

87250_Auto_Edited-check.docx

Name of Journal: *World Journal of Gastrointestinal Oncology*

Manuscript NO: 87250

Manuscript Type: META-ANALYSIS

Application of convolutional neural network-based endoscopic imaging in the diagnosis, treatment, and prognosis of patients with esophageal cancer or high-grade dysplasia: A systematic review with meta-analysis of the diagnostic accuracy and individual data

Zhang JQ *et al.* Diagnostic accuracy of CNN: A meta-analysis

Abstract

BACKGROUND

Esophageal cancer is the seventh-most common cancer type worldwide, accounting for 5% of death from malignancy. Development of novel diagnostic techniques has facilitated screening, early detection, and improved prognosis. Convolutional neural network (CNN)-based image analysis promises great potential for diagnosing and determining the prognosis of esophageal cancer, enabling even early detection of dysplasia.

AIM

To conduct a meta-analysis of the diagnostic accuracy of CNN models for the diagnosis of esophageal cancer and high-grade dysplasia (HGD).

METHODS

PubMed, EMBASE, Web of Science and Cochrane Library databases were searched for articles published up to November 30, 2022. We evaluated the diagnostic accuracy of using the CNN model with still image-based analysis and with video-based analysis for esophageal cancer or HGD, as well as for the invasion depth of esophageal cancer. The pooled sensitivity, pooled specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR) and area under the curve (AUC) were estimated, together with the 95% confidence intervals (CI). A bivariate method and hierarchical summary receiver operating characteristic method were used to calculate the diagnostic test accuracy of the CNN model. Meta-regression and subgroup analyses were used to identify sources of heterogeneity.

RESULTS

A total of 28 studies were included in this systematic review and meta-analysis. Using still image-based analysis for the diagnosis of esophageal cancer or HGD provided a pooled sensitivity of 0.95 (95%CI: 0.92-0.97), pooled specificity of 0.92 (0.89-0.94), PLR of

11.5 (8.3-16.0), NLR of 0.06 (0.04-0.09), DOR of 205 (115-365), and AUC of 0.98 (0.96-0.99). When video-based analysis was used, a pooled sensitivity of 0.85 (0.77-0.91), pooled specificity of 0.73 (0.59-0.83), PLR of 3.1 (1.9-5.0), NLR of 0.20 (0.12-0.34), DOR of 15 (6-38) and AUC of 0.87 (0.84-0.90) were found. Prediction of invasion depth resulted in a pooled sensitivity of 0.90 (0.87-0.92), pooled specificity of 0.83 (95%CI: 0.76-0.88), PLR of 7.8 (1.9-32.0), NLR of 0.10 (0.41-0.25), DOR of 118 (11-1305), and AUC of 0.95 (0.92-0.96).

CONCLUSION

CNN-based image analysis in diagnosing esophageal cancer and HGD is an excellent diagnostic method with high sensitivity and specificity that merits further investigation in large, multicenter clinical trials.

Key Words: Esophageal cancer; High-grade dysplasia; Convolutional neural network; Deep learning; Systematic review; Meta-analysis

Zhang JQ, Mi JJ, Wang R. Application of convolutional neural network-based endoscopic imaging in the diagnosis, treatment, and prognosis of patients with esophageal cancer or high-grade dysplasia: A systematic review with meta-analysis of the diagnostic accuracy and individual data. *World J Gastrointest Oncol* 2023; In press

Core Tip: This systematic review provides a meta-analysis of 28 studies evaluating the accuracy of convolutional neural network (CNN) models for diagnosing esophageal cancer and high-grade dysplasia, and for predicting the invasion depth of esophageal cancer. It also establishes a theoretical foundation for the clinical application of CNN models. Based on this meta-analysis, CNN-based image analysis may have great potential for diagnosing and estimating the prognosis of esophageal cancer, though further study is needed.

