A deep learning blockchain integration framework for UPJO diagnosis using ultrasound images

First A. Author, *Fellow, IEEE*, Second B. Author, and Third C. Author, Jr., *Member, IEEE*

***Abstract*—Ureteropelvic Junction Obstruction (UPJO) is a common hydronephrosis disease in children that can cause various symptoms and result in even progressive loss of renal function. Ultrasonography as a preliminary diagnostic step for UPJO has the nature of economical, radiationless, noninvasive, and high-noise. Artificial intelligence has been widely applied to medical fields and can greatly assistant for doctors' diagnostic ability. While transferring electronic medical data online demands high secure network environment, therefore leads to the development of blockchain technology.**

We build and test a framework that integrates a deep learning diagnosis model and blockchain technology. Our diagnosis model is a combination of an attention-based pyramid semantic segmentation network and a discrete wavelet transformation processed residual classification network. We also compare the performance between benchmark models and our models. For the blockchain system, we apply the IPFS protocol to build a secure and private sharing environment.

Our diagnosis model outperforms other benchmarks with an accuracy of 0.9177. VMware is used to generate virtual machines of Windows to simulate all the sites in the blockchain system. The performance of the integration of the diagnosis model and the IPFS is the same as when the model is trained alone locally. The generated accuracy of the model, the precision and recall of each classification are all satisfactory.

This framework can automatically grade the severity of UPJO by ultrasound images, guarantee secure medical data sharing, assistant for doctors' diagnostic ability, relieve patients' burden, and provide technical support for future federated learning, linkage of IoMT, trusted cooperation and development.

***Index Terms*—ultrasound image, UPJO, Hydronephrosis, SFU, deep learning, blockchain, IPFS, IoMT, DWT, CAD**

This paragraph of the first footnote will contain the date on which you submitted your paper for review. It will also contain support information, including sponsor and financial support acknowledgment. For example, "This work was supported in part by the U.S. Department of Commerce under Grant BS123456."

The next few paragraphs should contain the authors' current affiliations, including current address and e-mail. For example, F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: author@ boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

I. Introduction

Being a common kidney disease in children, hydronephrosis is generally congenital [1, 2], and has increasing morbidity above 1% [3, 4]. Ureteropelvic Junction Obstruction (UPJO), the obstruction at or along the pelvic-ureteral junction, being the main causation of children hydronephrosis [5], 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.

As a commonly adopted examination in the medical field, as well as a preliminary diagnostic step for UPJO, ultrasonography is economical, radiationless, meanwhile noninvasive [6]. Yet, since the ultrasonic images are high-noise, it is laborious for even trained sonographers to obtain enough information [7, 8].

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 [9-14]. 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 [15-16]. 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 [17]. As a shared ledger that provides a secure, decentralized, transparent, and trusted manner for data management, blockchain has been more extensively applied in medical data with the arrival of the 5G era [18-20].

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 that can diagnose the ultrasonic images of

UPJO, but adopt a blockchain protocol to integrate with to guarantee the secure transferring between different parities. The performance of this framework is validated and reported.

Our study has several contributions:

Our proposed deep learning diagnosis model 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.

We adopt a discrete wavelet transform (DWT) feature engineering method and replace the pooling layer in our deep learning network with it, increases the number of model input features and enhance the model performance.

The blockchain system 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 system 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 system 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 deep learning-Blockchain framework, which leaves the possibility for the linkage with the Internet of Medical Things (IoMT) [21] like portable and ultra-portable ultrasound scanners [22].

The organization of our paper is as follows for better understanding: in the METHODOLOGY section, we present the model architecture we proposed and introduce every element in detail. In the EXPERIMENT section, we demonstrate the experimental setting and network training. In the RESULTS section, we illustrate our performance and the comparison results with other networks. Finally, we discuss our whole study and the potential future direction.

II. METHODOLOGY

In this study, we develop a deep learning computer-aided diagnosis model that combined a semantic segmentation network and a classification network to grade the UPJO from hydronephrosis ultrasound images. In this section, we will introduce our pathological grading scheme of UPJO ultrasound images, and the architecture is shown in Fig. 1. This scheme combines an attention-based 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. Besides, we adopt a blockchain protocol to integrate with our deep learning-based computer-aided diagnosis (DL-CAD) model for secure transmission between separate parties. We then separately demonstrate these parts as following.

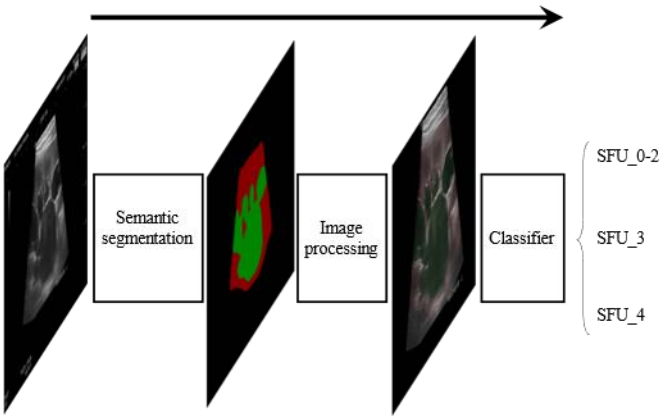


Fig. 1. The overall structure of our UPJO ultrasonic image diagnosis model. The input is the original image and the output is the grading result.

A. 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 [23] and an attention mechanism Convolutional Block Attention Module (CBAM) [24]. The overall structure is shown in Fig. 2. 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.

