AUC of 0.874. Zheng et al.[148] introduced a CNN named EBVNet, which was integrated with pathologists to predict the EBV status of GC slides. AUCs in the internal cross-validation and external dataset were 0.969 and 0.941, respectively. Thus, the introduction of human-machine fusion could enhance the predictive capability for EBV status. To assess the adjunctive role of DL in the diagnosis of GC by pathologists, Ba et al.[149] utilized a CNN with DeepLab v3 architecture[150] to analyze 110 WSI. The DL-assisted pathologists, in comparison to pathologists alone, achieved AUCs of 0.911 and 0.863, with corresponding sensitivities of 90.63 % and 82.75 %, and specificities of 78.23 % and 79.90 %, respectively. Furthermore, the DL-assisted approach resulted in reduced review time (22.68 s vs. 26.37 s), demonstrating that AI can enhance the diagnostic capabilities of pathologists.

4.3. Treatment response prediction

Currently, clinical treatment options for GC patients include surgery (open surgery and laparoscopic surgery), neoadjuvant chemotherapy (NACT), and chemotherapy. The application of AI in the clinical treatment of GC can provide indications of patients' response to treatment, screen potential beneficiaries, and assist doctors in designing more personalized treatment plans, reducing postoperative risks, and effectively improving patient prognosis. The summary of representative works on using AI to predict treatment response in GC patients is presented in Table 7.

The preferred treatment for GC is surgery supplemented with preoperative neoadjuvant therapy and postoperative adjuvant chemotherapy. Surgery is the primary treatment modality for GC. To ensure safe surgery, predicting complications and assessing surgical complexity are crucial. Wang et al.[151] developed a model using the light gradient boosting machine algorithm based on preoperative abdominal CT scans of 166 patients who underwent radical gastrectomy. This model predicts early postoperative complications in GC patients, achieving a prediction accuracy of 87.28 % and an AUC of 0.9232. It aids in personalized clinical decision-making for postoperative care. Similarly, Aoyama et al.[152] confirmed that dimpling lines—depressions formed between the pancreas and surrounding organs—are anatomical markers associated with postoperative pancreatic fistula, a serious complication of laparoscopic gastrectomy. They employed a HyperSeg[153] semantic segmentation model nested with a U-Net architecture to visualize dimpling lines in real-time during surgery, ensuring safety. Furthermore, Takeuchi et al.[154] implemented automatic prediction of the complexity of robotic distal gastrectomy using an AI model called TeCNO,[155] based on automatic phase recognition of 56 videos. By identifying surgical phases, particularly the extended duration of beginning phases, they predicted surgical complexity to optimize intraoperative decisions and forecast bleeding and complications.

NACT represents a promising therapeutic approach for patients with potentially resectable GC. Among patients with AGC, preoperative NACT not only improves the resection rate but also enhances patient survival by downstaging the primary tumor. Nevertheless, it is important to acknowledge that individual patient responses to NACT can vary. Hence, it becomes critical in clinical practice to accurately distinguish between potential responders and non-responders. Wang et al.[156] constructed a radiomic signature significantly associated with treatment response by extracting 20 features from CT images of 323 patients who underwent NACT from three hospitals. This radiomics signature, demonstrating an AUC of 0.679 for discriminating between respondents and non-respondents in the external validation cohorts, showed promise in predicting the treatment response to NACT in patients with LAGC. Similarly, Cui et al.[157] extracted 1125 handcrafted features and 1024 DL signatures extracted by a CNN model based on the DenseNet-121 architecture from each ROI in CT images of 719 LAGC patients from four hospitals. Subsequently, they developed a DLRN model by combining these handcrafted features and DL signatures with significant clinicopathological factors. The DLRN model demonstrated good discriminative performance in predicting a favorable response to NACT with AUCs of 0.829, 0.804, and 0.827 in the internal and two external validation cohorts, respectively. Unlike previous studies, to predict the response of metastatic lymph nodes to NACT in LAGC, Zhong et al.[158] developed and compared three models: a clinical model, DL model, and DL delta radiomic nomogram (DLDRN) based on the radiomics-delta signature, which was identified as the optimal predictor of metastatic lymph node response and was derived from CT images of 98 patients. The study findings demonstrated that DLDRN exhibited significantly superior predictive performance compared to the other two models (AUC, 0.94 [DLDRN] vs. 0.83 [clinical model] vs. 0.91[DL model]).

Postoperative chemotherapy for GC aims to eliminate residual cancer cells, reduce the risk of tumor recurrence, and prevent or delay cancer relapse and metastasis, thereby improving patient survival rates. Therefore, chemotherapy plays a crucial role in the overall treatment strategy for GC. Tan et al.[159] utilized a volumetric, fully convolutional neural network (V-net) DL algorithm to devise a semi-automatic segmentation method. Based on this semi-automatic segmentation approach and dual-energy CT images of 86 patients, they established a Delta radiomics model. Compared to the manual segmentation method, the semi-automatic delta radiomics model exhibited superior predictive capability for chemotherapy response in patients with far-advanced GC

Table 8
Radiomics and DL studies of the prognosis for GC.

Authors (year)	Type of task	Modeling	Evaluation
Zhang W. (2020)	Predict early recurrence in AGC	CT-based radiomic nomogram	AUC 0.826 (test set 1), AUC 0.806 (test set 2)
Zhang L. (2020)	Predict OS	DL model based on the architecture of residual CNN	C-index 0.78
Li. H. (2023)	Predict survival time	Semi-supervised confidence-aware multi-view co-training network	C-index 0.671±0.010
Jiang Y. (2021)	Predict OS and DFS	S-net, Nomogram	C-index 0.802(OS), C-index 0.792(DFS)
Li X. (2022)	Predict OS and DFS	SVM	AUC 0.852(OS), AUC 0.837(DFS)

Abbreviations: AGC, advanced gastric cancer; AUC, area under curve; CNN, convolutional neural network; CT, computed tomography; DFS, disease-free survival; DL, deep learning; GC, gastric cancer; OS, overall survival; SVM, support vector machine.

(mean AUC: 0.728 vs. 0.687 (in the testing cohort) and 0.828 vs. 0.749 (in the independent validation cohort)).

4.4. Prognosis

In the field of GC, prognostic prediction studies can mainly be classified into two categories based on treatment methods: one focuses on prognostic prediction for patients undergoing gastrectomy, while the other focuses on prognostic prediction for patients receiving adjuvant chemotherapy. Although medical technology is constantly advancing, the 5-year survival rate for the majority of GC patients remains poor, and the incidence of surgical complications is high.[160] In recent years, AI has shown promising results in prognostic prediction research for GC. Some representative works are shown in Table 8.

To prognosticate early recurrence following radical resection in patients with GC, Zhang et al.[161] developed a radiomic nomogram in their study by extracting radiomic features from preoperative diagnostic CT images of 669 patients across two centers. To reduce feature redundancy and ensure reproducibility and robustness, they used methods such as intraclass correlation coefficients (ICCs), coefficient of variation (CV), and consensus clustering. Multivariable logistic regression analysis was then employed to integrate both radiomic features and clinical risk factors into the nomogram. The selected radiomic features were significantly associated with early recurrence, further validating the strong prognostic efficacy of the nomogram. In another investigation conducted by Zhang et al.[162], CT images of 640 patients from three independent centers were collected. A DL model, utilizing a residual CNN, was developed to predict the OS risk in patients with GC. The performance of this DL model was compared with a radiomics model consisting of 24 features and a clinical model incorporating three crucial clinical variables. The findings demonstrated superior predictive performance of the DL model in assessing OS risk (external validation: DL vs. Clinical vs. Radiomics = 0.78 vs. 0.71 vs. 0.72). Furthermore, Li et al.[163] developed an image-based semi-supervised confidence-aware multi-view co-training network model for predicting patient survival time.

There are also some relevant studies on prognostic prediction for adjuvant chemotherapy. In order to identify patients who can benefit from adjuvant chemotherapy for GC, Jiang et al.[164] utilized a deep neural network called "S-net" and CT images from 1615 patients to construct the CT features. These CT features were combined with clinicopathologic factors to develop an integrated nomogram. The C-indices of the integrated nomogram, clinicopathologic nomogram, imaging signature, and TNM stage for DFS in the validation cohort were 0.792, 0.776, 0.719, and 0.736, respectively. The corresponding C-indices for OS in the validation cohort were 0.802, 0.781, 0.724, and 0.740. Therefore, the proposed integrated nomogram demonstrated superior discrimination performance. Li et al.[39] developed an SVM model based on nine clinicopathological features from 255 patients who underwent surgical resection. Through this model, GC patients were categorized into low, intermediate, and high-risk groups. The study results indicated that high-risk patients with TNM stage II and III had a higher likelihood of benefiting from adjuvant chemotherapy compared to low-risk patients.

