Identification of gastric cancer with convolutional neural networks: a systematic review

Yuxue Zhao[1] (ID) · Bo Hu[2] (ID) · Ying Wang[1] · Xiaomeng Yin[3] · Yuanyuan Jiang[4] · Xiuli Zhu[1] (ID)

Abstract

The identification of diseases is inseparable from artificial intelligence. As an important branch of artificial intelligence, convolutional neural networks play an important role in the identification of gastric cancer. We conducted a systematic review to summarize the current applications of convolutional neural networks in the gastric cancer identification. The original articles published in Embase, Cochrane Library, PubMed and Web of Science database were systematically retrieved according to relevant keywords. Data were extracted from published papers. A total of 27 articles were retrieved for the identification of gastric cancer using medical images. Among them, 19 articles were applied in endoscopic images and 8 articles were applied in pathological images. 16 studies explored the performance of gastric cancer detection, 7 studies explored the performance of gastric cancer classification, 2 studies reported the performance of gastric cancer segmentation and 2 studies analyzed the performance of gastric cancer delineating margins. The convolutional neural network structures involved in the research included AlexNet, ResNet, VGG, Inception, DenseNet and Deeplab, etc. The accuracy of studies was 77.3 – 98.7%. Good performances of the systems based on convolutional neural networks have been showed in the identification of gastric cancer. Artificial intelligence is expected to provide more accurate information and efficient judgments for doctors to diagnose diseases in clinical work.

Abbreviations

AI   artificial intelligence
CNN   convolution neural network
AUC   area under the curve

Yuxue Zhao and Bo Hu contributed equally to this work, so Yuxue Zhao and Bo Hu were co-first author.

✉ Xiuli Zhu
  15820022927@163.com

Extended author information available on the last page of the article

1 Introduction

Gastric cancer is the cancer of the cavity organs with the highest incidence [11], which is a serious threat to human health. The prognosis of gastric cancer is closely related to disease stage. The diagnosis and treatment of early gastric cancer is helpful to the recovery of patients, and the 5-year survival rate of patients can exceed 90% [59]. However, most patients have been already at an advanced stage when they are diagnosed [73]. Due to the limited treatment, the survival rate of advanced gastric cancer is low and the prognosis is poor [26]. With the advancement of medical technology and the enhancement of people's health awareness, the diagnosis and treatment of gastric cancer has become an urgent need for more patients. Therefore, to improve the accuracy of gastric cancer identification, especially early gastric cancer, has become the focus of current researches.

Gastric cancer is evaluated by endoscopy, pathological pictures, imaging examination, etc. First, endoscopy has been widely used in the detection of gastric cancer. Image-enhanced endoscopy, such as narrow-band imaging [57] and linked color imaging [52], can accurately analyze the surface structure. The studies showed that the application of these endoscopic methods could improve the accuracy of gastrointestinal tumor diagnosis [10, 27]. However, a study showed that 10% of upper gastrointestinal cancers were still missed by endoscopy [41]. Even if two experts participate in an endoscopy unit, there would be missed diagnosis [61]. The reason was that doctors' accurate diagnosis of gastroscopy images required years of experience accumulation. Secondly, histopathological image recognition is the gold standard of tumor diagnosis. The shortage of pathologists has led to a huge workload for pathologists and diagnostic errors [68]. Finally, imaging examination plays an important role in the evaluation of lymph node metastasis of gastric cancer. Imaging examination is mainly based on morphological characteristics of the lesions. For example, perigastric adipose tissue is rich so it's indistinguishable from lymph nodes. Coupled with the inexperience of doctors, missed diagnosis and misdiagnosis may occur. In particular, when the number of patients is large, the accuracy of diagnosis will inevitably decrease [12].

With the increasing requirements of more accurate detection, classification and segmentation or delineating margins, artificial intelligence (AI) is booming in medical applications. AI is to make machines think like people. Machine learning is one of the most important parts of AI. Compared with traditional machine learning, such as support vector machines and Bayesian networks, deep learning has better accuracy and flexibility, and can be more easily adapted to different fields and applications. Convolution neural network (CNN) is one of the most prominent algorithms for image processing in deep learning (Fig. 1). The basic structure of CNN includes convolution layer, pooling layer, full connection layer [69]. The convolutional layer performs feature extraction from a large amount of data. The pooling layer compresses the input feature map to extract the main features. And it can reduce the dimension of the extracted feature information, simplify the network calculation complexity, and improve the calculation speed. The function of the fully connected layer is to fit together all the features and send the output value to the classifier to get the final prediction result.

Classic CNN structures include AlexNet [29], ResNet [14], VGG [53], Inception [60] and DenseNet [20], etc. Network depth refers to the number of layers that need to update parameters through training, such as convolutional layers, fully connected layers, etc. The
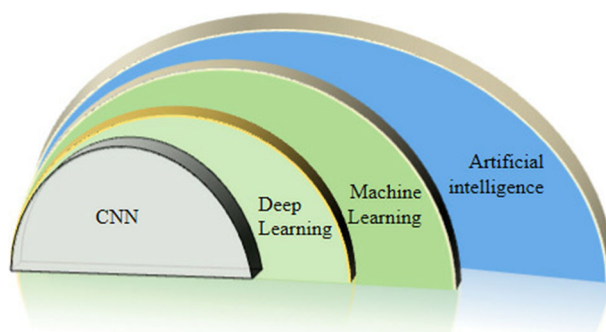
Fig. 1 The relationship between artificial intelligence, machine learning, deep learning and convolutional neural network

CNN: Convolutional neural network

basic structure of AlexNet includes 8 layers in total. There are 5 convolutional layers, followed by 3 fully connected layers. The accuracy of AlexNet has a big improvement compared to traditional methods. The kernel size of the first 4 convolutional layers are 11×11, 5×5, 3×3, 3×3, and 3×3. The kernel size of the first graphics processing unit is 3×3 and the kernel size of the second graphics processing unit is 5×5 in the fifth convolution layer. The channels of the 5 convolutional layers are 96, 256, 384, 384 and 256 respectively. The first two fully connected layers have 4096 neurons respectively, and the final output softmax has 1000 neurons.

The VGG network is developed on the basis of the AlexNet network. Compared with AlexNet, VGG uses the convolution kernel with 3×3, and the pooling kernel with 2×2. The developer changed the original three fully connected layers that were the same as AlexNet in the network into one 7×7 and two 1×1 convolutional layers in turn. The entire network is divided according to the number of layers. The most widely used is VGG-16, which is 13 convolutional layers and 3 fully connected layers.

GoogLeNet is a deep neural network model based on the inception module launched by Google. The inception module replaces the traditional operations of convolution and activation. Inception is characterized by the use of 1×1 convolution for dimensionality reduction, while convolution and re-aggregation can be performed on multiple scales. This network structure replaces all the fully connected layers with simple global average pooling, which greatly reduces the total number of parameters of the model. Subsequently, the model has been improving. Inception-v2, Inception-v3, Inception-v4 and other versions have been developed.

ResNet has been designed a residual block to allow us to train deeper networks and increase model performance. There are 5 different depth structures in PyTorch's official code, which are 18, 34, 50, 101 and 152 layers. The more common network is ResNet-50. It has a total of 50 layers of structure, including 49 convolutional layers and a fully connected layer. The size of the convolution kernel includes three types: 1×1, 3×3, and 7×7.

DenseNet is a convolutional neural network with dense connection. The emergence of DenseNet breaks away from the fixed thinking of deepening the number of network layers (ResNet) and widening the network structure (Inception) to improve network performance. Its advantage is that the network is narrower and has fewer parameters.