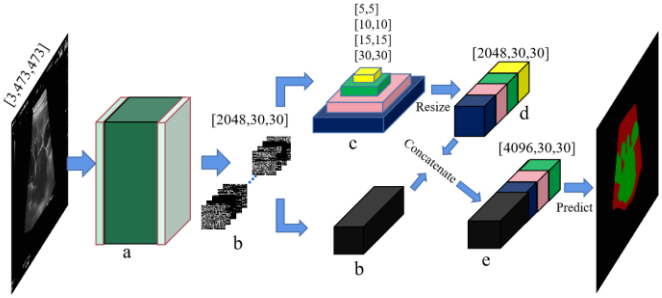


Fig. 2. The architecture of the A-PSPNet. The input is the original image, and the output is the segmentation result distinguished by color. a. The ResNet structure embedded in two attention layer CBAM. b. The basic feature of the output after a. c. The advanced feature of b after pyramid pooling. d. The result of unified dimension operation on c. e. The integrated structure of the features of b and d modules.

Using x denotes the input. The original input image is decomposed into basic feature representation $\mathcal{R}(x)$ and advanced feature representation $\mathcal{P}_i(x)$, where $i \in (1, 2, 3, 6)$. i 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 Fig. 2 as $\mathcal{H}$, where $\mathbb{C}(a, b)$ stands for the concatenate processing.

$$\mathcal{H} = \mathbb{C}\{\mathcal{R}(x), \mathbb{C}(\mathcal{P}_i(x) | i \in (1,\ 2,\ 3,\ 6))\}$$

After that, the concatenated feature layers are processed, predicted and colored to output the segmentation results. As shown in Fig. 3, 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 [25]. The final segmentation results are shown as the output in Fig. 2.

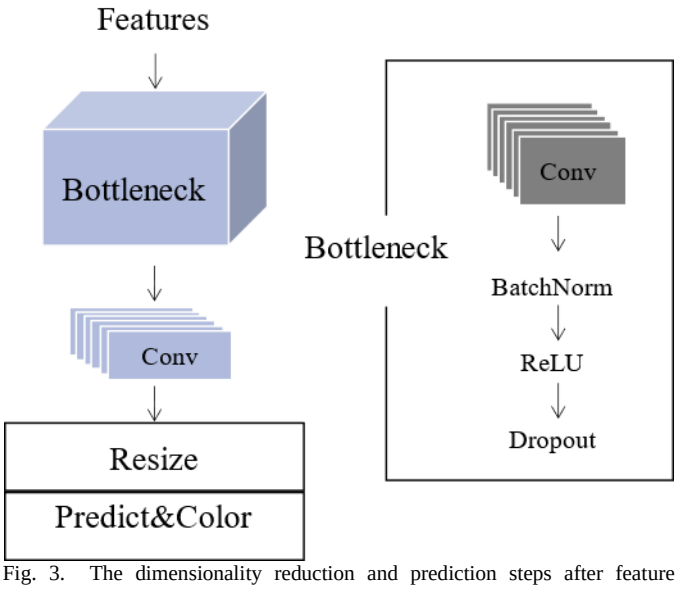


Fig. 3. The dimensionality reduction and prediction steps after feature concatenating.

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) [26], 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 $\Psi(x)$ and a scale function $\varphi(x)$. For 1D discrete sequence $f(n)$, its discrete wavelet series expansion coefficient is

$$W_\varphi(j_0,k)=\frac{1}{\sqrt{M}}\sum_n f(n)\,\varphi_{j_0,k}(n), j=0,1,\cdots,J-1$$

$$W_\Psi(j,k)=\frac{1}{\sqrt{M}}\sum_n f(n)\,\Psi_{j,k}(n),\ \ j\geq j_0$$

In the above equations, $\varphi_{j_0,k}(n)$ and $\Psi_{j,k}(n)$ are samples that use M equal intervals on the support region of the basis function on the J scale.

$$\varphi_{j,k}(x)=2^{j/2}\varphi(2^j x-k),\ \Psi_{j,k}(x)=2^{j/2}\Psi(2^j x-k)$$

k determines the position in a given direction, j determines the width of the sample, j_0 is any determined starting scale, while the obtained expanded set $\{\varphi_{j_0,k}(n)\}$ is a subset of $\{\varphi_{j,k}(n)\}$, and final transformation itself is composed of M 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 $M\times N$, its DWT is

$$W_\varphi(j_0,m,n)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)\,\varphi_{j_0,m,n}(x,y)$$

$$W_\Psi^i(j,m,n)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)\,\Psi_{j_0,m,n}^i(x,y), i=\{H,V,D\}$$

Finally, we get four half-sized output images: $W_\varphi, W_\Psi^H, W_\Psi^V, W_\Psi^D$, which represent average, horizontal, vertical, and diagonal information from the input source image.

Replace the above 2D-DWT process into the first pooling layer in our residual classification network, whose overall structure is shown in Fig. 4. Then the four half-sized images are passed through two BasicBlocks [25] 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 Fig. 4 are demonstrated in Fig. 5. The core idea is to add a residual term, the identity in Fig. 5, at the output end. Due to the introduction of identity shortcut connection which is also known as the residual term, it directly skips one or more layers then adds with the backbone, consequently suppresses the gradient disappearance phenomenon caused by the too deep network.

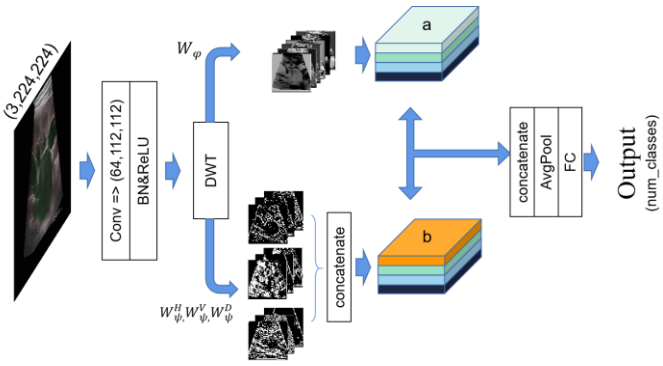


Fig. 4. The structure of the classifier. The image is divided into one average component and three detailed components through DWT, and then the output results are concatenate through neural network.