5. Discussion

Building upon the aforementioned content, it is evident that AI technologies, exemplified by radiomics and DL, have achieved notable research progress in the applications of screening, diagnosis, treatment, and prognosis for UGI cancers. The near future holds a promise that AI will significantly alleviate medical pressures, enhance medical efficiencies, and contribute to increased patient survival rates. Undoubtedly, the integration of AI with clinical medicine stands as a pivotal research direction, attracting an increasing number of researchers and fostering rapid development in this domain. Nevertheless, this field is characterized by abundant opportunities as well as specific limitations and challenges.[165]

5.1. Selection of AI algorithms in various scenarios

As introduced and analyzed above, different algorithms are suited for different scenarios. In the following section, this paper will summarize these findings.

In the context of screening for esophageal and gastric cancers, it is often easier to obtain large amounts of natural image data. Additionally, since screening may sometimes require models with localization capabilities, deep learning-based visual models are particularly well-suited for this clinical scenario. Models such as ResNet, VGG, and Vit have been proven effective in many studies.[65,71,74,118,119]

The diagnosis primarily involves three tasks: LNM, depth of invasion prediction, and distant metastasis prediction. For the task of LNM, current approaches predominantly rely on CT images for prediction. These include radiomic nomograms, multi-instance learning, and CNNs,[40,78,79,135] as well as combinations of radiomics and deep learning.[137] For depth of invasion prediction, similar to the screening task, endoscopic images are mainly used as training data, making DL-based visual models an effective choice.[75,76,133,134] Some studies have also used CT data, combining radiomic features and DL features to develop nomograms, achieving promising results.[132] Regarding the task of distant metastasis prediction, some studies emphasize the identification of effective predictive factors. For instance, Zhu et al.[34] used logistic univariate and multivariate regression analyses on radiomic features to identify clinically independent predictors. Other studies directly focus on using CT images for metastasis prediction, such as Dong et al.[137] who employed radiomics nomograms, and Jiang et al. who utilized Dense CNN models.[138]

The task categories for predicting treatment efficacy are more diverse, with the choice of algorithms depending on the specific scenario. Radiomic features derived from PET, CT, and MRI can be utilized to predict treatment responses for cCRT and definitive concurrent chemoradiotherapy (dCRT) in esophageal cancer.[84,85,87–90] For RP prediction,

feature extraction is often performed on dose-volume histograms, followed by either independent modeling or joint modeling with other features for prediction.[94,95] In terms of prognosis prediction for esophageal fistula, CNN models based on CT can be used as classifiers,[101] or logistic regression analysis can be applied using radiomic features.[100] For predicting the treatment response to NACT in gastric cancer, a radiomics-based approach similar to that used in esophageal cancer is primarily employed.[156–158] For postoperative complication prediction, Wang et al.[151] used a 3D convolutional neural network (3D-CNN) to extract image features from patients, applied LASSO for feature selection, and then modeled the prediction using light gradient boosting machine. In another study, Aoyama et al.[152] used U-net to predict indent lines, which serve as biomarkers for postoperative pancreatic fistula.

In prognosis, the primary algorithms involved are the CPH model, DL models, and radiomic nomograms. The CPH model offers strong interpretability and is more efficient and straightforward to implement, but it relies on relatively simple features, making it unsuitable for more complex scenarios.[31–33] DL models, on the other hand, excel in handling large-scale data and complex features.[106,162–164] Radiomic nomograms offer a more balanced approach compared to the former; however, in comparison to deep learning models, constructing features for radiomics involves higher labor costs.[36,107,161]

5.2. Scarcity of large-scale data

Currently, AI research applied to UGI cancers mainly consists of retrospective studies conducted on small datasets. However, this poses two problems. On one hand, the models' upper-performance limit is highly correlated with the amount of data. Small datasets will prevent models from fully capturing the complex patterns and variations in the data during training, thereby limiting the performance upper bound of models. On the other hand, the quality and completeness of data in retrospective studies are insufficient, resulting in relatively low levels of evidence in evidence-based medicine. Nowadays, multitudes of advanced AI technologies require a large amount of high-quality data to be effective. Insufficient data poses challenges to the development of AI in clinical fields. Therefore, how to increase the scale of training data and improve the quality of validation data is an important issue that must be addressed in the field of medicine.

To achieve this goal, one aspect is to solve the privacy issues of medical data. Unlike natural images, medical data is highly sensitive and contains patient personal health information, which is strictly regulated by laws and regulations, and is difficult to obtain for researchers. A potential solution is to use distributed AI technologies to overcome privacy restrictions, such as federated learning. Federated learning is a method for training models using distributed data from various devices or systems, rather than relying on a central server. In the medical field, federated learning can be used to train models on vast distributed patient datasets from different hospitals or clinics, which allows the sharing of information and knowledge between facilities while protecting patient data privacy and security. Currently, some works have explored the application of federated learning in related clinical problems such as skin cancer and lung cancer,[166–168] but its application in UGI cancers clinical problems is still relatively unexplored. The other aspect is to strongly advocate for conducting larger-scale prospective studies compared to retrospective research. Prospective studies, through the design of explicit research objectives and the collection of detailed, accurate, and more representative samples, can obtain higher levels of evidence and scientific reliability and efficiently drive the translation and application of research.

5.3. Popularization of early screening

Based on all the aforementioned studies, it can be concluded that among the four aspects of screening, diagnosis, treatment, and prognosis for UGI cancers, screening is the most crucial. In Korea and Japan, it has been found that approximately 50 % to 70 % of gastric cancers are early-stage cancers,[169] which are curable. Early detection of EC also can make subsequent treatments highly effective.[170] These all mean that with timely and accurate early screening, the focus of UGI cancers management will shift from treatment to screening. Therefore, early screening of UGI cancers is of utmost importance.

One approach to addressing this issue is to promote the widespread implementation of early screening. Invasive procedures such as endoscopy and the high costs associated with CT and MRI hinder the broad adoption of early screening. However, with the rapid development of AI technology, it is possible to integrate AI with non-invasive methods such as blood tests, urine tests, and others, making early screening for UGI cancers more convenient. Additionally, AI can improve imaging algorithms, reduce the environmental and equipment requirements for medical imaging, and make early screening for UGI cancers more accessible.

Another aspect is to enhance the accuracy of early-stage screening by promoting the use of AI in conjunction with gastroscopy, CT, and other examinations. Such integration can assist grassroots health workers in achieving better clinical diagnoses.

5.4. Clinical applications of AI

Currently, many AI algorithms have been applied in clinical research on EC and GC, each with great potential for translation and distinct characteristics. In the field of EC, Struyvenberg et al.[171] conducted a prospective study using principal component analysis to automatically screen multiple frames of volumetric laser endomicroscopy (VLE) images. They developed predictive models using support vector machine, random forest, and naive Bayes, which enabled rapid detection of Barrett's esophagus neoplasia, demonstrating the feasibility of real-time automated assessment during endoscopic examinations. Similarly, another prospective study[172] utilized FusionNet,[173] based on CNN architecture, to segment VLE images and DenseNet to generate score vectors. The study employed support vector machine, random forest, and K-nearest neighbors as classifiers to effectively classify high-grade dysplasia and early esophageal adenocarcinoma. In terms of GC, Yuan et al.[174] developed three different AI deep learning models based on the attentive pairwise interaction neural network, transformer architecture, and DeepLab v3+. They conducted a prospective multicenter clinical study for the diagnosis and screening of GC using tongue images. The AUC of these three models in independent external validation reached 0.83–0.88, effectively identifying GC patients. This study demonstrated the convenience and cost-effectiveness of using tongue imaging as a diagnostic tool for GC. Another study[175] used deep CNNs (including VGG16 and ResNet-50) and deep reinforcement learning to develop an AI system[176] for detecting local lesions and diagnosing GC. A randomized controlled trial confirmed that this AI system effectively reduces the missed diagnosis rate of GC and minimizes unnecessary biopsies. The reason these classic AI algorithms are applicable in the clinic is due to their extensive research, relative stability, and the ease with which researchers can interpret and understand their results. In the future, as AI algorithms continue to evolve and improve, we can anticipate more applications of large models and more advanced architectures in clinical research. This will bring more innovations to the medical field, helping doctors diagnose diseases more quickly and accurately.[177]

At present, there is limited practice in applying AI to clinical prognosis research for patients. This is mainly due to two reasons: Firstly, clinical prognosis requires long-term follow-up of patients, making the research cycle longer and increasing the difficulty and cost of prognosis studies. Secondly, clinical applications demand high accuracy from models, typically requiring over 90 % accuracy. However, current prognosis models often struggle to meet the satisfaction levels of both doctors and patients in practical clinical applications. Further research and development of AI prognosis models with higher accuracy and reliability are needed to successfully apply more AI technologies to patient clini-

cal prognosis research, bringing greater benefits and innovations to the medical field.

