Convergence and Technologies

Cancer Research

Identification of Metastatic Lymph Nodes in MR Imaging with Faster Region-Based Convolutional Neural Networks

Yun Lu[1,2], Qiyue Yu[1,2], Yuanxiang Gao[1], Yunpeng Zhou[1,2], Guangwei Liu[1,2], Qian Dong[1,2], Jinlong Ma[1], Lei Ding[1], Hongwei Yao[3], Zhongtao Zhang[3], Gang Xiao[4], Qi An[4], Guiying Wang[5], Jinchuan Xi[5], Weitang Yuan[6], Yugui Lian[6], Dianliang Zhang[7], Chunbo Zhao[7], Qin Yao[1], Wei Liu[1], Xiaoming Zhou[1], Shuhao Liu[1], Qingyao Wu[1], Wenjian Xu[1], Jianli Zhang[1], Dongshen Wang[1], Zhenqing Sun[1], Yuan Gao[1], Xianxiang Zhang[1], Jilin Hu[1], Maoshen Zhang[1], Guanrong Wang[1], Xuefeng Zheng[1], Lei Wang[8], Jie Zhao[1], and Shujian Yang[1]

Abstract

MRI is the gold standard for confirming a pelvic lymph node metastasis diagnosis. Traditionally, medical radiologists have analyzed MRI image features of regional lymph nodes to make diagnostic decisions based on their subjective experience; this diagnosis lacks objectivity and accuracy. This study trained a faster region-based convolutional neural network (Faster R-CNN) with 28,080 MRI images of lymph node metastasis, allowing the Faster R-CNN to read those images and to make diagnoses. For clinical verification, 414 cases of rectal cancer at various medical centers were collected, and Faster R-CNN–based diagnoses were compared with radiologist diagnoses using receiver operating characteristic curves (ROC). The area under the Faster R-CNN ROC was 0.912, indicating a more effective and objective diagnosis. The Faster R-CNN diagnosis time was 20 s/case, which was much shorter than the average time (600 s/case) of the radiologist diagnoses.

Significance: Faster R-CNN enables accurate and efficient diagnosis of lymph node metastases. *Cancer Res; 78(17); 5135–43. ©2018 AACR.*

Introduction

Rectal cancer is one of the most common digestive tract tumors, whose incidence, especially in China, has been rising year by year. Lymph node metastasis is the main and most common metastasis pathway in rectal cancer. The involvement of lymph nodes is not only an important factor in the selection of treatment options, but also an important factor in predicting prognosis. Therefore, accurate preoperative TNM staging is very important for the formulation of the treatment scheme. Many studies have shown that lymph node metastasis is not only meaningful for distant metastasis, but also has a predictive value for local recurrence. According to the National Comprehensive Cancer Network (NCCN) Guidelines, the patients' treatment methods are directly determined by whether their pelvic lymph nodes have metastasized, that is, whether they should be treated with radiotherapy and chemotherapy first and then surgery or surgery first followed by other therapies. Therefore, a complete and accurate clinical evaluation of lymph node metastasis is vital. At present, pelvic MRI has become the gold standard for determining lymph node metastasis (1). The traditional method is that radiologists read MRI scans to determine whether a patient has lymph node metastasis by observing the shape, boundary, and signal intensity of the lymph nodes. It is difficult for radiologists to make correct and timely judgements based on all the factors above in a short period of time, especially when they are faced with a large number of cases. Therefore, different radiologists can reach different conclusions from the same MRI image; in particular, it is difficult to obtain an accurate and objective evaluation of lymph node metastasis (2–5). Currently, the diagnostic criteria for metastatic lymph nodes are not identical, and lymph node size is taken as the diagnostic criteria in most cases; however, there has been no agreed-upon conclusion on the size threshold for metastatic lymph nodes. A more consistent view is that the size, shape, and signal of the lymph nodes can be the diagnostic basis (6, 7). The three indicators were quantified according to image processing theories. First, the shape of the border of a nonmetastatic lymph node is defined as a circle or oval whereas that of metastatic lymph nodes is a burr or lobular shape, and the degree of circularity is taken as a quantitative indicator of the shape of the metastatic lymph node. Second, the signal of a nonmetastatic lymph node is uniform while that of a metastatic lymph node is not; thus, the

[1]Affiliated Hospital of Qingdao University, Qingdao, China. [2]Shandong Key Laboratory of Digital Medicine and Computer Assisted Surgery, Qingdao, China. [3]Department of General Surgery, Beijing Friendship Hospital, Capital Medical University, & National Clinical Research Center for Digestive Diseases, Beijing, China. [4]Beijing Hospital & National Center of Gerontology, Beijing. China. [5]Fourth Hospital of Hebei Medical University, Hebei, China. [6]First Affiliated Hospital of Zhengzhou University, Zhenzhou, China. [7]Qingdao Municipal Hospital, Qingdao, China. [8]The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Y. Lu and Q. Yu contributed equally to this work.

Corresponding Authors: Y. Lu, Affiliated Hospital of Qingdao University, Shinan Jiangsu-road No. 16, Qingdao, Shandong 266071, China. Phone: 86-186-6180-2231; E-mail: luyun@qdyy.cn; and Q. Yu, yuqiyue@qdu.edu.cn

doi: 10.1158/0008-5472.CAN-18-0494

dispersion value of the lymph node images is taken as a second quantitative indicator of the signal of metastatic lymph nodes. Finally, based on the size difference between nonmetastatic and metastatic lymph nodes, the area of the lymph node image is taken as another quantitative indicator of lymph node metastasis (8).