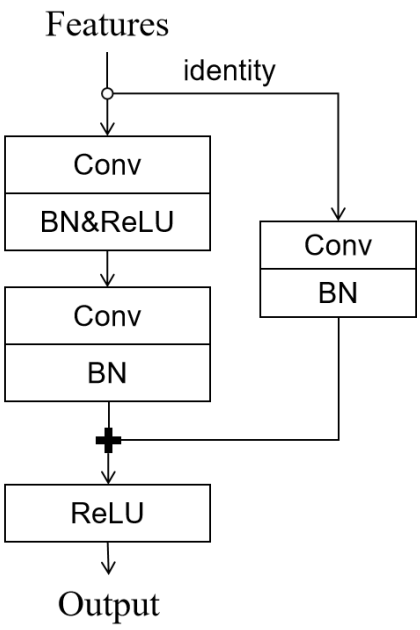


Fig. 5. Components of the BasicBlock. The identity branch is the residual term and is added to the trunk.

C. Deep learning-Blockchain framework

In the era of smart medicine, traditional medical institutions start to get assistance by deep learning-based computer-aided diagnosis (DL-CAD) system. In this background, sensitive data sharing and storing face the most challenges. As a classical protocol of blockchain, InterPlanetary File System (IPFS) [27] provides a secure, economical, tamper-resistant, and decentralized storage solution by using encryption hash. Taking RSA encryption 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 as shown in Fig. 6, the three ends, hospitals, patients, and our proposed DL-CAD model, together constitute a private IPFS. For example, in step 3, a hospital H_1 uploads the pathological image dataset to IPFS, get a unique CID_ H_1, and send it to the DL-CAD model. In step 4, the DL-CAD model uses this CID_ H_1 to download the desensitized dataset, return the diagnosis results to the hospital through IPFS and a new unique CID_M_1. 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 [28-32], 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 1. Notice, our proposed blockchain system can be extended to contain more sites, as long as this site has its own node.

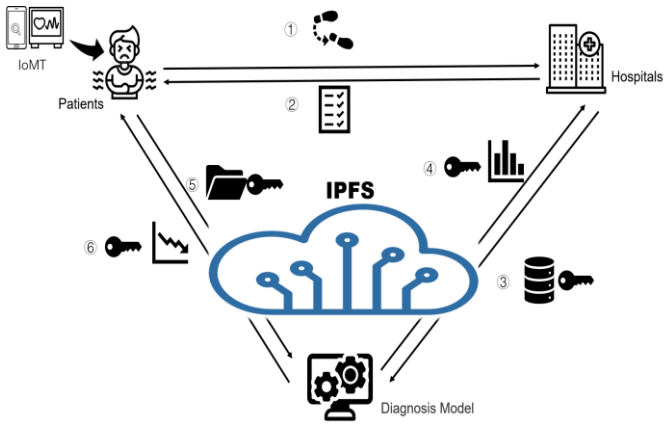


Fig. 6. The sketch map of the deep learning-blockchain framework.

III. Experiment

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 [33, 34]. 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. Segmentation process

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 [35] is used as the loss function, and the true value is defined as G_T, the predicted value is S_P, The combined loss value is L, and its calculation method is

$$L = 1 - \sum G_T \log S_P - \frac{2\sum G_T S_P}{\sum G_T^2 + S_P^2}$$

Taking Mean Intersection over Union (MIoU) [36] 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. Classification process

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:

$$\text{Output} = \text{F} \times (1-\alpha) + \text{S} \times \alpha$$

In the actual experiment, we set α 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 f and test set D with size N, accuracy is as follows

$$\text{Acc}(f;D) = \frac{1}{\text{N}}\sum (f(x_i) = label_i)$$

In order to evaluate the model more accurately, weigh the impact of missing diagnosis and misdiagnosis, and set the threshold as a diagnostic reference value, we also analyzed the ROC curve of sensitivity and specificity. Let F indicates the misdiagnosis, T indicates the correct diagnosis, the positive case is P, the negative case is N, the horizontal axis is the false positive case rate (FPR), the vertical axis is the true positive case rate (TPR). The more convex the ROC curve is and the closer it is to the upper left corner, the greater its diagnostic value is, hence evaluates the diagnostic accuracy. Besides, we add PR curve to show the quality of the model in a more intuitive way. Formulas are as follows.

$$\text{P} = \frac{\text{TP}}{\text{TP}+\text{FP}} \quad \text{R} = \frac{\text{TP}}{\text{TP}+\text{FN}}$$

$$\text{FPR} = \frac{\text{TP}}{\text{TP}+\text{FN}} \quad \text{FPR} = \frac{\text{FP}}{\text{TN}+\text{FP}}$$

Additionally, we add macro-average and micro-average evaluation indexes to the ROC curve:

$$\text{Macro} = \frac{2 \times \left(\frac{1}{\text{n}}\sum \text{P}_\text{i}\right) \times \left(\frac{1}{\text{n}}\sum \text{R}_\text{i}\right)}{\left(\frac{1}{\text{n}}\sum \text{P}_\text{i}\right) + \left(\frac{1}{\text{n}}\sum \text{R}_\text{i}\right)}$$

$$\text{Micro} = \frac{2 \times \left(\frac{\sum \text{TP}_\text{i}}{\sum \text{TP}_\text{i} + \sum \text{FP}_\text{i}}\right) \times \left(\frac{\sum \text{TP}_\text{i}}{\sum \text{TP}_\text{i} + \sum \text{FN}_\text{i}}\right)}{\left(\frac{\sum \text{TP}_\text{i}}{\sum \text{TP}_\text{i} + \sum \text{FP}_\text{i}}\right) + \left(\frac{\sum \text{TP}_\text{i}}{\sum \text{TP}_\text{i} + \sum \text{FN}_\text{i}}\right)}$$

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. 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.