5.5. Explainability of AI

We have previously outlined the fundamental concepts and principles of both radiomics and DL. A crucial distinction between the two lies in explainability. While DL can autonomously extract numerous features efficiently and associate them with the target task, it cannot offer a clear and explicit definition of features or explain the processing of these features. Radiomics primarily relies on manually predefined formulas to extract a plethora of features from medical images. Although it possesses interpretability, it exhibits high redundancy and is difficult to directly correspond to clinical objectives. The importance of explainability lies in facilitating doctors' comprehension and validation of AI decisions, fostering trust in AI systems, promoting effective collaboration, and aiding doctors in elucidating and communicating complex medical information to patients.

Several studies have delved into enhancing the explainability of AI in the medical domain. Ge et al.[178] and Suh et al.[179] elucidated the clinical significance and characteristics of features by considering feature weights. Kaji et al.[180] and Shickel et al.[181] employed attention mechanisms to characterize explainability. Additionally, some studies explored enhancing explainability through techniques such as knowledge distillation and dimensionality reduction.[182] Looking ahead, the widespread integration of AI in medicine demands not only mature technical support but also a commitment to transparency and explainability. However, due to existing theoretical constraints, research on the explainability of DL models is still in the early stages, and interdisciplinary research with clinical applications remains relatively limited. This field requires further strengthening of relevant theoretical research and increased collaboration between researchers and clinicians to facilitate reliable and sustainable development of AI in the medical and health sectors.

5.6. General medical AI

In recent years, there has been rapid development in large-scale AI models. These models, built upon vast amounts of data, large model parameters, and massive computational power, have demonstrated the ability to effectively learn general knowledge within the domain of expertise. They have surpassed traditional AI methods across various tasks, presenting new opportunities for the scalable implementation of AI. Among them, large language models and large multimodal models have exhibited powerful capabilities in the general domain.

Large language models refer to deep learning models trained on a large corpus of text data. By training on extensive text data, the model learns the grammar structures, contextual relationships, and semantic meanings among words, phrases, and sentences. This enables the model to understand, generate, and process the content and semantics of natural language. Examples include ChatGPT developed by OpenAI and LLaMA built by Meta. Large multimodal models refer to deep learning models trained on multiple modalities of information, such as text, images, videos, audio, and so on. By combining different types of data, large multimodal models can describe entities from various perspectives, leading to more comprehensive and accurate analysis results. Examples include Meta's ImageBind model[183] and GPT-4 V by OpenAI. Large models mentioned above, which leverage massive amounts of data and large model parameters, have achieved high accuracy in the general domain.

In the medical field, numerous clinical problems often involve multimodal information and multiple downstream tasks. For example, the diagnosis of GC may involve the use of endoscopy, CT scans as well as other modalities, and the specific diagnosis tasks are further divided into TNM staging diagnosis, typing diagnosis, and more. Currently, most studies only build models based on a single modality and for a single task. Integrating multi-modality and multi-task can enable models to better fuse various information, acquire richer knowledge, and also learn the relationships between multiple tasks to make better decisions. Therefore, constructing a general medical AI-oriented to multi-modality and multi-tasks is a highly promising research direction.

The development of large-scale AI models has provided new research ideas and paradigms for general medical AI. Building upon this foundation, some relevant studies have explored the possibilities of general medical AI. Wu et al.[184] conducted a comprehensive evaluation of GPT-4 V (vision) across 17 medical diagnosis tasks spanning various human body systems, highlighting its significant potential. In a separate study, Ma et al.[185] curated a medical image dataset encompassing 10 different modalities and introduced the MedSAM model, validating 86 internal validation tasks and 60 external validation tasks. Zhou et al.[186] introduced the IRENE framework, which processes unstructured clinical text, structured clinical information, and X-ray images, automatically extracting and integrating features to provide a conclusive diagnosis. Additionally, Zhou et al.[187] proposed RETFound, a foundational model for retinal images demonstrating outstanding performance in tasks such as eye disease diagnosis, prognosis, and predicting systemic diseases following fine-tuning.

Despite some progress, general medical AI models are still in the early exploratory stage due to factors such as data scarcity and the immaturity of multimodal pre-training techniques. However, with the advancement of pretraining techniques that enable models to leverage large-scale unlabeled data to enhance their language understanding and generation capabilities, along with the release of various open-source large-scale models, pretraining-based general medical AI models are expected to improve the accuracy of medical tasks such as efficacy prediction and prognosis forecasting. Furthermore, it is expected to achieve more remarkable advancements in the field of UGI cancers research.

6. Conclusions

In this review, we introduced and analyzed cutting-edge applications of AI in clinical tasks related to UGI cancers. Specifically, we reviewed the classic works applying radiomics and DL technology to EC and GC in the past 5 years. This study investigated the application of radiomics, DL, or a combination of both to address various clinical challenges in UGI cancers, encompassing screening, diagnosis, treatment response prediction, and prognosis, demonstrating the great potential of AI technology in solving problems related to UGI cancers. Additionally, we discussed challenges and noteworthy directions in this field, including the scarcity of large-scale data, the explainability of AI, the general medical model, and the popularization of early screening. We believe that through continuous exploration and verification, AI will significantly benefit patients with UGI cancers.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key R&D Program of China (grant number: 2023YFC2415200), National Natural Science Foundation of China (grant numbers: 82361168664, U24A20759, 82441018, 82372053, 92259302, 62027901, 82302296), Science and Technology Development Fund of Macao Special Administrative Region (grant number: 0006/2023/AFJ), Strategic Priority Research Program of the Chinese Academy of Sciences (grant number: XDB38040200), Beijing Natural Science Foundation (grant numbers: Z20J00105, JQ24048), Scientific and Technological Innovation Project of China Academy of Chinese Medical Sciences (grant number: CI2023C008YG), Key-Area Research and Development Program of Guangdong Province (grant num-

ber: 2021B0101420005), the Youth Innovation Promotion Association CAS (grant number: Y2021049), and China Postdoctoral Science Foundation under Grant (grant number: 2022M720357). The authors would like to acknowledge the instrumental and technical support of the multimodal biomedical imaging experimental platform, Institute of Automation, Chinese Academy of Sciences.