CNN has excellent performance in many fields, such as computer vision and natural language processing. Especially in the field of computer vision, CNN is the important model for image classification, image retrieval, object detection and semantic segmentation. For

example, agile CNN was used to diagnose the benign and malignant pulmonary nodules in chest CT images [72]. Meanwhile, CNN showed good diagnostic performance in the diagnosis of liver lesions [70] and breast ultrasound images [45]. CNN can also be used for pathological image classification. For example, fully connected CNN with extreme learning machine model was used to classify hepatocellular carcinoma nuclei [32]. CNN was also used for histopathological classification of osteosarcoma [43] and epithelial matrix pathological image classification [19].

In order to help doctors identify gastric cancer more accurately and in less time, a number of computer-aided diagnosis schemes have been developed [2, 25, 63]. Computer-aided diagnosis can help doctors reduce omissions and mischaracterizations of gastric tumor changes, thereby helping to solve the limitations of current doctors' lack of experience in examinations. Deep learning, especially CNN, has become a smarter and more accurate image processing technology. CNN is expected to become the mainstream application technology in the AI identification of gastric cancer.

The purpose of this system review is to summarize all current applications of CNN in gastric cancer identification and evaluate the performance of CNN in gastric cancer identification. This research can contribute to solving the knowledge gap in the development of CNN in gastric cancer identification.

2 Methods

2.1 Literature search

Literature retrieval was completed from Embase, Cochrane Library, PubMed, Web of Science database on 17th September, 2020 to find out the CNN study of gastric cancer's medical images . The PRISMA checklist was used for reporting the systematic review. We used medical subject headings and free-text words to search. The search terms in this article mainly include "machine learning", "artificial Intelligence", "convolutional neural network", "deep learning", "data mining", "algorithm", "tumor", "neoplasm", "carcinoma", "cancer", "lesion", "endoscope", "pathology", "computed tomography", "ultrasonography", "x-ray", "Magnetic Resonance Imaging", "stomach", "gastric", "digestive system", "diagnosis", "identification", "classification", "detection", "segmentation". There was no time limit on the publication of included articles.

To describe the proposed CNN performance in the identification of gastric cancer, we compared measures of precision, sensitivity, specificity, area under the curve (AUC) and accuracy (if available) in the models.

2.2 Selection criteria

The following inclusion and exclusion criteria were used to select the study. The inclusion criteria were:

1. Studies on the identification of gastric cancer images or lesions,
2. Using models based on CNN or the main components of CNN,
3. The diagnostic performance indexes of AI algorithm include precision, accuracy, sensitivity, specificity or AUC,
4. Studies of human subjects,

5.   Publications in English.

Guidelines, case reports, review articles, abstracts of conference papers, letters to editors and editorials were not included in this study.

2.3 Data extraction and quality assessment

The two authors (XY and YJ) extracted the data independently. The forms filled out by each author were compared, and dissenting opinions were resolved through review and discussion. In case of unresolved differences, the third evaluator acted as the final arbiter. The authors extracted following contents from the article: (a) author information (b) year of publication (c) study design (d) data set source (e) basic information about patients (f) number of images (g) number of lesions (h) CNN tasks (classification, detection, segmentation and delineating margins) (i) whether independent test sets are used (j) types of medical images or lesions.

  We carefully evaluated quality of the studies included in systematic review. The Quality Assessment on Diagnostic Accuracy Studies (QUADAS-2) [65] was used to evaluate quality of the included study. The tool consists of 4 areas: patient selection, index test, reference standard, and flow and timing [65]. Each area is assessed on the basis of high, low or unclear bias risks, and the first 3 areas are also assessed on the basis of high, low or ambiguous concerns about applicability [65]. Review Manager version 5.3 (RevMan for Windows 10, Nordic Cochrane Centre) was used to generate summary diagrams of methodology quality assessment.

3 Results

3.1 Search results

A total of 27 studies were included in this study. Figure 2 showed the PRISMA flow chart of research selection, which summarized the process of research and retrieval selection. A total of 6925 articles were selected for retrieval and deduplication in the database. 1071 irrelevant articles were deleted in the preliminary screening. The full texts of 98 articles were evaluated. Finally, 27 articles which met the inclusion criteria and were suitable for systematic evaluation were included (Fig. 2).

3.2 Characteristics of data

In all the studies, 19 articles [3, 6, 7, 15, 17, 22, 31, 34, 36–38, 40, 46, 51, 62, 64, 66, 71, 74] were about the application of CNN in gastric cancer diagnosis in the field of endoscopy and 8 articles [8, 18, 21, 33, 35, 48, 56, 58] were the application of CNN in the identification of gastric pathological images. A total of 99,777 patients, 1,422,523 images and 2605 lesions were included in this review. The average age of the patients involved in the dataset was 49.9 years old, and the range was 26 – 92 years old. Male patients accounted for 50.9% and female patients accounted for 49.1%. Most studies were based on images analysis and only one study [15] was based on lesions analysis. 2 studies [46, 51] analyzed both the images and the lesions.
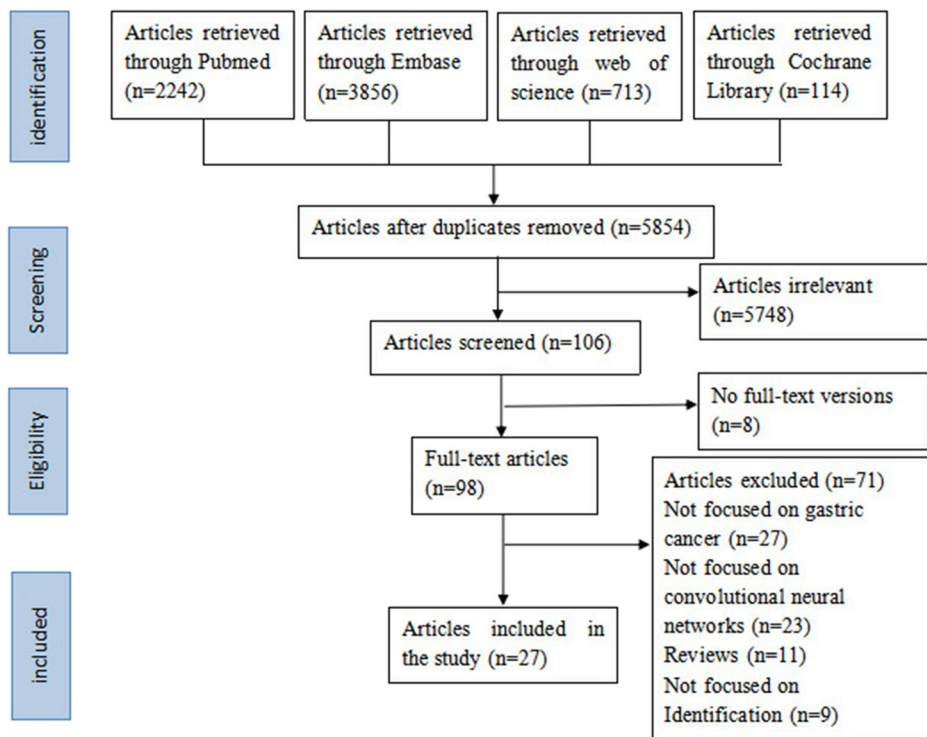
Fig. 2 PRISMA flowchart shows the searched articles using convolutional neural networks in gastric cancer image

The detailed features of the study were shown in Table 1. The data set with the best performance in each article were selected. The performance results were shown in Table 2.