In recent years, the development of deep learning technology made it possible to detect the target area in an image through image recognition technology and made it possible to classify images based on detected target features. This process is the same as in the diagnoses performed by radiologists, and it can provide a new solution to the aforementioned issues. Currently, medical image auto-recognition using deep learning technologies has mainly concentrated on the identification and marking of the focus as well as the automatic sketching and 3D reconstruction of the target area. These approaches both obtain knowledge and build networks from a large number of medical images through deep learning technology to make diagnoses of the focus areas. The deep learning algorithm developed by Stanford University is on par with professional dermatologists in the accurate recognition rate of skin cancer (9). Deep learning technologies have already surpassed senior physicians in terms of their diagnostic accuracy of image recognition related to lung, breast, prostate, and esophageal cancers, among others (10–13). However, only few research efforts have been applied to the accurate diagnosis of metastatic lymph nodes using deep learning technology. More specifically, there is no report on research regarding MRI imaging related to rectal cancer. Imaging diagnosis of lymph node metastasis is more sophisticated than substantial lump diagnosis and directly affects the N staging of patients; as a result, it is of greater significance for clinical treatment (14). For this purpose, we used the deep learning model faster region-based convolutional neural network (Faster R-CNN; ref. 15) as a means to achieve accurate assessment of lymph node metastasis (Fig. 1). Faster R-CNN integrates four basic procedures of target detection and identification, that is, feature detection, candidate regional generation, regional image classification, and location refinement, into a unified deep convolutional network. Sharing with convolutional neural networks occurs through image features, and Faster R-CNN conducts high performance parallel computing on a general-purpose graphics processing unit without repetition (16, 17). Thus, Faster R-CNN can achieve more efficient deep learning and auto-recognition of lymph node metastasis. In this article, first, MRI images of lymph nodes were collected from multiple large-scale medical institutions in China, and a large database was built. Then, we trained and tested Faster R-CNN and performed rigorous practical clinical verification using Faster R-CNN. This study is the first to propose that deep learning is better than general radiologists in the diagnosis and identification of metastatic lymph nodes in terms of both diagnostic quality and speed.

Patients and Methods

MRI database of metastatic lymph nodes

The GE Signa 3.0T HDX MR scanner and multichannel phased array coil were used for the establishment of the MRI database. Patients were in a supine position, and the magnetic field center was located at the superior border of the pubis. The main scanning parameters and sequences are shown in Table 1. Data include a preoperative pelvis MRI plain scan (351 cases) for patients whose MRI report clearly showed rectal cancer with pelvic lymph node metastasis from September 2011 to October 2017 at the Affiliated Hospital of Qingdao University. We obtained specifically written informed consent from the patients that we could use their data without revealing their privacy. The number of images for a single scan was usually 16 to 24, including 2 diffusion-weighted magnetic resonance images in the horizontal position (32–48 images), 1 T2 fat saturation scan in the horizontal position (16–24 images), and 1 T1 image in the horizontal position

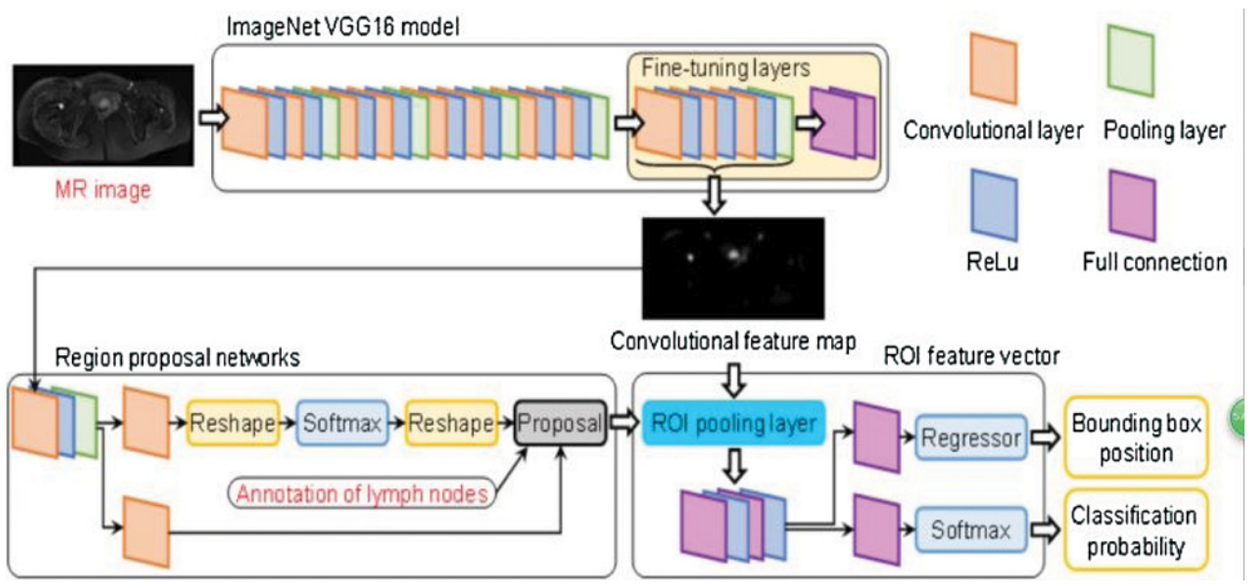

Figure 1.
Architecture of faster region-based convolutional neural networks.

Table 1. Sequence parameters of the MRI scan

Sequence		TR (ms)	TE (ms)	Layer space (mm)	Layer thickness (mm)	Matrix	FOV (cm)	NEX
OSAG	T2WI	2,000–4,000	60–120	1.0	6.0	320 × 256	32–44	1
OCOR	FS T2WI	2,000–4,000	60–120	2.0	5.0	320 × 192	36–44	2
AX	FS T2WI	2,000–4,000	60–120	1.5	6.0	320 × 192	36–44	2
AX	DWI (B = 700)	3,000–5,000	60–120	1.5	6.0	96 × 130	36–52	4

(16–24 images). MRI images with diagnostic annotations were input into Faster R-CNN. Each case had 48–72 images (Axial T2WI and DWI) and a total of 28,080 images on which senior radiologists had annotated the locations of metastatic lymph nodes for Faster R-CNN. The annotation samples of the metastatic lymph nodes are shown in Fig. 2A and B. On the basis of the data of the number of metastatic lymph nodes, N staging, and TNM staging in NCCN Guidance V3 2017, an image database was established for Faster R-CNN training and learning using all of the data above as well as the MRI images of each patient surveyed, with annotated lymph node images.