Author contributions

All authors contributed to writing the manuscript and prepared the figures and legends.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin*. 2024;74(3):229–263. doi:10.3322/caac.21834.
2. Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. *J Natl Cancer Cent*. 2024;4(1):47–53. doi:10.1016/j.jncc.2024.01.006.
3. Arnold M, Morgan E, Bardot A, et al. International variation in oesophageal and gastric cancer survival 2012–2014: differences by histological subtype and stage at diagnosis (an ICBP SURVMARK-2 population-based study). *Gut*. 2022;71(8):1532–1543. doi:10.1136/gutjnl-2021-325266.
4. Ilic M, Ilic I. Epidemiology of stomach cancer. *World J Gastroenterol*. 2022;28(12):1187–1203. doi:10.3748/wjg.v28.i12.1187.
5. An L, Zheng R, Zeng H, et al. The survival of esophageal cancer by subtype in China with comparison to the United States. *Int J Cancer*. 2023;152(2):151–161. doi:10.1002/ijc.34232.
6. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. 2018;391(10125):1023–1075. doi:10.1016/S0140-6736(17)33326-3.
7. Levy J, Gupta V, Amirazodi E, et al. Textbook Outcome and Survival in Patients With Gastric Cancer: an Analysis of the Population Registry of Esophageal and Stomach Tumours in Ontario (PRESTO). *Ann Surg*. 2022;275(1):140–148. doi:10.1097/SLA.0000000000003849.
8. Soerjomataram I, Cabasag C, Bardot A, et al. Cancer survival in Africa, central and south America, and Asia (SURVCAN-3): a population-based benchmarking study in 32 countries. *Lancet Oncol*. 2023;24(1):22–32. doi:10.1016/S1470-2045(22)00704-5.
9. American Cancer Society Cancer Facts & Figures 2024 https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2024/2024-cancer-facts-and-figures-acs.pdf.
10. Zhu Z, Gong Y, Xu H. Clinical and pathological staging of gastric cancer: current perspectives and implications. *Eur J Surg Oncol*. 2020;46(10):e14–e19. doi:10.1016/j.ejso.2020.06.006.
11. Sjoquist KM, Burmeister BH, Smithers BM, et al. Survival after neoadjuvant chemotherapy or chemoradiotherapy for resectable oesophageal carcinoma: an updated meta-analysis. *Lancet Oncol*. 2011;12(7):681–692. doi:10.1016/S1470-2045(11)70142-5.
12. Smyth EC, Nilsson M, Grabsch HI, Van Grieken NC, Lordick F. Gastric cancer. *Lancet*. 2020;396(10251):635–648. doi:10.1016/S0140-6736(20)31288-5.
13. Arnal MJD. Esophageal cancer: risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J Gastroenterol*. 2015;21(26):7933. doi:10.3748/wjg.v21.i26.7933.
14. De Gouw DJJM, Klarenbeek BR, Driessen M, et al. Detecting pathological complete response in esophageal cancer after neoadjuvant therapy based on imaging techniques: a diagnostic systematic review and meta-analysis. *J Thorac Oncol*. 2019;14(7):1156–1171. doi:10.1016/j.jtho.2019.04.004.
15. Fiorino C, Palumbo D, Mori M, et al. Early regression index (ERI) on MR images as response predictor in esophageal cancer treated with neoadjuvant chemoradiotherapy: interim analysis of the prospective ESCAPE trial. *Radiother Oncol*. 2024;194:110160. doi:10.1016/j.radonc.2024.110160.
16. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500–510. doi:10.1038/s41568-018-0016-5.
17. Arribas J, Antonelli G, Frazzoni L, et al. Standalone performance of artificial intelligence for upper GI neoplasia: a meta-analysis. *Gut*. 2021;70(8):1458–1468. doi:10.1136/gutjnl-2020-321922.
18. Hallinan JTPD, Venkatesh SK. Gastric carcinoma: imaging diagnosis, staging and assessment of treatment response. *Cancer Imaging*. 2013;13(2):212–227. doi:10.1102/1470-7330.2013.0023.
19. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230–243. doi:10.1136/svn-2017-000101.
20. Liu Z, Wang S, Dong D, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. *Theranostics*. 2019;9(5):1303–1322. doi:10.7150/thno.30309.
21. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500–510. doi:10.1038/s41568-018-0016-5.
22. Tajbakhsh N, Jeyaseelan L, Li Q, Chiang J, Wu Z, Ding X. Embracing imperfect datasets: a review of deep learning solutions for medical image segmentation. *Med Image Anal*. 2020;63:101693. doi:10.1016/j.media.2020.101693.
23. Gui J, Sun Z, Wen Y, Tao D, Ye J. A Review on Generative Adversarial Networks: algorithms, theory, and applications. *IEEE Trans on Knowl and Data Eng*. 2023;35(4):3313–3332. doi:10.1109/TKDE.2021.3130191.
24. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012;30(9):1234–1248. doi:10.1016/j.mri.2012.06.010.
25. Yuan Y, Chao M, Lo YC. Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. *IEEE Trans on Med Imaging*. 2017;36(9):1876–1886. doi:10.1109/TMI.2017.2695227.
26. Li X, Chen H, Qi X, Dou Q, Fu CW, H-DenseUNet Heng PA. Hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans on Med Imaging*. 2018;37(12):2663–2674. doi:10.1109/TMI.2018.2845918.
27. Pereira S, Pinto A, Alves V, Silva CA. Brain Tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans on Med Imaging*. 2016;35(5):1240–1251. doi:10.1109/TMI.2016.2538465.
28. Gillies RJ, Kinahan PE, Radiomics Hricak H. Images Are More than Pictures, They Are Data. *Radiology*. 2016;278(2):563–577. doi:10.1148/radiol.2015151169.
29. Bagherzadeh-Khiabani F, Ramezankhani A, Azizi F, Hadaegh F, Steyerberg EW, Khalili D. A tutorial on variable selection for clinical prediction models: feature selection methods in data mining could improve the results. *J Clin Epidemiol*. 2016;71:76–85. doi:10.1016/j.jclinepi.2015.10.002.
30. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749–762. doi:10.1038/nrclinonc.2017.141.
31. Gujjuri RM, Clarke JA, Elliott JA, et al. Predicting long-term survival and time-to-recurrence after esophagectomy in patients with esophageal cancer development and validation of a multivariate prediction model. *Ann of Surg*. 2023;277(6):971–978. doi:10.1097/SLA.0000000000005538.
32. Rahman SA, Walker RC, Maynard N, et al. The AUGIS survival predictor: prediction of long-term and conditional survival after esophagectomy using random survival forests. *Ann of Surg*. 2023;277(2):267–274. doi:10.1097/SLA.0000000000004794.
33. Cui J, Li L, Liu N, et al. Model integrating CT-based radiomics and genomics for survival prediction in esophageal cancer patients receiving definitive chemoradiotherapy. *Biomark Res*. 2023;11(1):44. doi:10.1186/s40364-023-00480-x.
34. Zhu C, Mu F, Wang S, Qiu Q, Wang S, Wang L. Prediction of distant metastasis in esophageal cancer using a radiomics-clinical model. *Eur J of Med Res*. 2022;27(1):272. doi:10.1186/s40001-022-00877-8.
35. Kawahara D, Murakami Y, Tani S, Nagata Yasushi. A prediction model for degree of differentiation for resectable locally advanced esophageal squamous cell carcinoma based on CT images using radiomics and machine-learning. *Br J Radiol*. 2021;94(1124):20210525. doi:10.1259/bjr.20210525.
36. Kawahara D, Nishioka R, Murakami Y, et al. A Nomogram based on pretreatment radiomics and dosiomics features for predicting overall survival for esophageal squamous cell cancer: multi-institutional study. *Int J Radiat Oncol Biol Phys*. 2023;117(2S):e470–e471. doi:10.1016/j.ijrobp.2023.06.1678.
37. Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. *Ann Oncol*. 2020;31(7):912–920. doi:10.1016/j.annonc.2020.04.003.
38. Jin X, Zheng X, Chen D, et al. Prediction of response after chemoradiation for esophageal cancer using a combination of dosimetry and CT radiomics. *Eur Radiol*. 2019;29(11):6080–6088. doi:10.1007/s00330-019-06193-w.
39. Li X, Zhai Z, Ding W, et al. An artificial intelligence model to predict survival and chemotherapy benefits for gastric cancer patients after gastrectomy development and validation in international multicenter cohorts. *Int J Surg*. 2022;105:106889. doi:10.1016/j.ijsu.2022.106889.
40. Wang Y, Liu W, Yu Y, et al. CT radiomics nomogram for the preoperative prediction of lymph node metastasis in gastric cancer. *Eur Radiol*. 2020;30(2):976–986. doi:10.1007/s00330-019-06398-z.
41. Taninaga J, Nishiyama Y, Fujibayashi K, et al. Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: a case-control study. *Sci Rep*. 2019;9(1):12384. doi:10.1038/s41598-019-48769-y.
42. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. 2018;9(4):611–629. doi:10.1007/s13244-018-0639-9.
43. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation Lecture Notes in Computer Science. Springer International Publishing; 2015:234–241. doi:10.1007/978-3-319-24574-4_28.
44. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014. doi:10.48550/ARXIV.1409.1556.
45. Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39(12):2481–2495. doi:10.1109/TPAMI.2016.2644615.
46. Li S, Tso GKF, Kaijian HE. Bottleneck feature supervised U-Net for pixel-wise liver and tumor segmentation. *Expert Syst with Appl*. 2020;145:113131. doi:10.1016/j.eswa.2019.113131.
47. Ghosh S, Chaki A, Santosh KC. Improved U-Net architecture with VGG-16 for brain tumor segmentation. *Phys and Eng Sci in Med*. 2021;44(3):703–712. doi:10.1007/s13246-021-01019-w.
48. Beeche C, Singh JP, Leader JK, et al. Super U-Net: a modularized generalizable architecture. *Pattern Recognit*. 2022;128:108669. doi:10.1016/j.patcog.2022.108669.
49. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *Commun ACM*. 2020;63(11):139–144. doi:10.1145/3422622.
50. Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. *2017 IEEE International Conference on Computer Vision (ICCV)*; 2017:2242–2251. doi:10.1109/ICCV.2017.244.

