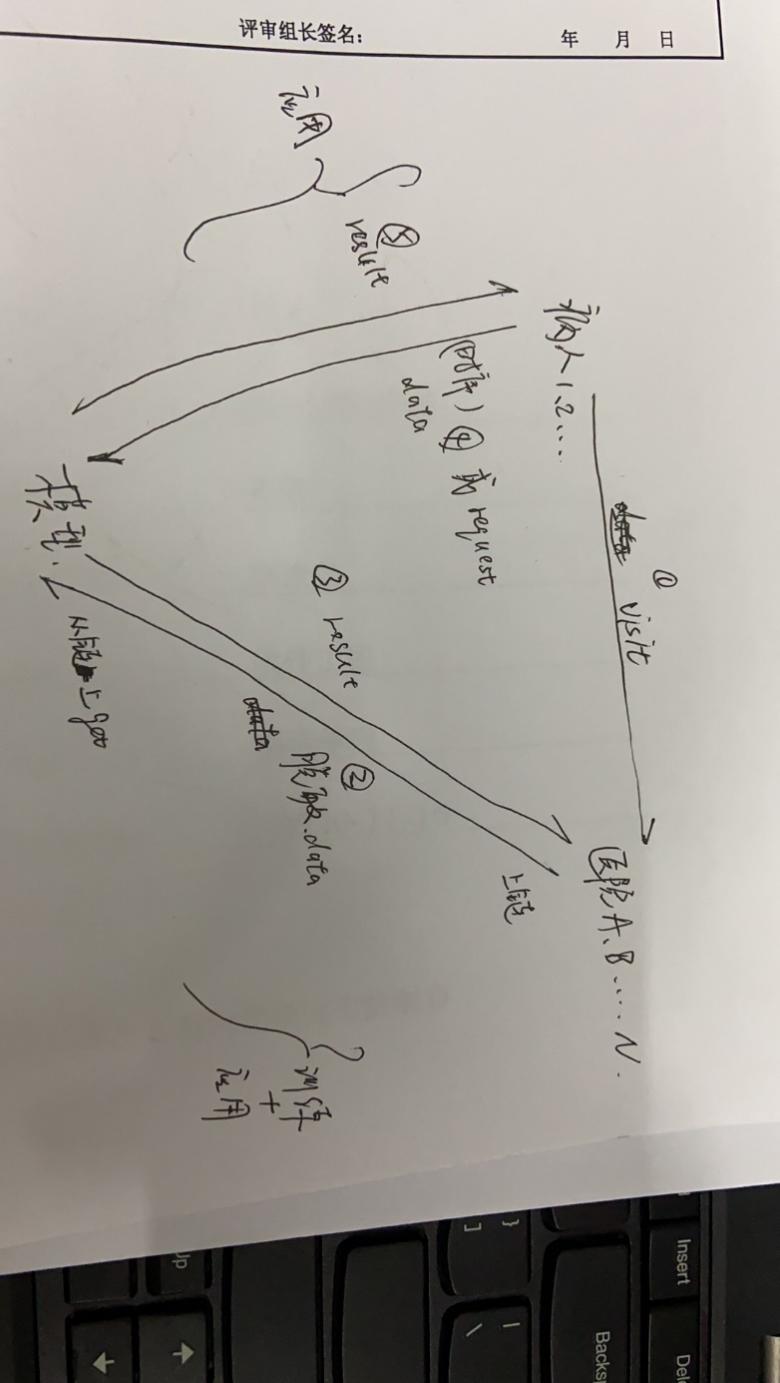
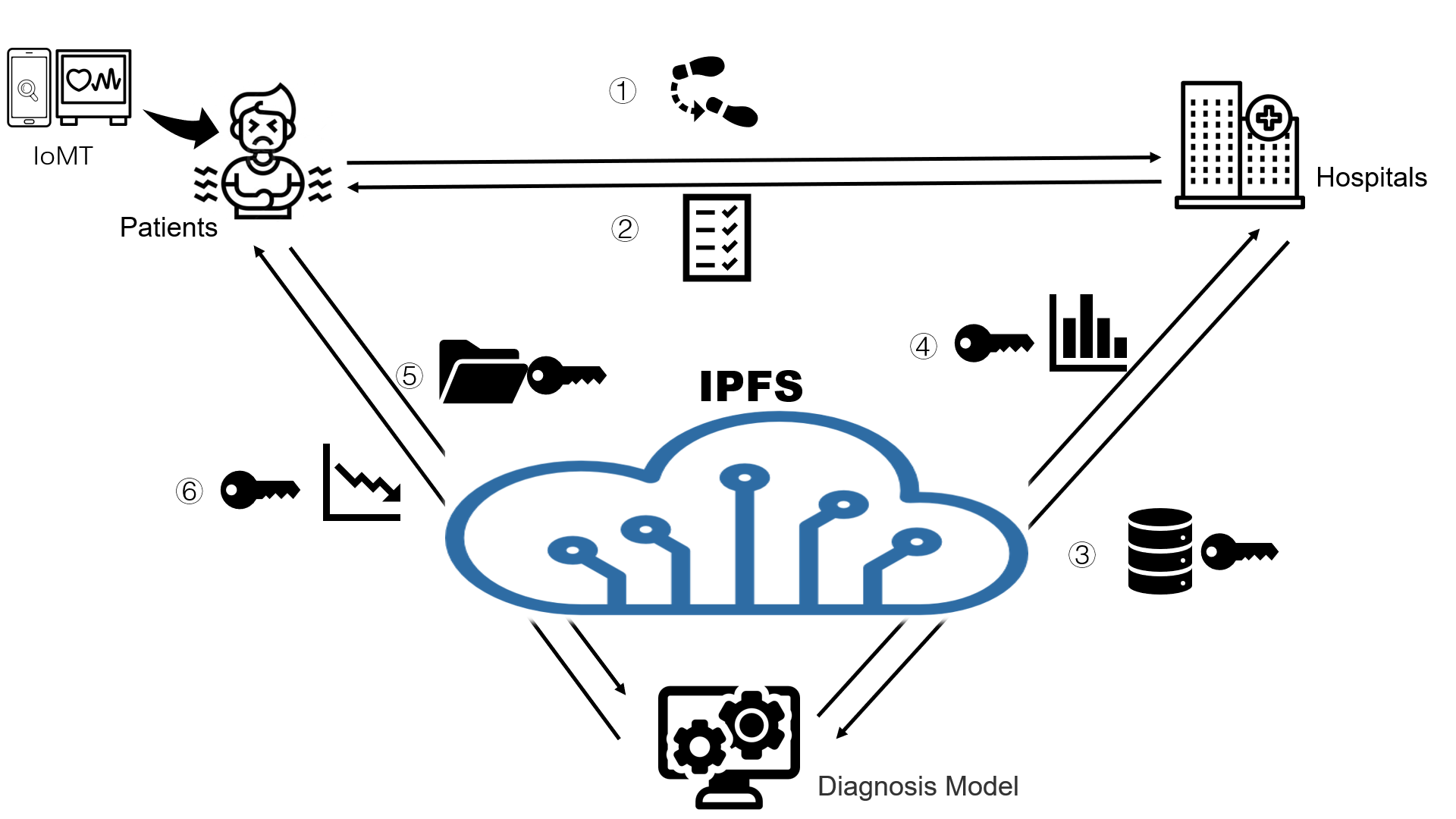
**在所有开始之前，需要先想明白要如何讲这个故事？因为这篇文章有两个重点，一个是机器学习方法帮助快速大量诊断病情（classification），第二个是区块链帮助数据&模型传输及共享，这两个点都需要照顾到。所以所有的section可以说都需要echo to/照顾到这两个点，一个都不能少，且存在重点的转移/侧重点不同。勿过度啰嗦，也勿忽视任意一个。还是一个section一个section地写吧，不要一个point一个point地写。整体逻辑链如下：**

