it is possible to obtain performance equal to development that uses only actual cases. In addition, we succeeded in verifying the performance of this development technique in breast cancer mass shadow CAD that uses mammograms, suggesting the applicability of the proposed development technique to multiple body areas.

References

[1] Perez P, Gangnet M, Blake A (2003) Poisson Image Editing, Proc. SIGGRAPH"03: 313–318.

[2] Thomas P, Tom, D (1984) Compositing digital images. ACM SIG-GRAPH Computer Graphics 18: 253–259.

Deep super-learning of polyp images for computer-aided detection in CT colonography

K. Umehara[1], J. Näppi[2], T. Hironaka[2], T. Ishida[1], D. Regge[3], H. Yoshida[2]

[1]Osaka University, Suita, Japan

[2]Massachusetts General Hospital and Harvard Medical School, Boston, United States

[3]Institute for Cancer Research and Treatment, Turin, Italy

Keywords Computer-aided detection · Deep learning · Colorectal cancer · CT colonography

Purpose

The use of computer-aided detection (CADe) yields a statistically significant improvement in the detection sensitivity of CT colonography (CTC). However, current CADe systems display many false-positive (FP) detections. It is desirable to reduce the number of FPs while maintaining high detection sensitivity.

Deep convolutional neural networks (DCNNs) have recently showed state-of-the-art performance in many classification tasks. Previously, we demonstrated that transfer learning provides an effective approach for robust application of DCNNs in CTC [1]. However, at a high detection accuracy, the differentiation of small polyps and non-polyps remained challenging.

Super-learner algorithm is a loss-based supervised ensemble-learning method that finds an optimal combination over a collection of prediction algorithms [2]. Thus, it provides a systematic approach for combining many predictors into one optimal predictor. In this study, we developed a deep super-learning (DESLA) classifier scheme based on the super-learner algorithm with multiple types of DCNNs to improve the polyp detection performance of CADe in CTC.

Methods

We sampled 154 CTC cases from a clinical multi-center trial that had been designed to evaluate CTC performance in daily clinical practice. No specific colon cleansing instruction was given. Fecal tagging by iodine was administered with or without barium in 34% of the cases. The CTC acquisitions were performed in dual positions with 120 kVp, 50 effective mA per second, and 2.5-mm section thickness. After same-day colonoscopy, expert radiologists correlated the colonoscopy findings with the CTC images.

Polyp candidates were detected from CTC data by use of a conventional fully automated CADe system. After colon extraction and shape-based detection of polyp candidates, an AdaBoost classifier reviewed radiomic shape and texture features of the polyp candidates to determine the final output of the CADe system.

We used three types of publically available DCNN models: CaffeNet [3], AlexNet [4], and GoogLeNet [5], which had been pre-trained to classify images with 1.3 million natural non-medical images from the ImageNet Large Scale Visual Recognition Challenge 2012 image set. The DCNNs were re-trained by transfer learning to identify polyps using virtual endoluminal (VE) images of the polyp candidates detected by our conventional CADe system. Nine different types of renderings were generated for each VE image (Fig. 1). A DESLA classifier scheme was developed by re-training eleven DCNNs on the rendered VE images, and by combining the re-trained DCNNs with the super-learner algorithm where a random forest classifier was used as the meta-classifier.

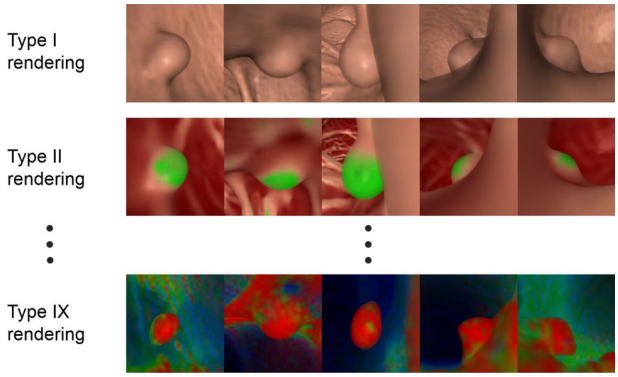

Fig. 1 Examples of the different types of virtual endoluminal renderings based on conventional rendering (Type I), shape-index feature (Type II), and translucency feature (Type IX)