51. Isola P, Zhu JY, Zhou T, Efros AA. Image-To-image translation with conditional adversarial networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2017:5967–5976. doi:10.1109/CVPR.2017.632.
52. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal*. 2019;58:101552. doi:10.1016/j.media.2019.101552.
53. Wang J, Wu QMJ, Pourpanah F. DC-cycleGAN: bidirectional CT-to-MR synthesis from unpaired data. *Comput Med Imaging Graph*. 2023;108:102249. doi:10.1016/j.compmedimag.2023.102249.
54. Zhang Y, Wang Q, Hu B. MinimalGAN: diverse medical image synthesis for data augmentation using minimal training data. *Appl Intell*. 2023;53(4):3899–3916. doi:10.1007/s10489-022-03609-x.
55. Dosovitskiy A., Beyer L., Kolesnikov A., et al. An Image is Worth 16×16 Words: transformers for Image Recognition at Scale. arXiv preprint arXiv:2010.11929, 2020. doi:10.48550/ARXIV.2010.11929.
56. Liu Z, Lin Y, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021:10012–10022. Accessed January 15, 2024. https://openaccess.thecvf.com/content/ICCV2021/html/Liu_Swin_Transformer_Hierarchical_Vision_Transformer_Using_Shifted_Windows_ICCV_2021_paper .
57. Chen J., Lu Y., Yu Q., et al. TransUNet: transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306, 2021. doi:10.48550/ARXIV.2102.04306.
58. Azad R, Kazerouni A, Heidari M, et al. Advances in medical image analysis with vision Transformers: a comprehensive review. *Med Image Anal*. 2024;91:103000. doi:10.1016/j.media.2023.103000.
59. Gheflati B, Rivaz H. Vision transformers for classification of breast ultrasound images. *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* IEEE; 2022:480–483. doi:10.1109/EMBC48229.2022.9871809.
60. Shi C, Xiao Y, Chen Z. Dual-domain sparse-view CT reconstruction with transformers. *Phys Med*. 2022;101:1–7. doi:10.1016/j.ejmp.2022.07.001.
61. Stewart BW, Wild C. *International Agency for Research on Cancer, World Health Organization, eds.* World Cancer Report 2014. International Agency for Research on Cancer; 2014.
62. Global Cancer Observatory. Esophageal Cancer Factsheet. 2020. https://gco.iarc.fr/today/data/factsheets/cancers/6-Oesophagus-fact-sheet.pdf. (accessed January 20, 2024).
63. National Cancer Institute. *Cancer Stat Facts: Esophageal Cancer*. 2023. https://seer.cancer.gov/statfacts/html/esoph.html. accessed January 20, 2024.
64. Morita FHA, Bernardo WM, Ide E, et al. Narrow band imaging *versus* lugol chromoendoscopy to diagnose squamous cell carcinoma of the esophagus: a systematic review and meta-analysis. *BMC Cancer*. 2017;17:1–14. doi:10.1186/s12885-016-3011-9.
65. Hou W, Wang L, Cai S, Lin Z, Yu R, Qin J. Early neoplasia identification in Barrett's esophagus via attentive hierarchical aggregation and self-distillation. *Med Image Anal*. 2021;72:102092. doi:10.1016/j.media.2021.102092.
66. Liu W, Yuan X, Guo L, et al. Artificial intelligence for detecting and delineating margins of early ESCC under WLI endoscopy. *Clin Transl Gastroenterol*. 2022;13(1):e00433. doi:10.14309/ctg.0000000000000433.
67. Kumagai Y, Toi M, Inoue H. Dynamism of tumour vasculature in the early phase of cancer progression: outcomes from oesophageal cancer research. *Lancet Oncol*. 2002;3(10):604–610. doi:10.1016/S1470-2045(02)00874-4.
68. Kuznetsov K, Lambert R, Rey JF. Narrow-Band Imaging: potential and limitations. *Endoscopy*. 2006;38(01):76–81. doi:10.1055/s-2005-921114.
69. Horie Y, Yoshio T, Aoyama K, et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc*. 2019;89(1):25–32. doi:10.1016/j.gie.2018.07.037.
70. Wang YK, Syu HY, Chen YH, et al. Endoscopic images by a single-shot Multibox detector for the identification of early cancerous lesions in the esophagus: a Pilot Study. *Cancers (Basel)*. 2021;13(2):321. doi:10.3390/cancers13020321.
71. Chou CK, Nguyen HT, Wang YK, et al. Preparing well for esophageal endoscopic detection using a hybrid model and transfer learning. *Cancers (Basel)*. 2023;15(15):3783. doi:10.3390/cancers15153783.
72. Feng Y, Wang B, Pan L, et al. Study protocol for artificial intelligence-assisted sponge cytology as pre-endoscopy screening for early esophegeal squmaous epithelial lesions in China. *BMC Cancer*. 2022;22(1):1105. doi:10.1186/s12885-022-10220-3.
73. Zhang P, She Y, Gao J, et al. Development of a deep learning system to detect esophageal cancer by barium esophagram. *Front Oncol*. 2022;12:766243. doi:10.3389/fonc.2022.766243.
74. Takeuchi M, Seto T, Hashimoto M, et al. Performance of a deep learning-based identification system for esophageal cancer from CT images. *Esophagus*. 2021;18(3):612–620. doi:10.1007/s10388-021-00826-0.
75. Tokai Y, Yoshio T, Aoyama K, et al. Application of artificial intelligence using convolutional neural networks in determining the invasion depth of esophageal squamous cell carcinoma. *Esophagus*. 2020;17(3):250–256. doi:10.1007/s10388-020-00716-x.
76. Shimamoto Y, Ishihara R, Kato Y, et al. Real-time assessment of video images for esophageal squamous cell carcinoma invasion depth using artificial intelligence. *J of Gastroenterol*. 2020;55(11):1037–1045. doi:10.1007/s00535-020-01716-5.
77. Peng G, Zhan Y, Wu Y, et al. Radiomics models based on CT at different phases predicting lymph node metastasis of esophageal squamous cell carcinoma (GASTO-1089). *Front in Oncol*. 2022;12:988859. doi:10.3389/fonc.2022.988859.
78. Wang Y, Zhu J, Guo D, et al. Deep learning for automatic prediction of lymph node station metastasis in esophageal cancer patients from contrast-enhanced CT. *Int J Radiat Oncol Biol Phys*. 2023;117(2):S55. doi:10.1016/j.ijrobp.2023.06.347.
79. Zhang ST, Wang SY, Zhang J, et al. Artificial intelligence-based computer-aided diagnosis system supports diagnosis of lymph node metastasis in esophageal squamous cell carcinoma: a multicenter study. *Heliyon*. 2023;9(3):e14030. doi:10.1016/j.heliyon.2023.e14030.
80. Rodrigues P, Vanharanta S. Circulating tumor cells: come together, right now, over metastasis. *Cancer Discov*. 2019;9(1):22–24. doi:10.1158/2159-8290.CD-18-1285.