Data analysis of metastatic lymph nodes in the database

Because the current intelligent platform cannot perform three-dimensional stereoscopic recognition, horizontal images of grade C patients were only labeled (AX FS T2WI and AX-DWI) for training. During the marking process, coronal and sagittal images were also used for reference (OSAG T2WI; OOR T2WI) to determine lymph node metastasis. In the diagnosis of lymph nodes, these images were performed on a slice-by-slice basis to distinguish them from continuous blood vessels. The features of blood vessels in MRI, though similar to metastatic lymph nodes, are discontinuous. In addition, assessments from two senior radiologists and 1 pelvic surgeon were chosen to annotate lymph nodes that had the following properties as metastatic lymph nodes: A minor axis equal to or more than 5 mm, unclear borders and irregular shapes on the corresponding T2 FS images and high signal intensity on the DWI images. In a case in which the diagnosis results from the three radiologists were inconsistent, the final conclusion was reached after discussion with a fourth radiologist. An MRI image database was established with a total of 28,080 MRI images of metastatic pelvic lymph nodes. According to the final diagnosis of the radiologists, the database covers five types of lymph nodes that frequently occur in clinical diagnosis and treatment (lymph nodes around the superior rectal artery, mesorectal lymph nodes, internal iliac lymph nodes, external iliac lymph nodes, and inguinal lymph nodes) for Faster R-CNN to learn and differentiate between normal lymph nodes (18, 19) and hyperplastic lymph nodes with inflammation.

Evaluation on the training effects of Faster R-CNN

A total of 28,080 MRI images of metastatic lymph nodes in the database were input into Faster R-CNN, which contained the labeling of lymph nodes and the division of five locations where metastatic lymph nodes often occur, and 80,000 iterations of a four-step process for training were conducted using Faster R-CNN. The experimental environment is as follows: CPU Intel i7 -7820X, dominant frequency 3.60 GHz, interface model LGA2066, memory 16 GB; mainboard model Gigabyte X299-UD4; and CUDA computing environment Gigabyte AORUS GTX 1080Ti with lower dominant frequency. The graphics card provided powerful acceleration effects for data calculations in experiments. The training parameters are shown in Table 2. The loss function values of the detection network, region proposal network (RPN; ref. 20) and total networks during training were output as shown in Fig. 3A–E. To evaluate the learning effect of Faster R-CNN, 500 T2 FS images and 500 DWI images were randomly selected from 28,080 images in the database as test data in the preliminary test, which were then entered in Faster R-CNN to complete the test. The recognition of Faster R-CNN of the test data and the annotations in the database were compared with obtain the mean average precision (mAP) of Faster R-CNN in the training process, as shown in Fig. 3F. mAP is an index for measuring recognition precision in target detection. In multiple categories of object detection, a curve of each category can be drawn according to recall and precision. AP is the area under the curve (AUC), and mAP is the mean value of multiple categories of APs. The precision increases when

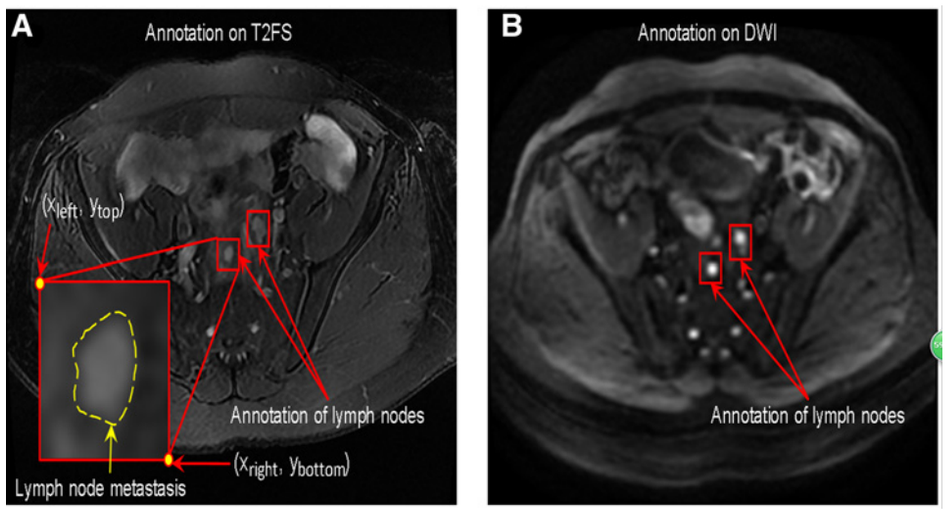

Figure 2.
Annotation samples of metastatic lymph nodes.

Table 2. Training parameters of Faster R-CNN

Iteration	80,000
Learning rate	0.001 before 60 K iterations 0.0001 after 60 K to 80 K iterations
Momentum	0.9
Weight decay	0.0005
Scale of anchor	8, 16, 32
Aspect ratio of anchor	1:1, 2:1

the index approaches 100%. According to Fig. 3F, the mAP for the test data after 20,000 iterations of training reached 98.5%, possibly due to errors in detection or training parameters. In the experiment, a regularization method was used. By limiting the weights, the model cannot fit the random noise in training data to avoid overfitting, and there was no overfitting during the 80,000 iterations of training. Therefore, Faster R-CNN was trained effectively for the MRI images of metastatic lymph nodes.

Ethics

The collection of samples and clinical information of the subjects was undertaken following informed consent and approval by the relevant ethics review board at the Affiliated Hospital of Qingdao University, in accordance with the tenets of the Declaration of Helsinki.

