Accepted Manuscript

Accurate Classification of Diminutive Colorectal Polyps Using Computer-aided Analysis

Peng-Jen Chen, Meng-Chiung Lin, Mei-Ju Lai, Jung-Chun Lin, Henry Horng-Shing Lu, Vincent S. Tseng

PII: S0016-5085(17)36251-0
DOI: 10.1053/j.gastro.2017.10.010
Reference: YGAST 61484

To appear in: *Gastroenterology*
Accepted Date: 9 October 2017

Please cite this article as: Chen P-J, Lin M-C, Lai M-J, Lin J-C, Lu HH-S, Tseng VS, Accurate Classification of Diminutive Colorectal Polyps Using Computer-aided Analysis, *Gastroenterology* (2017), doi: 10.1053/j.gastro.2017.10.010.

Accurate Classification of Diminutive Colorectal Polyps Using Computer-aided Analysis

Short Title: Deep Neural Network for Narrow-Band Imaging

Peng-Jen Chen,[1] Meng-Chiung Lin,[2,3] Mei-Ju Lai,[4] Jung-Chun Lin,[1] Henry Horng-Shing Lu,[5] Vincent S. Tseng[6]

[1]Division of Gastroenterology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan

[2]Department of Biological Science and Technology, National Chiao Tung University, Hsinchu, Taiwan

[3]Division of Gastroenterology, Taichung Armed Forces General Hospital, Taichung, Taiwan

[4]Department of Pathology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan

[5]Big Data Research Center and Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan

[6]Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan

Grant support: This research was partially supported by the Ministry of Science and Technology, Taiwan, under grant no. 101-2221-E-006-255-MY3.

Abbreviations used in this paper: CRC, colorectal cancer; DNN-CAD, computer-aided diagnosis with a deep neural network; NBI, narrow-band imaging; NICE, NBI International Colorectal Endoscopic classification; NPV, negative predictive value; PACS, picture archiving and communications system; PIVI, the Preservation and Incorporation of Valuable Endoscopic Innovations; PPV, positive predictive value; ROI, region of interest.

Corresponding Authors: Peng-Jen Chen, M.D. and Vincent S. Tseng, Ph.D.

Peng-Jen Chen, M.D.

Division of Gastroenterology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan

Address: No. 325, Sec. 2, Chenggong Rd., Neihu District, Taipei City 114, Taiwan

Email: endoscopy@ndmctsgh.edu.tw

Vincent S. Tseng, Ph.D.

Department of Computer Science, National Chiao Tung University, Taiwan

Address: No. 1001 University Road, Hsinchu 300, Taiwan

Email: vtseng@cs.nctu.edu.tw

Conflicts of interest: The authors disclose no conflicts.

Writing assistance: A native English speaker has edited the revised manuscript. The funding source for this assistance was the Ministry of Science and Technology, Taiwan, under grant no. 101-2221-E-006-255-MY3.

Author contributions:

Peng-Jen Chen, M.D.: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; statistical analysis; technical support; study supervision.

Meng-Chiung Lin, M.D.: acquisition of data; analysis and interpretation of data; drafting of the manuscript.

Mei-Ju Lai, M.D.: acquisition of data; analysis and interpretation of data; material support.

Jung-Chun Lin, M.D., Ph.D.: acquisition of data; analysis and interpretation of data.

Henry Horng-Shing Lu, Ph.D.: analysis and interpretation of data; critical revision of the manuscript for important intellectual content.

Vincent S. Tseng, Ph.D.: study concept and design; analysis and interpretation of data; critical revision of the manuscript for important intellectual content; statistical analysis; obtained funding; technical and material support; study supervision.

All authors had access to the study data and reviewed and approved the final manuscript.

Title: Accurate Classification of Diminutive Colorectal Polyps Using Computer-aided Analysis, With a Deep Neural Network, of Narrow-band Images

Abstract:

Background & Aims: Narrow-band imaging (NBI) is an image-enhanced form of endoscopy used to observed microstructures and capillaries of the mucosal epithelium that allows for real-time prediction of histologic features of colorectal polyps. However, NBI expertise is required to differentiate hyperplastic from neoplastic polyps with high levels of accuracy. We developed and tested a system of computer-aided diagnosis with a deep neural network (DNN-CAD) to analyze narrow-band images of diminutive colorectal polyps.