1. **大背景（background，此处写这个病和医疗方面的现状，之前专利和会议的内容可以借鉴），存在着什么样的问题（人力消耗大，其他诊断方式有辐射、费用高，医疗资源不平衡；医疗数据较敏感，数据不透明，诊断方式及结果不共享，每次换医院历史数据难以使用难以追踪）。这段全都作为introduction的第一部分，即background。也可以一起放在introduction的第一段。**
2. **deep learning/machine learning/AI被广泛地应用于各种方面including医疗方面。（如果肾积水+dl的文献不够用的话，前面可以先插一些其他病+dl。）在肾积水方面，文献xxxxx。**
3. **区块链是xxx，区块链和物联网的关系，它能够xxxx，它的xxx特性也因此使它被广泛用来xxxx。例如，文献xxxx。除此之外，物联网相关的其他技术，例如云计算、雾计算、边缘计算、iomt等物联网相关概念/方法/技术/applications也随着5g时代的推广/到来而被广泛地应用于医疗数据中。**
4. **在过去，同时integrate区块链和dl并且applied to medical data的不多，并且多数集中在xxxx（这里就把柳叶刀的相应部分写进来就行，它写的是“理论方面”）（如果贴切的话，可以举几个例子，文献xxx等。可以直接把柳叶刀那部分引用的文献拿过来，但是最好再对比一下）。我们build了一种方法以解决上述提到的问题。**
5. **几句话介绍一下方法，并讲一下contribution，最后说一下文章的arrangement。**



****

**（首先解决一个定位问题：我认为我们仍然是可以说multiparty的。虽然我的DL模型是根据唯一一个医院的数据来train和validate的，但是我的blockchain frame是在仿真的multiparty上验证过的，是没问题的，是可以推广到real word的multiparty的。）**

**方法简介：integrate了一个DL模型和一个区块链frame/platform，在肾积水超声图像上跑了实验并验证了结果。**

**Contribution：**

1. **我提出的DL模型能够快速地根据无辐射的便宜的超声图像classify肾积水，并达到了xx的acc、auc、roc，超过了常见的abc模型，能够为医生对肾积水的诊断省好多事儿，能给病人省钱。**
2. **我提出的blockchain frame能够帮助sensitive data的线上存储和传输，为未来可能的multiparty间的数据共享、联邦学习、合作等奠定了技术基础/提供了技术支持。**
3. **我使用的Blockchain platform能帮助病人安全地存储自己的信息，且能够在不去医院/换了个医院的情况下随时获取/追踪自己的相关病史（前面的部分，是已经去医院看过医生的数据能够随时获取。后面的部分，是未来自己在家里如果有设备的话可以随时上传一张照片就得到诊断结果&不需担心泄露自己的隐私&这个结果因为存在了线上所以日后如果又去看医生的话医生还能看到这个自己拍的片子）；moreover，DL和区块链的结合更是能够起到线上AI医生的效果，无需去医院，只要一张超声照片，既可得到准确率高达xxx的诊断。**

**以上这些，是逻辑链，但是也是introduction**

**Title：区块链平台，5g，医疗数据，iomt**

**title：区块链平台，5g，医疗数据，iomt。**

**A deep learning classification model of hydronephrosis ultrasound image based on blockchain platform. 一个基于区块链平台的肾积水超声图像深度学习分类模型。**

**a deep learning method/algorithm for upjo ultrosonic image classification integrated with blockchain frame/platform.** **结合区块链平台的upjo超声图像分类深度学习方法/算法。**

**deep learning-blockchain-integrated framework to diagnosis UPJO using ultrasound image**

**Authors：我，文棚嶒，李建强，公备，医生们，彭浩然**

**Abstract：**

**Introduction：**

As a common kidney disease in children, hydronephrosis is generally congenital [1], [2], and has increasing morbidity above 1% [3], [4], [5]. Ureteropelvic junction obstruction (UPJO), the obstruction at or along the pelvic-ureteral junction, being the main causation of children hydronephrosis, can result in quite a few symptoms like nausea and vomiting, abdominal pain, abdominal mass, urinary tract infection, hematuria, uremia, hypertension, and other progressive loss of renal function such as uremia or even renal rupture.

Been a commonly adopted examination in the medical field, as well as a preliminary diagnostic step for UPJO, ultrasonography is economical, radiationless, meanwhile noninvasive. Yet, since the ultrasonic images are high-noise, it is laborious for even trained sonographers to obtain enough information.

With the development of computing power, artificial intelligence (AI) technology has been applied to numerous regions, especially medical image analysis. It can greatly assist for the diagnostic ability of doctors, meanwhile alleviate the medical resources imbalance and the shortage in less-developed regions. \cite{he2019practical}, plus facilitate doctors' diagnostic abilities\cite{gulshan2016development,kermany2018identifying,esteva2017dermatologist,cheng2016computer}. As a subset of AI, deep learning can learn from source data, find out the important features to draw the conclusion automatically. If we can combine AI medical imaging technology and automatically distinguish the diseased regions during the ultrasound examination stage, we can not only support urologist surgeons for further diagnosis but also save lots of medical resources, manpower, money, and help the burdened patients.

In the interdisciplinary field of artificial intelligence and medical image processing, medical data is usually digitized (i.e., electronic medical record (EMR) or electronic health record (EHR)) and transmitted on the network. While in this kind of medical data sharing, a key challenge is the privacy protection of patients' sensitive information and the safe sharing of data by multi-party organizations. As a shared ledger that provides a secure, decentralized, transparent, and trusted manner for data management, blockchain is been more extensively applied in medical data with the arrival of the 5G era.

Previous literature that integrates deep learning and blockchain in the medical field always focuses on general and conceptual construction. We, however, not only propose a deep learning model, xxxnet, that can diagnose the ultrasonic images of UPJO, but create a blockchain framework to integrate with to guarantee the secure transferring between different parities. The performance of this framework and this integration is also validated and reported.

方法简介：integrate了一个DL模型和一个区块链frame/platform，在肾积水超声图像上跑了实验并验证了结果。

Contribution

Our proposed xxxnet can identify the UPJO severity according to the high-noised, cheap, nonradiative, and noninvasive ultrasound image, and has an accuracy of 0.9177, which is higher than other benchmark models.

在特征工程上，我们提出了一种结合离散小波变换的工程方法，替换池化层并增加模型特征数作为输入供算法和模型使用，且该方法不局限于在ResNet上实现变体，印证了新型特征对预测结果有正向且明显的影响。

The blockchain framework we formed ensures the online storage and transmission of sensitive data and provides technical support for further federated learning and cooperation among multi-parties.

The blockchain framework enables patients to safely store and track their medical history whenever and wherever even if they changed hospitals, meanwhile helps to build trust between institutions owing to its auditability and traceability.

Moreover, the combination of the deep learning model and blockchain framework can play the role of an online AI doctor. Without going to the hospital, without worrying about security, the patient can get a diagnosis with an accuracy rate of more than 91% by just uploading an ultrasound image to our AI-Blockchain platform, which leaves the possibility for the linkage with the Internet of Medical Things (IoMT) like portable and ultra-portable ultrasound scanners.