Results

Details of Faster R-CNN training

The architecture of Faster R-CNN includes detectors that consist of a feature extraction network of feature structures of images in ImageNet and region of interest eigenvectors, as well as RPNs. MRI images and a location marking dataset in the training database were used during the training process. ImageNet was initialized by the VGG16 model, and transfer training was performed. The weight initialization of the regional generation networks and

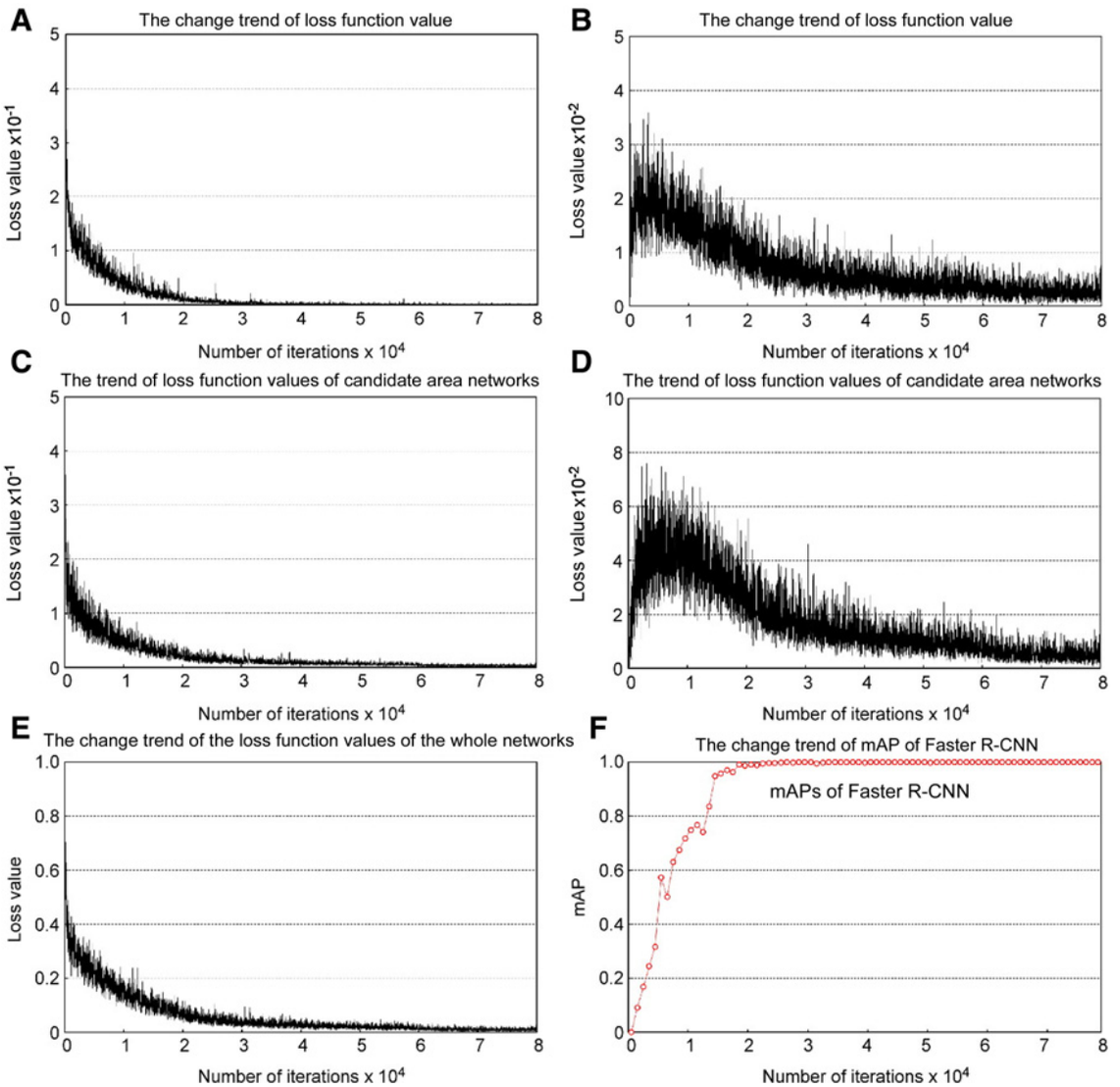

Figure 3.
Training loss function values for the detection network, RPN, and the whole network, as well as the mean average precision of Faster R-CNN.

region of interest eigenvectors shows a zero mean Gaussian distribution with a 0.01 deviation for the initial random weights. Other training parameters of Faster R-CNN are shown in Table 2. Detectors and regional generation networks perform end-to-end training by SGD and back propagation. The loss function values of the RPN and detectors, as well as the total loss function value of Faster R-CNN, were output at every iteration of training.

Four-step iteration training of Faster R-CNN

Step-1 MRI images and annotations of metastatic lymph nodes were used for training the RPN of Faster R-CNN. This network was initialized with an ImageNet-pre-trained model (VGG16) and fine-tuned end-to-end for the region proposal task.

Step-2 The detection network of Faster R-CNN was trained using the proposals generated by the step-1 RPN. This detection network was also initialized by the VGG16 model. In this training step, the RPN and detection networks did not share convolutional layers.

Step-3 The detection network was used to initialize the RPN training, but the shared convolutional layers were fixed, and only the layers of the RPN were fine-tuned. In this training step, the RPN and detection networks shared convolutional layers.

Step-4 Keeping the shared convolutional layers fixed, the detection network was fine-tuned. As such, both networks shared the same convolutional layers and were trained as a unified network.

Details of the Faster R-CNN evaluation

Faster R-CNN was evaluated after each iteration of training, and Faster R-CNN models after every 1,000 iterations of training were used to recognize the MRI images in the test database. The results were compared with the unknown mark (truth value) of the metastatic lymph nodes to obtain the mAP of the Faster R-CNN models. The calculation method for mAP is shown in Eq. A.

$$AveP = \sum_{k=1}^{n} P(k)\Delta r(k)$$
$$mAP = \frac{\sum_{q=1}^{Q} AveP(q)}{Q} \quad \text{(A)}$$

k is the number of images that have been identified, $P(k)$ is the precision at the cut-off k in the list, and $\Delta r(k)$ is the change in the recall from items $k-1$ to k. Q is the number of queries, where $AveP$ is the average precision, specifically referring to the area under PR curve. The calculation formula of AveP used the principle of calculus, taking the difference of recall ranging from k-1 to k as a small infinitesimal, which then multiplied the corresponding precision of k to obtain the area under the PR curve. mAP value, the mean average precision, is the average AP value of multiple validation sets (for individual categories). As an indicator to measure the detection accuracy in object detection, each category can draw a PR curve according to recall and precision. AP is the AUC, and mAP is the mean AP value of all categories.