Methods: We collected 1476 images of neoplastic polyps and 681 images of hyperplastic polyps, obtained from the picture archiving and communications system database in a tertiary hospital in Taiwan. Histologic findings from the polyps were also collected and used as the reference standard. The images and data were used to train the DNN. A test set of images (96 hyperplastic and 188 neoplastic polyps, smaller than 5 mm), obtained from patients who underwent colonoscopies from March 2017 through August 2017, was then used to test the diagnostic ability of the DNN-CAD vs endoscopists (2 expert and 4 novice), who were asked to classify the images of the test set as neoplastic or hyperplastic. Their classifications were compared with findings from histologic analysis. The primary outcome measures were diagnostic accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and diagnostic time. The accuracy, sensitivity, specificity, PPV, NPV, and diagnostic time were compared among DNN-CAD, the novice endoscopists, and the expert endoscopists. The study was designed to detect a difference of 10% in accuracy by a 2-sided McNemar test.

Results: In the test set, the DNN-CAD identified neoplastic or hyperplastic polyps with 96.3% sensitivity, 78.1% specificity, a PPV of 89.6%, and a NPV of 91.5%. Fewer than half of the novice endoscopists classified polyps with a NPV of 90% (their NPVs ranged from 73.9% to 84.0%). DNN-CAD classified polyps as neoplastic or hyperplastic in 0.45±0.07 sec—shorter than the time required by experts (1.54±1.30 sec) and nonexperts (1.77±1.37 sec) (both $P<.001$). DNN-CAD classified polyps with perfect intra-observer agreement (kappa score of 1). There was a low level of intra-observer and inter-observer agreement in classification among endoscopists.

Conclusions: We developed a system called DNN-CAD to identify neoplastic or hyperplastic colorectal polyps less than 5 mm; it classified polyps with a PPV of 89.6%, and a NPV of 91.5%, and in a shorter time than endoscopists. This deep-learning model has

potential for not only endoscopic image recognition but for other forms of medical image analysis, including sonography, computed tomography, and magnetic resonance images.

KEY WORDS: colon cancer detection, machine learning, cost-effectiveness, magnifying

Background and Aims

Colorectal cancer (CRC) is one of the most common malignancies worldwide and is currently the third leading cause of cancer death in Taiwan.[1] Most CRCs arise from preexisting adenomas, and the adenoma–carcinoma sequence offers an opportunity for the screening and prevention of CRC.[2] Colonoscopy with adenoma resection can reduce the incidence of CRC by as much as 80% and the associated mortality by 50%.[3, 4] However, endoscopic resection of hyperplastic polyps exacerbates medical costs, including those for resection and unnecessary pathologic evaluation, because malignant transformations are rare. Therefore, accurate diagnosis before endoscopic resection is important to avoid inappropriate resection, and, further, optical diagnosis has the potential to improve the cost-effectiveness and efficiency of colonoscopy.[5]

Narrow-band imaging (NBI) is a type of equipment-based image-enhanced endoscopy that has been used to observe the microstructures and capillaries of the mucosal epithelium and allows real-time histologic predictions based on colorectal polyps.[6] The NBI International Colorectal Endoscopic (NICE) classification is a diagnostic criterion for hyperplastic and adenomatous polyps using NBI.[7] An optical diagnosis of NBI in clinical practice requires expertise to differentiate polyp histology with high accuracy. Recent studies have reported that the results of optical diagnosis in a nonacademic setting were disappointing.[8, 9]

To overcome this limitation, computer-aided diagnosis (CAD) has been developed.[10-12] Although these computer-aided systems have excellent potential as diagnostic aids in colonoscopic examination, access to the technology is limited because of institution-specific and localized software implementations. We developed and validated a novel system of computer-aided diagnosis with a deep neural network (DNN-CAD) to analyze magnifying NBI of diminutive colorectal polyps.