IV. Results

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 Fig. 7.

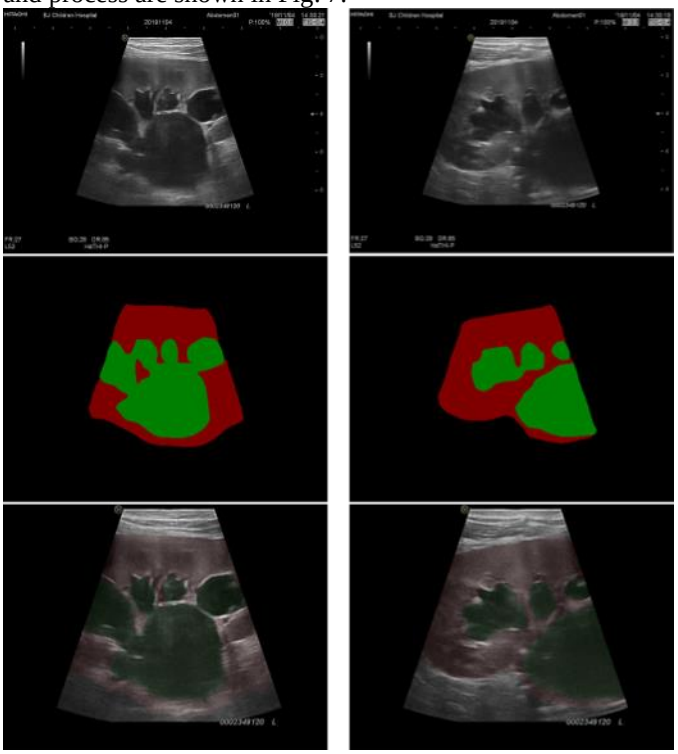


Fig. 7. Two examples of the specific results of the segmentation processing.

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 Fig. 8.

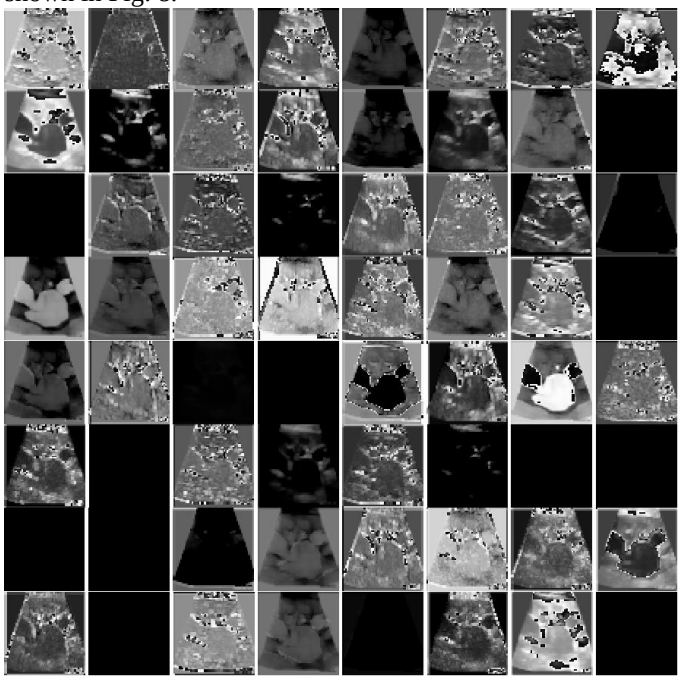
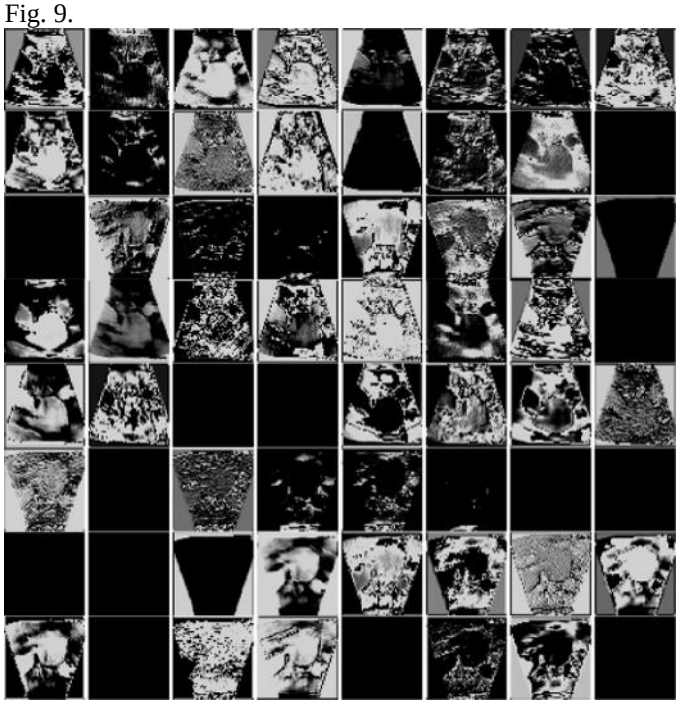
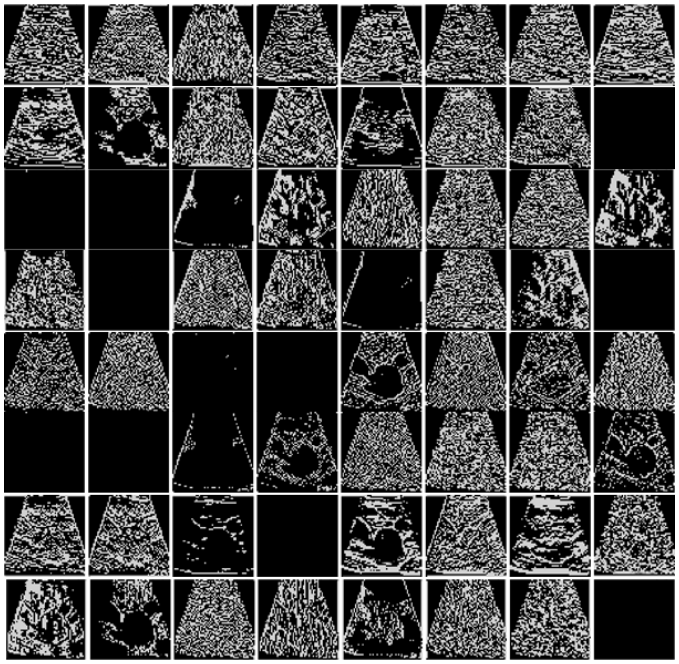