Detailed clinical verification of Faster R-CNN's auxiliary function in the MRI imaging diagnosis of pelvic lymph nodes

Receiver operating characteristic curves (ROC) and PRC tools were used in the evaluation for clinical verification, and the final evaluation index was obtained according to the AUC of the ROC. Before drawing the ROC and PRC curves, we obtained the true-positive rate (TPR), false-positive rate (FPR) and accuracy of Faster R-CNN for determining the metastatic lymph nodes under different threshold values between 0.0 and 1.0. TPR refers to the sensitivity in the ROC and the recall in the PRC; FPR refers to the specificity in the ROC (21). We define that true positive (TP) is the number of lymph nodes that are determined to be metastatic by both Faster R-CNN and radiologists; false positive (FP) is the number of lymph nodes that are determined to be metastatic by Faster R-CNN but nonmetastatic by doctors; true negative (TN) is the number of lymph nodes that are determined to be nonmetastatic by both Faster R-CNN and the doctors; and false negative (FN) is the number of lymph nodes that are determined to be nonmetastatic by Faster R-CNN but metastatic by doctors. The classic and well-known algorithm of the trapezoidal method was adopted for the calculation of the AUC of the ROC.

Clinical verification of Faster R-CNN's auxiliary function in the MRI imaging diagnosis of pelvic lymph nodes

In further testing, 414 patients with rectal cancer from 6 Chinese Medical Centers were studied to verify the adjuvant effect of Faster R-CNN on the diagnosis of lymph nodes in MRI imaging. The rectal cancer cases at multiple medical centers included 36,000 T2FS and DWI images of actual patients, and the final diagnosis was determined by multiple radiologists from different centers after rigorous analysis and on the basis of a final pathological diagnosis. The final diagnosis included the staging and quantity of metastatic lymph nodes, as well as their locations in MRI images (in accordance with five positions in FRCNN training). These image data were input to the trained Faster R-CNN to achieve the diagnosis results, including the number of metastatic lymph nodes and their locations on the MRI images. The diagnosis results were then compared with those of radiologists to analyze the sensitivity, specificity and accuracy of Faster R-CNN in diagnosing metastatic lymph nodes. The ROC and precision recall (PR) curves were obtained, as shown in Fig. 4A and B. The ROC curve was mainly analyzed, and the AUC was calculated by the trapezoidal method. As AUC approaches 1, a better diagnostic effect is revealed. AUC has low accuracy when it is between 0.5 and 0.7, a certain degree of accuracy when it is between 0.7 and 0.9, and a high accuracy when above 0.9. The resulting AUC was 0.912, which means that the diagnostic ability of Faster R-CNN equals that of the radiologists who have 20 years of diagnostic experience. Thus, Faster R-CNN–assisted diagnosis is extremely effective. In addition, the automatic recognition time of Faster R-CNN is 0.2 s per image, and the recognition time for patients who have 100 MRI images on average is 20 seconds, which surpassed the level (10 minutes) of a radiologist. Therefore, Faster R-CNN–assisted diagnosis is accurate and more effective and efficient than traditional diagnostic methods and has high feasibility and clinical value as well.

Discussion

Reasons for choosing Faster R-CNN as the auto-recognition method for metastatic lymph nodes

The typical examples of the traditional target detection algorithm are: (i) Haar features + Adaboost algorithm; (ii) Hog features + Svm algorithm; (iii) DPM algorithm. Most of them create a sample, extract the artificial features from the sample, make features into feature datasets, then use the training classifier of classification algorithm, and finally use the classifier to perform target detection. Compared with the deep learning algorithm, this algorithm takes too much time in training, most of the learned features are shallow features, the generalization ability is poor, and the detection effect for more complex scenes is very poor.

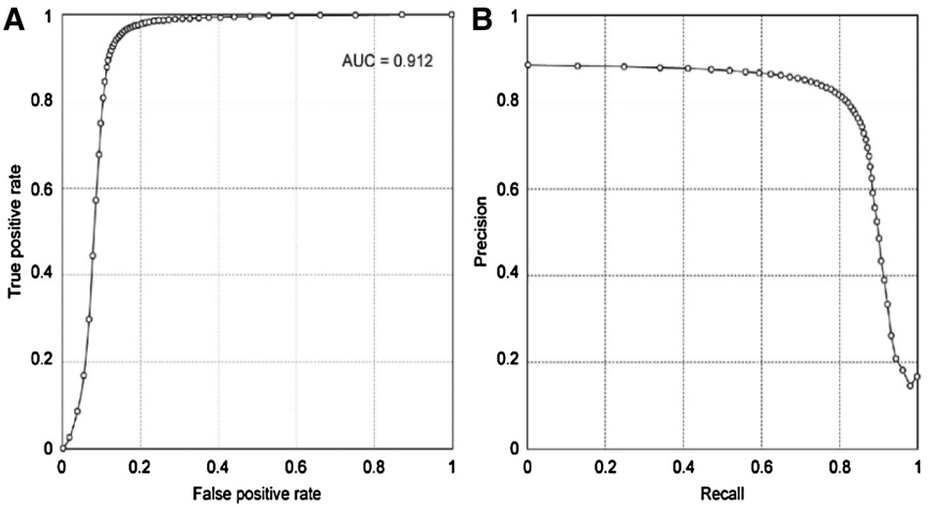

Figure 4.
ROC and PR curves of Faster R-CNN in diagnosing metastatic lymph nodes.