The CTC cases were divided randomly into an independent training dataset of 62 cases and a test dataset of 92 cases. Both datasets contained cases with and without fecal tagging. The test dataset contained 107 biopsy-confirmed adenomas and carcinomas ≥6 mm in size: 69 were ≥10 mm, and 38 were 6–9 mm in size. After training the CADe system to detect polyp candidates with the training dataset, the VE images of the polyp candidates were categorized manually into polyp and non-polyp classes and used to construct and train the DESLA classifier. Finally, the trained CADe system was used to detect polyp candidates from the test dataset, and the VE images of the polyp candidates were reviewed by the trained DESLA classifier to determine the final polyp detections.

Results

Figure 2 shows free-response receiver-operating characteristic (FROC) curves of the per-polyp detection performance of the three schemes: the standalone CADe scheme (CADe), CADe followed by the best-performing single DCNN (DCNN-CADe), and CADe followed by the proposed DESLA (DESLA-CADe) scheme. For 6–9 mm polyps at 2.9 FPs/patient on average, the per-polyp sensitivities of the CADe, DCNN-CADe, and DESLA-CADe schemes were 73.7, 81.6, and 86.8%, respectively (Fig. 2a). For large polyps, the detection sensitivities of the CADe, the DCNN-CADe, and the DESLA-CADe schemes were 85.5, 94.2, and 97.1%, respectively, at 3.9 FPs/patient on average (Fig. 2b).

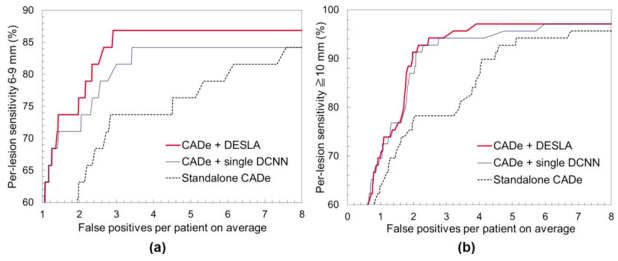

Fig. 2 FROC curves of the standalone CADe, DCNN-CADe, and DESLA-CADe schemes in the detection of polyps (a) 6–9 mm and (b) ≥10 mm in size

The DCNN-CADe yielded 93.5% sensitivity for ≥6 mm polyps at 17.0 FPs/patient on average, whereas with the DESLA-CADe scheme the number of FP detections was only 3.9 FPs/patient. This result indicates that the proposed DESLA scheme can reduce the number of FPs by 83% over that of using a single DCNN at a comparable detection sensitivity.

Conclusion

We developed a DESLA classifier scheme using transfer learning of publically available DCNN models for improving the detection performance of CADe for polyps in CTC. Our preliminary results indicate that the DESLA scheme can significantly improve polyp detection performance, both for large polyps and for small polyps. Therefore, the DESLA scheme provides an effective approach for improving the polyp detection performance of CADe in CTC.

References

[1] Näppi JJ, Hironaka T, Rogge D, Yoshida H (2016) Deep transfer learning of virtual endoluminal views for the detection of polyps in CT colonography, Proc. SPIE Int Soc Opt Eng 9785.

[2] van der Laan MJ, Polley EC, Hubbard AE (2007) Super Learner, U.C. Berkeley Division of Biostatistics Working Paper Series, Working Paper 222.

[3] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: convolutional architecture for fast feature embedding, Proc of the 22nd ACM International Conference on Multimedia, 675–678.

[4] Krizhevsky A, Sutskever I, Hinton G (2012) ImageNet classification with deep convolutional neural networks. The 26th Annual Conference on Neural Information Processing Systems (NIPS), 1–9.

[5] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions, Proc of the IEEE Conf. Comput. Vis. Pattern Recog., 1–9.