Materials and Methods

The study was approved by the Ethics Committee of Tri-Service General Hospital, Taiwan, and was conducted at the hospital's endoscopy center. All authors had access to the study data and reviewed and approved the final manuscript. Instruments used in this study included colonoscopes with an optical magnification function (CF-H260AZI, PCF-Q260AZI, CF-HQ290AZI; Olympus Optical Co., Ltd., Tokyo, Japan), and the EVIS LUCERAELITE Video System Center CV-290 (Olympus Medical Systems). Polyps detected by white-light colonoscopy were observed with NBI at the maximum magnification power.

Setup of the Image Classifier with a Deep Neural Network

The medical–engineering collaborative project between Tri-Service General Hospital and the Computer Science Department of National Chiao Tung University, initiated in 2016, resulted in marked improvements of our original software. The algorithm of the retraining model using TensorFlow,[13] comprises four steps: (i) image collection, (ii) model learning, (iii) creation of a classifier, and (iv) diagnostic output.

In 2016, two endoscopists with more than 10 years of colonoscopy experience reviewed colonoscopic images with NBI and full magnification from the picture archiving and communications system (PACS) database in Tri-Service General Hospital and, as the training set for TensorFlow, selected appropriate regions of interest (ROI) that contained high-quality images for visual inspection. Multiple ROIs in the same polyp were

collected to reduce the selection bias, and they were cropped from the endoscopic images captured from the video signal with a resolution of 1920×1080. These images were classified as hyperplastic or neoplastic polyps according to the NICE classification system (Figure 1). The histologic reports regarding these polyps were also collected. If the NICE classification of a polyp was not compatible with the histologic report, the histologic image was retrieved and reevaluated. Three gastrointestinal pathologists—two attending staff members (H.S.L., Y.J.P.) and one senior fellow (M.J.L.)—provided histologic assessments to determine the type of lesion. The final diagnosis was based on NBI images and adjusted by the histologic report. To set up the DNN-CAD, 1476 images of neoplastic polyps and 681 images of hyperplastic polyps were collected as the training set.

There were no statistically significant differences in age and sex between the hyperplastic group and neoplastic group in the NBI images. The training set excluded images with poor quality. Exclusion criteria included the presence of a staining artifact created by mucus, out-of-focus and insufficiently bright images, motion-blurred images, and the presence of histologic features of sessile serrated adenoma/polyps.

These two groups of images in the training set (i.e., neoplastic and hyperplastic polyps) were prepared to train the DNN for the classes of tumors we wanted to recognize using the Convolutional Neural Network in TensorFlow, followed by loading the pretrained Inception v3 model,[14] removing the old top layer, and training a new one based on the NBI photos we

collected. The 4000 steps of training process were completed after several hours of computation to build the graph used in the recognition system. Once the training process was completed, the recognition system can be used repeatedly. Figure 2 demonstrates the training accuracy of DNN-CAD, a graphical output generated by TensorBoard.[15]

Comparison Between DNN-CAD and Endoscopists

From March 2017 to August 2017, polyps smaller than 5 mm from patients with an appropriate indication for colonoscopy were collected. Exclusion criteria included age younger than 18 years, fulminant colitis, severe hematochezia, and poor bowel preparation. Patients who were unable to read or understand Chinese were also excluded. Written informed consent was obtained from each patient after the purpose of the study was explained, along with the risks, benefits, and alternatives to the procedure. Two experienced endoscopists captured and observed the magnifying NBI images of colorectal polyps at the optical maximum magnification. NBI images of resected polyps were collected from each patient. The histologic images were diagnosed by the three pathologists mentioned above.

A set of test images (96 hyperplastic and 188 neoplastic polyps) was prepared prospectively to assess the diagnostic ability of the DNN-CAD and the endoscopists. The NBI magnifying images obtained for the test set were selected according to the same image-quality criteria as was the training set, and the image resolution of the test set was the same as that of the training set. The endoscopic images of the test set were saved in JPEG (Joint Photographic Experts

Group) format. These images were then sequenced using a list of computer-generated random numbers.