Fig. 8. The feature map of a randomly selected ultrasonic image in a dataset after the pooling layer of the original residual network, with the dimension of 64×56×56.

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 Fig. 9.



a



b

Fig. 9. a. The feature map of approximate components after DWT with a dimension of $64 \times 56 \times 56$; b. feature maps that randomly selected from horizontal, vertical and diagonal detail components, whose total number is 192.

Compared with part a in Fig. 9 and Fig. 8, it can be found that the approximate component of DWT is an effective substitute for pooling operation, and the high-frequency features of Part b in Fig. 9 is added in the aspect of feature engineering, it can be seen from b in Fig. 9 that the new feature map contains dense high-frequency noise information and contour information.

We conduct 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. From where we observe that ResNet has a better accuracy among four benchmark networks, and our proposed model that based on ResNet outperforms all, which proves the correctness of us choosing ResNet.

TABLE I
CLASSIFICATION RESULTS OF BENCHMARKS AND OUR MODEL

Model	Accuracy	Epoch
AlexNet	0.8445	35
VGG	0.8636	27
GoogleNet	0.8475	38
ResNet	0.8994	38
Ours	**0.9177**	49

We also attempt to carry out the operation of replacing the pooling layer with DWT on classical benchmark networks. Since too many pooling layers will greatly increase the complexity of the network, we only select and test several networks with few pooling layers at important positions and VGG network is eliminated. The ACC changes during training are collected and all curves are smoothed. According to Fig. 10, replacing the pooling layer and increasing the number of features enhances the performance of all the networks, while

ResNet is the best, which confirms the feasibility of increasing the amount of feature information and the superiority of our proposed method.

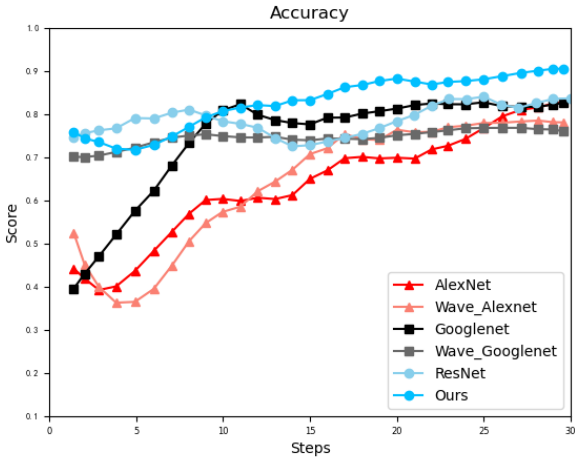
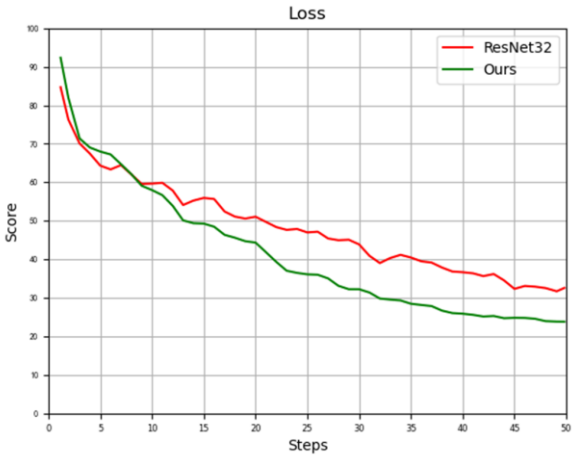


Fig. 10. Performance after replacing the pooling layer with DWT on classical benchmark networks

Our improved method is based on residual network. As shown in Fig. 11, compared with the original network, we can see that our model has a more stable training process and better overall performance than the original network from the vibration amplitude and frequency of the curve.

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.



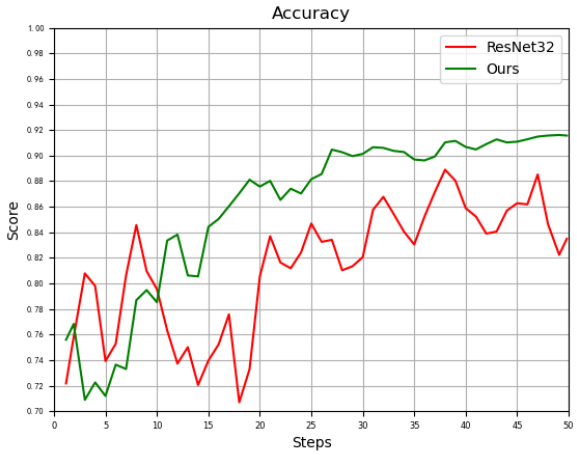


Fig. 11. The performance of the benchmark models versus our proposed model.