All studies evaluated by the QUADAS tool had a high risk of bias or were unclear in one area at least, except Ikenoyama et al. [22]. High-risk researches [3, 6–8, 17, 18, 21, 31, 33–38, 46, 48, 51, 56, 58, 66, 71] in one of the seven fields at least were rated as low methodological quality. Figure 3 showed the literature quality evaluation results. Besides, in order to better understand the inclusion study, we used Word Art (https://wordart.com/) to analyze the co-occurrence map of keywords, as shown in Fig. 4. We found that the hot keywords were artificial intelligence, convolutional neural network, deep learning, endoscopy, narrow band imaging, cancer, classification, diagnosis, which were consistent with the focus of identification of gastric cancer.

3.3 Detection (16 Studies)

There are 16 studies [7, 15, 17, 22, 31, 34, 38, 40, 46, 51, 56, 62, 64, 66, 71, 74] on the application of CNN in the detection of gastric cancer. 15 studies [7, 15, 17, 22, 31, 34, 38, 40, 46, 51, 62, 64, 66, 71, 74] were published in the field of gastric endoscopy. Among them, 4

Table 1 Studies on convolutional neural network and gastric cancer

Ref	Year	Study design	Location	Patients	Images or Lesions	Images	Lesions	Types of medical images or lesions	Task	Independent test set
An et al. [3]	2020	Retrospective	China	1095	Images	2488	NR	Endoscopic	Delineating Margins	No
Cho et al. [7]	2020	Retrospective	Korea	846	Images	3105	NR	Endoscopic	Detect invasion depth	Yes
Cho et al. [6]	2019	Retrospective	Korea	1269	Images	5017	NR	Endoscopic	Classification	Yes
Hirasawa et al. [15]	2018	Prospective	Japan	69	Lesions	NR	77	Endoscopic	Detection	Yes
Horiuchi et al. [17]	2019	Retrospective	Japan	NR	Images	2828	NR	Endoscopic	Detection	Yes
Ikenoyama et al. [22]	2020	Retrospective	Japan	2779	Images	16,524	NR	Endoscopic	Detection	Yes
Lee et al. [31]	2019	Retrospective	Korea	NR	Images	787	NR	Endoscopic	Detection	No
Li et al. [34]	2020	Retrospective	China	NR	Images	2429	NR	Endoscopic	Detection	Yes
Ling et al. [36]	2020	Retrospective	China	342	Images	4969	NR	Endoscopic	Delineating Margins	Yes
Liu et al. [37]	2018	Retrospective	China	NR	Images	3871	NR	Endoscopic	Classification	Yes
Lui et al. [38]	2019	Retrospective	Hong Kong, China	NR	Images	3000	NR	Endoscopic	Detection	Yes
Luo et al. [40]	2019	Retrospective	China	84,424	Images	1,036,496	NR	Endoscopic	Detection	Yes
Nagao et al. [46]	2020	Retrospective	Japan	1084	Images/Lesions	16,557	2434	Endoscopic	Detect invasion depth	Yes
Shibata et al. [51]	2020	Retrospective	Japan	135	Images/Lesions	1741	94	Endoscopic	Detection	No
Ueyama et al. [62]	2020	Retrospective	Japan	349	Images	7874	NR	Endoscopic	Detection	Yes
Wang et al. [64]	2019	Retrospective	China	NR	Images	104,864	NR	Endoscopic	Detection	Yes
Wu et al. [66]	2019	Retrospective	China	NR	Images	9351	NR	Endoscopic	Detection	Yes
Yoon et al. [71]	2019	Retrospective	Korea	800	Images	11,539	NR	Endoscopic	Detection/ Detect invasion depth	Yes
Zhu et al. [74]	2020	Retrospective	China	993	Images	993	NR	Endoscopic	Detect invasion depth	Yes
Qu et al. [48]	2018	Retrospective	Japan	NR	Images	48,000	NR	Pathological	Classification	No
Cho et al. [8]	2020	Retrospective	Korea	432	Images	803	NR	Pathological	Classification	Yes
Iizuka et al. [21]	2020	Retrospective	Japan	NR	Images	5103	NR	Pathological	Classification	Yes

Table 1 (continued)

| Ref	Year	Study design	Location	Patients	Images or Lesions	Images	Lesions	Types of medical images or lesions	Task	Independent test set
Sun et al. [58]	2019	Retrospective	China	NR	Images	500	NR	Pathological	Segmentation	No
Liang et al. [35]	2019	Retrospective	China	NR	Images	1900	NR	Pathological	Segmentation	No
Hu et al. [18]	2019	Retrospective	China	30	Images	65,328	NR	Pathological	Classification	No
Li et al. [33]	2019	Prospective	China	120	Images	48,000	NR	Pathological	Classification	No
Song et al. [56]	2020	Retrospective	China	4210	Images	6917	NR	Pathological	Detection	Yes

Note: NR: Not reported

Table 2 Studies of Images that provided evaluation index results

Ref	Precision (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
An et al. [3]	N/A	88.9	N/A	N/A	N/A
Cho et al. [7]	N/A	77.3	80.4	80.7	88.7
Cho et al. [6]	N/A	81.9	75.9	85.3	87.7
Hirasawa et al. [15]	N/A	98.6	92.2	N/A	N/A
Horiuchi et al. [17]	N/A	85.3	95.4	71.0	N/A
Ikenoyama et al. [22]	N/A	N/A	58.40	87.30	75.7
Lee et al. [31]	N/A	96.49	N/A	N/A	97
Li et al. [34]	N/A	90.91	91.18	90.64	N/A
Ling et al. [36]	N/A	88.1	N/A	N/A	N/A
Liu et al. [37]	99	96	99	N/A	N/A
Lui et al. [38]	N/A	91.0	97.1	85.9	91
Luo et al. [40]	N/A	92.8	94.2	92.3	N/A
Nagao et al. [46]	N/A	N/A	N/A	N/A	95.90
Shibata et al. [51]	N/A	N/A	96	N/A	N/A
Ueyama et al. [62]	N/A	98.70	98	100	N/A
Wang et al. [64]	N/A	N/A	79.622	78.48	N/A
Wu et al. [66]	N/A	92.5	94.0	91.0	N/A
Yoon et al. [71]	N/A	N/A	91.0	97.6	98.1
Zhu et al. [74]	N/A	89.16	76.47	95.56	94
Qu et al. [48]	86.9	N/A	N/A	N/A	96.3
Cho et al. [8]	N/A	78	100	56	100
Iizuka et al. [21]	N/A	N/A	N/A	N/A	98
Sun et al. [58]	91.6	N/A	N/A	N/A	N/A
Liang et al. [35]	N/A	91.09	N/A	N/A	N/A
Hu et al. [18]	N/A	94.38	94.99	93.76	N/A
Li et al. [33]	N/A	96.5	96.6	96.7	N/A
Song et al. [56]	N/A	87.3	99.6	84.3	98.6

Note: N/A: Not available

studies [7, 46, 71, 74] included the detection of invasion depth of gastric cancer. A study [56] was based on CNN to judge the benign and malignant pathological images.
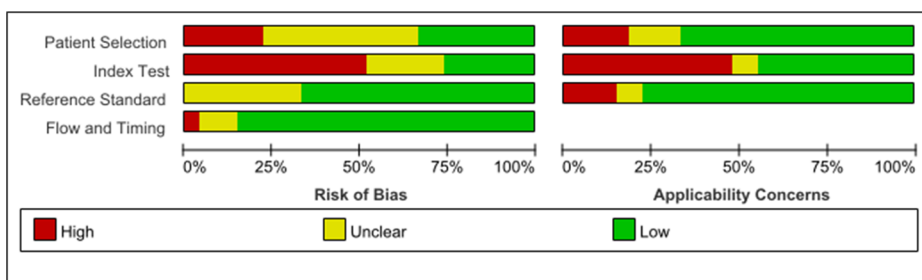
Fig. 3 The literature quality evaluation results. High: The relevant entry in each area is high risk in risk of bias; high concerns in applicability concerns; Unclear: The relevant entry in each area is unclear risk in risk of bias; unclear concerns in applicability concerns; Low: The relevant entry in each area is low risk in risk of bias; low concerns in applicability concerns. Cho 2020 a: Refers to the study whose first author is Bum-Joo Cho [7]. Cho 2020 b: Refers to the study whose first author is Kyung-Ok Cho [8]

Fig. 3 continued.