For a better understanding of this article, our organization is as below: In the second section, we briefly review the literature on feature extraction networks and attention mechanisms. In the third section, we present our overall network structure and the optimization method adopted. In the fourth section, we introduce the details of our data set, evaluation criteria, and network training. In the fifth section, we compare our model with other networks and perform ablation experiments, then explain the results. Finally, we discuss and summarize the whole paper.

**Method：（放一个整体的结构，然后小标题放流程。3000字）****methodology**

**Study design：**

In this study, we develop a (computer-aided/deep learning) diagnosis model that combined a semantic segmentation network and a classification network to grade the UPJO from hydronephrosis ultrasound images.

Besides, we construct a blockchain framework to integrate with our CAD/diagnosis model for secure transmission between separate parties.

We then separately demonstrate these parts as following.

CAD model

In this section, we will introduce our pathological grading scheme of UPJO ultrasound images, the architecture is shown in Figure 2. This scheme combines a semantic segmentation algorithm A- PSPNet with a classification algorithm Wavelet-CNN (hereinafter referred to as segmentation and classifier) to increase the available representation of the image and realize the final classification of UPJO.

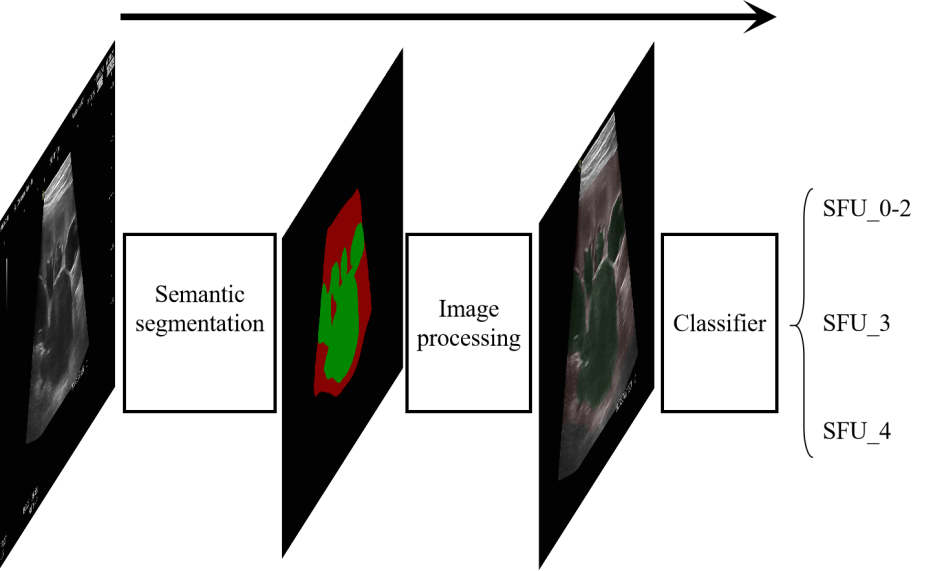


图 2. 图示为整体的肾超声图像病理分级系统，输入为原图，输出为病理分级结果

A- PSPNet: a semantic segmentation network

Since the ultrasonic image has high and dense noise, the accuracy of direct classification will be unsatisfactory. Hence it is necessary to supplement feature representations that are different from noise. Our segmentation network adopts a residual Pyramid Scene Parsing (PSPNet) network and an attention mechanism Convolutional Block Attention Module (CBAM). The overall structure is shown in Figure 3. The attention layer will help the network ignore unnecessary features and focus on important features to improve the segmentation accuracy. The parsing structure of PSPNet has multilayer perceptrons, which enable the network to obtain more global feature information, and then fuse with the basic features to help distinguish different regions.

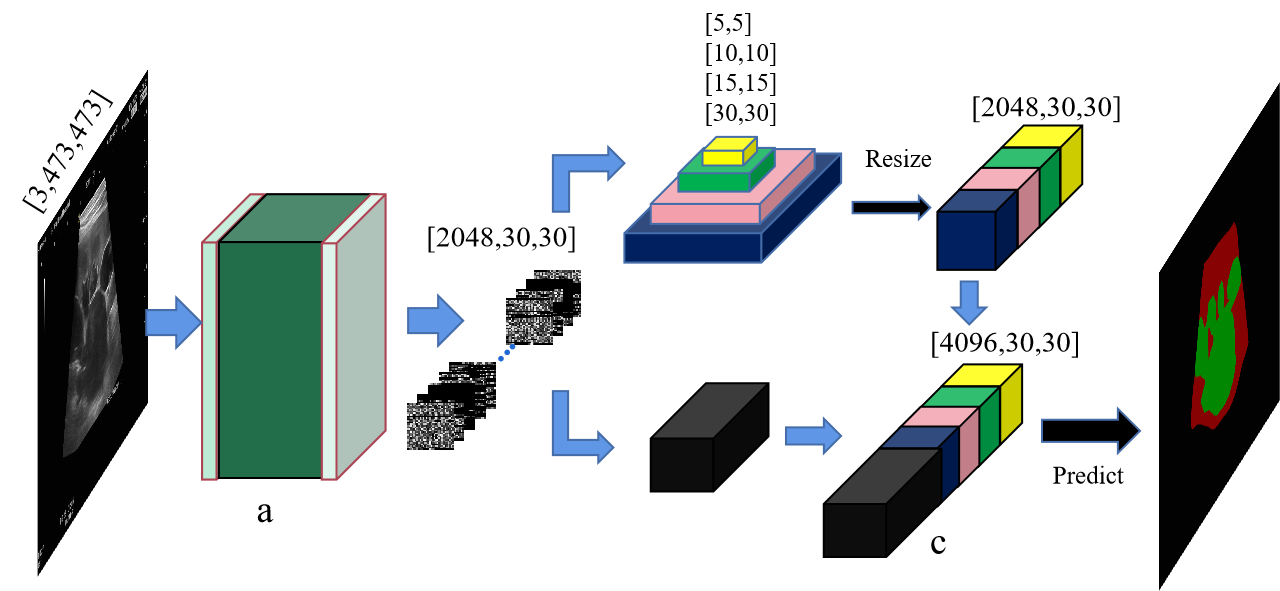


图 3.图示为分割器的网络结构图.输入为原图，输出以颜色区分的分割结果,a为嵌入注意力层的Resnet结构，b为金字塔池化结构，c模块为ab模块的输出特征的整合结构

Figure 3. The structure of the A-PSPNet. The input is the original image, and the output is the segmentation result distinguished by color. a. the Resnet structure embedded in the attention layer, b. the pyramid pooling structure. c. the integration structure of the output features of ab module

Using denotes the input. The original input image is decomposed into basic feature representation and advanced feature representation . represents the pooling scales in pyramid pooling and can further adjust the feature size parameters for concatenate. And the two kinds of feature maps are concatenate to form the c structure in Figure 2 as , where stands for the concatenate processing.

After that, the concatenated feature layers are processed, predicted and colored to output the segmentation results. As shown in Figure 4, the concatenated features are sent through a bottleneck layer to adjust the number of channels, then each pixel is predicted through a softmax function and colored. The final segmentation results are shown as the output in Fig. 3.