INTRODUCTION

In global data reported by the International Agency for Research on Cancer, esophageal cancer was the seventh-most common malignancy in incidence and sixth in mortality worldwide in 2020^[1]. Esophageal cancer has two main histological subtypes: Esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC)^[2]. ESCC is more common in Asian countries, accounting for approximately 87% of all cases of esophageal cancer, whereas EAC is more common in western countries and has been increasing in incidence recently^[3-5]. Diverse grades of dysplasia, especially high-grade dysplasia (HGD), are precancerous lesions known to progress to esophageal cancer^[3,6]. The majority of patient with esophageal cancer are diagnosed with advanced disease due to the lack of symptoms at earlier stages, resulting in a five-year survival rate of less than 20%^[7,8]. When diagnosed and treated early, however, the five-year survival rate can increase to more than 85%^[9,10]. Moreover, the choice of treatment modalities and the prognosis of esophageal cancer patients depend heavily on the predicted invasion depth^[11,12].

Traditional endoscopy is frequently used to detect esophageal cancer and to estimate its invasion depth. However, detection of the minor changes in the surrounding mucosa of early esophageal cancer using only white light imaging (WLI) endoscopy remains challenging^[13,14]. Although iodine staining provides greater accuracy in detection, it is employed infrequently during screening, as it causes discomfort, and allergies to iodine are not infrequent^[15,16]. Emerging endoscopic techniques such as narrow band imaging (NBI), blue-laser imaging (BLI), and post-processing imaging techniques such as i-scan and flexible spectral imaging color enhancement have greatly increased the rate of esophageal cancer detection, as has endocytoscopy, which is a novel endoscopic system that provides high-quality assessment of lesions *in vivo*, but these depend upon specialized training and experience for the endoscopist^[14,17,18]. Additionally, esophageal lesions frequently have irregular shapes and indistinct borders, resulting in variable performance even by experts due to the pressure to complete procedures quickly,

limiting the time available for diagnosis and the degree of confidence in the interpretation^[19].

Artificial intelligence (AI) in the medical industry is being utilized at an ever-increasing rate thanks to advances in deep learning (DL), one of its core branches^[20-22]. Convolutional neural network (CNN) is a DL model inspired by the biological mechanism of object perception in the animal brain^[23] with unique self-learning abilities that encode complex signals. After the original image is entered into the CNN model, the convolution layer automatically recognizes the color, texture, detailed features, and global features of the image according to the settings defined by the investigators. It then completes various diagnostic visual tasks, such as recognition of diabetic retinopathy or skin cancer^[24,25]. CNN can also assist in the diagnosis of gastrointestinal diseases by preserving the spatial relationship characteristics of endoscopic images (including the detection of colorectal polyps), *Helicobacter pylori* infection, and gastrointestinal cancer^[20,26-28]. Given the recent increasing use of endoscopy, CNN has been used extensively to diagnose esophageal cancer and premalignant lesions, as well as in predicting the invasion depth of esophageal cancer^[29-31].

For this systematic review we conducted a meta-analysis of the diagnostic test accuracy (DTA) achieved by the CNN model in diagnosing esophageal cancer and HGD, as well as its ability to predict the invasion depth of esophageal cancer.

MATERIALS AND METHODS

Literature search

This systemic review and meta-analysis were conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-analyses guidelines. Two investigators independently searched the PubMed, EMBASE, Web of Science, and Cochrane Library databases for all studies published before November, 2022 that used the CNN model to detect esophageal cancer and HGD. The following terms were used for the search: ("convolutional neural network" OR "convolutional neural networks" OR "computer-aided" OR "computer aided" OR "artificial intelligence" OR "machine

learning" OR "deep learning" OR "hierarchical learning" OR "computational intelligence" OR "machine intelligence" OR "computer reasoning") AND ("esophageal neoplasms" OR "esophageal neoplasm" OR "esophagus neoplasm" OR "esophagus neoplasms" OR "esophagus cancer" OR "esophagus cancers" OR "esophageal cancer" OR "esophageal cancers" OR "oesophagus cancer" OR "oesophageal cancers" OR "oesophagus neoplasm" OR "oesophageal neoplasms" OR "esophageal squamous cell carcinoma" OR "adenocarcinoma of esophagus" OR "oesophageal squamous cell carcinoma" OR "esophageal adenocarcinoma" OR "oesophageal adenocarcinoma" OR "Barrett's esophagus" OR "Barrett's oesophagus"). Only English-language articles were included. The author screened all articles and emailed the research author to obtain missing data or study material before excluding any relevant articles from the analysis. Repetitive studies, reviews, and meta-analyses, as well as non-relevant studies (as determined by reading the title, abstract and full text), were excluded from this meta-analysis. Studies with insufficient information or that did not meet the inclusion criteria were excluded. Two authors discussed any differences, and sought advice from a third author to reconcile any differences.