All articles evaluated the detection performance of CNN with one of the five indicators of precision, accuracy, sensitivity, specificity or AUC at least. 11 studies [7, 15, 17, 31, 34, 38, 40, 56, 62, 66, 74] described the accuracy, ranging from 77.3 to 98.6%. 14 studies [7, 15, 17,

Fig. 4 Keyword analysis of word cloud containing papers

22, 34, 38, 40, 51, 56, 62, 64, 66, 71, 74] described sensitivity, ranging from 58.40 to 99.6%. 12 studies [7, 17, 22, 34, 38, 40, 56, 62, 64, 66, 71, 74] described specificity, ranging from 71.0 to 100%. In the studies [7, 22, 31, 38, 46, 56, 71, 74] that provided AUC, the AUC was greater than or equal to 75.7% (Table 2).

Hirasawa et al. [15], Luo et al. [40], Li et al. [34], Cho et al. [7], Ikenoyama et al. [22] and Song et al. [56] used multi-center, multi-institutional data sets. Hirasawa et al. [15] retrospectively obtained a large number of images from two hospitals and two clinics in Japan to establish a gastric cancer image diagnosis system with an accuracy rate of 98.6%. Luo et al. [40] collected data from five hospitals in China for analysis to ensure sufficient and diverse data.

Narrow-band imaging was used in 4 studies [17, 34, 38, 62] and white light imaging was used in 8 studies [7, 31, 40, 51, 64, 66, 71, 74]. 3 studies [15, 22, 46] used images with standard white light, chromoendoscopy using indigo carmine spraying and narrow-band imaging. Song et al. [56] used the H&E-stained whole slide images for training and testing on CNN.

From the technical perspective, CNN structures in different studies have their own characteristics. Hirasawa et al. [15] and Ikenoyama et al. [22] used a Single Shot MultiBox Detector method without changing their algorithm. Single Shot MultiBox Detector is a deep CNN consisting of 16 or more layers. To train and test CNN, both studies used the Caffe deep learning framework with a global learning rate of 0.0001. In the study of Hirasawa et al. [15], CNN analyzed more than 2,000 test images in only 47 s. This high rate of recognizing and judging images far exceeded the recognition ability of human beings.

Horiuchi et al. [17] chose the GoogLeNet structure which was a 22-layer CNN to process the data. And the Caffe deep learning framework was used to train and verify the CNN system. The test speed of this model on images could reach 0.02 s/image. Li et al. [34] used the Inception-v3 model, which consisted of 11 Inception modules. An important feature of Inception-v3 was that it split a larger two-dimensional convolution into a smaller one-dimensional convolution, which improved computational efficiency of the model. The Keras deep learning framework was used to train the Inception-v3 model.

Lui et al. [38] constructed an image classifier based on pre-trained ResNet. The ResNet structure has 5 convolutional layers and 3 fully connected layers. Nagao et al. [46], Shibata

et al. [51], Ueyama et al. [62] and Zhu et al. [74] used the ResNet-50 structure to identify gastric cancer images or lesions. Nagao et al. [46], Ueyama et al. [62] and Zhu et al. [74] all adopted transfer learning methods to build the identification system. Shibata et al. [51] proposed a Mask R-CNN method, which introduced ResNet-50 as the backbone. The mask layer consisted of 7 convolutional layers.

Yoon et al. [71] used the VGG-16 model, which included 16 convolutional layers and a rectified linear unit. The research adapted the last layer of VGG-16 into a two-dimensional fully connected layer to classify the input endoscopic images. Wu et al. [66] aimed to establish an early gastric cancer diagnosis system for esophagogastroduodenoscopy without blind spots. This study combined VGG-16 and ResNet-50 to identify images. The TensorFlow deep learning framework was used to train, validate and test the system.

Cho et al. [7] used Inception-ResNet-v2 and the DenseNet-161 models to distinguish gastric cancer images. The results showed that in the external verification set, the average AUC values of Inception-ResNet-v2 and the DenseNet-161 could reach 0.769 and 0.887, respectively. Lee et al. [31] used Inception-v3, ResNet-50 and VGG-16 structures to detect gastric benign ulcers and cancers. Depth is important for many visual recognition tasks. The ResNet-50 architecture used in this study showed the best results. Wang et al. [64] combined the three CNN structures of AlexNet, GoogLeNet and VGG to construct a multi-column convolutional neural network. The results showed that the ensemble model was superior to other single classifiers in image analysis. The network structure was implemented by using the Caffe framework.

Luo et al. [40] developed a real-time artificial intelligence-assisted image recognition system, which was based on the DeepLab-v3 algorithm. It has been applied in clinical practice. Song et al. [56] used DeepLab-v3 architecture to help pathologists identify benign and malignant images of stomach's entire slide.

The performance of CNN in the detection of gastric cancer was compared with that of experts in 5 studies [22, 34, 38, 40, 66]. Luo et al. [40] showed that the accuracy of expert level can be achieved when using CNN-based diagnosis system to assist non-expert doctors in disease diagnosis. Li et al. [34] found that the specificity and accuracy of CNN's diagnostic system was not significantly different from that of experts, but CNN's diagnostic sensitivity was significantly higher than that of experts. Lui et al. [38] pointed out that the accuracy and AUC of the AI system were superior to all junior endoscopes. Wu et al. [66] found that the detection performance of Deep CNN was better than that of endoscopy at all levels. Ikenoyama et al. [22] compared the diagnostic ability of CNN with 67 endoscopists. The results showed that the sensitivity of CNN was significantly higher than endoscopists (by 26.5%; 95% confidence interval, 14.9 – 32.5%) and the detection time was greatly shortened.

3.4 Classification (7 Studies)

7 studies [6, 8, 18, 21, 33, 37, 48] used CNN to classify gastric cancer. 2 studies [6, 37] reported the classification of CNN in endoscopic images and 5 studies [8, 18, 21, 33, 48] reported the classification level of CNN in pathological images.

Liu et al. [37] and Qu et al. [48] reported the precision of 99% and 86.9%, respectively. In the 5 studies [6, 8, 18, 33, 37] reporting accuracy and sensitivity, the range of accuracy was 78 – 96.5% and the sensitivity was 75.9 – 100%. In the 4 studies [6, 8, 18, 33] reporting specificity, the ranged was 56 – 96.7%. In the 4 studies [6, 8, 21, 48] reporting AUC, the AUC range was 87.7 – 100%.

Cho et al. [6], Iizuka et al. [21] and Li et al. [33] used multicentre data sets. Cho et al. [6] collected 5017 white lights images to establish a classification model. The images of gastric cancer were divided into five categories: advanced gastric cancer, early gastric cancer, highgrade dysplasia, lowgrade dysplasia and non-neoplastic lesions. Iizuka et al. [21] used data from two hospitals and The Cancer Genome Atlas to the analyze whole slide images of stomach and colon. In addition, Li et al. [33] used prospective data sets from two Chinese hospitals.