Two experts with more than five years of colonoscopy experience and four novices with one year of colonoscopy experience, all of whom were blinded to histologic data, participated in this study and were asked to classify the images of the test set as neoplastic or hyperplastic. Experts were staff members in our Gastroenterology Department and wear glasses. Novices were fellows in Gastroenterology and wear glasses; they were all trained in Gastroenterology and had some experience of computer-based gaming. The annual colonoscopy volume of the experts was higher than 800, and the annual colonoscopy volume among novices was 50–100. All experts and novices had an adenoma detection rate >30% and a polyp detection rate >40% because the novices were supervised by senior endoscopists while performing colonoscopy. Both experts and novices assessed the images of the test set in digital format on a retina display on a laptop. The diagnostic time was defined as the duration between the test image appearing on the screen and the answer being given.

DNN-CAD scanned each image of the test set saved in JPEG format on the hard drive and analyzed it. A string including the prediction of polyp histology and the diagnostic time between opening the image file and displaying the prediction were generated and displayed on the computer monitor and saved onto the hard drive for statistical analysis.

Outcome Measures and Sample Size

The main outcome measures were diagnostic accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and diagnostic time. The accuracy, sensitivity, specificity, PPV, NPV, and diagnostic time were compared among DNN-CAD, the novice endoscopists, and the expert endoscopists.

The study was designed to detect a difference of 10% in accuracy by a two-sided McNemar test. With a power of 80% and a significance level of 5%, assuming 20% discordant pairs, a sample size of 255 was required to evaluate the diagnostic performance of DNN-CAD and the endoscopists. We collected 284 colorectal polyps from 193 patients as the test set.

Statistical Analysis

A two-sided McNemar test with significance level of .05 was used to compare differences in accuracy, sensitivity, and specificity, and a P value < .05 was considered statistically significant. Interobserver and intraobserver agreement of the endoscopists and intraobserver agreement of the DNN-CAD were evaluated using Cohen's kappa coefficient. All calculations were performed using SPSS v.20 (IBM, Chicago, IL, USA).

Results

Between March 2017 and August 2017, we created the test set of 284 polyps from 193 patients we prospectively enrolled in Endoscopic Center of the Tri-Service General Hospital. Table 1 shows the characteristics of the 284 polyps in the test set.

Among the 284 polyps, 56 (19.7%) were located in the cecum and ascending colon, 43 (15.1%) in the transverse colon, 12 (4.2%) in the descending colon, 51 (18.0%) in the sigmoid colon, and 122 (43.0%) in the rectum. According to the Paris classification[16] of polyp morphology, 97 (34.2%) of the test set were of the protruded type (Paris classification Is and Isp), and 187 (65.8%) were of the slightly elevated type (Paris classification IIa). Regarding histologic evaluation, 96 colorectal lesions (33.8%) were diagnosed as hyperplastic, whereas 188 (66.2%) polyps were neoplastic, including 117 (41.2%) tubular adenomas with low-grade dysplasia and 71 (25.0%) tubular adenomas with high-grade dysplasia.

Image Analysis by Deep Neural Network

The diagnostic performance of DNN-CAD and endoscopists are shown in Table 2. Among the 284 colorectal polyps, DNN-CAD correctly classified the neoplastic histology in 181 of the 188 neoplastic polyps (sensitivity of 96.3%) and hyperplastic histology in 75 of the 96 hyperplastic polyps (specificity of 78.1%). Among the 71 polyps of high-grade dysplasia,

DNN-CAD demonstrated a sensitivity of 100.0%, and 94.0% were correctly classified in the subgroup of 117 tubular adenomas with low-grade dysplasia.