Faster-RCNN is a new target detection algorithm proposed by Ross B. Girshick in 2016 after the accumulation of RCNN and Fast-RCNN. The target detection is mainly driven by the regional proposal and the region-based convolutional neural network. RCNN and Fast-RCNN that are proposed previously have already realized good optimization on the calculation of region-based convolutional neural networks. Therefore, the region proposal calculation has become a bottleneck for the target detection. Compared with other deep learnings, Faster-RCNN introduced a region proposal network and built RPN by adding some additional convolutional layers. The network shares the convolutional features of the full image with the region-based detection network. RPN is a fully convolutional network and can predict the target boundary and target scores at each position simultaneously. After the end-to-end training, RPN can generate a high-quality region proposal, which is used for detection by FAST R-CNN. The RPN and Fast R-CNN are further merged into a single network by sharing the convolutional features, thereby significantly reducing the computational cost of the region proposal, and realizing further optimization.

In the MRI imaging diagnosis process, medical radiologists will detect, classify and select metastatic lymph nodes that exist in the whole MRI image before they determine the location (s) of the metastatic lymph nodes. This diagnostic process is exactly the same as that of Faster R-CNN. In the automatic identification field of deep learning, R-CNN has become a mainstream method because of its introduction of optional searching. Fast R-CNN integrated regions of interest pooling into R-CNN to make the size of its recognition target flexible. In addition, by integrating a regional generation network into Fast R-CNN, Faster R-CNN effectively reduces the enormous repetitive computing and achieves real-time target recognition. Faster R-CNN is the best technology in the field of deep learning automatic recognition, and what we need to achieve is the detection, labeling and classification of lymph nodes, namely, the detection and classification of targets. Although the detection of targets mainly includes the following four steps (generation of proposal regions, extraction and classification of features, adjustment of locations), our faster R-CNN network is developed from R-CNN and fast R-CNN. Initially, the target detection problem was treated as a regression problem, but the effect was unsatisfactory. Then, due to the good effect of CNN network on classification problems, the target detection problem was transformed into a classification problem. R-CNN drew on region proposal to obtain images in a local area that may be a target. The images were input in CNN network, respectively, to obtain the features of the target area. Then, a classifier is added to the target area to determine whether it belongs to a specific target or a background. However, there are many repetitions in the extraction of regions, and the features in the overlapped portion will be extracted multiple times. Therefore, there were a large number of repeated calculations in the R-CNN network. Fast R-CNN mapped the proposal region onto the feature map of the last conv layer of CNN. In this way, the feature of each image only needs to be extracted once, which greatly increased the speed. However, region proposal will limit the speed of fast R-CNN. Faster R-CNN also used the CNN network for region proposal. Hereinto, the region proposal network is actually a Fast-RCNN. The region proposal input from the Fast-RCNN network is fixed (i.e., a picture is divided into n*n regions. Each region gives 9 different proposals of the ratio and scale). The output is a determination of whether the fixed proposal of the input belongs to the background or the foreground and the regression of aligned position. The output of region proposal network is then input to the second Fast-RCNN for finer classification and regression of the location of the Boundingbox. Therefore, the faster R-CNN network is a target detection algorithm that is continuously being optimized and adjusted while taking into account both the calculation speed and the detection accuracy, which is why we chose Faster R-CNN as an automatic identification method (22–23). The team created 70,000 images in preliminary experiments. 60,000 images were treated as training data sets, and 10,000 images as test data sets. The training datasets were used for artificial intelligence training. All parameters in the network were adjusted to optimal values to monitor the recognition effect. The results showed that the recognition accuracy rate was 99.94%, as shown in Fig. 5.

Transfer learning helps to achieve high-precision recognition using small amounts of data

In some previous studies, hundreds of thousands of images were used as training data sets for deep neural networks (DNN), but this study used only 28,080 MRI images and location

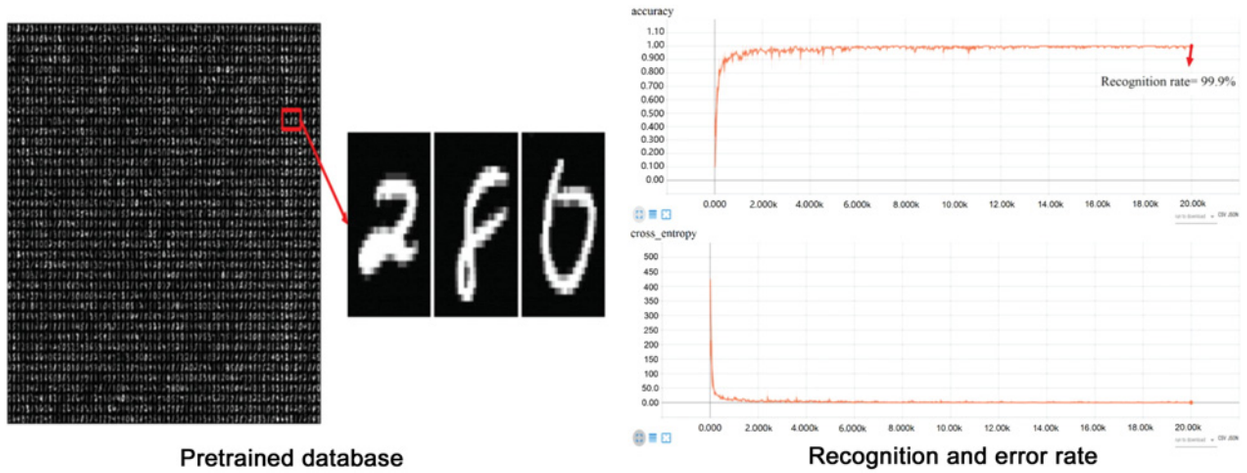

Figure 5.
Pre-trained data model and recognition and error rate curves of the test data.

markings. The reason for this approach is the application of transfer learning technology in deep learning (24), which can profoundly alleviate the problems caused by insufficient data resources. The intelligent platform we established is an automatic recognition system that simulates a radiologist, and it was required to achieve training different from normal images, so for each patient's MRI, a complete process from the first image slide to the last is needed, which contained images that labeled metastatic lymph nodes and normal images. Thus, the data were not all images containing metastatic lymph nodes. According to the Faster R-CNN architecture used in this article, the feature structure of images in ImageNet is used as a feature extraction network in Faster R-CNN, and its initialization is the VGG16 model. ImageNet has 10 million image data and object tags, and VGG16 is a DNN model based on ImageNet's image feature extraction and classification, which is quite mature (25). This study performed transfer training for the feature extraction network on the basis of VGG16. Therefore, the training of the feature extraction is based on the large dataset of ImageNet, and high precision feature extraction and ultimately high precision recognition can be achieved with a relatively small amount of new data.