Cho et al. [6] used 3 CNN structures of Inception-v4, ResNet-152 and Inception-ResNet-v2. Inception-RESNET-v2 was the best structure for gastric cancer classification with an accuracy of 81.9%. Liu et al. [37] proposed a new transfer learning deep CNN method. 4 typical pre-trained CNNs were selected to construct the classifier, namely VGG-16, Inception-v3, Inception-ResNet-v2 and ResNet-50. All the original network architecture of the CNN model in the article was kept before the first fully connected layer and a new 3 layers are added: the global average convergence layer and 4 fully connected layers. The results showed that the CNN's classification performance of fine-tuning pre-training proposed was better than that of traditional manual features and trained CNN. The transfer learning of CNN structure had a strong potential in the classification of narrow-band imaging images of the stomach. Qu et al. [48] proposed a CNN scheme using two-level fine-tuning and introduced an intermediate data set related to the target, which could solve the problem of insufficient data. The program used VGG-16, AlexNet and Inception-v3 methods. Cho et al. [8] established an automatic classifier using whole slide images from 432 patients from The Cancer Genome Atlas. AlexNet, ResNet-50 and Inception-v3 architectures to classify whole slide images were used in the study.

Iizuka et al. [21] used Inception-v3 network as the architecture of the classification model. And a depth multiplier of 0.35 was used to reduce the number of parameters, thus achieving a more streamlined structure. Hu et al. [18] proposed a Spectral-Spatial-CNN method to classify the hyperspectral images of gastric cancer. Its performance was better than that of artificial neural network and support vector machine. Li et al.[33] used ResNet-34 combined with a spectral spatial classification method to classify gastric cancer tissues, including a convolutional layer and 4 residual block residual modules.

Cho et al. [6], Cho et al. [8], Iizuka et al. [21] and Li et al. [33] used TensorFlow to train, test and verify the model. The deep learning model proposed by Hu et al. [18] was implemented by the Keras framework.

3.5 Segmentation (2 Studies)

Sun et al. [58] and Liang et al. [35] reported the application of CNN in the segmentation of gastric cancer pathological images. Sun et al. [58] proposed a multi-scale embedded network to segment pathological images. ResNet-101 was selected as the backbone network of the segmentation model. The pixel-level accuracy reached 91.60% and mean Intersection over Union was 82.65%. Liang et al. [35] used a pathological image segmentation model based on the combination of full convolution network and the patch-based method. The model framework consisted of 30 convolutional layers, 7 pooling layers, 7 down-sampling layers and an output layer. This study proposed the reiterative learning method, which enabled full convolution network to achieve good segmentation effect in the case of using the gastric cancer pathological data set without manual annotation. The accuracy of training on sequential

patches in the 1th iteration was about 5% better than patch model. The recall rate of training on mixed patches in the 1th iteration was 5.29% higher than patch model.

3.6 Delineating margins (2 Studies)

A real-time monitoring system was developed by Ling et al. [36]. This system was used to accurately identify the differentiation state of early gastric cancer and delineate the edge in magnifying narrow-band imaging endoscopy. The study used multi-center data sets from Nanjing Gulou Hospital and Wuhan University People's Hospital in China. An et al. [3] established a system based on fully convolutional neural network for sketching the edges of early gastric cancer in indigo carmine chromoendoscopy or white light endoscopy. 2 studies had all introduced Unet++ for more accurate segmentation of the image range. The accuracy of the system was 85.7% for chromoendoscopy images and 88.9% for white light endoscopy images. The resection area predicted by the system covered all the lesion areas of all patients, while the resection area based on magnifying narrow-band imaging only covered all cancer areas of 80.00% of the patients.

4 Discussion

The lack of access to primary health care and the correctness of diagnosis represent a major challenge for the global health care system [54, 55]. In recent years, the application of machine learning algorithms in medicine has been more and more extensive, from simpler traditional machine learning methods to newly developed deep learning methods [23, 44, 47]. Both endoscopy and pathological examination are operator-dependent, and the diagnosis process is completely subjective [16, 49]. However, artificial intelligence-assisted inspection may help to provide another opinion and avoid reliance on operators in diagnostic tests [49]. The application of the algorithm has a great driving force for improving the reliability of clinical diagnosis and promoting the development of medical and health undertakings.

Our systematic review shows that the CNN models of gastric cancer perform well in detection, classification, segmentation and delineating margins. There is no uniform standard and evaluation standard for the specific CNN model design. It is an ideal research process to compare the precision, accuracy, sensitivity, specificity or AUC of CNN with other model methods and data sets in the studies. In all studies, the precision, accuracy, sensitivity, specificity or AUC were greater than 70% except Ikenoyama et al. [22] and Cho et al. [8]

All articles were published between 2018 and 2020. This shows that CNN has developed rapidly in the identification of cancer in recent years, and the application of deep learning in medical decision-making is more extensive. The countries involved in the research were concentrated in China, Japan,South Korea and Hong Kong, China. We infer that more studies on gastric cancer in these countries are related to the higher incidence of gastric cancer than in other countries.

9 studies [6, 15, 21, 22, 33, 34, 36, 40, 56] used multi-center, multi-institutional data sets. At the same time, most of the researches processed the images that did not conform to the standards. During model training and testing, the image quality of model recognition is high. An important prerequisite for the application of AI in medical decision-making is sufficient and high-quality medical images. A number of data sets for pathological image studies are derived from The Cancer Genome Atlas [9, 28]. However, there is no uniform and public

large-scale image database in the endoscopic examination of gastric cancer. Data acquisition is still a major challenge.

CNN has great potential in image recognition. Compared with traditional machine learning, it has better discrimination performance. Developers can process images by using a more mature network structure, such as VGG. CNN can also be combined with transfer learning and parameter fine-tuning to make the network structure and data more coordinated, which can alleviate the problem of insufficient training data and over-fitting. Since AlexNet was proposed in 2012, CNN has ushered in explosive development. With the emergence of networks such as VGG, Inception, and ResNet, CNN's network depth is getting larger and larger to ensure accuracy. And it is more inclined to use smaller convolution kernels to ensure calculation speed. CNN is gradually developing towards lightweight and rapid development. The emergence of MobileNet [30] which can meet needs of mobile terminals makes CNN's development prospects broader.

The application of AI in gastrointestinal diseases is increasingly mature. A review published by Min et al. [42] evaluated the performance of deep learning in the disease detection through gastrointestinal endoscopy and compared it with healthcare professionals. The results of CNN's endoscope diagnosis were encouraging, but its diagnostic ability was highly dependent on the quality and quantity of training data. A meta-analysis showed that machine learning performs well in the diagnosis of helicobacter pylori infection based on patients and images [4]. Another meta-analysis showed that AI was more accurate in the diagnosis of esophageal and gastric neoplastic lesions, with AUC > 0.88 [39]. Many diseases are increasingly inclined to use artificial intelligence to assist diagnosis. For example, during the global novel coronavirus pneumonia epidemic, some scholars used deep learning algorithm to identify chest CT images [67]. In the test set, AUC could be achieved with the accuracy, sensitivity and specificity of 0.819, 0.760, 0.811 and 0.615, respectively [67].

At present, AI technology is widely used, but it still has some limitations. Currently, one of the limitations of deep learning is interpreting and understanding questions related to how AI models make decisions, known as "black box" questions [1]. The emergence of explainable AI offers hope for solving this problem [50]. Secondly, there are moral and safety issues such as the privacy of patients data and how to determine who is responsible for patients' misdiagnosis or incorrect treatment.