Comparison Between DNN-CAD and Endoscopists

Table 2 compares the diagnostic performance of DNN-CAD and the endoscopists. The Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI) initiatives reported in the American Society of Gastrointestinal Endoscopy guidelines suggest[17] that a histologic assessment of diminutive polyps ($\leq$ 5 mm) should provide a 90% NPV for adenoma detection. With a PPV of 89.6% and an NPV of 91.5%, the level of performance using DNN-CAD met the "leave in situ" criteria proposed by the PIVI initiatives. However, three of the novice endoscopists failed to meet the PIVI criteria and achieved an NPV of 68.5%–84.0%. The diagnostic abilities of the nonexperts were not satisfactory. With regard to the time required for a diagnosis for each polyp in the test set, DNN-CAD achieved a diagnostic time of 0.45$\pm$0.07 seconds, which was shorter than the time required by both the experts and the nonexperts (1.54$\pm$1.30 seconds and 1.77$\pm$1.37 seconds, respectively; both $P < 0.001$).

Most of the nonexperts achieved moderate interobserver agreement (kappa scores of 0.48–0.77), and the experts had substantial interobserver agreement (kappa scores of 0.67). The intraobserver agreement of experts and nonexperts was moderate or higher (kappa scores of

0.77–0.81 in the expert group and 0.57–0.84 in the nonexpert group). DNN-CAD achieved perfect intraobserver agreement (kappa score of 1) (Table 3).

Discussion

We retrospectively reviewed the polyp images and the corresponding pathology for the training set (1476 neoplastic polyps vs 681 hyperplastic polyps). The ratio of identified polyps, 1476 vs 681, is comparable with the colonoscopic findings in our daily practice. Thus, we prospectively enrolled 284 polyps and pathology images for the test set to evaluate the diagnostic ability of CAD and colonoscopists. DNN-CAD classified the images as hyperplastic polyps or adenoma based on > 2000 training images retrieved from PACS in the Tri-Service General Hospital. DNN is one of the best machine-learning methods for computer vision and image classification.

Removing neoplastic polyps reduces the risk of developing colorectal cancer.[18] Endoscopic observation that can differentiate between neoplastic and hyperplastic colorectal polyps is highly beneficial because removal of hyperplastic lesions without malignant potential not only increases the duration of the colonoscopic procedure, but it also results in unnecessary costs and a risk of procedure-related bleeding or perforation.[19] Correct optical diagnosis that differentiates adenomas from hyperplastic polyps should improve the cost-effectiveness of colonoscopy.

The NBI system can be used to evaluate the capillary pattern and microstructure of the mucosa with high accuracy. When a colorectal lesion is a neoplastic polyp, vascularization occurs, and the size as well as the density of blood vessels increase; however, capillaries in

the epithelium of a hyperplastic polyp cannot be visualized.[20, 21] In our study, the images were obtained by experienced endoscopists and were focused clearly under maximal magnification. The ROIs from the gathered endoscopic images were selected and extracted manually with obvious NBI features. DNN-CAD can detect possible NBI features of polyps and provides a more objective judgment that may assist nonexpert endoscopists to achieve more accurate diagnoses.

The PIVI initiatives state that a "leave-in-place" approach is acceptable for a diminutive hyperplastic polyp when endoscopic judgment can achieve > 90% NPV for adenomas.[17] Kuiper et al.[9] reported that optical diagnosis in a nonacademic setting proved to be disappointing and achieved a sensitivity of 77.0% and a specificity of 78.8%; thus, the clinical benefits of optical diagnosis remain limited to expert endoscopists. To overcome this limitation, CAD, with its quick response and good reproducibility, has gained attention as a clinical tool for nonexperts.[10-12] In our study, DNN-CAD achieved an NPV of 91.5% for adenomas when the computer diagnosis was made. DNN-CAD could be a powerful assistant tool for endoscopists who are capable of capturing endoscopic images of high quality.

In this study, DNN-CAD for magnifying NBI images of diminutive colorectal polyps had a good diagnostic ability to identify neoplastic polyps, with an accuracy of 90.1%, a sensitivity of 96.3%, and a specificity of 78.1%. In comparison with classifications made by novice endoscopists, DNN-CAD is higher in terms of sensitivity, specificity, and NPV. Three of the

nonexperts correctly identified only 65.6%–77.1% of nonneoplastic polyps, whereas DNN-CAD achieved a specificity of 78.1%. The expert endoscopists identified 65.6%–77.1% of nonneoplastic polyps correctly; therefore, computer-aided analysis is not inferior to that of experts. The diagnostic concordance between DNN-CAD and the two experienced endoscopists was 87.7% (249/284) and 88.3% (251/284). Although human observers might unconsciously take account of additional characteristics other than vascular patterns, machine learning by a DNN itself recognizes several features to classify these two groups of NBI images. Thus, the DNN we used appears to be a suitable core system for computer-aided diagnosis.

