IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 18, 2019, accepted June 17, 2019, date of publication June 20, 2019, date of current version July 10, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924207

Automatic Kidney Lesion Detection for CT Images Using Morphological Cascade Convolutional Neural Networks

HUI ZHANG[1], YURONG CHEN[1], YANAN SONG[1], ZHENLIN XIONG[1], YIMIN YANG[2], (Member, IEEE), AND Q. M. JONATHAN WU[3]

[1]Department of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China
[2]Computer Science Department, Lakehead University, Thunder Bay Campus, Thunder Bay, ON P7B 5E1, Canada
[3]Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada

Corresponding authors: Hui Zhang (zhanghuihby@csust.edu.cn) and Yurong Chen (chenyurong1998@outlook.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61601061, Grant 61841103, and Grant 81401490, in part by the National Key Research and Development Program of China under Grant 2018YFB1308200, in part by the Scientific Research Fund of the Hunan Provincial Education Department under Grant 17C0046, in part by the Hunan Key Laboratory of Intelligent Robot Technology in Electronic Manufacturing under Grant IRT2018009, and in part by the Hunan Key Project of Research and Development Plan under Grant 2018GK2022 and Grant 2018JJ3079.

ABSTRACT The CT scan image is one of the most useful tools for diagnosing and locating lesions in the kidney. It can provide precise information about the location and size of lesions in many medical applications. Manual and traditional medical testings are labor-consuming and time-costing. Nowadays, detecting lesions in CT automatically is an integral assignment to the paramount importance of clinical diagnosis. Computer-aided diagnosis (CAD) is needed to develop and improve medical testing efficiency. However, it is still a tremendous challenge to the extant low precision and incomplete detection algorithm. In this paper, we proposed a lesion detection tool using multi intersection over union (IOU) threshold based on morphological cascade convolutional neural networks (CNNs). For improving the detection of small lesions (1–5 mm) and increasing the stableness of network, we proposed two morphology convolution layers and modified feature pyramid networks (FPNs) in the faster RCNN and combined four IOU threshold cascade RCNNs. In this lesion detection task, the modified CNN was trained in pytorch framework. The experiments were conducted in CT kidney images of DeepLesion that are published by hospitals' picture archiving and communication systems (PACSs). Finally, our method achieved AP of 0.840 and AUC of 0.871, and the results demonstrated that our proposed detector is an outstanding tool for detecting lesions in CT and outperformed in the data set.

INDEX TERMS Kidney lesion detect, deep learning, morphology, RCNN.

I. INTRODUCTION

Nowadays, overwork that brings out human immunity reducing, excessive salt intake, and worse external environmental factors such as cold, damp are easy to further cause kidney disease. As the research of the Kidney International [1] shows that the population of the chronic kidney disease (CKD) is beyond 2.8 hundred million and still increase rapidly. And it is reported the morbidity and mortality of renal diseases has doubled in the last year by National Institute of Health, and the incidence of renal disease tends younger and younger. The nephropathy will be a more and more serious problem, so the exact diagnose is of great significant for discovering and treating renal disease. It is significant for renal function diagnose to take precaution before symptoms. And understanding the location and extent of kidney damage can help to diagnose and guide treatment. In clinical diagnosis, doctors and radiologists usually detect kidney disease via computed tomography (CT). CT scans is one of the most common tools used for the screening, diagnosis and treatment of lesions. It can detect small calcification, stones or negative stones that cannot be clearly detect by standard X-ray examination. CT can determine the location, extent and hematoma of the kidney injury, as well as postoperative complications. And compared with standard X-rays, CT scans are more detailed.

The associate editor coordinating the review of this manuscript and approving it for publication was Jiachen Yang.

The most common kinds of kidney lesions are that renal cyst, renal stone and hydronephrosis. For quickly finding and locating those renal lesions, comparing with the time-consuming and laborious traditional diagnosis in CT scans, the algorithms for automatic detection of lesions can ameliorate those situations: (1) Saving time of diagnosis, with the number of patients increasing rapidly, the medical testing needs more and more time to get the result, so it can be efficient and effective with using CAD to help diagnose. (2) Improving the precision of testing, the lesion detection algorithm can help radiologists and doctors, especially those do not have much experience to decide where are lesions. (3) Reducing costs of hospitals and schools, undoubtedly, training a preeminent radiologist and doctor needs large money and time.

Nowadays, more and more people and laboratory are interesting in detection of medical images. And all of them accommodate their methods to deal with medical images tasks. Jiang et al.[2] fixing structure and training procedure that could be applied to resolve a medical imaging problem. Noll et al. [3] utilized basic kidney shape information to detect the kidney position. Zhang et al. [4] developed an efficient Hessian based Difference of Gaussians (HDoG) detector to identify the glomeruli. Greenspan et al. [5] proposed an early detection of acute renal transplant rejection using diffusion-weighted MRI. Akkasaligar and Biradar [6] using wavelet thresholding and Wavelet decomposition method to process of kidney ultrasound images. van Ravesteijn et al. [7] presented a CAD system for computed tomography colonography that orders the polyps according to clinical relevance. Grigorescu et al. [8] proposed a novel method for automated detection and segmentation of all large lesions based on a local measure for protrusion and clustering based on geodesic distance. Zhou and Qi [9] focuses on the adaptive imaging for lesion detection. Moon et al. [10] developed for analyzing automated breast ultrasound images (ABUS) based on multi-scale images detection. Li [11] presents Gaussian Proposal Networks (GPNs) to detect lesion bounding ellipses. Zhang et al. [12] predicted the subsequent involvement regions of a tumor. Khalifa et al. [13] proposed a novel framework for the classification of acute rejection versus nonrejection status of renal transplants. Moveover, Zheng et al. [14] proposed a CNN based approach that to improve the effectiveness of detection nodules. Tang et al. [15] based on Mask RCNN proposes a hard negative example mining strategy to detect lesions. Shin et al. [16] fine-tuned CNN models pre-trained from natural image data set to medical image tasks. Arnaud et al. [17] designed a fully automated method that can both localization and characterization of lesion. Kolachalama et al. [18] using Deep Neural Networks to detect pathological fibrosis renal. Danaee et al. [19] modified a tool to detection the cancer. Turco et al. [20] dominantly focused on polycystic kidney disease. Ben-Cohen et al. [21] detected in computed tomography (CT) examinations, using both a global context with a fully convolutional network (FCN). Cuingnet et al. [22] proposed an automatic detection and segmentation of kidneys in 3D CT images using random forests. Zhang et al. [23] did for identification of small structures (blobs) from medical images to quantify clinically relevant features. Mahapatra et al. [24] designed an automatic detected disease tissues algorithm. However, all those networks are not suitable for the lesion detection of kidney, due to those mentioned methods cannot deal neatly with the problem of sparse and small lesions in kidney and cannot achieve a high precision and comprehensive detector. However, unlike the common image classification and object detection, the detection kidney of medical image needs resolves lots of extra problems [25]. Such as, lacking a large-scale annotated image data set and only those has professional knowledge can label the image annotation. CAD is the most popular and interesting hot area as the deep learning becoming more and more comprehensive, especially in image processing. Detection lesions tend to be one of the most popular research. In a word, convolutional neural network (CNN) is shown a valid tool for a range of computer vision tasks, including medical imaging detection, which is indispensable for the algorithm to diagnosis precise and comprehensive. Contrast to the traditional diagnosis ways of CT, it has a better performance with high accuracy and spends less time to detect the lesion, and it will relieve the work of the radiologist. CAD [26] tools help the radiologist especially the young and inexperienced doctor to deal with recognition lesions in medical image.