Study selection

The inclusion criteria were: (1) Analysis of a CNN model utilizing still images or video to diagnose esophageal cancer or HGD; (2) Analysis of a CNN model for predicting the invasion depth of esophageal cancer or HGD, or for identifying intrapapillary capillary loops (IPCLs) of esophageal cancer and HGD; (3) Prospective or retrospective studies; (4) Cases of histologically-proven esophageal cancer and HGD; and (5) Studies published in English. The exclusion criteria were: (1) Reviews or meta-analyses; (2) Proceedings, letters or comments; (3) Experimental studies; (4) Animal studies; or (5) Studies with incomplete data.

Data extraction

Two authors independently extracted information from the identified reports, and resolved disagreements through extensive discussion to reach a consensus. The authors extracted the following information from each eligible study: First author, publication year, continent, scale (single center or multicenter), external validation (yes/no), study format, case type (image or patient), real-time (yes/no), histological type, image type, quality (see below), number of patients or endoscopic images, and algorithms for CNN models. The rates of true positivity, false positivity, false negativity, and true negativity for the CNN models and endoscopists in diagnosing esophageal cancer and HGD were also extracted, together with the prediction of invasion depth of esophageal cancer or HGD, or identification of IPCLs of esophageal cancer or HGD.

Quality assessment

The Quality for Assessment of Diagnostic Studies (QUADAS) score was used to determine the quality of the included studies. This score was assessed in four parts, comprising patient selection, index test, reference standard, and flow and timing, with the first three parts utilized for applicability assessment^[32]. Each part was graded by two authors as having a high, low, or unclear risk of bias.

Outcome measures

The primary outcomes determined were the pooled diagnostic accuracy, pooled sensitivity, pooled specificity, positive likelihood ratio (PLR) and negative likelihood ratio (NLR) of the CNN models for diagnosing esophageal cancer and HGD, for predicting the invasion depth of esophageal cancer and HGD, and for identifying IPCLs in esophageal cancer or HGD. The area under the curve (AUC) was used to measure the accuracy of the CNN. The secondary outcome was the performance of endoscopists compared with that of CNN models for detection of esophageal cancer or HGD using the same still images and videos.

Statistical analyses

The main statistical data processing for this DTA meta-analysis used the bivariate method and the hierarchical summary receiver operating characteristic (HSROC) method to calculate the pooled sensitivity, pooled specificity, PLR, NLR, diagnostic odds ratio (DOR) and area under the receiver operating characteristic curve (AUROC) of CNN models and of endoscopists to detect esophageal cancer or HGD. This approach also considers the correlation between specificity and sensitivity. Heterogeneity was analyzed using the HSROC method to determine the correlation coefficient between logit-transformed sensitivity and specificity and the asymmetry parameter β . A β value of 0 serves as the standard for evaluating the symmetry of the ROC, with the HSROC curve inspected visually for signs of heterogeneity. Regression and subgroup analyses were used to determine the source of heterogeneity. The 95% confidence interval (CI) of AUROC was calculated and compared within each subgroup. A statistically significant difference between two subgroups was indicated by a non-overlapping 95%CI of the AUROC. STATA software version 15.1 (College Station, Texas, United States) with the installed packages MIDAS and METANDI was used to perform the main statistical analysis. Meta-DiSc 1.4 (XI Cochrane Colloquium, Barcelona, Spain) was used for the subgroup analysis of data with a small sample size. The figures for methodological quality assessment and the HSROC curve for small sample size data were drawn using RevMan 5.3 (The Nordic Cochrane Centre, Copenhagen, Denmark). Publication bias was analyzed using Deeks' test^[33]. A P value < 0.05 was considered statistically significant. The statistical methods of the study were reviewed by professor Ming-Cheng Li from Beihua University.