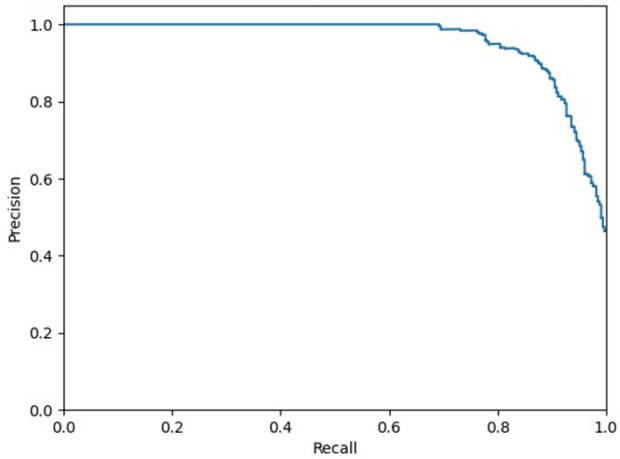


Fig. 12 Micro-averaged over all classes: Average precision score=0.96

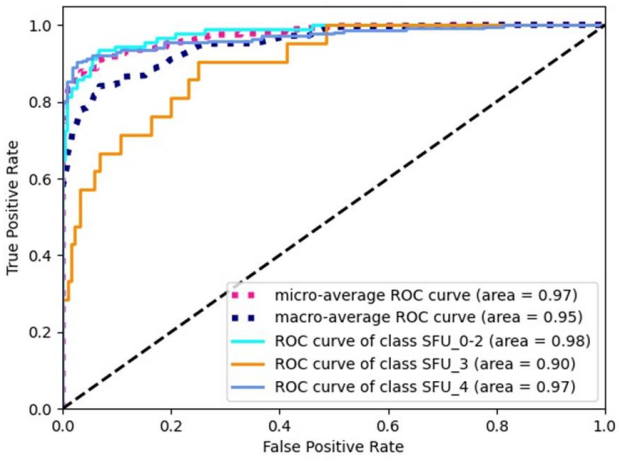


Fig. 1 Classifier performance indicators: ROC

As shown in the PR curve in Fig. 12, the average precision of our final model for different recall points is AP = 0.96, which proves that the diagnostic ability of the model is quite excellent.

According to the ROC curve in Fig.13, the worst diagnostic result is of SFU_3, yet it still reaches 0.90. On the whole, macro-average ROC= 0.97, macro-average ROC = 0.95, which again proves the superiority of our classifier.

We also simulate it when medical institutions test the DL-CAD model performance on blockchain system. Three classes of cases, 600 original ultrasound images for saving training cost, 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 Fig. 14. The accuracy is 0.9111, which is almost the same as when training the whole dataset on the local diagnosis model alone without transmission on the IPFS (0.9177). Therefore, it can be concluded that the data transmission process through IPFS will not affect the final diagnosis results, which further confirms the reliability and effectiveness of the DL-CAD model.

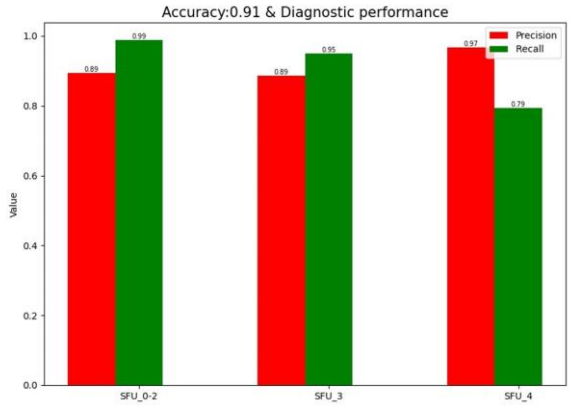


Fig. 14. The performance of each classification.

V. Discussion

In this study, we proposed a framework that integrated a deep learning diagnosis model and blockchain technology to recognize and grade UPJO on the basis of ultrasound images. The accuracy of our diagnosis model outperformed benchmark networks and reached 0.9177, which indicates that our model can highly assistant for doctors and save their labors. The integration of blockchain protocol enables secure and desensitized data sharing for institutions and patients, allows them to access the history information whenever and wherever. However, limitations still exist in our study. For example, the framework we proposed is capable to contain more than one site at each end (the patient end the hospital end), but the real-world dataset that we can use to verify is only one since there is no useable public UPJO ultrasound image dataset and the data sharing between institution is restricted. This however demonstrates the urgent demands of our kind of works from another point of view. Besides, although we mentioned IoMT as portable and ultra-portable ultrasound scanners at patients' side, we did not really include any specific IoMT in our framework, plus the development of ultra-portable ultrasound scanners for the kidney is also not sophisticated enough.

In the future, we could link actual IoMT to our system. Because patients of UPJO are always young children, they can be terrified to go to hospitals and less cooperative with doctors. The changes of their bodies and organs are also constant and unpredictable. Therefore a household ultra-portable ultrasound scanners can greatly relieve parents' burden and struggle of visiting doctors, the fast and automatic diagnosis allows parents to monitor children's situation timely, and the blockchain technology facilitates the storage as well as the accessibility of personal medical history at any time. On the other hand, the proposed method provides technical support for further cooperation among multi-parties and perhaps federated learning. Lastly, this system can also be generalized for the safe diagnosis of other diseases.