Based on our data set (DeepLesion) [27] and the actual situation that lesions are most small and distribute sparsely, we use faster rcnn as the baseline. Our proposed network is multi stages network which improved from two stage network, but more accurate than two stages network. To deal with this problem and to improve detection feature maps, we use an effective and exact residual neural network (ResNet) as backbone for obtaining kidney lesion feature maps. Because the shallow layers made from more high-resolution convolutional feature maps that include more spatial information and the deep layers containing more low-resolution feature maps that including more semantic information [30]. In conclusion, our method focuses on improving those following three points: (1) Using two kinds of kernel of convolution image to achieve the effect of mathematical morphology image processing. After dilating and eroding, the small goals in CT image can be detected undemanding. (2) Combine modified feature pyramid networks P2, P3, P4, P5, P6, P7 with ResNet101 in multi IOU thresholds cascade network at the Faster RCNN baseline. (3) Various data augment methods, including flipping, resizing (the image size 512x512 is sub-sampled by factors of 2 down to 256x256 and up to 1024x1024 and cropping image with the center of kidney. The framework of the proposed method is shown in Fig. 1, which mainly makes of two steps, that is different size of kidney feature maps acquiring for localization and classification and multi-IOU cascade RCNN. In the first stage, getting several different sizes feature maps via two kinds of morphology convolution layers and backbone with

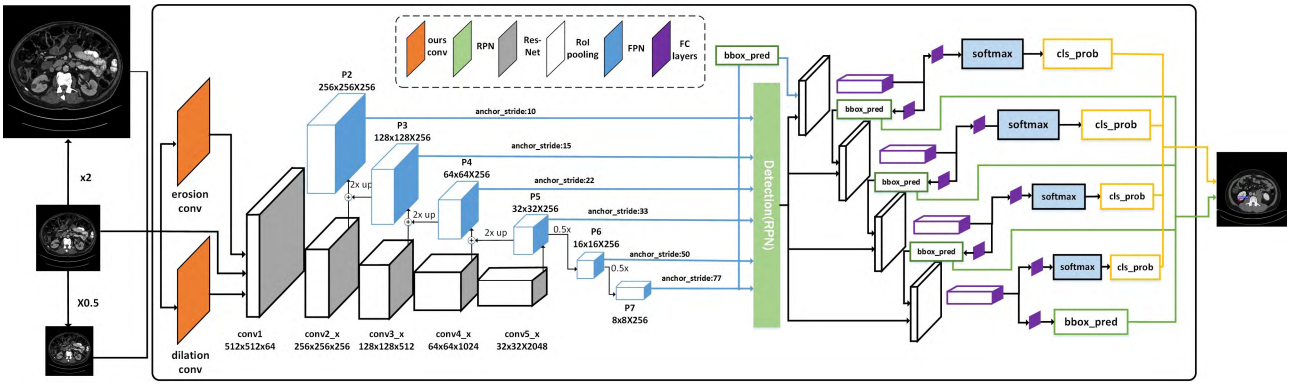


FIGURE 1. The Framework of the proposed kidney lesion detection method which mainly includes two steps: the acquisition of six different sizes lesions detection feature maps and multi-IOU cascade RCNN.

feature pyramid networks. Subsequently, the region proposal network and four IOU thresholds will deal with those feature maps with diverse anchor size and then output class prediction score and four coordinate of boundary box. Without doubt, our experiment shows a high accuracy, high speed and with no human involvement program of kidney lesions detection in CAD.

II. METHOD

Given an input (2D CT images), the network of Fig. 1 detects renal lesions following those steps. In the first step, the input image via two kinds of convolution kernels of morphological operations [28] (eroding and dilating) will generates erosion and dilation image layers (Section III-A); In the second step, six different sizes feature maps are exacted through the backbone including ResNet and FPN from three paths images (the original image, erosion layer and dilation layer) (Section III-B); In the final step, the RPN and four IOU thresholds are cascaded to deal with those different sizes feature maps to give the class prediction score and coordinates regression (Section III-C).

A. MORPHOLOGICAL OPERATIONS

Morphology is a theory and technique for the analysis and processing of geometrical structures and commonly suitable to digital image process. Morphological image processing is a group of non-linear operations associates with the shape or morphology. The basic morphological operations include dilation, erosion, opening and closing. Morphological operations depend on the correlation of pixel values, not on their absolute numerical values. So, morphological operations are quite suitable to the processing of binary images, such as lesions in the kidney. There are many functions of morphological operations. (1) noise elimination; (2) image independent elements isolation and image adjacent elements joins; (3) the maximum or minimal region finding in the image. The application of our data set is that, reduce some noise of the CT scan and make the lesion more obvious. In the Fig.2, we show the comparison of effect of dilation and erosion.

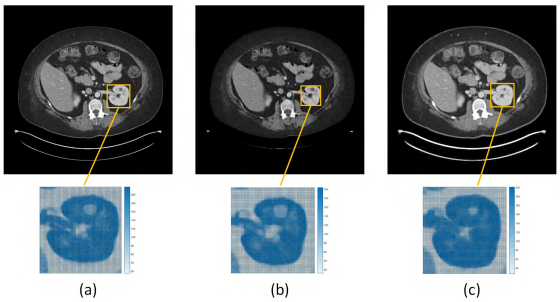


FIGURE 2. Outputs of the morphological operations layer. (a) The original image, (b) dilation, (c) erosion.

1) DILATION

The process of dilation is that use a kernel (called B) to convolve the image (called A), and the center of the kernel B is called anchor point. When the kernel B makes convolution with image A, assign the maximum pixel value of the area B to the anchor point. The image is convolved with a kernel which can be of any shape and size, and we used a 3x3 square. At its center, we call it an anchor point. The dilation is an operation of finding the local maximum. When the kernel B is convolved with the image, find the maximum value of the pixel of the region covered by the kernel B and give this to the anchor point. This will cause the highlighted area in the image to grow gradually. The dilation formula is shown as below:

$$dst(x, y) = \max_{(x',y'):element(x',y')=0} src(x + x', y + y') \quad (1)$$

Replacing the value of (x, y) with the maximum value in the surrounding area (x+x', y+y'), images will be more highlighting and the black area will shrink. The application of CT scans as the FIG. 2(a) shows, the lesion is a black area in the white kidney, after dilation, the lesion becomes more contrast and centralized [29]. The dilation can be understood as the center of B (an anchor point) sliding along the outer boundary of A. For the highlight area, the dilation replaces outside no highlight area pixel with inside highlight area

pixel. After dilation, the image is larger than the original image.