Despite some challenges, the prospect of AI is very exciting. The biggest feature of AI is efficient calculation and accurate decision-making. Technically, CNN's predictions can be automated. The creation of a telemedicine service platform and even AI automatic diagnosis services are very likely to be realized. With the continuous accumulation of available data, the performance and reliability of AI will continue to be improved. It can be used as a tool for identification of endoscopic pictures or pathological slices and used to discover diseases and predict the depth of invasion. It can also be used to classify the type of cancer and to delineate the edges of pathological slices. The CNN-based identification technology for gastric cancer is expected to become the mainstream deep learning technology in the next few decades. We believe that doctors and pathologists can use the results provided by CNN as a reference to solve the diagnosis of difficult cases in the near future. At the same time, it is no longer necessary to examine the images of all patients one by one, but only to identify those unclear and difficult images, which will greatly reduce the burden of work. AI can also play an important role in areas where the economy is underdeveloped and medical resources are relatively scarce. The diversification of online medical methods, such as online consultation, has led to the rapid development of telemedicine services. Based on the telemedicine service platform, CNN can identify patient images in areas where lacking professional doctors, greatly

improving the medical level of remote areas. We hope to see that clinicians, pathologists and artificial intelligence work together to support healthcare.

The advantages of this review included the careful selection of research reports on CNN algorithms evaluated in gastric cancer images. We avoided other redundant studies that evaluate non-CNN-based learning algorithms and articles that were not original studies. We compared the data sources of researches, the number of images and the performance of each model, etc., and strictly evaluated the quality of the research. In addition, previous studies [13, 24, 69] have reviewed the application of AI to gastric diseases, but this research was the first to focus entirely on the application of CNN in gastric cancer. At the same time, this study elaborated on the details of various CNNs, and fully excavated CNN's information in gastric cancer detection, classification, segmentation and delineating margins, etc. This can provide more comprehensive help for future CNN-based research on gastric cancer identification.

There are also limitations in this study. First of all, this study only summarized the results of gastric cancer imaging published by CNN in English, and did not include excellent articles in other languages. Secondly, it is difficult to make simple comparisons of performance due to the different difficulty of images or lesions evaluation. Third, images included in research came from different countries and regions, and the research objects were from different races, so the images were heterogeneous. At the same time, there were differences in methods and evaluation indicators. No meta-analysis is carried out on the article. Fourth, most of the current studies were retrospective and mainly based on still images. The recent guideline [5] points out that AI needs to be validated in multicenter randomized controlled trials before routine clinical use, so prospective and multicenter studies are needed in the future.

5 Conclusions

In the summary of CNN's research on medical images, it finds that CNN is a good tool for identifying gastric cancer. All the studies showed good identification performance and the model accuracy of all studies was 77.3 – 98.7%. CNN is expected to become an important tool to help doctors and pathologists improve the accuracy and efficiency of disease identification.

Author contributions Conception and design: Yuxue Zhao, Bo Hu, Ying Wang, Xiuli Zhu. Data analysis and interpretation: Yuxue Zhao, Bo Hu, Ying Wang, Xiaomeng Yin, Yuanyuan Jiang, Xiuli Zhu. Article drafting: Yuxue Zhao, Bo Hu, Xiaomeng Yin, Yuanyuan Jiang. Critical revision of the article for important intellectual content: Xiuli Zhu. All authors discussed the results and approved the final manuscript.

Declarations

Conflict of interest None to declare.

References

1. Acs B, Rantalainen M, Hartman J (2020) Artificial intelligence as the next step towards precision pathology. 288(1):62–81. https://doi.org/10.1111/joim.13030
2. Ali H, Yasmin M, Sharif M, Rehmani MH (2018) Computer assisted gastric abnormalities detection using hybrid texture descriptors for chromoendoscopy images. Comput Methods Programs Biomed 157:39–47. https://doi.org/10.1016/j.cmpb.2018.01.013
3. An P, Yang DM, Wang J, Wu LL, Zhou J, Zeng Z, Huang X, Xiao Y, Hu S, Chen YY, Yao F, Guo MW, Wu Q, Yang YN, Yu HG (2020) A deep learning method for delineating early gastric cancer resection margin under chromoendoscopy and white light endoscopy. Gastric Cancer 23(5):884–892. https://doi.org/10.1007/s10120-020-01071-7
4. Bang CS, Lee JJ (2020) Artificial intelligence for the prediction of helicobacter pylori infection in endoscopic images: systematic review and meta-analysis of diagnostic test accuracy. 22(9):e21983. https://doi.org/10.2196/21983
5. Bisschops R, East JE, Hassan C, Hazewinkel Y, Kamiński MF, Neumann H, Pellisé M, Antonelli G, Bustamante Balen M, Coron E, Cortas G, Iacucci M, Yuichi M, Longcroft-Wheaton G, Mouzyka S, Pilonis N, Puig I, van Hooft JE, Dekker E (2019) Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. Endoscopy 51(12):1155–1179. https://doi.org/10.1055/a-1031-7657
6. Cho BJ, Bang CS, Park SW, Yang YJ, Seo SI, Lim H, Shin WG, Hong JT, Yoo YT, Hong SH, Choi JH, Lee JJ, Baik GH (2019) Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. Endoscopy 51(12):1121–1129. https://doi.org/10.1055/a-0981-6133
7. Cho BJ, Bang CS, Lee JJ, Seo CW, Kim JH (2020) Prediction of submucosal invasion for gastric neoplasms in endoscopic images using deep-learning. J Clin Med 9(6):14. https://doi.org/10.3390/jcm9061858
8. Cho KO, Lee SH, Jang HJ (2020) Feasibility of fully automated classification of whole slide images based on deep learning. Korean J Physiol Pharmacol 24(1):89–99. https://doi.org/10.4196/kjpp.2020.24.1.89
9. Cooper LA, Demicco EG (2018) PanCancer insights from The Cancer Genome Atlas: the pathologist's perspective. 244(5):512–524. https://doi.org/10.1002/path.5028
10. Dohi O, Majima A, Naito Y, Yoshida T, Ishida T, Azuma Y, Kitae H, Matsumura S, Mizuno N, Yoshida N (2020) Can image-enhanced endoscopy improve the diagnosis of Kyoto classification of gastritis in the clinical setting? 32(2):191–203. https://doi.org/10.1111/den.13540
11. Florea A, Sangaré L, Lowe K (2020) A multinational assessment of gastric, esophageal, and colorectal cancer burden: a report of disease incidence, prevalence, and fatality. 51 (3):965-971. https://doi.org/10.1007/s12029-019-00328-4
12. Gao Y, Zhang ZD, Li S, Guo YT, Wu QY, Liu SH, Yang SJ, Ding L, Zhao BC, Li S, Lu Y (2019) Deep neural network-assisted computed tomography diagnosis of metastatic lymph nodes from gastric cancer. Chin Med J 132(23):2804–2811. https://doi.org/10.1097/cm9.0000000000000532
13. Gonçalves WGE, Dos Santos MHP, Lobato FMF, Ribeiro-Dos-Santos Â, de Araújo GS (2020) Deep learning in gastric tissue diseases: a systematic review. 7(1):e000371. https://doi.org/10.1136/bmjgast-2019-000371
14. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778
15. Hirasawa T, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T, Ohnishi T, Fujishiro M, Matsuo K, Fujisaki J, Tada T (2018) Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric Cancer 21(4):653–660. https://doi.org/10.1007/s10120-018-0793-2
16. Hoogenboom SA, Bagci U, MB W (2020) Artificial intelligence in gastroenterology. The current state of play and the potential. How will it affect our practice and when? Tech Innov Gastrointest Endosc 22(2):42–47. https://doi.org/10.1016/j.tgie.2019.150634
17. Horiuchi Y, Aoyama K, Tokai Y, Hirasawa T, Yoshimizu S, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T (2020) Convolutional neural network for differentiating gastric cancer from gastritis using magnified endoscopy with narrow band imaging. Dig Dis Sci 65(5):1355–1363. https://doi.org/10.1007/s10620-019-05862-6
18. Hu BL, Du J, Zhang ZF, Wang Q (2019) Tumor tissue classification based on micro-hyperspectral technology and deep learning. Biomed Opt Express 10(12):6370–6389. https://doi.org/10.1364/boe.10.006370
19. Huang Y, Zheng H, Liu C, Ding X, Rohde G (2017)Epithelium-stroma classification via convolutional neural networks and unsupervised domain adaptation in histopathological images. IEEE J Biomed Health Inform 21(6):1625–1632. https://doi.org/10.1109/jbhi.2017.2691738