81. Akashi T, Okumura T, Terabayashi K, et al. The use of an artificial intelligence algorithm for circulating tumor cell detection in patients with esophageal cancer. *Oncol Lett*. 2023;26(1):1–9. doi:10.3892/ol.2023.13906.
82. Zhou Z, Yu L, Tian S, et al. Local-global multiple perception based deep multimodality learning for sub-type of esophageal cancer classification. *Biomed Signal Process Control*. 2022;77:103757. doi:10.1016/j.bspc.2022.103757.
83. Herskovic A, Martz K, Al-Sarraf M, et al. Combined chemotherapy and radiotherapy compared with radiotherapy alone in patients with cancer of the esophagus. *N Engl J Med*. 1992;326(24):1593–1598. doi:10.1056/NEJM199206113262403.
84. Li X, Gao H, Zhu J, et al. 3D deep learning model for the pretreatment evaluation of treatment response in esophageal carcinoma: a prospective study (ChiCTR2000039279). *Int J Radiat Oncol Biol Phys*. 2021;111(4):926–935. doi:10.1016/j.ijrobp.2021.06.033.
85. An D, Li B, Cao Q, Yin W. PH-0720: delta-radiomics based on MRI predicts response to concurrent chemoradiotherapy in esophageal cancer. *Radiotherapy and Oncology*. 2020;152:S408. doi:10.1016/S0167-8140(21)00742-8.
86. Van Hagen P, Hulshof MCCM, Van Lanschot JJB, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med*. 2012;366(22):2074–2084. doi:10.1056/NEJMoa1112088.
87. Murakami Y, Kawahara D, Tani S, et al. Predicting the local response of esophageal squamous cell carcinoma to neoadjuvant chemoradiotherapy by radiomics with a machine learning method using 18F-FDG PET images. *Diagnostics (Basel)*. 2021;11(6):1049. doi:10.3390/diagnostics11061049.
88. Beukinga RJ, Wang D, Karrenbeld A, et al. Addition of HER2 and CD44 to 18F-FDG PET-based clinico-radiomic models enhances prediction of neoadjuvant chemoradiotherapy response in esophageal cancer. *Eur Radiol*. 2021;31(5):3306–3314. doi:10.1007/s00330-020-07439-8.
89. Liu Y, Wang Y, Ma Z, et al. A machine learning method to predict pathological complete response of esophageal cancer after neoadjuvant chemoradiotherapy with clinicohematological markers and MR radiomics: a multi-center study. *Int J Radiat Oncol Biol Phys*. 2023;117(2):e318. doi:10.1016/j.ijrobp.2023.06.2355.
90. Lu S, Wang C, Liu Y, et al. The MRI radiomics signature can predict the pathologic response to neoadjuvant chemotherapy in locally advanced esophageal squamous cell carcinoma. *Eur Radiol*. 2024;34(1):485–494. doi:10.1007/s00330-023-10040-4.
91. Chufal KS, Ahmad I, Dwivedi A, et al. Deep learning using Pre-NACRT imaging can predict pathological response in esophageal cancer. *Radiother and Oncol*. 2021;161:S1530–S1531. doi:10.1016/S0167-8140(21)08253-0.
92. Wang Q, Yue H, Zhou X, et al. Develop a deep radiomics model for predicting the response to neoadjuvant chemoradiotherapy (nCRT) in patients with locally advanced esophageal cancer using three-stage longitudinal CT images. *Int J Radiat Oncol Biol Phys*. 2023;117(2):e491. doi:10.1016/j.ijrobp.2023.06.1722.
93. Tonison JJ, Fischer SG, Viehrig M, et al. Radiation pneumonitis after intensity-modulated radiotherapy for esophageal cancer: institutional data and a systematic review. *Sci Rep*. 2019;9(1):2255. doi:10.1038/s41598-018-38414-5.
94. Puttanawarut C, Sirirutbunkajorn N, Khachonkham S, Pattaranutaporn P, Wongsawat Y. Biological dosiomic features for the prediction of radiation pneumonitis in esophageal cancer patients. *Radiot Oncol*. 2021;16:1–9. doi:10.1186/s13014-021-01950-y.
95. Sheng L, Zhuang L, Yang J, et al. Radiation pneumonia predictive model for radiotherapy in esophageal carcinoma patients. *BMC Cancer*. 2023;23(1):988. doi:10.1186/s12885-023-11499-6.
96. Shinoda M, Ando N, Kato K, et al. Randomized study of low-dose *versus* standard-dose chemoradiotherapy for unresectable esophageal squamous cell carcinoma (JCOG0303). *Cancer Sci*. 2015;106(4):407–412. doi:10.1111/cas.12622.
97. Hihara J, Hamai Y, Emi M, et al. Role of definitive chemoradiotherapy using docetaxel and 5-fluorouracil in patients with unresectable locally advanced esophageal squamous cell carcinoma: a phase II study: definitive CRT using docetaxel and 5-FU. *Dis Esophagus*. 2016;29(8):1115–1120. doi:10.1111/dote.12433.
98. Sun X, Han S, Gu F, et al. A retrospective comparison of taxane and fluorouracil-based chemoradiotherapy in patients with inoperable esophageal squamous cell carcinoma. *J Cancer*. 2016;7(9):1066–1073. doi:10.7150/jca.13547.
99. Kawakami T, Tsushima T, Omae K, et al. Risk factors for esophageal fistula in thoracic esophageal squamous cell carcinoma invading adjacent organs treated with definitive chemoradiotherapy: a monocentric case-control study. *BMC Cancer*. 2018;18:1–7. doi:10.1186/s12885-018-4486-3.
100. Li Z, Gong J, Zhao LN. Clinical-radiomics nomogram for risk prediction of esophageal fistula in patients with esophageal squamous cell carcinoma treated by IMRT or VMAT. *Int J Radiat Oncol Biol Phys*. 2023;117(2):e315. doi:10.1016/j.ijrobp.2023.06.2348.
101. Xu Y, Cui H, Dong T, et al. Integrating clinical data and attentional CT imaging features for esophageal fistula prediction in esophageal cancer. *Front in Oncol*. 2021;11:688706. doi:10.3389/fonc.2021.688706.
102. Dubecz A, Gall I, Solymosi N, et al. Temporal trends in long-term survival and cure rates in esophageal cancer: a SEER database analysis. *J Thorac Oncol*. 2012;7(2):443–447. doi:10.1097/JTO.0b013e3182397751.
103. Jung JO, Crnovrsanin N, Wirsik NM, et al. Machine learning for optimized individual survival prediction in resectable upper gastrointestinal cancer. *J Cancer Res Clin Oncol*. 2023;149(5):1691–1702. doi:10.1007/s00432-022-04063-5.
104. DeFreitas MR, Toronka A, Nedrud MA, et al. CT-derived body composition measurements as predictors for neoadjuvant treatment tolerance and survival in gastroesophageal adenocarcinoma. *Abdom Radiol (NY)*. 2023;48(1):211–219. doi:10.1007/s00261-022-03695-y.