RESULTS

Literature search and screening results

A total of 2045 studies were identified initially using the screening search strategy described. Of these, 655 were excluded because they were duplicate studies, 133 because they were meta-analyses or reviews, and 1205 because they were deemed irrelevant based on their titles and abstracts. The remaining 52 studies were examined

in full and 24 were rejected because they contained insufficient data, or were comments and proceedings. Finally, the authors identified 28 studies that met the inclusion and exclusion criteria for this systematic review and meta-analysis^[17,20,29,31,34-57]. The flowchart for the search procedure is shown in Figure 1.

Quality assessment of the included literature

The QUADAS-2 tool was used to assess the quality and bias risk of the included studies. Most studies were rated as having low bias risk in all parts. Of the 28 studies, only three failed to indicate whether the selected patients were continuous or chosen at random^[41,44,45], hence their risk for patient selection was not clear. The studies included were of excellent quality, as shown in Figures 2A and B.

Meta-analysis of CNN for the diagnosis of esophageal cancer or HGD based on still images

A total of 19 still image-based studies assessed the diagnostic value of CNN for esophageal cancer or HGD^[17,20,34-50], providing extractable data from a total of 20867 images. Twelve studies assessed images from Asian populations^[17,20,34,37-39,45-50], five from European and American populations^[35,41-44], and two from multiregional populations^[36,40]. Thirteen studies provided CNN data using WLI^[17,20,34,36,39,41-44,46-49] and 10 applied advanced imaging technologies, including NBI, BLI, etc.^[17,20,35,37,38,40,42,45,48,50]. The histological types examined were sorted into three categories: Esophageal cancer (including ESCC and EAC)^[17,20,34-38,40,41,46,47], Barrett's neoplasia (including HGD and EAC)^[39,42-45,49,50], and esophageal squamous cell neoplasia (including HGD and ESCC)^[48]. One study diagnosed esophageal cancer IPCLs by CNN based on still images^[38]. This study was also included in the meta-analysis. For the 19 still image-based studies describing the diagnosis of esophageal cancer or HGD, the pooled sensitivity was 0.95 (95%CI: 0.92-0.97), pooled specificity was 0.92 (0.89-0.94), PLR was 11.5 (8.3-16.0), NLR was 0.06 (0.04-0.09), DOR was 205 (115-365), and AUC was 0.98 (0.96-0.99) (Table 1). The studies included in this analysis showed heterogeneity ($P =$

0.000, $I^2 = 0.98$). However, the HSROC shape was symmetric, and the following results negated the impact of the threshold effect: A correlation coefficient between logit-transformed sensitivity and specificity of $r = -0.225$, and an asymmetric β parameter with a nonsignificant P -value of 0.431 (Figure 3A).

A coupled forest plot of sensitivity and specificity is shown in Figure 3B. Meta-regression analysis of these data revealed that histological type was the only significant source of heterogeneity ($P = 0.01$) when the publication year ($P = 0.26$), continent ($P = 0.65$), scale ($P = 0.61$), external validation ($P = 0.94$), study type ($P = 0.84$), case type ($P = 0.10$), real-time ($P = 0.90$), image type ($P = 0.07$), quality ($P = 0.10$) and number of cases ($P = 0.22$) were included in the analysis. The data were also subjected to subgroup analysis (Table 2 and Figure 3C). Three still image-based studies compared the diagnostic performance of endoscopists with that of CNN models^[20,34,50]. The CNN models showed higher sensitivity than did endoscopists [0.96 (95%CI: 0.92-0.98) vs 0.87 (95%CI: 0.81-0.91)] (Figure 4A-D). Using the same still image dataset, the diagnostic performance of the CNN models was marginally better than that of endoscopists, as shown by the plot of the HSROC curve (Figure 5).

Meta-analysis of CNN models for the diagnosis of esophageal cancer or HGD based on videos

Eight video-based studies reported the diagnostic value of CNN for esophageal cancer or HGD^[43,47,51-56], evaluating a total of 1262 videos. Six studies provided images from Asian populations^[47,51,53-56] and two from European populations^[43,52]. Five studies provided data using WLI^[43,47,53-55] and five using advanced imaging technology including NBI, BLI, etc.^[51-54,56]. The histological types analyzed in these studies included ESCC^[47,51,53-56] and Barrett's neoplasia (including HGD and EAC)^[43,52]. Results from these eight video-based studies showed a pooled sensitivity of 0.85 (95%CI: 0.77-0.91), pooled specificity of 0.73 (0.59-0.83), PLR of 3.1 (1.9-5.0), NLR of 0.20 (0.12-0.34), DOR of 15 (6-38), and AUC of 0.87 (0.84-0.90) (Table 3).