VI. Conclusion

We developed a deep learning-based model to diagnose UPJO using ultrasound images, then we integrated this model with a constructed blockchain technology IPFS, and finally established a deep learning-blockchain framework to ensure secure data transferring and fast UPJO diagnosis. Experiments were conducted and verified the superiority and reliability of our system. In the future, more institutions can be included to extend this cooperation, IoMT like portable and ultra-portable ultrasound scanners can be linked to facilitate the convenient household diagnosis to relieve patients' burden, finally, our framework can also provide technical support for possible federated learning and generalized diagnosis of other diseases.

Appendix

Appendixes, if needed, appear before the acknowledgment.

Acknowledgment

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank ... ." Instead, write "F. A. Author thanks ... ." In most cases, sponsor and financial support acknowledgments are placed in the unnumbered footnote on the first page, not here.

References

[1] Siegel MJ: Pediatric Sonography, 4th ed. Philadelphia: Lippincott Williams & Wilkins 2010.

[2] J. M. Smith, D. M. Stablein, R. Munoz, D. Hebert, and R. A.McDonald, "Contributions of the transplant registry: the 2006 annual report of the North American Pediatric Renal Trials and Collaborative Studies (NAPRTCS)," Pediatric Transplantation, vol. 11, no. 4, pp. 366–373, 2007.

[3] B. A. Warady and V. Chadha, "Chronic kidney disease inchildren: the global perspective," Pediatric Nephrology,vol. 22, no. 12, pp. 1999–2009, 2007.

[4] H. T. Nguyen, C. D. A. Herndon, C. Cooper et al., "The Society for Fetal Urology consensus statement on the evaluation and management of antenatal hydronephrosis," Journal of Pediatric Urology, vol. 6, no. 3, pp. 212–231, 2010.

[5] Hashim, H. , and C. Woodhouse . "Ureteropelvic Junction Obstruction." European urology supplements: official journal of the European Association of Urology (2012). EUROPEAN UR OLOGY S UP PL EM EN TS 11 (2012) 25–32

[6] Ucar, A. K. , and S. Kurugoglu . "Urinary Ultrasound and Other Imaging for Ureteropelvic Junction Type Hydronephrosis (UPJHN)." Frontiers in Pediatrics 8(2020):546. DOI:10.3389/fped.2020.00546

[7] None. "AIUM practice guideline for the performance of an ultrasound examination in the practice of urology." Journal of Ultrasound in Medicine 31.1(2012):133. DOI: 10.7863/jum.2012.31.1.133

[8] Amole, Kuranga, and Oyejola. "American Institute of Ultrasound in Medicine. (2012) AIUM Practice Guideline for the Performance of an Ultrasound Examination in the Practice of Urology. Journal of Ultrasound in Medicine. 31, pp. 133-144.

[9] Smail, L. C. , Dhindsa, K. , Braga, L. H. , Becker, S. , & Sonnadara, R. R. . (2020). Using deep learning algorithms to grade hydronephrosis severity: toward a clinical adjunct. Frontiers in Pediatrics, 8, 1-.

[10] Turco, S. , Frinking, P. , Wildeboer, R. , Arditi, M. , & Mischi, M. . (2020). Contrast-enhanced ultrasound quantification: from kinetic modeling to machine learning. Ultrasound in Medicine & Biology, 46(3).

[11] Shokoohi, H. , Lesaux, M. A. , Roohani, Y. H. , Liteplo, A. , Huang, C. , & Blaivas, M. . (2019). Enhanced point-of-care ultrasound applications by integrating automated feature-learning systems using deep learning. Journal of Ultrasound in Medicine, 38.

[12] Dhindsa, K. , Smail, L. C. , Mcgrath, M. , Braga, L. , & Sonna Da Ra, R. R. . (2018). Grading Prenatal Hydronephrosis from Ultrasound Imaging Using Deep Convolutional Neural Networks. 2018 15th Conference on Computer and Robot Vision (CRV). IEEE Computer Society.

[13] Blum, E. S. , Porras, A. R. , Biggs, E. , Tabrizi, P. R. , Sussman, R. D. , & Sprague, B. M. , et al. (2018). Early detection of ureteropelvic junction obstruction using signal analysis and machine learning: a dynamic solution to a dynamic problem. J Urol, S0022534717777774.

[14] Jianxing He et al. "The practical implementation of artificial intelligence technologies in medicine". In: Nature medicine 25.1 (2019), pp. 30–36

[15] Hillestad, R. , Bigelow, J. , Bower, A. , Girosi, F. , Meili, R. , & Scoville, R. , et al. (2005). Can electronic medical record systems transform health care? potential health benefits, savings, and costs. Health Affairs, 24(5), 1103-1117. DOI: 10.1377/hlthaff.24.5.1103

[16] Vimalachandran, P. , W. Hua , and Y. Zhang . "Securing Electronic Medical Record and Electronic Health Record Systems Through an Improved Access Control." Springer International Publishing Springer International Publishing, 2015. DOI: 10.1007/978-3-319-19156-0_3

[17] McDonald, & C., J. . (1997). The barriers to electronic medical record systems and how to overcome them. Journal of the American Medical Informatics Association, 4(3), 213.

[18] Selvaraj S , Sundaravaradhan S . Challenges and opportunities in IoT healthcare systems: a systematic review[J]. SN Applied Sciences, 2020, 2(1). DOI: 10.1007/s42452-019-1925-y

[19] Rahmadika, S. , Firdaus, M. , Jang, S. , & Rhee, K. H. . (2021). Blockchain-enabled 5g edge networks and beyond: an intelligent cross-silo federated learning approach. Security and Communication Networks, 2021(11), 1-14. DOI: 10.1155/2021/5550153

[20] Kumar, R. , Khan, A. A. , Zhang, S. , Wang, W. Y. , & Kumar, J. . (2020). Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging. DOI: 10.13140/RG.2.2.35590.80961

[21] Joyia, G. J. , Liaqat, R. M. , Farooq, A. , & Rehman, S. . (2017). Internet of medical things (iomt): applications, benefits and future challenges in healthcare domain. Journal of Communications, 12(4), 240-247.