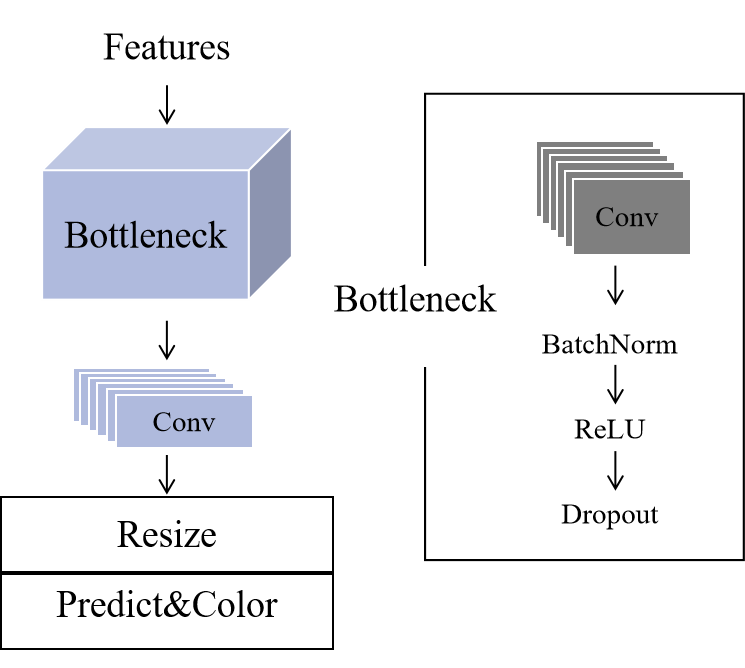


图 4 图示为特征堆叠完成后的降维和预测步骤 The dimensionality reduction and prediction steps after feature concatenating

b Classifier

As mentioned previously, we add a classifier to acquire the UPJO pathological grading. In traditional classification networks, the pooling layer employs the down-sampling to decompose images to reduce the total amount of calculation, which however results in the loss of information. Since both wavelet transformation and its inverse operation are reversible thus guarantee no information loss, to reduce the complexity without losing information, Discrete Wavelet Transformation (DWT), which is deemed as an effective operator to decompose images and obtain the desired frequency-domain information, is utilized to the classifier to replace the traditional pooling layer.

For better understanding of the process in the two-dimensional (2D) wavelet transformation, we start with the one-dimensional (1D) ones. First, we define two basic functions, a wavelet function and a scale function . For 1D discrete sequence , its forward discrete wavelet series expansion coefficient is

In the above equations, and are samples that use equal intervals on the support region of the basis function on the scale.

,

determines the position in a given direction, determines the width of the sample, is any determined starting scale, while the obtained expanded set is a subset of , and final transformation itself is composed of coefficients.

Similar to 1D DWT, a 2D DWT use 2D scale function and wavelet function. We first take 1D transformation of row of the 2D array, and then take 1D transformation of column of the last step result. For one 2D image of size , its DWT is

Finally we get four half-sized output images: , which represent average, horizontal, vertical, and diagonal information from the input source image.

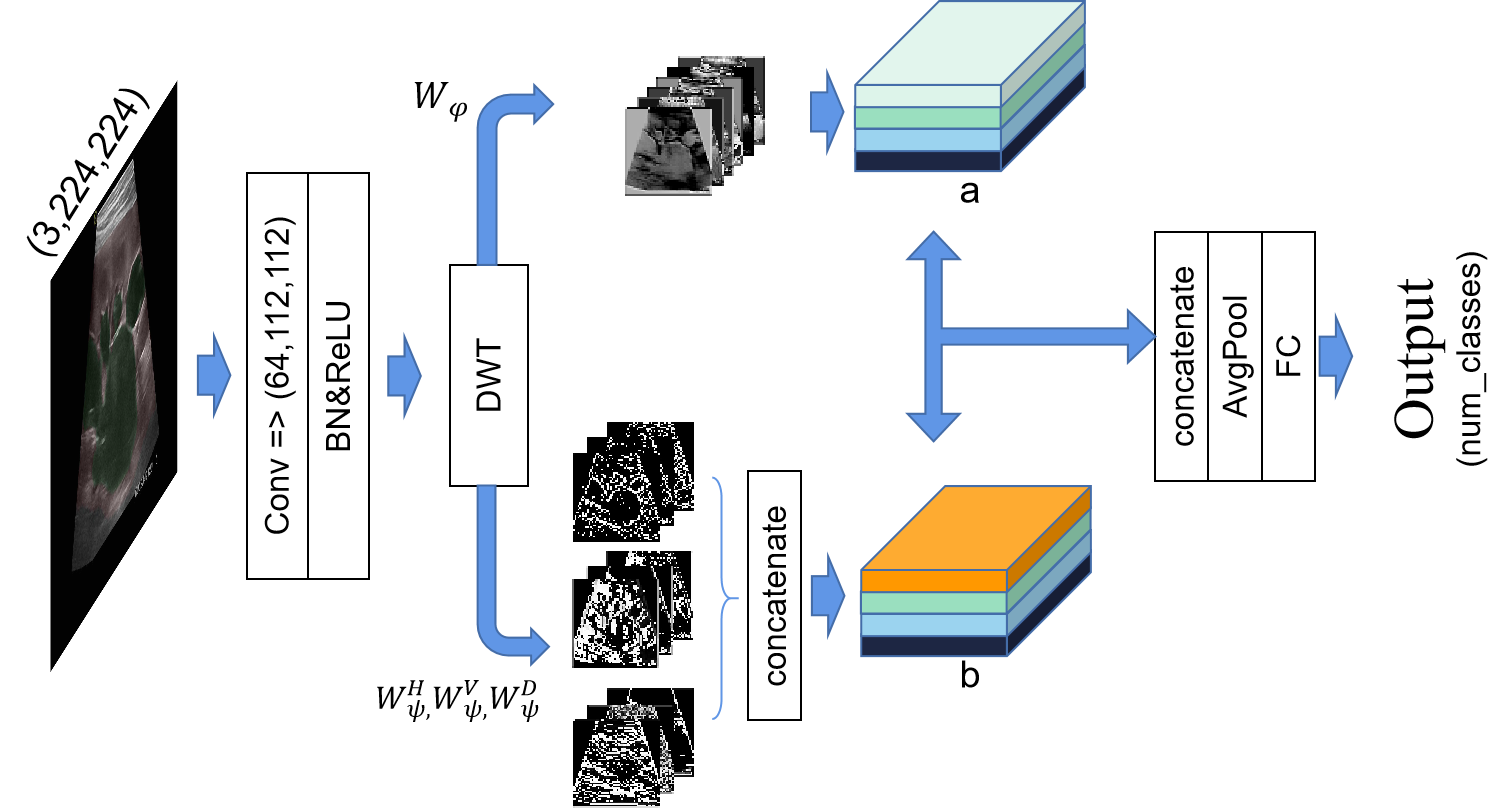


图 5 图示为分类器结构，图像经过离散小波变换并分成近似分量和三个细节分量共两路，然后并行通过神经网络最后合并输出结果

Figure 5. The structure of the classifier. The image is divided into one average component and three detailed components through DWT, and then the output results are concatenate through neural network

Replace the above 2D-DWT process into the first pooling layer in our residual classification network, whose overall structure is shown in Figure 5. Then the four half-sized images are passed through two BasicBlocks a and b, separately, where a and b are only different from the dimensions of its first layer. After that, the two components are concatenated, and the final classification results are output through the average pooling dimension reduction and full connection layer. Components of a and b in figure 5 are demonstrated in Fig. 6.