2) EROSION

Erosion and dilation are the opposite operations, and erosion is to find the local minimum. We use the kernel B make convolution with image A, assign the minimum pixel value of the area B to the anchor point. A 3x3 square kernel with the anchor point we choose. The erosion is an operation of finding the local minimum that is, when the kernel B is convolved with the image, the minimum value of the pixel of the region covered by the kernel B, and then the minimum value is assigned to the pixel specified by the anchor point. This will cause the dark area in the image to grow gradually. The erosion formula is shown as below:

$$dst(x, y) = \min_{(x',y'):element(x',y')=0} src(x + x', y + y') \quad (2)$$

find the surrounding area (x+x', y+y') of point (x, y) of minimum value and then assign to (x, y). After erosion, images will allow thicker lines to get skinny. The application in CT scans as the FIG. 2(c) shows, the area of the lesion will expand larger and the kidney will shrink. The erosion can be known as the kernel B sliding along the inter boundary of A. For the highlight area, the dilation substitute for inside highlight area pixel with outside no highlight area pixel. After erosion, the image is smaller than the original image. In a sum, we added two kind of kernel (dilation kernel and erosion kernel) before the inputting image to backbone, after morphological operations, there are three paths of imputing image.

B. ACQUIRING FEATURE MAPS

Faster RCNN [30] is proposed from 2015, but it still the most exact and accurate detection network. It is still the basic of many target detection algorithms. The original Faster RCNN mainly involves four parts: (1) convolution layers: using a set of basic convolution (conv)+ rectified linear unit (relu)+ pooling layers to extract the feature maps of the input image. The feature maps will be used for subsequent RPN layers and fully connected layers. (2) Region propose network (RPN): the RPN network is mainly used to generate region proposals. First, a bunch of anchor box is generated. After cutting and filtering, it is judged by softmax classifier that the anchors area belong to the foreground or the background, that is, it is a goal or not a goal, so this is a dichotomy classifier. At the same time, another path from the bounding box regression corrects the anchor box, forming a more precise proposal. (3) Region of interest (RoI) Pooling: this layer uses the feature map generated by the RPN and the feature map obtained by the last layer of backbone to obtain a fixed-size proposal feature map. Then the full connection (FC) layers operation can be used for target recognition and location. (4) Classifier: this layer will fully connect the fixed-size feature map from the RoI Pooling layer and softmax is used to classify specific categories. At the same time, the bounding box regression operation is completed by Smooth L1 Loss Layer [30] to obtain the precise location of the goal.

We modified the first and third step to help us apply and suit for our data. The first step is made up of two parts, that is, ResNet101 and six layers (P2, P3, P4, P5, P6, P7) feature pyramid network. The original backbone of Faster RCNN is VGG16, but our experiment showed that the VGG16 has a lower performance than ResNet101. The VGG16 contains 16 convolution layers and fully connect layer. However, when the network becoming deeper, the overfitting will be a problem that needs to be solved. ResNet101 deals with the problem by shortcut connection as the network's depth increasing. ResNet101 includes 33 building blocks and each block has three layers, adding the first convolution layer and fully connect layer, finally 101 layers. Compared with VGG16's 16 layers, ResNet101 is a quite deeper bottleneck architecture network. As our flow chart shows, the ResNet backbone can be summarized as five bottlenecks that is conv1, conv2_x, conv3_x, conv4_x and conv5_x. The conv1 layer uses a 7x7 kernel to convolute the inputting image and the stride is 2, and its depth is 64, so the conv1 output 512x512x64 layers. After conv1 layer, a 3x3 max polling with stride 2 is adopted to down sampling and select features which are easy to distinguish and reduce some parameters. The conv2_x-conv5_x have some same configurations. That includes three layers: (1) a 1x1 kernel convolution; (2) a 3x3 kernel convolution; (3) a 1x1 kernel convolution. The first 1x1 convolution kernel is made use of reducing the number of channels to 1/4, corresponding, the third 1x1 convolution kernel is employed to recover the number of channels. Those two convolutional layers effective decrease the computing cost and parameters. The second convolution kernel doesn't change the number of channels. The conv2_x has three bottlenecks; con3_x has 4 this structure; con4_x has 23 building blocks; and con5_x has 3 bottlenecks. Different building blocks or bottlenecks connect to the resent. The con2_x outputs 256x256x256 dimension feature maps, the reason why divides 2 is that the stride of each first convolution layer is 2 and for improving the application ability, the number of each feature map channels multiple 2. Alike, the conv3_x outputs 128x128x512; the conv4_x outputs 64x64x1024, and the conv5_x outputs 32x32x2048, that is, the final feature map shrinks 32 times of the imputing image. By the way, the batch normalization (BN) layer also used in the network. However, the original Faster RCNN only takes the conv4_x feature maps as the boundary box coordinates regression and conv5_x feature maps as the input of the RPN network, which results of those small goals cannot be detected and located effectively, due to deeper layer feature maps loss plentiful small goals spatial and semantic informations.

For fixing this question, we modified feature prime networks of ResNet101 [31] to keep the information of shallow layers feature maps. These convolution layers will be used to produce FPN [32]. We designed three kinds of FPN networks of 4 layers (P2, P3, P4, P5), 5 layers (P2, P3, P4, P5, P6)

and our 6 layers (P2, P3, P4, P5, P6, P7), as shown as TAB1. From the result, we finally choose the 6 layers FPN network. The P2 layer is composed of corresponding conv2_x by 1x1 convolution kernel plus P3 layer with up sampling, and P2 dimension is 256x256x256. P3 is made up of corresponding conv3_x by 1x1 convolution kernel add P4 with a 2x up sampling, and P3 dimension is 128x128x256. Similarly, P4 is 64x64x256 and P5 is 32x32x256. For the last two layers, P6 is obtained by convolving P5 with 3x3, stride = 2, that is, the down sampling, so the size of P6 is: 16x16x256. The same operation for P7, so P7 is 8x8x256. The reason to design the P2-P4 is that helps the network to find and location the small lesion, because of the shallow layer containing more small goals information. And as the backbone becoming deeper, the small objects information will dismiss because of the convolution and polling, and there are only those big targets in the deep feature map layer. On the other hand, the cause for us to add P6 and P7 layers is that help detect the large lesions. And according to the size of the anchor box, we cautiously choose 6 anchor base size for each feature pyramid network layers via K-means clustering algorithm. And though this program, we got six anchor sizes: [10, 11], [15, 16], [22, 23], [33, 33], [50 49], [77, 78]. Obviously, even the biggest anchor size is not the large goal. Because the official Pascal VOC data set has various classes that has to including all anchor sizes but those anchor sizes are not suitable for our data set. After we change those anchor sizes, we also change the ratio of the anchor box from [0.5, 1.0, 2.0] to [0.9, 1.0, 1.1] to apply our data set in testing process. Each size of feature map layer only uses one size of anchor box with three ratios for coordinate regression. Such as for P2, only adopts anchor size 10 with 0.9, 1.0, 1.1 ratios to predict the goals region. And P3 uses a 15 anchor size with 3 ratios sliding window (3x3 convolution kernel) and two 1x1 convolution kernels to get the lesion region proposal; P4 with a 22 anchor base size; P5 with a 33 anchor size; P6 50; P7 77. Later, all those different sizes of feature maps are handled by the RPN network.