The studies included in the video-based analysis exhibited heterogeneity ($P = 0.000$, $I^2 = 0.93$). Furthermore, the HSROC curve shape was symmetric (Figure 6A), the correlation coefficient between logit-transformed sensitivity and specificity was observed to be $r = 0.277$, and an asymmetric β parameter with a nonsignificant P -value (0.630) was obtained. Thus, the observed heterogeneity was not due to the threshold effect. Coupled forest plots for sensitivity and specificity are shown in Figure 6B. Meta-regression and subgroup analyses revealed no obvious sources of heterogeneity (Table 4 and Figure 6C) from the publication year ($P = 0.86$), continent ($P = 0.73$), scale ($P = 0.55$), histological type ($P = 0.73$), external validation ($P = 0.94$), study type ($P = 0.89$), real-time ($P = 0.13$), image type ($P = 0.76$), or number of cases ($P = 0.76$). The type of case and quality were consistent among the eight studies, so no relevant meta-regression analysis was indicated. Because only one video-based study compared the diagnostic performance of the CNN model to that of endoscopists^[53], no data analysis was performed for the endoscopists.

Meta-analysis of CNN for predicting the invasion depth of esophageal cancer

Three studies used the CNN model to predict the invasion depth of esophageal cancer and gave precise data^[29,31,57]. One differentiated between pathological intraepithelial (pEP)-submucosal microinvasive (SM1) (pEP-SM1) and pathological submucosal deep invasive (pSM2/3) cancers^[29], one reported the diagnostic performance of CNN for pEP-SM1 and pEP-muscularis mucosa cancer^[31], and one reported the diagnostic performance of CNN for pEP-SM1 cancer^[57]. The pooled sensitivity was 0.90 (95%CI: 0.87-0.92), pooled specificity was 0.83 (0.76-0.88), PLR was 7.8 (1.9-32.0), NLR was 0.10 (0.41-0.25), DOR was 117.76 (10.63-1304.7), and AUC was 0.95 (0.92-0.96) (Table 5). The HSROC curve and coupled forest plots of sensitivity and specificity are shown in Figures 7A-C, respectively. Two studies compared the diagnostic performance of endoscopists *vs* CNN models for predicting the invasion depth of esophageal cancer^[31,57]. However, because only one of these provided specific data, an analysis was not performed on the diagnostic performance of the endoscopists.

Evaluation of publication bias

Deeks' funnel plot of 19 still image-based studies showed ³ a symmetrical shape with respect to the regression line (Figure 8A). The asymmetric test revealed no significant publication bias ($P = 0.07$). Furthermore, Deeks' funnel plot of eight video-based studies also showed a symmetrical shape with respect to the regression line (Figure 8B), with no significant publication bias ($P = 0.55$).

DISCUSSION

Esophageal cancer is a malignant neoplasm with early, rapid metastasis and a poor prognosis, but endoscopy can provide early diagnosis and therapy^[58]. Endoscopists can find it challenging to accurately diagnose esophageal cancer and HGD when relying solely on their own skills, but AI may have clinical applicability to achieve greater accuracy^[59]. CNN is a branch of DL that uses a special learning method to develop image recognition capabilities through training datasets. Recently, CNN has been applied to the analysis of endoscopic images and videos, showing rapid progress and developing progressively into a crucial auxiliary tool for endoscopists^[60]. Additionally, CNN has been used to recognize the geometry of IPCLs to gauge the invasion depth of esophageal cancer, as well as to help medical professionals build treatment regimens^[61,62]. This systematic review and meta-analysis demonstrates that the CNN method can reliably identify esophageal cancer and HGD, providing great clinical applicability. The current meta-analysis found that CNN was effective at identifying esophageal cancer based on still image data, with ¹ values for pooled sensitivity, pooled specificity, PLR, NLR, DOR and AUC of 0.95, 0.92, 11.5, 0.06, 205, and 0.98, respectively. The still image dataset demonstrates the ability of CNN to identify uncertain lesions discovered during endoscopy, and CNN showed higher sensitivity than endoscopists. It might therefore reduce the rate of missed diagnosis of esophageal cancers and neoplasms and help endoscopists find lesions that are easily overlooked.