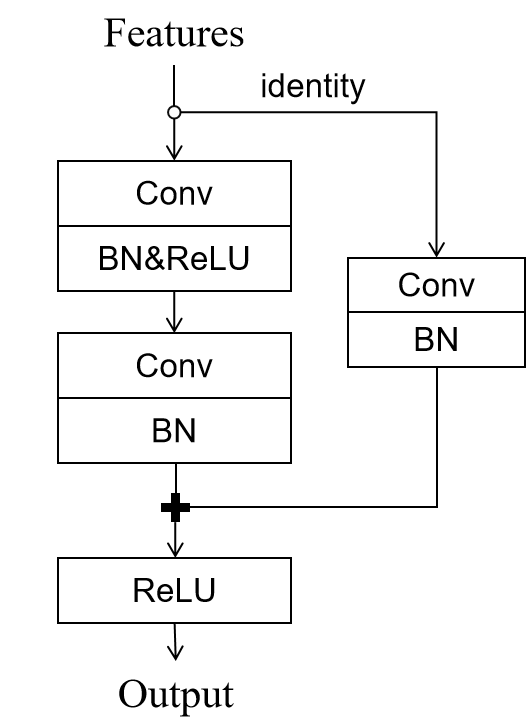


图 6. 图示为BasicBlock的基本构成单元，identity分支为残差项并与主干相加

Figure 6. Components of the BasicBlock. The identity branch is the residual term and is added to the trunk

The core idea is to add a residual term,the identity in Fig 6, at the output end，由于引入了identity shortcut connection,也称残差项，其直接跳过一个或多个层， 与主干相加以达到抑制网络由于过深所引起的梯度消失现象的效果。

**Blockchain-enabled AI platform**

Data Desensitization

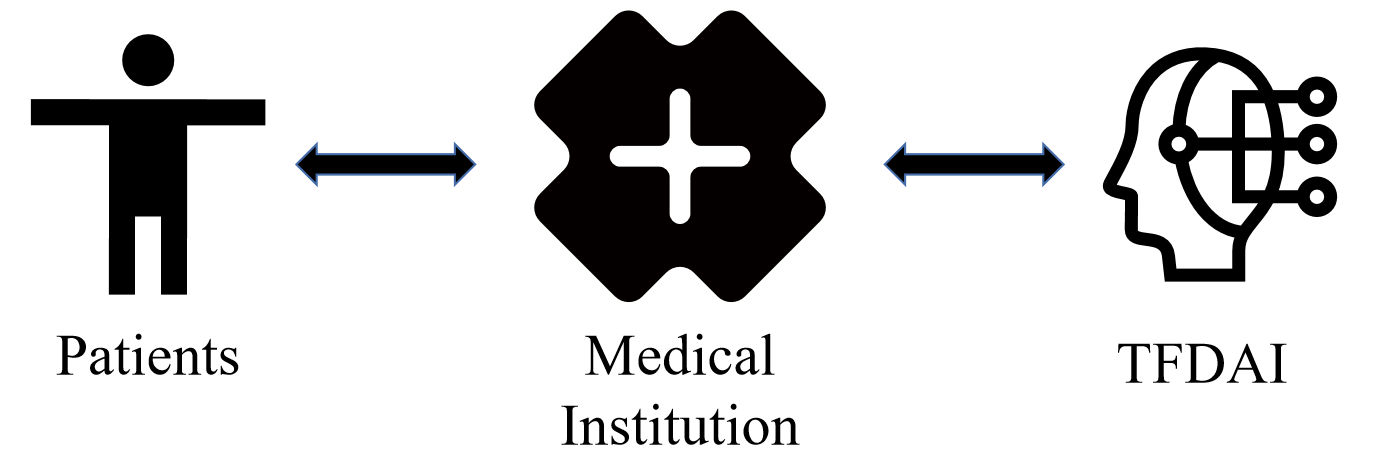
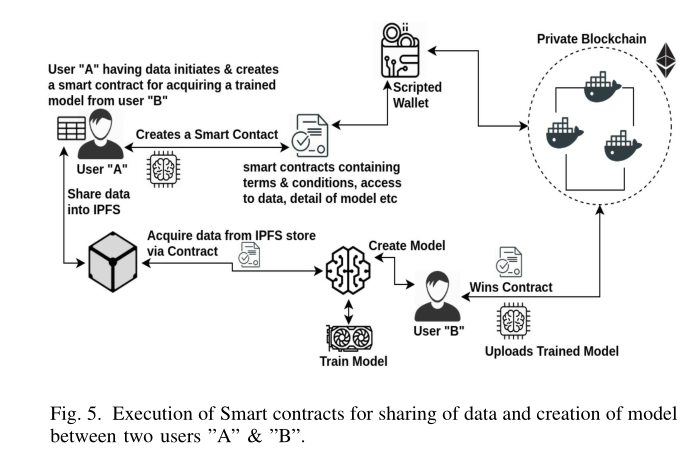
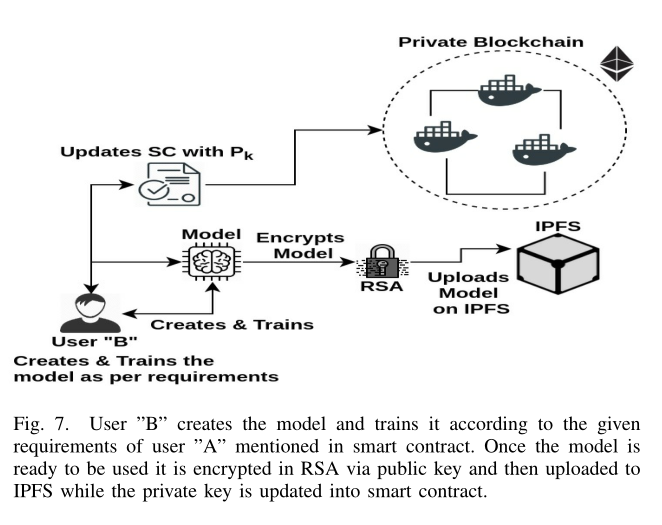
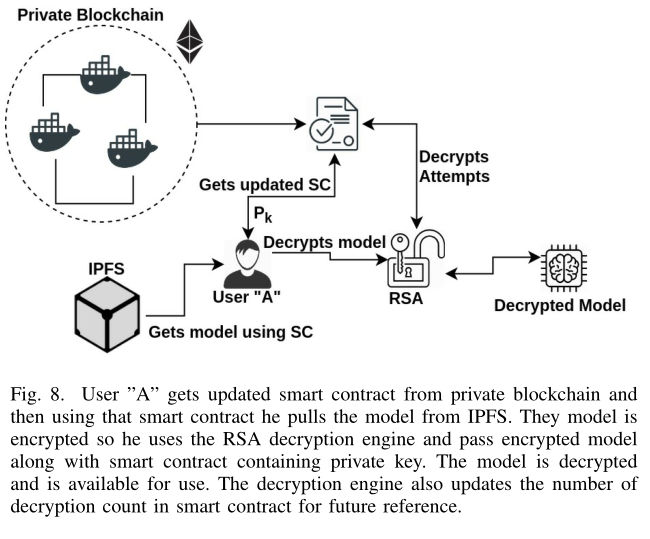


图 1

In the era of smart medicine, traditional medical institutions start to get assistance by computer-aided diagnosis system (CAD). In this background, sensitive data sharing and storing face the most challenges. As a classical protocol of blockchain, InterPlanetary File System (IPFS) provides a secure, economical, tamper-resistant, and decentralized storage solution by using encryption hash. Taking RSA algorithm for example. First, it generates a pair of public key and private key, which are used to encrypt information to obtain ciphertext. Then, the hexadecimal encoding will replace the sensitive information in the file. The decryption process is to decode and decrypt the password with the private key to obtain plaintext, and cover the sensitive information in the original file.