105. Lin Z, Cai W, Hou W, et al. CT-Guided Survival Prediction of Esophageal Cancer. *IEEE J Biomed Health Inform*. 2022;26(6):2660–2669. doi:10.1109/JBHI.2021.3132173.
106. Gong J, Zhang W, Huang W, et al. CT-based deep learning model for predicting local recurrence-free survival in esophageal squamous cell carcinoma patients received concurrent chemo-radiotherapy: a multicenter study. *Int J Radiat Oncol Biol Phys*. 2022;114(3):S121–S122. doi:10.1016/j.ijrobp.2022.07.566.
107. Yu N, Wan Y, Zuo L, et al. MRI and CT radiomics features to predict overall survival of locally advanced esophageal cancer after definite chemoradiotherapy. *Int J Radiat Oncol Biol Phys*. 2022;114(3):E169–E170. doi:10.1016/j.ijrobp.2022.07.1051.
108. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660.
109. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA: Cancer J Clin*. 2016;66(2):115–132. doi:10.3322/caac.21338.
110. Zong L, Abe M, Seto Y, Ji J. The challenge of screening for early gastric cancer in China. *Lancet*. 2016;388(10060):2606. doi:10.1016/s0140-6736(16)32226-7.
111. Cao R, Tang L, Fang M, et al. Artificial intelligence in gastric cancer: applications and challenges. *Gastroenterol Rep*. 2022;10:goac064. doi:10.1093/gastro/goac064.
112. Lee DH, Kim SH, Joo I, Han JK. CT Perfusion evaluation of gastric cancer: correlation with histologic type. *Eur Radiol*. 2018;28(2):487–495. doi:10.1007/s00330-017-4979-5.
113. Kaise M. Advanced endoscopic imaging for early gastric cancer. *Best Pract Res Clin Gastroenterol*. 2015;29(4):575–587. doi:10.1016/j.bpg.2015.05.010.
114. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88. doi:10.1016/j.media.2017.07.005.
115. Niikura R, Aoki T, Shichijo S, et al. Artificial intelligence *versus* expert endoscopists for diagnosis of gastric cancer in patients who have undergone upper gastrointestinal endoscopy. *Endoscopy*. 2021;54(08):780–784. doi:10.1055/a-1660-6500.
116. Oura H, Matsumura T, Fujie M, et al. Development and evaluation of a double-check support system using artificial intelligence in endoscopic screening for gastric cancer. *Gastric Cancer*. 2021;25(2):392–400. doi:10.1007/s10120-021-01256-8.
117. Tang D, Wang L, Ling T, et al. Development and validation of a real-time artificial intelligence-assisted system for detecting early gastric cancer: a multicentre retrospective diagnostic study. *EBioMedicine*. 2020;62:103146. doi:10.1016/j.ebiom.2020.103146.
118. He X, Wu L, Dong Z, et al. Real-time use of artificial intelligence for diagnosing early gastric cancer by magnifying image-enhanced endoscopy: a multicenter diagnostic study (with videos). *Gastrointest Endosc*. 2022;95(4):671–678.e4. doi:10.1016/j.gie.2021.11.040.
119. Hu H, Gong L, Dong D, et al. Identifying early gastric cancer under magnifying narrow-band images with deep learning: a multicenter study. *Gastrointest Endosc*. 2021;93(6):1333–1341.e3. doi:10.1016/j.gie.2020.11.014.
120. Ling T, Wu L, Fu Y, et al. A deep learning-based system for identifying differentiation status and delineating the margins of early gastric cancer in magnifying narrow-band imaging endoscopy. *Endoscopy*. 2021;53(05):469–477. doi:10.1055/a-1229-0920.
121. Gong L, Wang M, Shu L, et al. Automatic captioning of early gastric cancer using magnification endoscopy with narrow-band imaging. *Gastrointest Endosc*. 2022;96(6):929–942.e6. doi:10.1016/j.gie.2022.07.019.
122. Zhu SL, Dong J, Zhang C, Huang YB, Pan W, Cao S. Application of machine learning in the diagnosis of gastric cancer based on noninvasive characteristics. *PLoS ONE*. 2020;15(12):e0244869. doi:10.1371/journal.pone.0244869.
123. Zhang B, Cheng L, Niu Y, et al. Identification Tool for Gastric Cancer Based on Integration of 33 Clinical Available Blood Indices Through Deep Learning. *IEEE Access*. 2022;10:106081–106092. doi:10.1109/access.2022.3172477.
124. Elsayad AS, Desouky AIE, Salem MM, Badawy M. A Deep Learning H2O Framework for Emergency Prediction in Biomedical Big Data. *IEEE Access*. 2020;8:97231–97242. doi:10.1109/access.2020.2995790.
125. Tang J, Yuan Q, Wen X, Usman M, Tay ACY, Wang L. Label-free surface-enhanced Raman spectroscopy coupled with machine learning algorithms in pathogenic microbial identification: current trends, challenges, and perspectives. *Interdiscip Med*. 2024;2(3):e20230060. doi:10.1002/INMD.20230060.
126. Tang JW, Li F, Liu X, et al. Detection of Helicobacter pylori Infection in Human Gastric Fluid Through Surface-Enhanced Raman Spectroscopy Coupled With Machine Learning Algorithms. *Lab Invest*. 2024;104(2):100310. doi:10.1016/j.labinv.2023.100310.
127. Si YT, Xiong XS, Wang JT, et al. Identification of chronic non-atrophic gastritis and intestinal metaplasia stages in the Correa's cascade through machine learning analyses of SERS spectral signature of non-invasively-collected human gastric fluid samples. *Biosens Bioelectron*. 2024;262:116530. doi:10.1016/j.bios.2024.116530.
128. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA: Cancer J Clin*. 2017;67(2):93–99. doi:10.3322/caac.21388.
129. Amin MB, Edge SB, Greene FL, et al. *AJCC Cancer Staging Manual*. 8th ed. New York: Springer; 2017.
130. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. *Lancet*. 2016;388(10060):2654–2664. doi:10.1016/S0140-6736(16)30354-3.
131. Joshi SS, Badgwell BD. Current treatment and recent progress in gastric cancer. *CA: Cancer J Clin*. 2021;71(3):264–279. doi:10.3322/caac.21657.
132. Sun RJ, Fang MJ, Tang L, et al. CT-based deep learning radiomics analysis for evaluation of serosa invasion in advanced gastric cancer. *Eur J Radiol*. 2020;132:109277. doi:10.1016/j.ejrad.2020.109277.
133. Goto A, Kubota N, Nishikawa J, et al. Cooperation between artificial intelligence and endoscopists for diagnosing invasion depth of early gastric cancer. *Gastric Cancer*. 2023;26(1):116–122. doi:10.1007/s10120-022-01330-9.
134. Nagao S, Tsuji Y, Sakaguchi Y, et al. Highly accurate artificial intelligence systems to predict the invasion depth of gastric cancer: efficacy of conventional white-light imaging, nonmagnifying narrow-band imaging, and indigo-carmine dye contrast imaging. *Gastrointest Endosc*. 2020;92(4):866–873.e1. doi:10.1016/j.gie.2020.06.047.
135. Zhao L, Han W, Niu P, et al. Using nomogram, decision tree, and deep learning models to predict lymph node metastasis in patients with early gastric cancer: a multi-cohort study. *Am J Cancer Res*. 2023;13(1):204–215.
136. Mirniaharikandehei S, Heidari M, Danala G, Lakshmivarahan S, Zheng B. Applying a random projection algorithm to optimize machine learning model for predicting peritoneal metastasis in gastric cancer patients using CT images. *Comput Methods Programs Biomed*. 2021;200:105937. doi:10.1016/j.cmpb.2021.105937.
137. Dong D, Tang L, Li ZY, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. *Ann Oncol*. 2019;30(3):431–438. doi:10.1093/annonc/mdz001.
138. Jiang Y, Liang X, Wang W, et al. Noninvasive Prediction of Occult Peritoneal Metastasis in Gastric Cancer Using Deep Learning. *JAMA Netw Open*. 2021;4(1):e2032269. doi:10.1001/jamanetworkopen.2020.32269.
139. Zurlo IV, Basso M, Strippoli A, et al. Treatment of Locally Advanced Gastric Cancer (LAGC): back to Lauren's Classification in Pan–Cancer Analysis Era? *Cancers (Basel)*. 2020;12(7):1749. doi:10.3390/cancers12071749.
140. Ma J, Shen H, Kapesa L, Zeng S. Lauren classification and individualized chemotherapy in gastric cancer. *Oncol Lett*. 2016;11(5):2959–2964. doi:10.3892/ol.2016.4337.
141. zhen Qiu M, yan Cai M, sheng Zhang D, et al. Clinicopathological characteristics and prognostic analysis of Lauren classification in gastric adenocarcinoma in China. *J Transl Med*. 2013;11:1–7. doi:10.1186/1479-5876-11-58.
142. Pernot S, Terme M, Radosevic-Robin N, et al. Infiltrating and peripheral immune cell analysis in advanced gastric cancer according to the Lauren classification and its prognostic significance. *Gastric Cancer*. 2020;23(1):73–81. doi:10.1007/s10120-019-00983-3.
143. Tang CT, Zeng L, Yang J, Zeng C, Chen Y. Analysis of the Incidence and Survival of Gastric Cancer Based on the Lauren Classification: a Large Population-Based Study Using SEER. *Front Oncol*. 2020;10:1212. doi:10.3389/fonc.2020.01212.
144. Wang XX, Ding Y, Wang SW, et al. Intratumoral and peritumoral radiomics analysis for preoperative Lauren classification in gastric cancer. *Cancer Imaging*. 2020;20:1–10. doi:10.1186/s40644-020-00358-3.
145. Wang S, Dong D, Zhang W, et al. Specific Borrmann classification in advanced gastric cancer by an ensemble multilayer perceptron network: a multicenter research. *Med Phys*. 2021;48(9):5017–5028. doi:10.1002/mp.15094.
146. Muti HS, Heij LR, Keller G, et al. Development and validation of deep learning classifiers to detect Epstein-Barr virus and microsatellite instability status in gastric cancer: a retrospective multicentre cohort study. *Lancet Digit Health*. 2021;3(10):e654–e664. doi:10.1016/s2589-7500(21)00133-3.
147. Lee SH, Lee Y, Jang HJ. Deep learning captures selective features for discrimination of microsatellite instability from pathologic tissue slides of gastric cancer. *Int J Cancer*. 2023;152(2):298–307. doi:10.1002/ijc.34251.
148. Zheng X, Wang R, Zhang X, et al. A deep learning model and human-machine fusion for prediction of EBV-associated gastric cancer from histopathology. *Nat Commun*. 2022;13(1):2790. doi:10.1038/s41467-022-30459-5.
149. Ba W, Wang S, Shang M, et al. Assessment of deep learning assistance for the pathological diagnosis of gastric cancer. *Mod Pathol*. 2022;35(9):1262–1268. doi:10.1038/s41379-022-01073-z.
150. Song Z, Zou S, Zhou W, et al. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun*. 2020;11(1):4294. doi:10.1038/s41467-020-18147-8.
151. Wang W, Sheng R, Liao S, et al. LightGBM is an Effective Predictive Model for Postoperative Complications in Gastric Cancer: a Study Integrating Radiomics with Ensemble Learning. *J Imaging Inform Med*. 2024;37(6):3034–3048. doi:10.1007/s10278-024-01172-0.
152. Aoyama Y, Matsunobu Y, Etoh T, et al. Artificial intelligence for surgical safety during laparoscopic gastrectomy for gastric cancer: indication of anatomical landmarks related to postoperative pancreatic fistula using deep learning. *Surg Endosc*. 2024;38(10):5601–5612. doi:10.1007/s00464-024-11117-x.
153. Nirkin Y., Wolf L., Hassner T. HyperSeg: patch-Wise Hypernetwork for Real-Time Semantic Segmentation. arXiv preprint arXiv:2012.11582, 2020. doi:10.48550/ARXIV.2012.11582.
154. Takeuchi M, Kawakubo H, Tsuji T, et al. Evaluation of surgical complexity by automated surgical process recognition in robotic distal gastrectomy using artificial intelligence. *Surg Endosc*. 2023;37(6):4517–4524. doi:10.1007/s00464-023-09924-9.
155. Czempiel T, Paschali M, Keicher M, et al. TeCNO: surgical Phase Recognition with Multi-stage Temporal Convolutional Networks. In: Martel AL, Abolmaesumi P, Stoyanov D, et al., eds. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Vol 12263. Lecture Notes in Computer Science. Springer International Publishing; 2020:343–352. doi:10.1007/978-3-030-59716-0_33.
156. Wang W, Peng Y, Feng X, et al. Development and Validation of a Computed Tomography–Based Radiomics Signature to Predict Response to Neoadjuvant Chemotherapy for Locally Advanced Gastric Cancer. *JAMA Netw Open*. 2021;4(8):e2121143-e2121143. doi:10.1001/jamanetworkopen.2021.21143.
157. Cui Y, Zhang J, Li Z, et al. A CT-based deep learning radiomics nomogram for predicting the response to neoadjuvant chemotherapy in patients with locally advanced gastric cancer: a multicenter cohort study. *EClinicalMedicine*. 2022;46:101348. doi:10.1016/j.eclinm.2022.101348.
158. Zhong H, Wang T, Liu X, Tian Y, Zhou Y. ASO Author Reflections: deep-Learning Radiomics Nomogram Based on Enhanced CT to Predict the Effect of Neoadjuvant Chemotherapy on Metastatic Lymph Nodes in Locally Advanced Gastric Cancer. *Ann Surg Oncol*. 2023;31(1):454–455. doi:10.1245/s10434-023-14508-x.