C. MODIFY CASCADE FOUR IOU THRESHOLDS NETWORK

In the Faster RCNN, the intersection over union is used to classify the positive sample or negative sample when the RPN network inspect proposes the region of goals and ground truth and also the testing process. The IOU threshold of the official Faster RCNN code is 0.5, that means when the sample area of overlap dividing area of union is great than 0.5, this sample will be defined as positive sample. On contrast, if less than 0.5, this sample will be classified as negative sample. We set the rate of positive samples to negative samples to 1:3, that in one batch of 256 samples training at the same time, the 1/3 samples are chosen as positive samples and the 2/3 samples are randomly selected with negative samples. But if there is not enough positive samples in one batch of training, the cascade RCNN [33] algorithm will complement with others negative samples, which may easy result for over fitting. In Faster RCNN, the IOU threshold is 0.5 to prevent using negative samples to replenish positive samples, but it will take a question that lots of noise samples that have huge differences with the ground truth are classified as positive samples, that is, when the threshold is taken as 0.5, there will be more false detections, because the threshold of 0.5 will cause more background in the positive sample, which is the reason for more false detection. So that, the algorithm trained numerous of wrong samples and made the detector had a low accuracy and worse performance. Obviously, enhancing the IOU threshold [34] is seemed right method to improve the precision of the detector with dismissing some wrong positive samples and can reduce false detections. However, the detection effect is not necessarily the best. The main reason is that the higher the IOU threshold, the less the number of positive samples, so the greater risk of over fitting.

To deal with this dilemma, we added and improved cascade RCNN. We proposed a four IOU (0.5, 0.55, 0.6, 0.7) thresholds cascade RCNN, as shown in Fig. 3. As the flow chart shown, each branch RCNN has different positive samples and negative samples though different IOU thresholds. The input of the next network is the output of the last network plus the samples proposed from region proposed network with the set unique IOU threshold. The interpretation of cascade four IOU thresholds RCNN is that after the first-time boundary box coordinates regression, the IOU of the samples and the ground truth will be increased. For example, if a sample's IOU is 0.52, and after the first box coordination's regression, the IOU increased to 0.57, and then this sample is sent to the second cascade RCNN; the IOU will be improved further and closer to the Grand Truth. After four IOU thresholds (0.5, 0.55, 0.6, 0.7) cascade RCNN network, the IOU of the sample is improved a lot. That is, there are multiple headers with increasing IOU thresholds, and each level uses the box after the previous level of refinement as input. This ensures that each header can get enough positive samples, and the quality of the positive samples can be upgraded step by step.

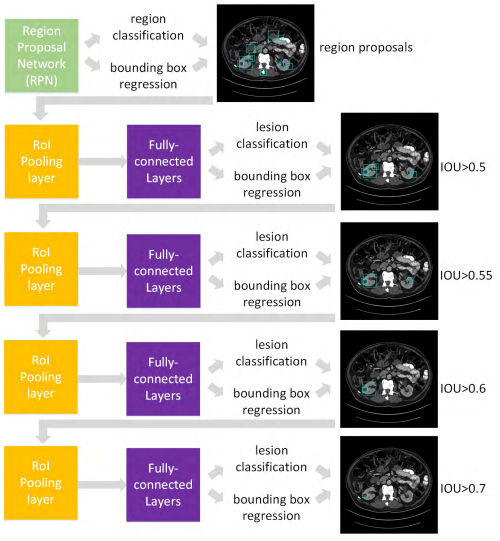


FIGURE 3. The architecture of modified cascade RCNN with four different IOU thresholds.

This operation is also working during testing. In the test, used the average of four header outputs as the final score of this proposal, which prove that the result will be further improved. The improvement we did is to develop the more cascade RCNN networks and find the IOU threshold between 0.55 to 0.65 is the most suitable for our data set, so we add one 0.55 IOU threshold RCNN between 0.5 and 0.6, the result shown the AP improved.

In conclusion, we modified the network as follows: (1) We joined two morphology operations (dilation and erosion) before the inputting image in backbone by two different convolution kernels and change the size of the network inputting image size to 1024x1024 for detecting small goals undemanding. After morphological operations, there are three paths of imputing image to the backbone. (2) Used and improve FPN. Based on ResNet101, we compared the different size of FPN, including 4 layers FPN, 5 layers FPN and our 6 layers FPN. The 4 layers and 5 layers FPN making up of P2, P3, P4, P5 and P6 show a much lower precision in detecting medium and large goals and also a litter bit lower precision in detecting small goals, shown as TAB1. As the network flow chart shows, the P6 and P7 are added to help locating the large target. But compared with ResNet101 that does not contain shallow feature maps location and semantic information, the FPN whether 4 layers and 5 layers FPN or our 6 layers FPN are have a greater performance in small, medium and large goals detection, those different size feature maps will be sent to RPN to deal with different anchor box sizes and strides which we use K-means to cluster those lesions' annotation bounding box size. (3) Modify the cascade RCNN via skillfully adding an IOU threshold RCNN. Obviously, after we improve feature pyramid network and combining with ResNet101, the accuracy increased, but there still are some false detections and miss detections, as shown in the FIG.1. We adopted the cascade multi IOU threshold RCNN to refine and correct the sample's class prediction scores and box coordinates. After the RPN output the region with a given IOU, the first step will revise the bounding box coordinates so that IOU between the sample and the ground truth can be augmenting and the sample will be closer to the annotation. The effect of IOU threshold between 0.5 to 0.7 is best, we carefully choose 0.5, 0.55, 0.6 and 0.7 for cascade four RCNN. The output of the class prediction score is chosen the biggest from the four RCNN and the bounding box regression is from the last RCNN. The result proven our change is valid.

TABLE 1. The comparison of AP small, medium, large between different layers FPN and Res101.