The meta-analysis of video data revealed that the CNN model performs exceptionally well for the diagnosis of esophageal cancer and HGD. The pooled sensitivity, pooled specificity, PLR, NLR, DOR and AUC were 0.85, 0.73, 3.1, 0.20, 15, and 0.87, respectively. Despite having good diagnostic performance, the meta-analysis results for the CNN model based on video data showed that CNN was slightly less accurate when used on video images than on static images as the dataset. This reduced accuracy may arise because the performance of CNN on video images is influenced by a variety of factors, including poor insufflation, bleeding, blurring, focus, angle, surgical procedure, patient participation, and image quality. However, video more accurately mimics the endoscopic procedure performed by the endoscopist, which serves as a valuable benchmark for the operational performance of CNN. Esophageal lesions can be overlooked due to the endoscope passing through the esophagus too quickly, or because of insufficient expertise by the endoscopist. CNN can assist endoscopists to correctly identify and further diagnose esophageal lesions. Refinement and expansion of the training dataset should improve CNN performance in the identification of video-based lesions^[37,38].

The robustness of the diagnostic performance of the CNN model can be seen in the subgroup analysis, in that no appreciable differences in its performance were observed across different subgroups. Moreover, the diagnostic efficacy of the CNN model did not differ significantly according to continent, histology, or case type. Thus, we conclude that CNN based on still images can be applied to a wide range of gastrointestinal diseases and endoscopic functions^[63,64]. Importantly, CNN models based on WLI and other advanced imaging modalities show similarly excellent diagnostic performance. Advanced imaging modes such as NBI and BLI can improve detection of the surface structure and microvascular morphology of lesions, which is one of the standard ways to diagnose esophageal cancer. It is worth noting that the more advanced endocytoscopy can recognize the histological structure of the pre-cancer epithelium with the help of intraprocedural coloration, so called “virtual histology”^[65]. Application of the CNN model to these methods may compensate for interobserver variability^[20,66].

Advances in real-time diagnostic capabilities have also increased the importance of CNN in clinical practice. CNN requires a recognition speed of at least 25 frames per second, while current methods can frequently achieve more than 30 to 60 frames/s without a latency period^[52,54,55]. The identification speed of CNN may therefore reduce the time needed for diagnosis and increase the speed of endoscopic procedures. Determining diagnosis and selecting the appropriate treatment strategy depend upon accurate endoscopic prediction of the invasion depth of esophageal cancer. Endoscopic resection should be the treatment of choice in esophageal lesions that affect only the EP-SM1 because there is a low chance of lymph node metastases for this extent of disease. Lesions that invade SM2-SM should be removed surgically or with chemoradiotherapy due to the increased risk of lymph node metastasis^[11,67,68]. Based on this meta-analysis, the CNN model is ideally suited for predicting the invasion depth of esophageal cancer, with a pooled sensitivity of 0.90 (95%CI: 0.88-0.93), pooled specificity of 0.83 (0.76-0.88), and AUC of 0.95 (0.92-0.96). Two prior studies compared the diagnostic performance of endoscopists to that of CNN models for predicting the invasion depth of esophageal cancer. Tokai *et al*^[57] concluded the CNN model was more accurate than were endoscopists, while Nakagawa *et al*^[31] reported that CNN performed similarly to experienced endoscopists. Morphological changes in IPCLs, which are microvascular structures on the surface of esophageal cancer, are closely associated with the invasion depth of the tumor. Only one of the studies examined reported the diagnostic performance of CNN for identifying IPCLs in esophageal cancer^[38]. Using a CNN model based on still images, this study found a mean diagnostic precision of 89.2% at the lesion level and 93.0% at the pixel level.