159. Tan J wen Wang L, Chen Y, et al. Predicting Chemotherapeutic Response for Far-advanced Gastric Cancer by Radiomics with Deep Learning Semi-automatic Segmentation. *J Cancer*. 2020;11(24):7224–7236. doi:10.7150/jca.46704.
160. Tegels JJ. Improving the outcomes in gastric cancer surgery. *World J Gastroenterol*. 2014;20(38):13692. doi:10.3748/wjg.v20.i38.13692.
161. Zhang W, Fang M, Dong D, et al. Development and validation of a CT-based radiomic nomogram for preoperative prediction of early recurrence in advanced gastric cancer. *Radiother Oncol*. 2020;145:13–20. doi:10.1016/j.radonc.2019.11.023.
162. Zhang L, Dong D, Zhang W, et al. A deep learning risk prediction model for overall survival in patients with gastric cancer: a multicenter study. *Radiother Oncol*. 2020;150:73–80. doi:10.1016/j.radonc.2020.06.010.
163. Li H, Wang S, Liu B, et al. A multi-view co-training network for semi-supervised medical image-based prognostic prediction. *Neural Netw*. 2023;164:455–463. doi:10.1016/j.neunet.2023.04.030.
164. Jiang Y, Jin C, Yu H, et al. Development and Validation of a Deep Learning CT Signature to Predict Survival and Chemotherapy Benefit in Gastric Cancer: a Multicenter, Retrospective Study. *Ann Surg*. 2021;274(6):e1153–e1161. doi:10.1097/SLA.0000000000003778.
165. Mori M, Palumbo D, De Cobelli F, Fiorino C. Does radiomics play a role in the diagnosis, staging and re-staging of gastroesophageal junction adenocarcinoma? *Updates Surg*. 2023;75(2):273–279. doi:10.1007/s13304-022-01377-4.
166. Subashchandrabose U, John R, Anbazhagu UV, Venkatesan VK, Ramakrishna MT. Ensemble Federated Learning Approach for Diagnostics of Multi-Order Lung Cancer. *Diagnostics (Basel)*. 2023;13(19):3053. doi:10.3390/diagnostics13193053.
167. du Terrail JO, Leopold A, Joly C, et al. Federated learning for predicting histological response to neoadjuvant chemotherapy in triple-negative breast cancer. *Nat Med*. 2023;29(1):135–146. doi:10.1038/s41591-022-02155-w.
168. ul Ain Q, Khan MA, Yaqoob MM, et al. Privacy-Aware Collaborative Learning for Skin Cancer Prediction. *Diagnostics (Basel)*. 2023;13(13):2264. doi:10.3390/diagnostics13132264.
169. Hatta W, Gotoda T, Koike T, Masamune A. History and future perspectives in Japanese guidelines for endoscopic resection of early gastric cancer. *Dig Endosc*. 2020;32(2):180–190. doi:10.1111/den.13531.
170. Recent Topics and Perspectives on Esophageal Cancer in Japan. *JMA J*. 2018;1(1):30–39. doi:10.31662/jmaj.2018-0002.
171. Struyvenberg MR, Van Der Sommen F, Swager AF, et al. Improved Barrett's neoplasia detection using computer-assisted multiframe analysis of volumetric laser endomicroscopy. *Dis Esophagus*. 2020;33(2):doz065. doi:10.1093/dote/doz065.
172. Van Der Putten J, Struyvenberg M, De Groof J, et al. Deep principal dimension encoding for the classification of early neoplasia in Barrett's Esophagus with volumetric laser endomicroscopy. *Comput Med Imaging Graph*. 2020;80:101701. doi:10.1016/j.compmedimag.2020.101701.
173. Quan TM, Hildebrand DGC, Jeong WK. FusionNet: a Deep Fully Residual Convolutional Neural Network for Image Segmentation in Connectomics. *Front Comput Sci*. 2021;3:613981. doi:10.3389/fcomp.2021.613981.
174. Yuan L, Yang L, Zhang S, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. *eClinicalMedicine*. 2023;57:101834. doi:10.1016/j.eclinm.2023.101834.
175. Wu L, Shang R, Sharma P, et al. Effect of a deep learning-based system on the miss rate of gastric neoplasms during upper gastrointestinal endoscopy: a single-centre, tandem, randomised controlled trial. *Gastroenterol Hepatol*. 2021;6(9):700–708. doi:10.1016/S2468-1253(21)00216-8.
176. Wu L, He X, Liu M, et al. Evaluation of the effects of an artificial intelligence system on endoscopy quality and preliminary testing of its performance in detecting early gastric cancer: a randomized controlled trial. *Endoscopy*. 2021;53(12):1199–1207. doi:10.1055/a-1350-5583.
177. Xie H, Jia Y, Liu S. Integration of artificial intelligence in clinical laboratory medicine: advancements and challenges. *Interdiscip Med*. 2024;2(3):e20230056. doi:10.1002/INMD.20230056.
178. Ge W, Huh JW, Park YR, Lee JH, Kim YH. Turchin A. An Interpretable ICU Mortality Prediction Model Based on Logistic Regression and Recurrent Neural Networks with LSTM units. *AMIA Annu Symp Proc*. 2018;2018:460–469.
179. Suh J, Yoo S, Park J, et al. Development and validation of an explainable artificial intelligence-based decision-supporting tool for prostate biopsy. *BJU Int*. 2020;126(6):694–703. doi:10.1111/bju.15122.
180. Kaji DA, Zech JR, Kim JS, et al. An attention based deep learning model of clinical events in the intensive care unit. *PLoS ONE*. 2019;14(2):e0211057. doi:10.1371/journal.pone.0211057.
181. Shickel B, Loftus TJ, Adhikari L, Ozrazgat-Baslanti T, Bihorac A, DeepSOFA Rashidi P. A Continuous Acuity Score for Critically Ill Patients using Clinically Interpretable Deep Learning. *Sci Rep*. 2019;9(1):1879. doi:10.1038/s41598-019-38491-0.
182. Yang G, Ye Q, Xia J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond. *Inf Fusion*. 2022;77:29–52. doi:10.1016/j.inffus.2021.07.016.
183. Girdhar R., El-Nouby A., Liu Z., et al. Imagebind: one embedding space to bind them all. arXiv preprint arXiv:2305.05665, 2023. doi:10.48550/ARXIV.2305.05665.
184. Wu C., Lei J., Zheng Q., et al. Can GPT-4V(ision) Serve Medical Applications? Case Studies on GPT-4V For Multimodal Medical Diagnosis. arXiv preprint arXiv:2310.09909, 2023. Doi:10.48550/ARXIV.2310.09909.arXiv:2310.09909.
185. Ma J, He Y, Li F, Han L, You C, Wang B. Segment anything in medical images. *Nat Commun*. 2024;15(1):654. doi:10.1038/s41467-024-44824-z.
186. Zhou HY, Yu Y, Wang C, et al. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. *Nat Biomed Eng*. 2023;7(6):743–755. doi:10.1038/s41551-023-01045-x.
187. Zhou Y, Chia MA, Wagner SK, et al. A foundation model for generalizable disease detection from retinal images. *Nature*. 2023;622(7981):156–163. doi:10.1038/s41586-023-06555-x.