When a file is uploaded to the IPFS, this file is divided into several blocks and stored on different network nodes. Each block gets a unique ID through hash operation to facilitate network nodes to identify and de-duplicate. This specific ID is also required when people want to access this file on the IPFS.

In our blockchain framework, the three paries, hospitals, patients, and CAD model (our proposed diagnosis model), together constitute a private IPFS. The hospital uploads the pathological image dataset to IPFS, get a unique CIDH, and send it to CAD. The CAD uses this CIDH to download the desensitized dataset, return the diagnosis results to the hospital through IPFS and a new unique CIDA. This framework is also helpful for individuals, for example, to upload a single pathological image and safely get a diagnosis suggestion if they have an Internet of Medical Things (IoMT) like a portable or ultra-portable ultrasound scanner, or to check their relevant medical history if they ever changed a hospital. The detailed demonstration of this framework and examples of how it works can be found in the supplementary video xxx. Notice, our proposed blockchain framework can be extended to contain more sites, as long as this site has its own node.

**Experiment：（先介绍数据，然后deeplearning用的什么机器，然后小标题dl的具体实验，最后再加一个区块链的验证的实验设置。1500字）**

a dataset

All of the experiments are conducted on the outpatient dataset of Beijing Children's Hospital, Capital Medical University from September 2019 to February 2021. This dataset includes 229 patients in total, containing 3289 images taken by experienced sonographer in standard positions (transverse and coronal sections of the kidney). Among them, 1850 images of 17 patients are annotated into hydronephrosis area, kidney area and background area by professionals using LabelMe. All the 3289 images are labelled into 5 grades, SFU 0-4, according to the Classification System of Congenital Hydronephrosis designs by Society of Fetal Urology. All the experiments were performed using Intel CPU and NVIDIA GPU with 8GB memory.

In clinical practice, SFU 0 is healthy, SFU 1-2 is mild and does not need surgery, SFU 3 is medium and needs further observation and diagnosis, and SFU 4 is severe and needs surgery. To better fit the clinical requirement, we further divide the 3289 to be three grades, SFU 0-2, SFU 3, SFU 4. Stages of experiments are demonstrated separately as below.

b 分割阶段

The 1850 annotated images are divided into training set and test set with the ratio of 9:1. Batch training is adopted, and the batch size is set to 8. The training generation is 200. Adam optimizer is added to the training process. The initial learning rate is set to 0.005, and the learning rate is updated by the scheme with the ratio of 0.9 and the frequency of 1 generation. The combination of Dice Loss and Cross Entropy Loss is used as the loss function, and the ture value is defined as , the predicted value is , The combined loss value is , and its calculation method is

Taking Mean Intersection over Union (MIoU) as the evaluation standard of the model results, the actual input of the whole training network are images that are reduced and filled to size 474 × 473, the output segmentation image size is 1024 × 768。

c 分类阶段

After obtained a segmentation network with good performance, all the remaining un-annotated images are sent through the segmentation network again, and their segmentation result S is merged into the original image F. The merging algorithm is as follows:

In the actual experiment, we set to be 0.1. After 3289 labeled images are segmented, unnecessary regions are cropped according to the principle of keeping organ regions as much as possible, then the result sized 810 × 608 are derived. Divide these results into training set and test set with the ratio of 9:1. Comparative experiments are also conducted to prove the superiority of our classification network in this stage and accuracy is taken as the evaluation standard. the calculation method of classification model f and test set D with size n is as follows For classification model and test set with size , accuracy is as follows

All the parameters are set the same except initial learning rate. Batch training with batch size of 32 is adopted, training generation is set to 50, and Adam optimizer is added to update the learning rate with the scheme of 0.9 ratio and 1 generation frequency. In order to solve the problem of imbalanced data sets, the ratio of 0.3:0.6:0.1 is added to cross entropy loss as the weight, which is used as the loss function of the training process. The input of the whole training network are 224×224 size images.

For the wavelet transform in our target network, Haar wavelet transform and standard decomposition method are adopted. Firstly, one-dimensional wavelet is used to transform the pixel values of each row of the image, and then each column, and the approximate component and detail component are generated, in which the approximate component well replaces the output of pooling layer. In order to verify this effect, we conduct a comparative experiment, that is, we use the approximate component of wavelet transform to replace the output of pooling layer on other networks and train them. The training generation is 30, and other parameters are the same.

d simulation of the blockchain framework

We use VMware to generate two virtual machines of Windows system to simulate medical institutions and patients, and build IPFS and related environment. In our simulation, the file name of a file is its privacy information. The diagnostic results of the image set generated by CAD model include the classification results and the confidence level with the range of 0:1. A validation module is also added, that is, medical institutions can choose to upload marked test datasets to test the performance of CAD model, including accuracy, precision and recall of each class.

**Results：（dl的实验结果，和区块链的验证结果。1500字）**

Experiments are carried out and the model that has the highest MIoU is selected as out segmentation network. The segmentation results are combined with the original image and then clipped as the input of our classifier. The specific results and process are shown in Figure 7.



图 7.图示为分割器的处理过程 the specific results of the segmentation processing

In the classifier, we add the wavelet transform without changing the feature dimension, obtain the approximate component which is consistent with the effect of twice down-sampling, and the detail component with high-frequency information, which increases the available feature information of the whole network from the initial 64×56×56 to 256×56×56. The characteristic diagram after the original pooling process is shown in Figure 8.

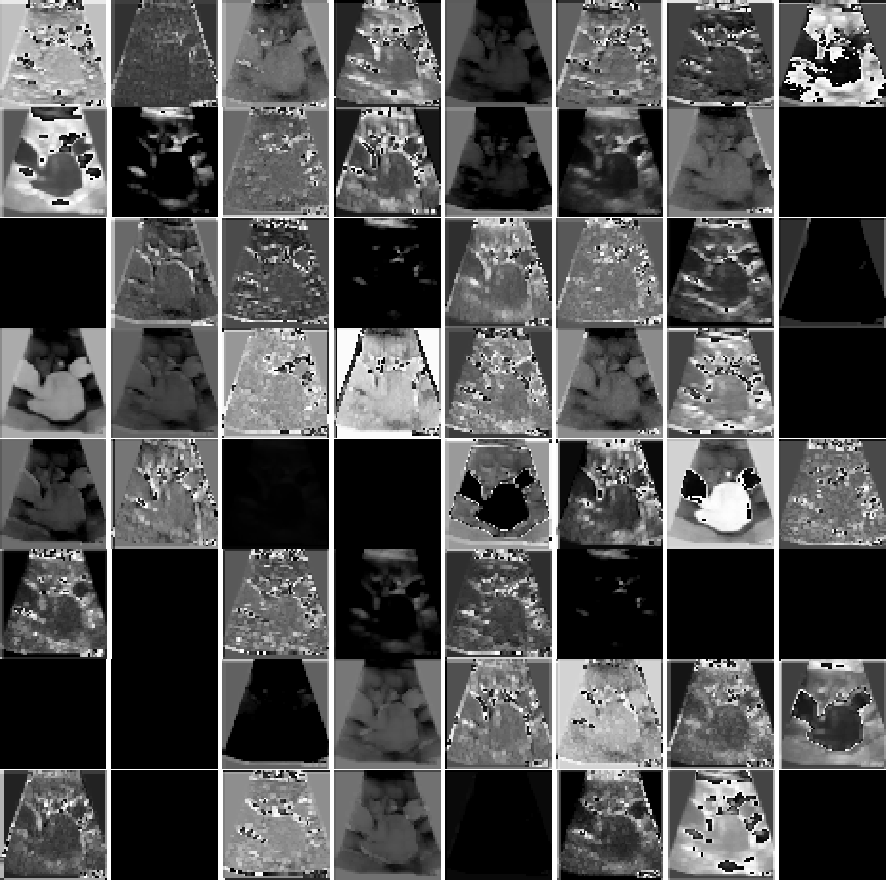


图 8.图示为随机选取的一张数据集中的超声波图像，并在原残差网络池化层过后的特征图，维度为64×56×56 Figure 8. the feature map of a randomly selected ultrasonic image in a dataset after the pooling layer of the original residual network, with the dimension of 64×56×56