In our study, the interobserver agreement was not high in comparison with the results of other studies.[22, 23] The nonexperts achieved mostly moderate interobserver agreement (kappa scores of 0.48–0.77), and the experts achieved substantial interobserver agreement (a kappa score of 0.67). For the interobserver variability due to the subjective interpretation of the NBI classification system, a human learning curve in the diagnosis of vascular pattern exists.

Because NBI-based diagnosis requires training and experience, an objective diagnosis is necessary. In our study, we used the up-to-date neural network model to develop DNN-CAD. It achieved perfect intraobserver agreement (kappa score of 1), but nonexperts had variable intraobserver agreement (kappa scores of 0.57–0.84). Because it has good and consistent diagnostic performance for identifying neoplastic lesions, DNN-CAD can remove some of

the diagnostic subjectivity and be a useful aid for nonexperts in managing colorectal polyps with improved cost-effectiveness.

TensorFlow, an open-source framework, features DNN that use data flow graphs.[24] It accepts sets of images and corresponding textual labels as input data and constructs a neural network to classify these images. Our DNN system analyzed the NBI endoscopic images. The script in this study loaded the pretrained Inception v3 model, removed the old top layer, and trained a new one on the endoscopic images we had collected. An advantage of transfer learning is that lower layers that have been trained to distinguish between some objects can be reused for many recognition tasks without any alteration.

CAD of colon polyps has been developed using a support vector machine.[10-12] Although these computer-aided models of image classification show potential in colonoscopy, access to the technology is limited because of institution-specific and localized software implementations. In our study, we created DNN-CAD with Google's TensorFlow API because the classifier training is easy with TensorFlow, which is a widely available library for machine learning to implement and execute large-scale DNN.[24] Moreover, the classification is fast, which is a requirement for clinical applications. TensorFlow, released in November 2015, offers open-source frameworks for public use[25-27] and has demonstrated potential for endoscopic image analysis in our study.

Another strength of this study is that DNN-CAD uses an open-source framework and can be used in several operating systems and mobile devices without platform limitation.[13] At its initial start-up, DNN-CAD requires several hours to complete the 4000 steps of the training process in order to generate the graph used in the recognition system. Once the training process is done, the recognition system can be used repeatedly. This core recognition system can be utilized on several platforms and can analyze JPEG images in real time by capturing images from the video signal output of the endoscopic video system. In addition, highly magnifying endoscopy significantly improved the rates of high-confidence, NBI-based optical diagnoses of diminutive and small colorectal polyps, as compared with non-magnifying endoscopy.[28]

There are some limitations to this study. First, the DNN-CAD diagnosis was based on high-quality images, and bias might occur with poor-quality images such as out-of-focus images or blurred images caused by mucus during real-time colonoscopy. In addition, acquiring magnified images during colonoscopy requires steady movement by endoscopists. Although the colonoscopist should try to achieve good-quality images because human eyes cannot analyze poor-quality images with confidence, it is possible to design a DNN-CAD that includes poor-quality images, and users can increase the number of poor images in the training set during set-up. Second, maximum magnification power was used for all diagnoses in our study; however, the magnifying endoscope is not commercially available in most western countries. Several studies in western countries showed that optical diagnosis of diminutive

polyps using NBI was comparable to histology in expert hands and required a learning curve, but NBI-assisted optical diagnosis cannot currently be recommended for routine use outside of expert centers.[5, 29, 30] In the future, it may be possible to build a version of DNN-CAD without magnification by adjusting the size of the ROI and retraining the DNN using images without magnification.