Selection of evaluation tools for clinical verification of Faster R-CNN's auxiliary function in the MRI imaging diagnosis of pelvic lymph nodes

In this article, ROC and PRC were used as evaluation tools for the clinical verification of Faster R-CNN because they are excellent evaluation tools in the field of DNN detection and classification (26, 27). These two methods will make a more objective and impartial evaluation of the same samples (28). The AUC indexes in most studies are obtained according to the ROC (29); thus, the AUC value, which is the final evaluation index in the article, is also mainly calculated with the ROC.

Research extension: recognition of pelvic MRI images with high resolution

MRI equipment with high resolution has been gradually popularized among Chinese medical institutions, and has replaced traditional plain scan MRI equipment since 2017 to obtain more accurate imaging diagnoses using clearer MRI images. The problem is that a database of MRI images with high resolution cannot be established, because the equipment has not been used for long, and not much data exist. For this reason, plain scan MRI images were collected before 2017 for the construction of the database. Then, to explore the feasibility of clinical usage in the future, 1,000 MRI images from 50 recent clinical cases were collected, and recognition experiments were conducted using the achievements of this article. The trained Faster R-CNN had a zero-recognition rate for the high-resolution MRI images. According to radiologists, high-resolution MRI images not only are different in resolution but also have darker MRI image signals for metastatic lymph nodes compared with those of plain scan images, and their features are essentially different. The difference was manifested in T2WI. The internal signal of high-resolution MRI lymph nodes was quite uneven, and the density was also higher than that of ordinary plain scans, showing more irregular shapes and more pronounced burrs in the edges. In DWI, both showed a high signal, with little difference. However, to determine lymph nodes requires a combination of the two imaging methods. This experiment showed that the research achievement here does not apply to the recognition of high-resolution MRI images. In the future, a high-resolution MRI image database will be established, and transfer learning of Faster R-CNN will be further studied to achieve high-precision automatic recognition of pelvic metastatic lymph nodes in high-resolution MRI images.

Conclusion

MRI imaging diagnoses of metastatic pelvic lymph nodes rely heavily on a radiologist's subjective judgement and thus lack objectivity and repeatability in addition to having a low level of diagnosis accuracy. To solve this problem, the identification of metastatic lymph nodes in MRI with Faster R-CNN was studied. This work helps radiologists to make more accurate diagnoses of metastatic lymph nodes and is of profound importance and clinical value in the TNM staging of rectal cancer, especially for the selection of patient treatment programs. In this article, 28,080 plain scan MRI images of metastatic lymph nodes were collected, and the locations of metastatic lymph nodes were marked according to the radiologists' diagnoses to establish a large database.

Therefore, auto-recognition of metastatic lymph nodes was achieved using Faster R-CNN, which was validated by tests. MRI image data from multiple medical centers were collected to perform clinical verification. The verification results showed that the AUC value of Faster R-CNN was 0.912, which equals the level of senior radiologists, whose diagnosis time was 20 s/person. Thus, Faster R-CNN is of good clinical use and is feasible. This method is more accurate and efficient than the traditional diagnosis method.

However, there are still limitations to our research. For the diagnosis of metastatic lymph nodes, gold standard pathology was not applied, but a more uniform imaging diagnostic criteria was adopted in the establishment of the training database. This approach is because the study is aimed to make comparisons between trained Faster R-CNN and radiologists, not pathologists. In other words, this is a deep learning model simulating the diagnoses conducted by radiologists. In the future, the gold standard, pathology, may be incorporated into our research to achieve the integration of artificial intelligence imaging and pathology. In addition, due to insufficient data, training and automatic recognition of high-resolution MRI images for metastatic pelvic lymph nodes could not be achieved at this time. In the future, we will mainly focus on the establishment of a high-resolution MRI image database, as well as on transfer learning for high-resolution MRI images, for the diagnosis of pelvic metastatic lymph nodes. Moreover, when there occurred inconsistent views in the judgment of lymph node metastasis in the research, joint consultations of experts were sought. Such "difficult" nodules will be analyzed through further research and development to enable deep learning platforms to identify normal, metastatic, and "suspect" lymph nodes in more detail. Finally, although the results of the study indicated that the deep learning platform is similar to or better than the average diagnostic skill of radiologists, this research aims to develop an assistive tool to aid radiologists to more effectively and accurately determine lymph node metastases, not to be a substitute for doctors.

Data availability

The data that support the findings of this study are available on request from the corresponding author Yun Lu (cloudylucn@126.com). The data are not publicly available due to them containing information that could compromise research participant privacy/consent.

Code availability

Faster R-CNN Python code is available on GitHub at https://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf.

Clinical trial management public platform

The Clinical Trial Management Public Platform, organized by the Chinese Clinical Trial Registry: a primary registry of International Clinical Trial Registry Platform World Health Organization, is a professional platform for providing data storage and data analysis for global medical clinical trial projects. In the clinical verification, 414 cases of patients with rectal cancer from 6 Chinese medical centers were used to validate Faster R-CNN's auxiliary effectiveness. These medical centers uploaded their data to this Public Platform, and the registration number is ChiCTR-DDD-17013842.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: Y. Lu, Q. Dong, Q. Yao
Development of methodology: Y. Lu, Q. Yu, Q. Yao, W. Xu
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Y. Lu, Y. Gao, Y. Zhou, G. Liu, J. Ma, G. Xiao, Q. An, G. Wang, J. Xi, W. Yuan, D. Zhang, C. Zhao, W. Liu, X. Zhou, Y. Gao, X. Zheng, L. Wang, S. Yang
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y. Lu, Y. Gao, L. Ding, W. Yuan, D. Zhang, Q. Wu, G. Wang
Writing, review, and/or revision of the manuscript: Y. Lu, Y. Gao, G. Liu, W. Yuan, D. Wang
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y. Lu, Z. Sun, Y. Gao, G. Liu, J. Ma, H. Yao, Z. Zhang, Y. Lian, Q. Yao, S. Liu, D. Wang, J. Hu, M. Zhang, J. Zhao
Study supervision: Y. Lu, Q. Dong, H. Yao, Z. Zhang, D. Zhang, Q. Yao, J. Zhang, D. Wang, X. Zhang, J. Hu