Network	P2	P3	P4	P5	P6	P7	AP Small	AP Medium	AP Large
Our FPN	✓	✓	✓	✓	✓	✓	0.390	0.625	0.800
4 layers FPN	✓	✓	✓	✓			0.370	0.606	0.692
5 layers FPN	✓	✓	✓	✓	✓		0.401	0.623	0.724
ResNet101							0.346	0.531	0.700

III. EXPERIMENTS AND RESULTS

A. DATASET AND MATERIALS

The DeepLesion [27] is the largest medical CT scan image data sets is published by National Institutes of Health (NIH) in Journal of Medical Imaging, aiming to help people improve the machine learning in CAD. In total, there are 4,427 unique patients collecting from 2010 in hospitals picture archiving and communication system (PACS). But we only focus on the kidney lesion detection, after discarded some noise samples, we picked out 956 kidney images from the DeepLesion and form a kidney data set. We labeled 1120 lesions annotation with renal stone, kidney cyst and hydronephrosis, and the renal cyst is the most part. The ground truth of the dataset is annotated by radiologists in clinically meaningful medical images. The slice intervals of the CT study in the data set range between 0.25 and 22.5 mm. About 48.3 percentages of them are 1 mm and 48.9 percentages are 5 mm. After NIH covert the DICOM file to 16-bit PNG file, we used MATLAB to transform the original images with no contrast to readable images by subtracting 32768 pixels. The CT scan images are saved by non-contrast pictures for high fidelity stored. The most common lesions of kidney are that renal cyst, renal stone and hydronephrosis. In our work, we focus on the lesion of kidney, in this data set, we include various types of lesion, in Fig. 4. But results from a small amount of renal stone and hydronephrosis, we use unified all annotation (lesion). As we all know, the number of the data set is a vital factor of deep learning, so it is not satisfied to train an excellent detector only with those original images. After we divided 156 images into the testing set, we used various data augments, including flipping, cropping, resizing. Flipping is the most popular data augments way to double the data. A lot of experiment showed that is an efficient method.

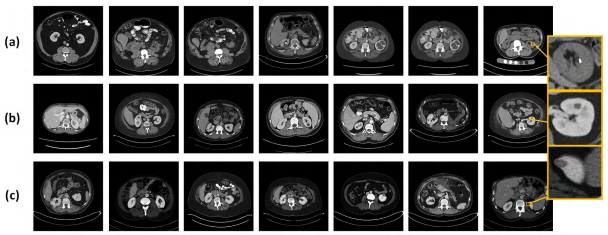


FIGURE 4. Examples of (a) kidney stone, (b) hydronephrosis, and (c) renal cyst, respectively.

Except flipping, for helping detect the small goal, we resized the image from 512x512 to 256x256, because the same lesion in the 512x512 image occupies more pixels and space than 256x256 images, that is, after we training the network with 256x256 and 512x512 images, the same kind of lesions only testing in the 512x512 images also can learn the small goals in small images which can improve the micro objects detection. In the same way, resizing 512x512 to 1024x1024 can increase the accuracy detection of large goals. And cropping the kidney in images for focusing on detecting

those big goals. For the objectivity, all those augment data only used for training.

B. EVALUATION

To quantify the effect and utility of the parameter and network proposed [35], a plentiful experiment is conducted. We divide the data set into a trainval set and a test set. The trainval set contains 800 original CT images, 800 flip CT images and 800 crop CT images, totally, 2400 images. And the valuation set occupied 20 percentage of the trainval set. The test set including 156 CT images are randomly selected by program to evaluate the performance of the algorithm, but the test set does not change the learning algorithm or parameters accordingly. The pixel of original CT scan images is 512x512. And the size of most lesions only ranges from 10x11 to 77x78, all those lesions are not the large target. So, we adjusted the anchor base size for suiting our data set.

1) HARDWARE AND SOFTWARE SETUP

Experiments were conducted on a Workstation with Intel Core i5, 2.7GHz CPU, 8GB RAM under Ubuntu 18, and a NVIDIA GTX 1080 video processing card with 8GB memory. Cascade RCNN was deployed in pytorch 1.0 framework, and Faster RCNN was conducted tensorflow-gpu 1.8. And both frameworks are based on python3.5, cuda 9.0 and cudnn 7.1.4. MATLAB 2018a is used for image pro-processing and post-experiments, such as image pixel operation, plotting, mathematical formulation, and experimental analysis.

2) TRAINING PROCESSES

Firstly, we labeled 956 CT volumes, then use data augments (flip, crop, resize). Secondly, the Faster RCNN was implemented using ResNet101 and FPN in tensorflow framework, and the cascade RCNN was trained in ResNet101 pre-trained file of ImageNet in pytorch framework. In the end, when the loss is stable, stop training and use 156 volumes without annotations to test the result. According to our hardware, we carefully choose those parameters to get better performance. The original Faster RCNN we set the training learning rate as 0.001; train momentum as 0.9; training weight decay as 0.001 after 30000 iterations; training gamma as 0.1; training batch size as 128 with mini_batch 2; IOU threshold to use as a bounding box regression training example as 0.5; anchor strides as 4, 8 and 16; and its training time is 8 hours and one second can test 8 images. The Faster RCNN+ means Faster RCNN with data augments with ResNet101 and Faster RCNN with 5 layers FPN have same training learning rate, train momentum, training weight decay and training gamma with the original Faster RCNN; training batch size is 256 with mini_batch 2; in different, the for the 5 layers FPN with different anchor stride for P2-P6 is 4, 8, 16, 32, 64 and 3 ratios (0.9, 1.0, 1.1) and the FPN channels is 256; and training time are both 10 hours and one second can test 8 images and 5 images. The Faster RCNN++ is that Faster RCNN+ plus morphological operations, and parameters of Faster RCNN++ with two branch Res101 and 5 layers FPN are same as Faster RCNN+; and their training time is 10 hours and 11 hours, and one second can test 8 images and 5 images. The cascade RCNN++ means the original cascade RCNN adds data augments and morphological operations;; and its training time is 22 hours, and one second can test 3 images. The cascade RCNN using 5 layers (P2, P3, P4, P5, P6) FPN as backbone and cascade three different IOU thresholds (0.5, 0.6, 0.7) RCNN; and its training time is 18 hours, and one second can test 3 images. The layers sequence of our proposed network is two morphological convolution layers, ResNet101 with six layers FPN backbone to extract feature maps and then four cascade RCNN used to classified and calculate boundary box coordinates. And the kernels size of two morphological and backbone convolution layers are both 3x3 with the pad of 1 and the stride of 2. And the kernel size of backbone and RPN pooling layers are both 2x2 with the stride of 2. The optimization algorithm we used is Stochastic Gradient Descent (SGD).