20. Huang G, Liu Z, Laurens V, Weinberger KQ (2017) Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 2261-2269. https://doi.org/10.1109/CVPR.2017.243

21. Iizuka O, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M (2020) Deep learning models for histopathological classification of gastric and colonic epithelial tumours. Sci Rep 10(1):11. https://doi.org/10.1038/s41598-020-58467-9

22. Ikenoyama Y, Hirasawa T, Ishioka M, Namikawa K, Yoshimizu S, Horiuchi Y, Ishiyama A, Yoshio T, Tsuchida T, Takeuchi Y, Shichijo S, Katayama N, Fujisaki J, Tada T (2020) Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists. Dig Endosc 10. https://doi.org/10.1111/den.13688

23. Jeyaraj PR, Samuel Nadar ER (2019)Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. J Cancer Res Clin Oncol 145(4):829–837. https://doi.org/10.1007/s00432-018-02834-7

24. Jin P, Ji X, Kang W, Li Y, Liu H, Ma F, Ma S, Hu H, Li W, Tian Y (2020) Artificial intelligence in gastric cancer: a systematic review. 146(9):2339–2350. https://doi.org/10.1007/s00432-020-03304-9

25. Kanesaka T, Lee TC, Uedo N, Lin KP, Chen HZ, Lee JY, Wang HP, Chang HT (2018)Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. Gastrointest Endosc 87(5):1339–1344. https://doi.org/10.1016/j.gie.2017.11.029

26. Katai H, Ishikawa T, Akazawa K, Isobe Y, Miyashiro I, Oda I, Tsujitani S, Ono H, Tanabe S, Fukagawa T, Nunobe S, Kakeji Y, Nashimoto A (2018)Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001-2007). Gastric Cancer 21(1):144–154. https://doi.org/10.1007/s10120-017-0716-7

27. Kim JW (2018) Usefulness of narrow-band imaging in endoscopic submucosal dissection of the stomach. Clin Endosc 51(6):527–533. https://doi.org/10.5946/ce.2018.186

28. Kosaraju SC, Hao J, Koh HM, Kang M (2020) Deep-Hipo: Multi-scale receptive field deep learning for histopathological image analysis. Methods 179:3–13. https://doi.org/10.1016/j.ymeth.2020.05.012

29. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. Proceedings of the Advances in Neural Information Processing Systems, 1097–1105

30. Kulkarni U, Gurlahosur SMM, Bhogar SV (2021) Quantization friendly MobileNet (QF-MobileNet) architecture for vision based applications on embedded platforms. Neural Nnetw 136:28–39. https://doi.org/10.1016/j.neunet.2020.12.022

31. Lee JH, Kim YJ, Kim YW, Park S, Choi YI, Kim YJ, Park DK, Kim KG, Chung JW (2019) Spotting malignancies from gastric endoscopic images using deep learning. Surg Endosc Other Interv Tech 33(11):3790–3797. https://doi.org/10.1007/s00464-019-06677-2

32. Li S, Jiang H, Pang W (2017) Joint multiple fully connected convolutional neural network with extreme learning machine for hepatocellular carcinoma nuclei grading. Comput Biol Med 84:156–167. https://doi.org/10.1016/j.compbiomed.2017.03.017

33. Li YP, Deng LY, Yang XH, Liu Z, Zhao XP, Huang FR, Zhu SQ, Chen XD, Chen ZQ, Zhang WM (2019) Early diagnosis of gastric cancer based on deep learning combined with the spectral-spatial classification method. Biomed Opt Express 10(10):4999–5014. https://doi.org/10.1364/boe.10.004999

34. Li L, Chen Y, Shen Z, Zhang X, Sang J, Ding Y, Yang X, Li J, Chen M, Jin C, Chen C, Yu C (2020) Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. Gastric Cancer 23(1):126–132. https://doi.org/10.1007/s10120-019-00992-2

35. Liang QK, Nan Y, Coppola G, Zou KL, Sun W, Zhang D, Wang YN, Yu GZ (2019) Weakly supervised biomedical image segmentation by reiterative learning. IEEE J Biomed Health Inform 23(3):1205–1214. https://doi.org/10.1109/jbhi.2018.2850040

36. Ling TS, Wu LL, Fu YW, Xu QW, An P, Zhang J, Hu S, Chen YY, He XQ, Wang J, Chen X, Zhou J, Xu YM, Zou XP, Yu HG (2020) A deep learning-based system for identifying differentiation status and delineating the margins of early gastric cancer in magnifying narrow-band imaging endoscopy. Endoscopy :9. https://doi.org/10.1055/a-1229-0920

37. Liu X, Wang C, Bai J, Liao G (2020)Fine-tuning pre-trained convolutional neural networks for gastric precancerous disease classification on magnification narrow-band imaging images. Neurocomputing 392:253–267. https://doi.org/10.1016/j.neucom.2018.10.100

38. Lui TKL, Wong KKY, Mak LLY, To EWP, Tsui VWM, Deng Z, Guo J, Ni L, Cheung MKS, Leung WK (2020) Feedback from artificial intelligence improved the learning of junior endoscopists on histology prediction of gastric lesions. Endosc Int Open 8(2):E139-e146. https://doi.org/10.1055/a-1036-6114

39. Lui TKL, Tsui VWM, Leung WK (2020) Accuracy of artificial intelligence-assisted detection of upper GI lesions: a systematic review and meta-analysis. Gastrointest Endosc 92(4):821-830e829. https://doi.org/10.1016/j.gie.2020.06.034

40. Luo H, Xu G, Li C, He L, Luo L, Wang Z, Jing B, Deng Y, Jin Y, Li Y, Li B, Tan W, He C, Seeruttun SR, Wu Q, Huang J, Huang DW, Chen B, Lin SB, Chen QM, Yuan CM, Chen HX, Pu HY, Zhou F, He Y, Xu RH (2019)Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. Lancet Oncol 20(12):1645–1654. https://doi.org/10.1016/S1470-2045(19)30637-0

41. Menon S, Trudgill N (2014) How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis. Endosc Int Open 2(2):E46-50. https://doi.org/10.1055/s-0034-1365524

42. Min J, Kwak M, Cha J (2019) Overview of deep learning in gastrointestinal endoscopy. Gut Liver 13(4):388–393. https://doi.org/10.5009/gnl18384

43. Mishra R, Daescu O, Leavey P, Rakheja D, Sengupta A (2018) Convolutional neural network for histopathological analysis of osteosarcoma. J Comput Biol 25(3):313–325. https://doi.org/10.1089/cmb.2017.0153

44. Miyaki R, Yoshida S, Tanaka S, Kominami Y, Sanomura Y, Matsuo T, Oka S, Raytchev B, Tamaki T, Koide T, Kaneda K, Yoshihara M, Chayama K (2015) A computer system to be used with laser-based endoscopy for quantitative diagnosis of early gastric cancer. J Clin Gastroenterol 49(2):108–115. https://doi.org/10.1097/mcg.0000000000000104

45. Moon WK, Lee YW, Ke HH, Lee SH, Huang CS, Chang RF (2020)Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. Comput Methods Programs Biomed 190:105361. https://doi.org/10.1016/j.cmpb.2020.105361

46. Nagao S, Tsuji Y, Sakaguchi Y, Takahashi Y, Minatsuki C, Niimi K, Yamashita H, Yamamichi N, Seto Y, Tada T, Koike K (2020) 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 92(4):866. https://doi.org/10.1016/j.gie.2020.06.047