In conclusion, DNN-CAD, our automated diagnostic system, provides accurate and consistent diagnostic performance for colorectal polyps. DNN-CAD is promising in computer-based recognition and is not inferior to the human eyes of experts. TensorFlow, an open-source API that we used to construct DNN-CAD, is not an institution-specific or localized software implementation. The deep learning model has clinical potential to develop CAD not only in endoscopic image recognition but also in a variety of medical imaging diagnoses, including sonography, computed tomography, and magnetic resonance imaging. Further studies and clinical applications of DNN are needed to further investigate its performance.

References

1. Liou JM, Lin JT, Huang SP, et al. Screening for colorectal cancer in average-risk Chinese population using a mixed strategy with sigmoidoscopy and colonoscopy. Dis Colon Rectum 2007;50:630–640.

2. Walsh JM, Terdiman JP. Colorectal cancer screening: scientific review. JAMA 2003;289:1288–1296.

3. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med 1993;329:1977–1981.

4. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012;366:687–696.

5. Patel SG, Schoenfeld P, Kim HM, et al. Real-time characterization of diminutive colorectal polyp histology using narrow-band imaging: implications for the resect and discard strategy. Gastroenterology 2016;150:406–418.

6. Tanaka S, Sano Y. Aim to unify the narrow band imaging (NBI) magnifying classification for colorectal tumors: current status in Japan from a summary of the consensus symposium in the 79th Annual Meeting of the Japan Gastroenterological Endoscopy Society. Dig Endosc 2011;23(Suppl 1):131–139.

7. **Hewett DG**, **Kaltenbach T**, Sano Y, et al. Validation of a simple classification system

for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. Gastroenterology 2012;143:599–607.

8. Ladabaum U, Fioritto A, Mitani A, et al. Real-time optical biopsy of colon polyps with narrow band imaging in community practice does not yet meet key thresholds for clinical decisions. Gastroenterology 2013;144:81–91.

9. Kuiper T, Marsman WA, Jansen JM, et al. Accuracy for optical diagnosis of small colorectal polyps in nonacademic settings. Clin Gastroenterol Hepatol 2012;10:1016–1020.

10. Kominami Y, Yoshida S, Tanaka S, et al. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. Gastrointest Endosc 2016;83:643–649.

11. Tischendorf JJ, Gross S, Winograd R, et al. Computer-aided classification of colorectal polyps based on vascular patterns: a pilot study. Endoscopy 2010;42:203–207.

12. Misawa M, Kudo SE, Mori Y, et al. Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy. Gastroenterology 2016;150:1531–1532.

13. Tensorflow. Available at http://www.tensorflow.org/. Accessed 25 April 2016.

14. Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for

computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016;2818-2826.

15. TensorBoard. Available at http://www.tensorflow.org/versions/r0.7/how_tos/summaries_and_tensorboard/index.html. Accessed 25 April 2016.

16. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. Gastrointest Endosc 2003;58:S3–S43.

17. Rex DK, Kahi C, O'Brien M, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc 2011;73:419–422.

18. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525–532.

19. ASGE Technology Committee, Abu Dayyeh BK, Thosani N, et al. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE PIVI thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc 2015;81:502.e1–502.e16.

20. Su MY, Hsu CM, Ho YP, et al. Comparative study of conventional colonoscopy,

chromoendoscopy, and narrow-band imaging systems in differential diagnosis of neoplastic and nonneoplastic colonic polyps. Am J Gastroenterol 2006;101:2711–2716.

21. Chiu HM, Chang CY, Chen CC, et al. A prospective comparative study of narrow-band imaging, chromoendoscopy, and conventional colonoscopy in the diagnosis of colorectal neoplasia. Gut 2007;56:373–379.

22. Rogart JN, Jain D, Siddiqui UD, et al. Narrow-band imaging without high magnification to differentiate polyps during real-time colonoscopy: improvement with experience. Gastrointest Endosc 2008;68:1136–1145.

23. Sikka S, Ringold DA, Jonnalagadda S, et al. Comparison of white light and narrow band high definition images in predicting colon polyp histology, using standard colonoscopes without optical magnification. Endoscopy 2008;40:818–822.

24. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. Available at: http://download.tensorflow.org/paper/whitepaper2015.pdf. Accessed 14 April 2016.

25. Frome A, Corrado GS, Shlens J et al. DeVISE: a deep visual-semantic embedding model. Available at: http://research.google.com/pubs/archive/41869.pdf. Accessed 15 April 2016.

26. Rosenberg C. Improving Photo Search: A step across the semantic gap. Available at:

http://googleresearch.blogspot.com/2013/06/improving-photo-search-step-across.html. Accessed 15 April 2016.

27. Szegedy C, Wei L, Yangqing J, et al. Going deeper with convolutions. Paper presented at; 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7–12 June 2015; Boston, USA.

28. Iwatate M, Sano Y, Hattori S, et al. The addition of high magnifying endoscopy improves rates of high confidence optical diagnosis of colorectal polyps. Endosc Int Open 2015;3:E140–E145.

29. Rex DK. Narrow-band imaging without optical magnification for histologic analysis of colorectal polyps. Gastroenterology 2009;136:1174–1181.

30. Rees CJ, Rajasekhar PT, Wilson A, et al. Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study. Gut 2017;66:887–895.

Author names in bold designate shared co-first authorship.

Table 1. Characteristics of Neoplastic and Hyperplastic Polyps in the Test Set

	Total	Neoplastic (n = 188)	Hyperplastic (n = 96)
Location			
Rectum	122	67	55
Sigmoid colon	51	31	20
Descending colon	12	8	4
Transverse colon	43	35	8
Ascending colon	45	38	7
Cecum	11	9	2
Morphology			
Is and Isp		73	24
IIa		115	72
Pathology			
High grade dysplasia		71	
Low grade dysplasia		117	

Table 2. Diagnostic Performance of Deep Neural Network and Humans in Differentiating Neoplastic and Hyperplastic Colorectal Diminutive Polyps.

	Sensitivity	Specificity	Accuracy	PPV	NPV	Diagnostic time
	n (%)	n (%)	n (%)	n (%)	n (%)	seconds
DNN-CAD	181/188 (96.3)	75/96 (78.1)	256/284 (90.1)	181/202 (89.6)	75/82 (91.5)	0.45±0.07
Expert 1	183/188 (97.3)	74/96 (77.1)	183/284 (90.5)	183/205 (89.3)	74/79 (93.7)	1.68±1.35*
Expert 2	184/188 (97.9)	63/96 (65.6)*	247/284 (87.0)	184/217 (84.8)	63/67 (94.0)	1.39±1.24*
Novice 1	183/188 (97.3)	67/96 (69.8)	250/284 (88.0)	183/212 (86.3)	67/72 (93.1)	1.54±1.07*
Novice 2	176/188 (93.6)	63/96 (65.6)*	239/284 (84.2)*	176/209 (84.2)	63/75 (84.0)	2.09±1.95*
Novice 3	154/188 (81.9)*	74/96 (77.1)	228/284 (80.3)*	154/176 (87.5)	74/108 (68.5)	2.04±1.20*
Novice 4	158/188 (84.0)*	85/96 (88.5)	74/284 (85.6)	158/169 (93.5)	85/115 (73.9)	1.42±0.90*

DNN-CAD, computer-assisted diagnosis with deep neural network; PPV, positive predictive value; NPV, negative predictive value.

*Significant difference compared with DNN-CAD.

Table 3. Intraobserver Agreement and Interobserver Agreement of Endoscopists

	κ value					
	Expert 1	Expert 2	Novice 1	Novice 2	Novice 3	Novice 4
Expert 1	0.81					
Expert 2	0.67	0.77				
Novice 1			0.84			
Novice 2			0.77	0.72		
Novice 3			0.54	0.48	0.57	
Novice 4			0.53	0.54	0.51	0.66

Figure 1. (A) NBI demonstrating brown microvessels surrounding branched white structures, compatible with neoplastic polyps; (B) NBI showing no microvessels, compatible with hyperplastic polyps.

Figure 2. TensorBoard disclosing the training process of the deep neural network to classify neoplastic and hyperplastic polyps.