And the number of channels of FPN is 256; the anchor size of each layer is 4, 8, 16, 32, 64; use SGD optimizer and with a 0.02 initial learning rate and 0.0001 weight decay. The configuration of our network is that 6 layers FPN (P2, P3, P4, P5, P6, P7) with the set anchor stride that is 10 for P2; 15 for P3; 22 for P4; 33 for P5; 50 for P6; 77 for P7, those box sizes are clustered by K means, and the four unique IOU thresholds (0.5, 0.55, 0.6, 0.7) are cascaded for get a better performance detector; the learning rate is 0.005 and the weight decay is 0.0001 with 0.9 momentum to ensure the loss will not be nan during training. The training iterations of Faster RCNN with ResNet101 is 60000 iterations achieving best performance; Faster RCNN+ with ResNet101 is 45000 iterations; Faster RCNN+ with FPN is 30000 iterations; Faster RCNN++ with ResNet101 is 55000 iterations; Faster RCNN++ with FPN is 30000 iterations; Cascade RCNN++ with FPN is 45000 iterations; Our network with FPN is 20 hours with 40000 iterations. Our network has drawback in training cost and test speed shown in frames per second (FPS) as TAB2, although it is slower than the One-stage detection algorithm, it does not matter in medical images test, due to it does not need on-time detections.

3) RESULTS

Comparison our modified the cascade RCNN and others networks' results as following. We improved the cascade RCNN network, we compared 4 layers, 5 layers and 6 layers feature pyramid networks of the performance of detecting small, medium and large goals as shown in TAB1. Without doubt, our modified 6 layers FPN outperformed 4 layers, 5 layers FPN and ResNet101 (without FPN), particularly in the detection and location of large objects, because the added layers (P6 and P7) are deep layers feature maps that contain more information about large targets. The index to evaluate our network and others are AP, AUC, Spec, Sens, and F1. In the detection task, samples are classified by dichotomy to positive and negative samples. The true positive (TP) is that a sample is positive and also is predicted to a positive

TABLE 2. Comparing AUC, AP, (Specificity) Spec, (Sensitivity) Sens F1 and FPS of our proposed kidney lesion detection method with other backbones and other networks.

	backbone	AUC	AP	Spec	Sens	F1	FPS
Ours	FPN	**0.871**	**0.840**	0.855	**0.838**	**0.839**	3
Cascade RCNN++	FPN	0.821	0.819	**0.876**	0.746	0.781	10
Faster RCNN++	FPN	0.838	0.826	0.764	0.811	0.818	5
Faster RCNN++	ResNet	0.767	0.739	0.770	0.744	0.741	8
Faster RCNN+	FPN	0.815	0.778	0.868	0.776	0.777	5
Faster RCNN+	ResNet	0.745	0.718	0.795	0.750	0.734	8
Faster RCNN	ResNet	0.700	0.675	0.762	0.723	0.698	8

sample. A sample is positive, but is predicted to be a negative that is false negative (FN). And we define a sample is false positive (FP) when the sample is a negative sample, but classified as positive. The true negative (TN) is meaning the sample is a negative sample, and also classified as negative.

Firstly, we introduced the sensitive(3) and the specificity(4), those formulas are following:

$$sens = \frac{TP}{(TP + FN)} \tag{3}$$

$$spec = \frac{TN}{(FP + TN)} \tag{4}$$

The recall is called sensitive in medicine diagnose, that is the proportion of samples that are classified to be positive among all samples that are positive. In the medical testing, people commonly define those high-risk things such as lesions as positive categories. The aftereffect of omitting lesions (positive samples) is quite damaging, such as misdiagnosis, which may lead to delayed treatment of patients. In medicine diagnosis, try our best to reduce the rate of missed diagnosis, that is, to improve the sensitive is more important than to improve the precision. The specificity is that among the all negative sample, the ratio of how much negative samples can be predicted. This definition is similar with sensitive, and the difference between the two is only the object is different, the sensitive is for the positive example, and the specificity is for negative examples. The specificity is also considered an important indicator in medical testing. That is, the false detection rate is high as the specificity low. Therefore, in the medicine diagnose, specificity and sensitivity need to be considered at the same time. The precision(5) is calculated by this formula:

$$prec = \frac{TP}{(TP + FP)} \tag{5}$$

$$F1 = \frac{2 * P * R}{(P + R)} \tag{6}$$

Based on the recall and precision results, we plotted RP curve with recall as the transverse axis, precision as the longitudinal axis to measure the generalization ability of those network, shown as Fig.5, and the area under the RP curve is the average precision (AP) to balanced evaluate recall and precision. From the PR curve, it is obvious that our modified network has a better performance than others, that is, more

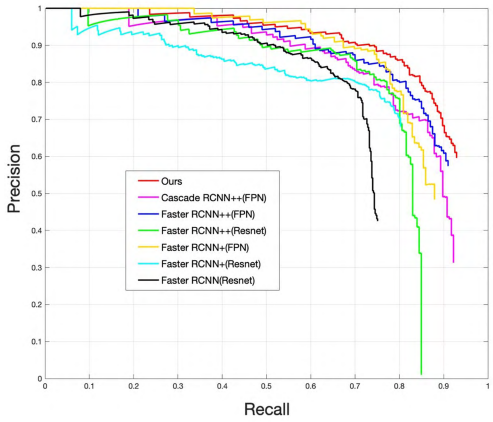


FIGURE 5. P-R curves of various methods on the test set of kidney lesion CT images.

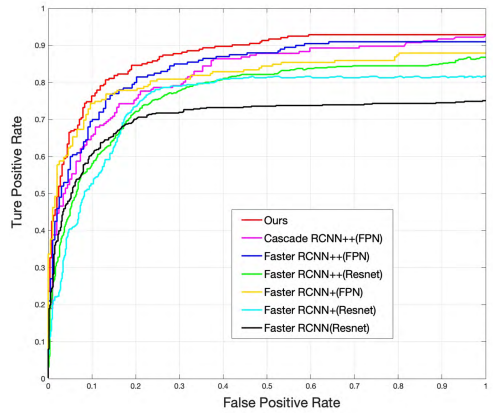


FIGURE 6. ROC curves of various methods on the test set of kidney lesion CT images.

TABLE 3. Comparing the detection confidence of case CT images of various algorithms.