After changing to DWT, the path of original pooling layer is replaced by approximate component, and the path of detail component is added. The specific characteristics are shown in Figure 9.

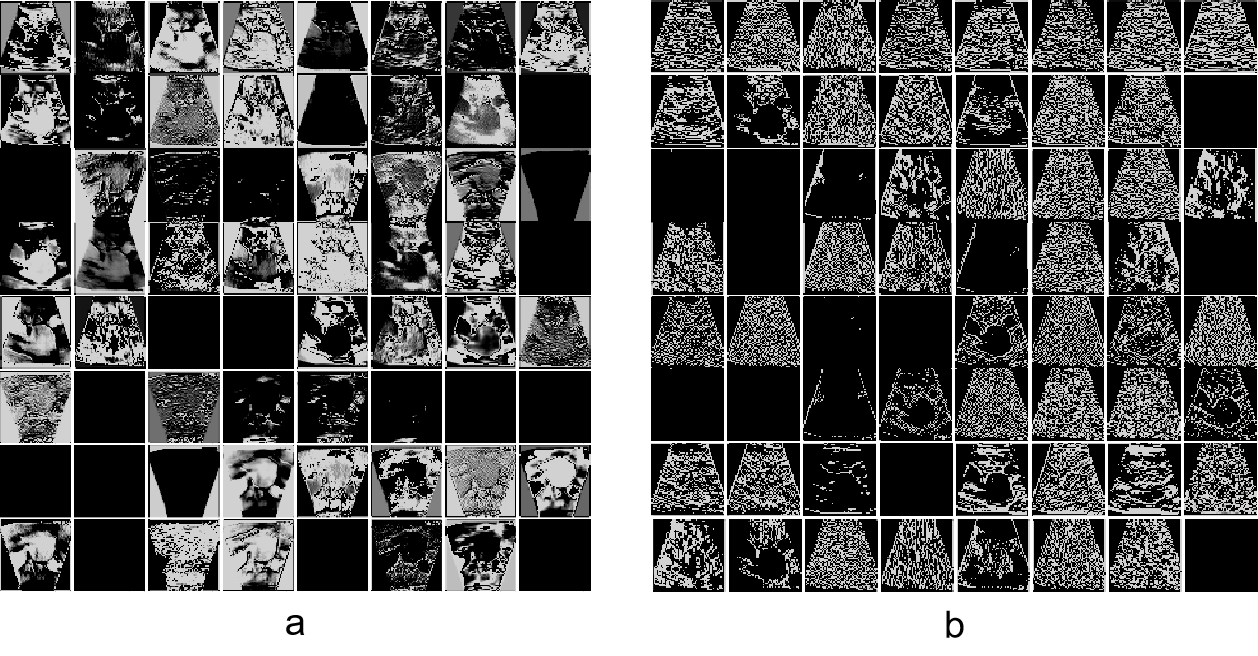


图 9.图中a表离散小波变换后的近似分量的特征图，维度为64×56×56；为方便对比，图b为水平、垂直、对角细节分量随机抽取的64张特征图，原总数为192张

Figure 9. a. The feature map of approximate components after DWT with a dimension of 64×56×56； b. feature maps that randomly selected from horizontal, vertical and diagonal detail components

对比图8与图9中的a部分，从特征图是上可以发现小波变换的近似分量很好的替代了池化操作，并在特征工程的维度上增加了图9 b部分的高频特征，It can be seen from b in Fig.9 that the new feature map contains dense high-frequency noise information and contour information.

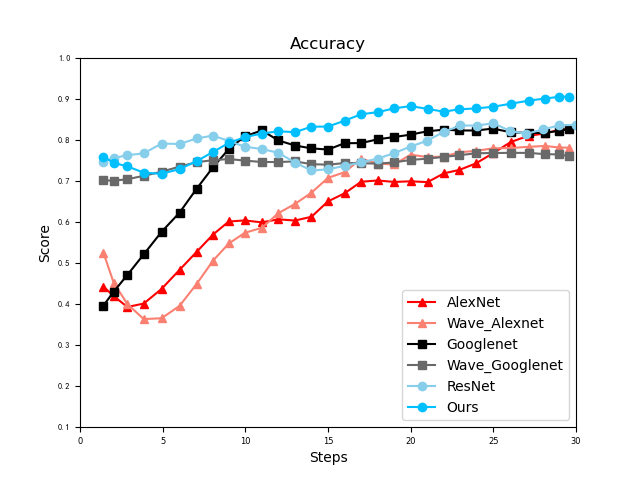


图 10

另外，我们尝试了在一些经典模型上使用我们的方法替换池化层并增加高频分量的特征，值得注意的是这里没有实验VGG网络，是因为VGG网络的池化层数过多会极大地增加网络复杂度，因此我们只挑选并实验了池化层数极少且位置关键的几个经典模型，收集训练过程的acc变化结果并平滑所有曲线。整体结果从图10可以看出，在各模型上，替换池化层并增加特征数的方法使得新网络都能优于原网络，且在ResNet上的性能表现最佳，由此印证了该方法的可靠性与优越性。

In addition, we do a comparative experiment of network classification performance, and the specific results are shown in Table 1. The table presents the optimal accuracy model and the training epoch in the training process of different networks, which further confirms that our method can effectively improve the prediction ability of the network by increasing the amount of feature information.

Table 1

|  |  |  |  |
| --- | --- | --- | --- |
|  | accuracy | F1-score | epoch |
| AlexNet | 0.8445 |  | 35 |
| VGG | 0.8636 |  | 27 |
| GoogleNet | 0.8475 |  | 38 |
| ResNet | 0.8994 |  | 38 |
| **Ours** | **0.9177** |  | 49 |

Our improved method is based on residual network. As shown in Figure 10, compared with the original network, our loss value in the training process descends faster, and the overall process is more stable. In the first 30 generations, the accuracy and loss value of the original network training model on the test set tend to decrease at the same time, increasing the possibility of network over-fitting, thus further confirms the superiority of our network structure. Based on the above results, a reasonable replacement of pooling layer by DWT will increase the number of features, make the training process more stable and the performance of the model more excellent.