The present DTA meta-analysis has demonstrated the powerful detection efficiency of the CNN model for esophageal cancers and neoplasms. This analysis has several limitations that should be considered. First, the included studies did not contain sufficient information to allow evaluation of the overall diagnostic accuracy of endoscopists or endocytoscopy. Second, despite a recent increase in the number of CNN studies that predict the invasion depth of esophageal cancer, only three studies

met our inclusion and exclusion criteria to be included in this meta-analysis. Accurate prediction of the invasion depth, which is the foundation for early diagnosis and treatment, is vital for the further development of CNN for this disease. Third, there was insufficient data to allow comparison of the diagnostic abilities of endoscopic physicians with CNN models. Although the majority of current studies reported high diagnostic accuracy for the CNN model, some aspects, such as the use of video datasets and prediction of the invasion depth, require additional supporting evidence. Fourth, the CNN training procedure used in this meta-analysis has not been standardized, and the training dataset cannot be recorded or used in subgroup analysis. A large, multi-center cohort analysis is indicated to validate the use of CNN for esophageal cancer and HGD, and to compare its diagnostic ability with that of endoscopists. Follow-up studies that use the same video datasets are also needed.

CONCLUSION

In conclusion, the CNN model has excellent potential for accurately diagnosing esophageal cancers and HGD. It is anticipated to develop into an important diagnostic tool for endoscopists, showing promise for predicting the invasion depth of esophageal cancer.

ARTICLE HIGHLIGHTS

Research background

The development of convolutional neural network (CNN) model as a novel diagnostic technology has promoted the screening, early detection, and improved prognosis of esophageal cancer and high-grade dysplasia (HGD).

Research motivation

Explore the diagnostic value of CNN model for esophageal cancer and HGD, and provide basis for its clinical application.

Research objectives

Conduct a meta-analysis of the diagnostic accuracy of CNN models for the diagnosis of esophageal cancer and HGD.

Research methods

We searched for relevant studies in various search engines, evaluated the diagnostic accuracy of CNN models, and calculated the diagnostic test accuracy with a bivariate method and hierarchical summary receiver operating characteristic method. Meta-regression and subgroup analyses were used to identify sources of heterogeneity.

Research results

After processing 28 items of still image-based and video-based analysis in statistics, CNN models have been proven to have high accuracy and diagnostic efficiency in diagnosing esophageal cancer or HGD and predicting the invasion depth of esophageal cancer.

Research conclusions

CNN-based image analysis in diagnosing esophageal cancer and HGD is an excellent diagnostic method with high sensitivity and specificity.

Research perspectives

A thorough evaluation of the accuracy of diagnosis in esophageal cancer and HGD requires further investigation. Large-scale trials are needed to assess performance and predict clinical values.

5%

SIMILARITY INDEX

PRIMARY SOURCES

- | | | |
|----------|---|-----------------|
| 1 | www.frontiersin.org
<small>Internet</small> | 50 words — 1% |
| <hr/> | | |
| 2 | www.omicsdi.org
<small>Internet</small> | 30 words — 1% |
| <hr/> | | |
| 3 | www.jmir.org
<small>Internet</small> | 21 words — < 1% |
| <hr/> | | |
| 4 | Chang Seok Bang, Jae Jun Lee, Gwang Ho Baik.
"Computer-aided diagnosis of esophageal cancer
and neoplasms in endoscopic images: a systematic review and
meta-analysis of diagnostic test accuracy", <i>Gastrointestinal
Endoscopy</i> , 2020
<small>Crossref</small> | 20 words — < 1% |
| <hr/> | | |
| 5 | Si Min Zhang, Yong Jun Wang, Shu Tian Zhang.
"Accuracy of artificial intelligence-assisted
detection of esophageal cancer and neoplasms on endoscopic
images: A systematic review and meta-analysis", <i>Journal of
Digestive Diseases</i> , 2021
<small>Crossref</small> | 20 words — < 1% |
| <hr/> | | |
| 6 | Yoshikazu Yoshifuku, Yoji Sanomura, Shiro Oka,
Mio Kurihara et al. "Evaluation of the visibility of
early gastric cancer using linked color imaging and blue laser
imaging", <i>BMC Gastroenterology</i> , 2017
<small>Crossref</small> | 20 words — < 1% |

7	spandidos-publications.com	Internet	18 words — < 1%
8	www.researchgate.net	Internet	17 words — < 1%
9	www.wjgnet.com	Internet	16 words — < 1%
10	bsdwebstorage.blob.core.windows.net	Internet	15 words — < 1%
11	journals.sbmj.ac.ir	Internet	15 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 15 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 15 WORDS