	Ours	Cascade RCNN++	Faster RCNN++ (FPN)	Faster RCNN++ (ResNet)	Faster RCNN+ (FPN)	Faster RCNN+ (ResNet)	Faster RCNN (ResNet)
1	1.00	0.99	0.96	0.96	0.99	0.99	1.00
2	0.94/1.00	0.66/0.97	0/0.84	0/0.94	0/0.71	0/0.59	0/0.97
3	0.96	0.66	0.69	0.92	0.79	0	0
4	1.00/0	0.96/0	0.91/0.77	0.75/0	0.97/0	0.803/0	0.99/0.72
5	0.67	0.91	0.85	0	0	0	0.94
6	0.99/0.99	0.98/0.96	0.79/0	0.91/0.83	0.99/0.95	0.99/0	0.99/0.99
7	1.00	0.99	0.99	0.85	0.99	0.94	0
8	0.92/0.52	0.73/0	0.85/0	0/0	0.65/0	0.69/0	0/0

protruding at the top right. Similarly, in order to evaluate the receiver operating characteristic plotted receiver operating characteristic (ROC) curve with false positive rate (FPR) as the transverse axis and true positive rate (TPR) as the longitudinal axis in Fig.6, and from this ROC curve, it shows that our network (red line) is most outstanding in top left. And the area under ROC curve is the AUC which quantifies the classification ability of ROC curve. The classification ability is closely related to probability, threshold value. The better the classification ability (the greater the AUC), the more

FIGURE 7. Lesion detection results for sample kidney CT images of various methods. (a) The original image. (b) Ground truth. (c) Our network. (d) Cascade RCNN++. (e) Faster RCNN++(FPN). (f) Faster RCNN++(Resnet). (g) Faster RCNN+(FPN). (h) Faster RCNN+(Resnet). (i) Faster RCNN(Resnet).

reasonable the output probability, the more credible the result of sorting. We also used F1- Measure (6) to comprehensive measure those algorithms. All those comparison are shown in TAB2. In a sum, our modified model has achieved greater results of AUC, AP, Sensitive and F1. Among them, our proposed method with 0.871 of AUC, 0.840 of AP, 0.838 of sensitive and 0.839 of F1-Measure outperforms than others model. And the Faster RCNN++ with FPN has a better Specificity of 0.876. For visualizing the performance of those algorithms, we selected 8 CT scan volumes and test the

inference process to present our results, as shown in Fig. 7. We displayed all confidence of each lesions detection result shown as tab3. Without doubt, our modified model has a comprehensive and accurate score among them.

IV. DISCUSSION AND CONCLUSION

In sum, medical image lesion detection by RCNN shows advantages in timesaving and labor-saving in location and analysis lesions. Previous researches proven deep learning is a practicable algorithm in CAD, but those cannot meet the requirement of accuracy and comprehensiveness. In this study, we proposed morphological cascade convolutional neural networks, which is a high precise and robust detector of kidney lesions in CT scan image. In our network, two kinds of morphological operations were proposed firstly to make small goals more conspicuous and easier detect. Secondly, we modified a six layers FPN to generate different sizes feature maps for overall location and classification with our set anchor sizes and ratios. Finally, we developed a four IOU threshold cascade RCNN to archive high precision detection. The validation experiments were conducted on CT image, and the results proven our proposed cascade RCNN is greater outperforming in detecting lesions. However, there also some questions in medical image test. Our network has a low precision when there are many and complicate goals such as polycystic kidney. And in some complex images and lesions, there are some misdetections and false detections that need to improve. Despite those, comparing with those state-of-art algorithms, our modified network is an efficient and accurate detector, and is also of huge application value in other lesion or organ detection.

REFERENCES

[1] Y. Xie, B. Bowe, and H. A. Mokdad, H. Xian, Y. Yan, T. Li, G. Maddukuri, C. Tsai, T. Floyd, and Z. Al-Aly, "Analysis of the global burden of disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016," *Kidney Int.*, vol. 94, no. 3, pp. 567–581, Sep. 2018.

[2] J. Jiang, P. Trundle, and J. Ren, "Medical image analysis with artificial neural networks," *Comput. Med. Imag. Graph.*, vol. 34, no. 8, pp. 617–631, Dec. 2010.

[3] M. Noll, X. Li, and S. Wesarg, "Automated kidney detection and segmentation in 3D ultrasound," in *Proc. Workshop Clin. Image-Based Procedures*, Sep. 2014, pp. 83–90.

[4] M. Zhang, T. Wu, and K. M. Bennett, "A novel Hessian based algorithm for rat kidney glomerulus detection in 3D MRI," *Proc. SPIE, Med. Imag., Image Process. Int. Soc. Opt. Photon.*, vol. 9413, Mar. 2015, Art. no. 94132N9.

[5] M. Shehata, F. Khalifa, A. Soliman, M. Ghazal, F. Taher, M. A. El-Ghar, A. C. Dwyer, G. Gimel'farb, R. S. Keynton, and A. El-Baz, "Computer-aided diagnostic system for early detection of acute renal transplant rejection using diffusion-weighted MRI," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 2, pp. 539–552, Feb. 2019.

[6] P. T. Akkasaligar and S. Biradar, "Diagnosis of renal calculus disease in medical ultrasound images," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC)*, Dec. 2017, pp. 1–5.

[7] V. F. van Ravesteijn, C. van Wijk, F. M. Vos, R. Truyen, J. F. Peters, J. Stoker, and L. J. van Vliet, "Computer-aided detection of polyps in CT colonography using logistic regression," *IEEE Trans. Med. Imag.*, vol. 9, no. 7, pp. 120–131, Jan. 2019.

[8] S. E. Grigorescu, S. T. Nevo, M. H. Liedenbaum, R. Truyen, J. Stoker, L. J. van Vliet, and F. M. Vos, "Automated detection and segmentation of large lesions in CT colonography," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 675–684, Mar. 2010.

[9] J. Zhou and J. Qi, "Adaptive imaging for lesion detection using a zoom-in PET system," *IEEE Trans. Med. Imag.*, vol. 30, no. 1, pp. 119–130, Jan. 2011.

[10] W. K. Moon, Y.-W. Shen, M. S. Bae, C.-S. Huang, J.-H. Chen, and R.-F. Chang, "Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1191–1200, Jul. 2013.

[11] Y. Li, "Detecting lesion bounding ellipses with Gaussian proposal networks," 2018, *arXiv:1902.09658*. [Online]. Available: https://arxiv.org/abs/1902.09658

[12] L. Zhang, L. Lu, R. M. Summers, E. Kebebew, and J. Yao, "Convolutional invasion and expansion networks for tumor growth prediction," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 638–648, Feb. 2018.

[13] F. Khalifa, G. M. Beache, M. A. El-Ghar, T. El-Diasty, G. Gimel'farb, M. Kong, and A. El-Baz, "Dynamic contrast-enhanced MRI-based early detection of acute renal transplant rejection," *IEEE Trans. Med. Imag.*, vol. 32, no. 10, pp. 1910–1927, Oct. 2013.

[14] S. Zheng, J. Guo, X. Cui, R. N. J. Veldhuis, M. Oudkerk, and P. M. A. van Ooijen, "Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection," *IEEE Trans. Med. Imag.*, to be published.

[15] Y.-B. Tang, K. Yan, Y. Tang, J. Liu, J. Xiao, and R. M. Summers, "ULDOR: A universal lesion detector for CT scans with pseudo masks and hard negative example mining," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, to be published.

[16] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.

[17] A. Arnaud, F. Forbes, N. Coquery, N. Collomb, B. Lemasson, and E. L. Barbier, "Fully automatic lesion localization and characterization: Application to brain tumors using multiparametric quantitative MRI data," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1678–1689, Jul. 2018.