47. Pham HHN, Futakuchi M, Bychkov A, Furukawa T, Kuroda K, Fukuoka J (2019) Detection of lung cancer lymph node metastases from whole-slide histopathologic images using a two-step deep learning approach. Am J Pathol 189(12):2428–2439. https://doi.org/10.1016/j.ajpath.2019.08.014

48. Qu J, Hiruta N, Terai K, Nosato H, Murakawa M, Sakanashi H (2018) Gastric pathology image classification using stepwise fine-tuning for deep neural networks. J Healthc Eng 2018:13. https://doi.org/10.1155/2018/8961781

49. Rantalainen M, Klevebring D, Lindberg J, Ivansson E, Rosin G, Kis L, Celebioglu F, Fredriksson I, Czene K, Frisell J, Hartman J, Bergh J, Grönberg H (2016)Sequencing-based breast cancer diagnostics as an alternative to routine biomarkers. Sci Rep 6:38037. https://doi.org/10.1038/srep38037

50. Samek W, Wiegand T, Müller KR (2017) Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models

51. Shibata T, Teramoto A, Yamada H, Ohmiya N, Saito K, Fujita H (2020) Automated detection and segmentation of early gastric cancer from endoscopic images using mask R-CNN. Appl Sci 10(11). https://doi.org/10.3390/app10113842

52. Shinozaki S, Osawa H, Hayashi Y, Lefor AK, Yamamoto H (2019) Linked color imaging for the detection of early gastrointestinal neoplasms. 12:1756284819885246. https://doi.org/10.1177/1756284819885246

53. Simonyan K, Zisserman A (2015) Going deeper with convolutions. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) abs/1409.1556, 1–9

54. Singh H, Meyer A, Thomas E (2014) The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. BMJ Qual Saf 23(9):727–731. https://doi.org/10.1136/bmjqs-2013-002627

55. Singh H, Schiff G, Graber M, Onakpoya I, Thompson M (2017) The global burden of diagnostic errors in primary care. BMJ Qual Saf 26(6):484–494. https://doi.org/10.1136/bmjqs-2016-005401

56. Song ZG, Zou SM, Zhou WX, Huang Y, Shao LW, Yuan J, Gou XN, Jin W, Wang ZB, Chen X, Ding XH, Liu JH, Yu CK, Ku C, Liu CC, Sun Z, Xu G, Wang YF, Zhang XQ, Wang DD, Wang SH, Xu W, Davis RC, Shi HY (2020) Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. Nat Commun 11(1):9. https://doi.org/10.1038/s41467-020-18147-8

57. Sumiyama K (2017) Past and current trends in endoscopic diagnosis for early stage gastric cancer in Japan. Gastric Cancer 20(Suppl 1):20–27. https://doi.org/10.1007/s10120-016-0659-4

58. Sun MY, Zhang GH, Dang H, Qi XQ, Zhou XG, Chang Q (2019) Accurate gastric cancer segmentation in digital pathology images using deformable convolution and multi-scale embedding networks. IEEE Access 7:75530–75541. https://doi.org/10.1109/access.2019.2918800

59. Suzuki H, Oda I, Abe S, Sekiguchi M, Mori G, Nonaka S, Yoshinaga S, Saito Y (2016) High rate of 5-year survival among patients with early gastric cancer undergoing curative endoscopic submucosal dissection. Gastric Cancer 19(1):198–205. https://doi.org/10.1007/s10120-015-0469-0

60. Szegedy C, Wei L, Jia Y, Sermanet P, Rabinovich A (2015) Going deeper with convolutions. 2015 IEEEConference on Computer Vision and Pattern Recognition (CVPR), 1–9
61. Toyoizumi H, Kaise M, Arakawa H, Yonezawa J, Yoshida Y, Kato M, Yoshimura N, Goda K, Tajiri H (2009) Ultrathin endoscopy versus high-resolution endoscopy for diagnosing superficial gastric neoplasia. Gastrointest Endosc 70(2):240–245. https://doi.org/10.1016/j.gie.2008.10.064
62. Ueyama H, Kato Y, Akazawa Y, Yatagai N, Komori H, Takeda T, Matsumoto K, Ueda K, Matsumoto K, Hojo M, Yao T, Nagahara A, Tada T (2020) Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. J Gastroenterol Hepatol (Australia). https://doi.org/10.1111/jgh.15190
63. Wang S, Zhu Y, Yu L, Chen H, Lin H, Wan X, Fan X, Heng PA (2019) RMDL: Recalibrated multi-instance deep learning for whole slide gastric image classification. Med Image Anal 58:101549. https://doi.org/10.1016/j.media.2019.101549
64. Wang H, Ding S, Wu D, Zhang Y, Yang S (2019) Smart connected electronic gastroscope system for gastric cancer screening using multi-column convolutional neural networks. Int J Prod Res 57(21):6795–6806. https://doi.org/10.1080/00207543.2018.1464232
65. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155(8):529–536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009
66. Wu L, Zhou W, Wan X, Zhang J, Shen L, Hu S, Ding Q, Mu G, Yin A, Huang X, Liu J, Jiang X, Wang Z, Deng Y, Liu M, Lin R, Ling T, Li P, Wu Q, Jin P, Chen J, Yu H (2019) A deep neural network improves endoscopic detection of early gastric cancer without blind spots. Endoscopy 51(6):522–531. https://doi.org/10.1055/a-0855-3532
67. Wu X, Hui H, Niu M, Li L, Wang L, He B, Yang X, Li L, Li H, Tian J, Zha Y (2020) Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: A multicentre study. Eur J Radiol 128:109041. https://doi.org/10.1016/j.ejrad.2020.109041
68. Xu Y, Jia Z, Wang L, F Z YA (2017) Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. BMC Bioinform 18(1):281. https://doi.org/10.1186/s12859-017-1685-x
69. Yang YJ, Bang CS (2019) Application of artificial intelligence in gastroenterology. World J Gastroenterol 25(14):1666–1683. https://doi.org/10.3748/wjg.v25.i14.1666
70. Yasaka K, Akai H (2018) Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary Study. 286(3):887–896. https://doi.org/10.1148/radiol.2017170706
71. Yoon HJ, Kim S, Kim JH, Keum JS, Oh SI, Jo J, Chun J, Youn YH, Park H, Kwon IG, Choi SH, Noh SH (2019) A lesion-based convolutional neural network improves endoscopic detection and depth prediction of early gastric cancer. J Clin Med 8(9):10. https://doi.org/10.3390/jcm8091310
72. Zhao X, Liu L, Qi S (2018) Agile convolutional neural network for pulmonary nodule classification using CT images. 13(4):585–595. https://doi.org/10.1007/s11548-017-1696-0
73. Zhu LF, Ma P, Hu YL, Feng Y, Li P, Wang H, Guo YB, Mao QS, Xue WJ (2019)HCCR-1 is a Novel Prognostic Indicator for Gastric Cancer and Promotes Cell Proliferation. J Cancer 10(15):3533–3542. https://doi.org/10.7150/jca.22462
74. Zhu Y, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, Zhang YQ, Chen WF, Yao LQ, Zhou PH, Li QL (2019) Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. Gastrointest Endosc 89(4):806. https://doi.org/10.1016/j.gie.2018.11.011

Affiliations

Yuxue Zhao[1] ⬤ · Bo Hu[2] ⬤ · Ying Wang[1] · Xiaomeng Yin[3] · Yuanyuan Jiang[4] · Xiuli Zhu[1] ⬤

1   School of Nursing, Department of Medicine, Qingdao University, No. 15, Ningde Road, Shinan District, Qingdao 266073, China

2   Department of Thoracic Surgery, Qingdao Municipal Hospital, Qingdao, China