[22] Krishna, K. D. , Akkala, V. , Bharath, R. , Rajalakshmi, P. , & Mohammed, A. M. . (2015). FPGA based preliminary CAD for kidney on IoT enabled portable ultrasound imaging system. 2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom). IEEE. DOI: 10.1109/HealthCom.2014.7001851

[23] Hengshuang Zhao et al. "Pyramid scene parsing network". In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, pp. 2881–2890.

[24] Sanghyun Woo et al. "Cbam: Convolutional block attention module". In: Proceedings of the European conference on computer vision (ECCV). 2018, pp. 3– 19.

[25] 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.

[26] B Vidakovic. (2008). Discrete Wavelet Transformation. John Wiley & Sons, Ltd.

[27] Benet, J.. (2014). Ipfs - content addressed, versioned, p2p file system. Eprint Arxiv.

[28] Smith, J. A. , and O. C. Jensen . "PORTABLE ULTRASOUND SCANNER.", US20110118562. 2011.patent

[29] Bharath, R. , D Chandrashekar, Akkala, V. , D Krishna, Ponduri, H. , & Rajalakshmi, P. , et al. (2016). Portable ultrasound scanner for remote diagnosis. International Conference on E-health Networking. IEEE.

[30] Portable ultrasound scanners. Shrinking size, growing market. Health Devices. 2002 Aug;31(8):279-94. PMID: 12530359.

[31] Portable ultrasound scanners. Health Devices. 2008 Aug;37(8):221-45. PMID: 19058439.

[32] Portable ultrasound scanners. Health Devices. 2003 Aug;32(8):285-320. PMID: 14513789.

[33] K. Fernbach, M. Maizels, and J. J. Conway, "Ultrasound grading of hydronephrosis: introduction to the system used by the Society for Fetal Urology," Pediatric Radiology, vol. 23, no. 6, pp. 478–480, 1993.

[34] A. Onen, "An alternative grading system to refine the criteria for severity of hydronephrosis and optimal treatment guidelines in neonates with primary UPJ-type hydronephrosis," Journal of Pediatric Urology, vol. 3, no. 3, pp. 200–205, 2007.

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

[36] Rahman, M. A., & Yang, W.. (2016). Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation. International Symposium on Visual Computing. Springer, Cham.

First A. Author (M'76–SM'81–F'87) and all authors may include biographies. Biographies are often not included in conference-related papers. This author became a Member (M) of IEEE in 1976, a Senior Member (SM) in 1981, and a Fellow (F) in 1987. The first paragraph may contain a place and/or date of birth (list place, then date). Next, the author's educational background is listed. The degrees should be listed with type of degree in what field, which institution, city, state, and country, and year the degree was earned. The author's major field of study should be lower-cased.

The second paragraph uses the pronoun of the person (he or she) and not the author's last name. It lists military and work experience, including summer and fellowship jobs. Job titles are capitalized. The current job must have a location; previous positions may be listed without one. Information concerning previous publications may be included. Try not to list more than three books or published articles. The format for listing publishers of a book within the biography is: title of book (publisher name, year) similar to a reference. Current and previous research interests end the paragraph.

The third paragraph begins with the author's title and last name (e.g., Dr. Smith, Prof. Jones, Mr. Kajor, Ms. Hunter). List any memberships in professional societies other than the IEEE. Finally, list any awards and work for IEEE committees and publications. If a photograph is provided, it should be of good quality, and professional-looking. Following are two examples of an author's biography.

Second B. Author was born in Greenwich Village, New York, NY, USA in 1977. He received the B.S. and M.S. degrees in aerospace engineering from the University of Virginia, Charlottesville, in 2001 and the Ph.D. degree in mechanical engineering from Drexel University, Philadelphia, PA, in 2008.

From 2001 to 2004, he was a Research Assistant with the Princeton Plasma Physics Laboratory. Since 2009, he has been an Assistant Professor with the Mechanical Engineering Department, Texas A&M University, College Station. He is the author of three books, more than 150 articles, and more than 70 inventions. His research interests include high-pressure and high-density nonthermal plasma discharge processes and applications, microscale plasma discharges, discharges in liquids, spectroscopic diagnostics, plasma propulsion, and innovation plasma applications. He is an Associate Editor of the journal *Earth*, *Moon*, *Planets*, and holds two patents.

Dr. Author was a recipient of the International Association of Geomagnetism and Aeronomy Young Scientist Award for Excellence in 2008, and the IEEE Electromagnetic Compatibility Society Best Symposium Paper Award in 2011.

Third C. Author, Jr. (M'87) received the B.S. degree in mechanical engineering from National Chung Cheng University, Chiayi, Taiwan, in 2004 and the M.S. degree in mechanical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2006. He is currently pursuing the Ph.D. degree in mechanical engineering at Texas A&M University, College Station, TX, USA.

From 2008 to 2009, he was a Research Assistant with the Institute of Physics, Academia Sinica, Tapei, Taiwan. His research interest includes the development of surface processing and biological/medical treatment techniques using nonthermal atmospheric pressure plasmas, fundamental study of plasma sources, and fabrication of micro- or nanostructured surfaces.

Mr. Author's awards and honors include the Frew Fellowship (Australian Academy of Science), the I. I. Rabi Prize (APS), the European Frequency and Time Forum Award, the Carl Zeiss Research Award, the William F. Meggers Award and the Adolph Lomb Medal (OSA).