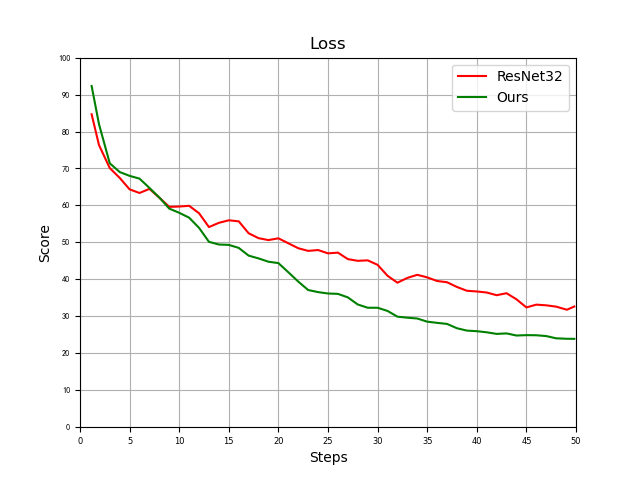


图 11

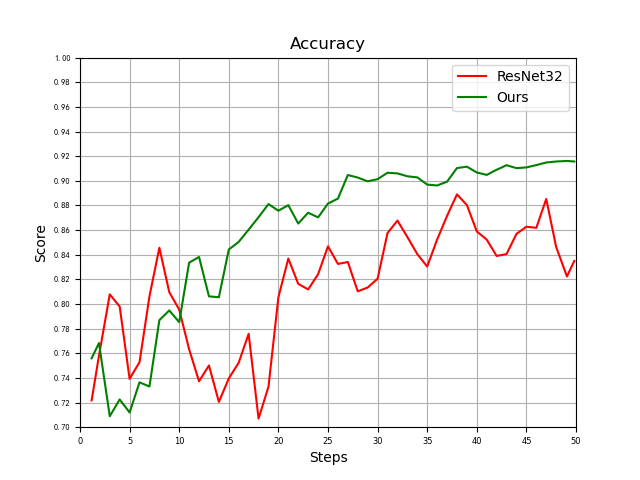


图 12

从图12中的acc曲线可以看出，从曲线的震荡幅度与频率来看，我们的模型相比原网络的训练过程更加稳定，整体性能更加优越。

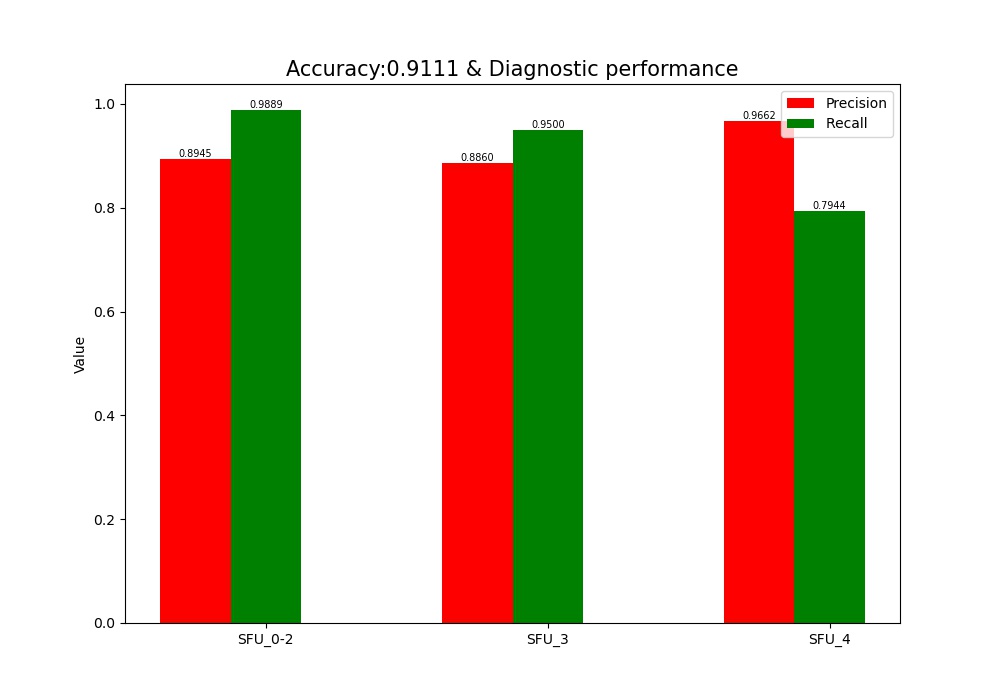


图 13

We simulate the performance results when medical institutions test the CAD model. Three classes of cases, including 600 original ultrasound images, are adopted and securely transmitted through IPFs. The generated performance results including the accuracy of the model, the precision and recall of each classification are shown in Figure 13., 其Accuracy为0.9111,相比CAD本地非上链的整体数据集，其训练结果(非上链：0.9177)几乎相等，因此可以看出，数据通过IPFS的传输过程后并不会影响最后的诊断结果，印证了医疗机构测试CAD模型的可靠性与有效性，并且CAD能够给予医疗机构所有分级的全面的诊断性能的结果，包括精确度与召回率，以帮助医疗机构更清晰地认识CAD模型的特点，从而更好地配合相关的工作。

**Discussion（写个至少500字，此处可以有至少10个文献）：**

**（未来可以如何扩充，比如我现在的multiparty是用的仿真数据，但是我们也验证了这个区块链frame的robustness，所以我们相信真实数据也没问题的，但是这个可以写成一个limitation，就是说现在share自己数据的医疗机构还是少数（是不是之前找过公开数据集但是没有找到？如果是的话也就能更好/更有说服力的说明以下两件事①为什么我只用了一个机构的数据②sharing现在还很不到位 大家还都很不放心 sharing这件事还有很大的发展空间）；**

**比如因为病人可以有掌上超声之类的机器（文献123），所以也可以随时上传图片并查看结果之类的。limitation同时也可以是未来的方向。）（如果病人在别的医院看了病拍了b超，可以追溯历史。如果病人有便携式检测机器，可以诊断病情。如果未来有multiparty机构，可以共享结果、可以联邦学习）（比如未来可以封装成APP？）**

**（Contribution里提到的一些情况可以作为discussion里的一些情况。①我提出的DL模型能够快速地根据无辐射的便宜的超声图像classify肾积水，并达到了xx的acc、auc、roc，超过了常见的abc模型，能够为医生对肾积水的诊断省好多事儿，能给病人省钱。②我提出的blockchain frame能够帮助sensitive data的线上存储和传输，为未来可能的multiparty间的数据共享、联邦学习、合作等奠定了技术基础/提供了技术支持。③我使用的Blockchain platform能帮助病人安全地存储自己的信息，且能够在不去医院/换了个医院的情况下随时获取/追踪自己的相关病史（前面的部分，是已经去医院看过医生的数据能够随时获取。后面的部分，是未来自己在家里如果有设备的话可以随时上传一张照片就得到诊断结果&不需担心泄露自己的隐私&这个结果因为存在了线上所以日后如果又去看医生的话医生还能看到这个自己拍的片子）；moreover，DL和区块链的结合更是能够起到线上AI医生的效果，无需去医院，只要一张超声照片，既可得到准确率高达xxx的诊断。）**

**Conclusion（200-300字）：**

Abstract，contribution，paper arrangement，discussion，conclusion这几块儿可以互相借鉴，最后一起写就行。整体逻辑理顺了啥都有了。

**Funding：可以先不写，发公备时问他，然后最后排版时加上就行**

**Bibliography：可以先不写，最后排版时写就行**

**Reference**

//ComboLoss

Xu L, Xiang J. ComboLoss for Facial Attractiveness Analysis with Squeeze-and-Excitation Networks[J]. arXiv preprint arXiv:2010.10721, 2020.

//ResNet

He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.