Acknowledgments

We are deeply grateful to all of the participants and to the doctors who contributed to this project. Y. Lu received financial support from the Twelfth Five-Year Plan for Science and Technology Support of China (Grant No. 2013BAI01B03), the Major Science and Technology Project of Independent Innovation of Qingdao, China (Grant No. 14-6-1-6-zdzx) and the Key Research and Development Plan (tackle hard-nut problems in science and technology) of Shandong Province (Grant No. 2018GSF118206).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 16, 2018; revised May 15, 2018; accepted July 9, 2018; published first July 19, 2018.

References

1. Wáng YXJ, Idée JM. A comprehensive literatures update of clinical researches of superparamagnetic resonance iron oxide nanoparticles for magnetic resonance imaging. Quant Imaging Med Surg 2017;7:88–122.
2. Muthusamy VR, Chang KJ. Optimal methods for staging rectal cancer. Clin Cancer Res 2007;13:6877s–6884s.
3. Al-Sukhni E, Messenger DE, Charles Victor J, McLeod RS, Kennedy ED. Do MRI reports contain adequate preoperative staging information for end users to make appropriate treatment decisions for rectal cancer? Ann Surg Oncol 2013;20:1148–1155.
4. Tytherleigh MG, Ng VV, Pittathankal AA, Wilson MJ, Farouk R. Preoperative staging of rectal cancer by magnetic resonance imaging remains an imprecise tool. ANZ J Surg 2008;78:194–198.
5. Zhou J, Zhan S, Zhu Q, Gong H, Wang Y, Fan D, et al. Prediction of nodal involvement in primary rectal carcinoma without invasion to pelvic structures: accuracy of preoperative CT, MR, and DWIBS assessments relative to histopathologic findings. Lid - e92779. PloS ONE 2014;9:e92779.
6. Saklani AP, Bae SU, Clayton A, Kim NK. Magnetic resonance imaging in rectal cancer: a surgeon's perspective. World J Gastroenterol 2014;20:2030.

7. Al-Sukhni E, Milot L, Fruitman M, Beyene J, Victor JC, Schmocker S, et al. Diagnostic accuracy of MRI for assessment of T category, lymph node metastases, and circumferential resection margin involvement in patients with rectal cancer: a systematic review and meta-analysis. Ann Surg Oncol 2012;19:2212–2223.
8. Beets-Tan RGH, Beets GL. Local staging of rectal cancer: a review of imaging. J Magn Reson Imaging 2011;33:1012–1019.
9. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–118.
10. Nakamura N. Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules with use of artificial neural networks. Radiology 2000;3:214.
11. Chen DR, Chang RF, Huang YL. Computer-aided diagnosis applied to US of solid breast nodules by using neural networks. Radiology 1999;213: 407–412.
12. Chen CM, Chou YH, Han KC, Hung GS, Tiu CM, Chiou HJ, et al. Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks. Radiology 2003;226:504–514.
13. Suzuki K, Li F, Sone S, Doi K. Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network. IEEE Trans Med Imaging 2005;24:1138–1150.
14. Ishihara S, Kawai K, Tanaka T, Kiyomatsu T, Hata K, Nozawa H, et al. Oncological outcomes of lateral pelvic lymph node metastasis in rectal cancer treated with preoperative chemoradiotherapy. Dis Colon Rectum 2017;60:469–476.
15. Ren S, He K, Girshick R, Sun J. Faster R-CNN. Towards real-time object detection with region proposal networks. Adv Neural Inform Processing Syst 2015;91–99.
16. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. Procs. IEEE Conf. on Computer Vision and Pattern Recognition2016;779–788.
17. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. Overfeat: integrated recognition, localization and detection using convolutional networks. arXiv Preprint ArXiv: 2013:1312.6229.
18. Kim YM, Suh JH, Cha HJ, Jang SJ, Kim M-J, Yoon S, et al. Additional lymph node examination from entire submission of residual mesenteric tissue in colorectal cancer specimens may not add clinical and pathologic relevance. Hum Pathol 2007;38: 762–767.
19. Brown G, Richards CJ, Bourne MW, Newcombe RG, Radcliffe AG, Dallimore NS, et al. Morphologic predictors of lymph node status in rectal cancer with use of high-spatial-resolution MR imaging with histopathologic comparison. Radiology 2003;227:371–377.
20. Sarikaya D, Corso J, Guru K. Detection and localization of robotic tools in robot-assisted surgery videos using deep neural networks for region proposal and detection. IEEE Trans Med Imaging 2017;36: 1542–1549.
21. Zhu Y.TP, TN, FP and FN Tables for dierent methods in dierent parameters. 2015.
22. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 2017;39:1137–1149.
23. Chollet F. Deep learning with separable convolutions. arXiv preprint arXiv 2016;1610.02357.
24. Weiss K, Khoshgoftaar TM, Wang D. Transfer learning techniques. In big data technologies and applications. In: Furht B, Villanustre F. editors. Springer International Publishing, Cham, 2016:53–99.
25. Xie X, Han X, Liao Q, Shi G. Visualization and Pruning of SSD with the base network VGG16. Procs. 2017 Intl. Conf. on Deep Learning Technologies. (ACM, Chengdu, China, 2017), 90–94.
26. Hajian-Tilaki K.Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med 2013; 4:627–635.
27. Park SH, Goo JM, Jo C. Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean J Radiol 2004;5:11–18.
28. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE 2015;10:e0118432.
29. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit 1997;30:1145–1159.