Breast mass classification using the fundamental deep learning approach

K. G. Kim[1], Y. Kim[1], C. Park[1]

[1]Gachun University, Biomedical Engineering, Incheon, South Korea

Keywords Convolutional neural network · Deep learning · Biomedical engineering · Breast data

Purpose

Breast cancer is the most commonly diagnosed cancer and the leading cause of cancer death among women worldwide, accounting for approximately 1.7 million cases and 521,900 deaths in 2012 [1]. The American Cancer Society, in turn, recommends women over the age of 40 to get breast cancer screening mammograms on a regular basis for the purpose of early detection [2]. However, it is difficult for radiologists to detect and analyze masses due to their variation in shape, size, and boundary as well as their low signal to noise ratio, resulting in unnecessary biopsies or missed masses [3]. A computer-aided diagnosis (CAD) system has been traditionally used in breast mass classification. However, according to one study on the effectiveness of CAD, it shows no significant improvements in the sensitivity for invasive breast cancer [4]. To solve this problem, convolutional neural networks (CNNs) based on deep learning approaches are being developed by many researchers to be used in clinical practice. Unlike traditional CAD systems that use pre-determined features, CNNs determine the most relevant features from data in order to classify images as normal tissue or malignant masses [5]. A CNN commonly includes the convolutional layers, the pooling layers, followed by fully connected layers. The convolutional layers consist of a set of learnable filters that are convolved with the input image.

A CNN commonly includes the convolutional layers, the pooling layers, followed by fully connected layers. The convolutional layers consist of a set of learnable filters that are convolved with the input image. The pooling layers reduce the size of the input and max-pooling is commonly used. Fully connected layers have full connections to all activations in the previous layer and calculate the final output with a soft-max function.

Methods

The two layered CNN architecture used in this study consists of 3 stages of convolutional layers, ReLU (rectified linear unit) activation layers, and max pooling layers, followed by fully connected layers. A dropout layer with dropout factor of 0.75 was added before the fully connected layers to prevent overfitting. An optimum number of iterations had to be determined since a small number of iterations results in less training and too large a number of iterations results in high error rates. Observing the iterations versus accuracy graph in Fig. 1, it was confirmed that to keep increasing the number of iterations does not further increase the testing accuracies. Therefore, the number of iterations was set to 50,000 where the curve reached a plateau, and the batch size was set to 30 for all datasets. Each filter or kernel in convolutional layers extracts particular features from the images. Before the number of filters was increased, the model used 32 filters in the first convolutional layer and then 64 filters in the second layer, extracting 2048 features from one image. The number of convolutional filters was then increased to 64 and 128, to see if it enables the model to extract more features and show better performance. 2.5 Image sizes Since the CNN model used in this study was modified from MNIST classification model, 28 by 28 was the default setting that could be used as input image sizes. To see if increasing image sizes enables the network to extract smaller and more detailed features and ultimately to show better performance in breast mass classification, input image sizes of 64 by 64, 128 by 128, and 256 by 256 were compared.

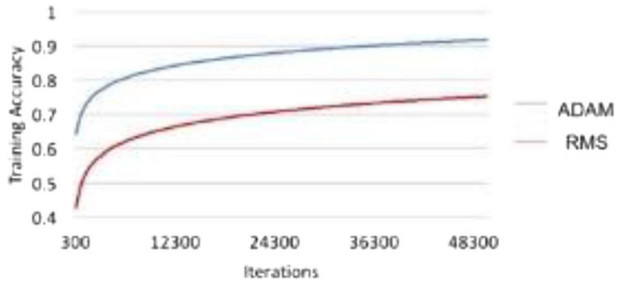

Fig. 1 Iteration versus Training accuraccy comapring RMS and ADAM optimizer

Results

The main goal of this paper was to build the optimal model for breast mass classification by applying various methods that influence the performance of Convolutional Neural Network (CNN). The proposed model achieved the accuracy of 0.887, sensitivity of 0.903, and specificity of 0.869 for normal tissue versus malignant mass classification with augmented data, more convolutional filters, and ADAM optimizer.

Conclusion

Therefore, it is verified that breast mass classification using CNN has potential to be a better assisting tool than a CAD system in providing a consistent second opinion to a radiologist by reducing false-positive and false-negative diagnoses. A limitation of this method, however, was that it only considered malignant masses