[18] V. B. Kolachalama, P. Singh, C. Q. Lin, D. Mun, M. E. Belghasem, J. M. Henderson, J. M. Francis, D. J. Salant, and V. C. Chitalia, "Association of pathological fibrosis with renal survival using deep neural networks," *Kidney Int. Rep.*, vol. 3, no. 2, pp. 464–475, Mar. 2018.

[19] P. Danaee, R. Ghaeini, and D. A. Hendrix, "A deep learning approach for cancer detection and relevant gene identification," in *Proc. Pacific Symp. Biocomput.*, vol. 22. 2016, pp. 219–229.

[20] D. Turco, S. Severi, R. Mignani, V. Aiello, R. Magistroni, and C. Corsi, "Reliability of total renal volume computation in polycystic kidney disease from magnetic resonance imaging," *Acad. Radiol.*, vol. 22, no. 11, pp. 1376–1384, Nov. 2015.

[21] A. Ben-Cohen, E. Klang, A. Kerpel, E. Konen, M. M. Amitai, and H. Greenspan, "Fully convolutional network and sparsity-based dictionary learning for liver lesion detection in CT examinations," *Neurocomputing*, vol. 275, pp. 1585–1594, Jan. 2018.

[22] R. Cuingnet, R. Prevost, D. Lesage, L. D. Cohen, B. Mory, and R. Ardon, "Automatic detection and segmentation of kidneys in 3D CT images using random forests," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent.*, Oct. 2012, pp. 66–74.

[23] M. Zhang, T. Wu, S. C. Beeman, L. Cullen-McEwen, J. F. Bertram, J. R. Charlto, E. Baldelomar, and K. M. Bennett, "Efficient small blob detection based on local convexity, intensity and shape information," *IEEE Trans. Med. Imag.*, vol. 35, no. 4, pp. 1127–1137, Apr. 2016.

[24] D. Mahapatra, P. J. Schüffler, J. A. W. Tielbeek, J. C. Makanyanga, J. Stoker, S. A. Taylor, F. M. Vos, and J. M. Buhmann, "Automatic detection and segmentation of Crohn's disease tissues from abdominal MRI," *IEEE Trans. Med. Imag.*, vol. 32, no. 12, pp. 2332–2347, Dec. 2013.

[25] K. Yan, X. Wang, L. Lu, L. Zhang, A. P. Harrison, M. Bagheri, and R. M. Summers, "Deep lesion graphs in the wild: Relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database," in *Proc. IEEE CVPR*, Jun. 2018, pp. 9261–9270.

[26] D.-T. Lin, C.-C. Lei, and S.-W. Hung, "Computer-aided kidney segmentation on abdominal CT images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 59–65, Jan. 2006.

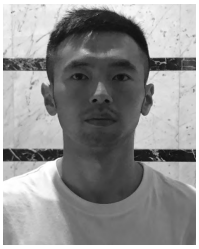
[27] K. Yan, X. Wang, L. Lu, and R. M. Summers, "DeepLesion: Automated mining of large-scale lesion annotations and universal lesion detection with deep learning," *Proc. SPIE*, vol. 5, no. 3, Jul. 2018, Art. no. 036501.

[28] L. Liu, Z. Jia, J. Yang, and N. K. Kasabov, "SAR image change detection based on mathematical morphology and the K-means clustering algorithm," *IEEE Access*, vol. 7, pp. 43970–43978, 2019.

[29] S. Hossain, S. Najeeb, A. Shahriyar, Z. R. Abdullah, and M. A. Haque, "A pipeline for lung tumor detection and segmentation from CT scans using dilated convolutional neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1348–1352.
[30] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.
[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE CVPR*, Jun. 2016, pp. 770–778.
[32] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE CVPR*, Jul. 2017, pp. 2117–2125.
[33] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into high quality object detection," in *Proc. IEEE CVPR*, Jun. 2018, pp. 6152–6162.
[34] K. Wang, Y. Dong, H. Bai, Y. Zhao, and K. Hu, "Use fast R-CNN and cascade structure for face detection," in *Proc. IEEE VCIP*, 2016, pp. 1–4.
[35] D. Shen, G. Wu, and H. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, Jun. 2017.

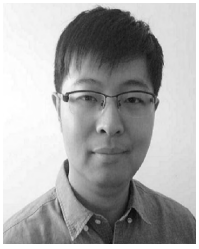
HUI ZHANG received the B.S., M.S., and Ph.D. degrees in pattern recognition and intelligent system from Hunan University, Changsha, China, in 2004, 2007, and 2012, respectively. He is currently an Assistant Professor with the College of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha. He is currently a Visiting Scholar with the CVSS Laboratory, Department of Electrical and Computer Engineering, University of Windsor. His research interests include machine vision, deep learning, and medical image processing.

YURONG CHEN was born in Liling, Hunan, China, in 1998. He received the B.S. degree from the School of Electrical and Information Engineering, Changsha University of Science and Technology. He is currently pursuing the degree with the Swanson School of Engineering, University of Pittsburgh. His main research interests include machine learning and medical image processing.



YANAN SONG received the bachelor's degree from the Henan University of Technology, in 2017. He is currently pursuing the degree with the School of Electrical and Information Engineering, Changsha University of Science and Technology. His main research interests include rail visual inspection and image processing.

ZHENLIN XIONG received the bachelor's degree from Hubei Engineering University, in 2017. He is currently pursuing the degree with the School of Electrical and Information Engineering, Changsha University of Science and Technology. His main research interest includes medical image processing.



YIMIN YANG (S'10–M'13) received the Ph.D. degree in electrical engineering from Hunan University, China, in 2013. He is currently an Assistant Professor with the Computer Science Department, Lakehead University, Thunder Bay, ON, Canada. From 2014 to 2018, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. He has authored or coauthored more than 40 refereed papers. His research interests include artificial neural networks, data processing, and robotics. He was a recipient of the Outstanding Ph.D. Thesis Award of Hunan Province, and the Outstanding Ph.D. Thesis Award Nominations of the Chinese Association of Automation, China, in 2014 and 2015, respectively. He has been serving as a Reviewer for international journals of his research field, a Guest Editor of multiple journals, and a Program Committee Member of some international conferences.

Q. M. JONATHAN WU received the Ph.D. degree in electrical engineering from the University of Wales, Swansea, U.K., in 1990. He was with the National Research Council of Canada for ten years beginning in 1995, where he became a Senior Research Officer and a Group Leader. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. He is a Visiting Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. He has published more than 300 peer-reviewed papers in computer vision, image processing, intelligent systems, robotics, and integrated microsystems. His current research interests include 3-D computer vision, active video object tracking and extraction, interactive multimedia, sensor analysis and fusion, and visual sensor networks. He holds the Tier 1 Canada Research Chair in Automotive Sensors and Information Systems. He is an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS and the *Cognitive Computation*. He has served on technical program committees and international advisory boards for many prestigious conferences.