

WebRTC Based Invariant Scattering Convolution Network for Automated Validation of Ultrasonic Videos for IoT Enabled Tele-Sonography

R Bharath, and P Rajalakshmi
WiNet Research Lab, Department of Electrical Engineering
Indian Institute of Technology Hyderabad, India
email: { ee13p0007, raji }@ iith.ac.in

Abstract—Tele-sonography works on inherent assumption that the transmitted medical ultrasound videos scanned from remote patients contain the representative data for doing the diagnosis. Due to the high subjectivity involved in scanning and semi-skilled nature of the operating person, this assumption may not always be valid. The remotely scanned ultrasound video contains a lot of redundant information, which is not useful for diagnosis. Transmitting redundant and large volumes of medical data to the expert end for analysis may lead to faulty diagnosis, associated with high transmission cost, and also poses serious challenges on data storage, processing, infrastructure, etc. Addressing these issues, we propose a novel WebRTC based framework to detect representative frames in the ultrasound video and transmit only those frames to the remote sonographer for getting a diagnosis. Detection of representative frames in ultrasound video is done with invariant scattering convolution network. The entire framework is developed using WebRTC, which enables the browser to browser communication thus reducing the computation on end ultrasound scanner and ensures ubiquitous and secured connectivity between technician and the sonographer. The proposed video validation algorithm achieved an accuracy of 96.5% in classifying the representative frames and nonrepresentative frames in the ultrasound video.

I. INTRODUCTION

Internet of Things (IoT) has been revolutionized the healthcare, which prompted to connect and integrate every medical device to the Internet [1] [2]. According to business insider survey, by the year 2020, 34 billion devices are going to connect the Internet leading to hyper connectivity [3]. These IoT medical devices generate and exchange a huge amount of data among them, and causes a serious challenge on data storage, bandwidth requirements, data processing, etc. Most of the data generated by the medical IoT devices are not useful for diagnosis, which is resulted due to the improper acquisition and continuous monitoring of data from remote patients [4]. Monitoring large volumes of medical data in cloud demands for more medical experts, and the experts need to go through large volumes of data which all are not useful for diagnosis. Hence, there is a necessity to devise automated algorithms at the end devices to filter insignificant data before transmitting it to the cloud for analysis. Motivated by the problem, we proposed a Web Real-Time Communication (WebRTC) based diagnostically driven compression algorithm for IoT enabled tele-sonography [5].

Non-invasive medical imaging like ultrasound scanning, Magnetic Resonance Imaging (MRI), Computed Tomogra-

phy (CT), Positron Emission Tomography (PET), etc. covers a wide spectrum of diagnosis in healthcare. The services of medical imaging have been limited to centralized hospitals due to high form factor and need for expertise to operate the device at remote locations. Unlike MRI, CT and PET, ultrasound scanners are portable, capable of real-time imaging and considered to be safest diagnosing modality. Ultrasound scanning is capable of diagnosing almost all organs in the body including kidney, liver, heart, fetus monitoring, etc. Advancement in computing platforms like FPGAs, DSPs, GPUs, etc. have reduced the size of the ultrasound scanning to the portable level making the device flexible to use in Point of Care (PoC) diagnostics such as in ambulances, military, remote healthcare, etc [6]–[9]. PoC diagnosis is highly beneficial to infants, pregnant women, senior citizens and the patients who suffer from mobility problems. Ultrasound scanning is not used to its potential in PoC diagnostics due to lack of sonographers.

Tele-sonography is used to address lack of sonographers, where a nonexpert can scan the patient and send the scanned ultrasound data to the sonographer present elsewhere via wired/wireless communication technologies for diagnosis [10]–[12]. Recently, smartphone-based ultrasound scanners are available for clinical practices, thus brings an inherent advantage of high mobility and data connectivity which further improved its access for extensive use in tele-sonography. Voice over IP (VoIP) along with smartphone-based ultrasound scanners is a potential resource for tele-sonography, which can unleash the potentiality of ultrasound scanning to use in remote and PoC diagnostics. In [13], we analyzed the performance of popular VoIP services like Skype, Facebook and WebRTC for real-time tele-sonography, and found that these services are viable for real-time tele-sonography provided stable network conditions.

In tele-sonography, an inherent assumption is made that the scanned ultrasound video consists of representative data (data having sufficient information) for doing the diagnosis. But, due to the subjectivity involved in scanning and semiskilled nature of the operating person, the scanned ultrasound video may contain a lot of frames which are not useful for diagnosis. Transmitting all these data to the expert end will not be useful for diagnosis and involves high cost due to transmission of huge data. Additionally, the expert

sitting in the cloud has to go through the huge volume of data, which is less representative for doing the diagnosis. If the transmitted data is not representative enough, it may also lead to faulty diagnosis, or the expert can request to rescan the patient which is going to be tedious and time-consuming.

In this paper, we are proposing a novel algorithm for automatic validation of scanned ultrasound videos before transmitting it to the expert for analysis. If the scanned ultrasound video contains less number of representative frames, then the non-expertise can go for rescanning until he obtains the representative data for diagnosis. The proposed algorithm is based on a supervised learning algorithm. Scattering coefficients [14] are used as features for representing each frame, while linear SVM [15] classifier is used to classify each frame as valid or invalid based on the extracted features. The entire framework is implemented using the WebRTC [16]. The graphical representation of proposed framework for WebRTC based video validation algorithm for tele-sonography is shown in Fig. 1. Upon request, the application can read each frame in the incoming ultrasound video, detect valid frames and sends the valid data to the remote sonographer via data channel.

II. INVARIANT SCATTERING CONVOLUTION NETWORK BASED ULTRASOUND VIDEO VALIDATION

A. Problem formulation

In general, the ultrasound diagnosis is done by freezing the representative frames in the ultrasound video and analyzing it. The ultrasound video consists of frames which are useful and not useful for diagnosis. The frames with low contrast and the regions which appear homogeneous throughout the image without any structure are regarded as the invalid or nonrepresentative frames. The invalid images occurs due to improper adjustment of gain knobs, focusing depth, inadequate gel between transducer and skin, etc. Sonographers cannot infer any information from these frames, and hence it is regarded as invalid frames. The valid frames are considered as images with good contrast and consist of edges or structures corresponding to the shape of an organ. The organs are composed of soft tissues, blood vessels, etc., representing

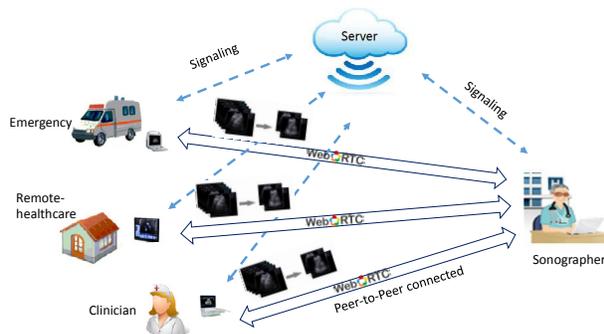


Fig. 1: Architecture of WebRTC based automated validation of ultrasound videos for IoT enabled tele-sonography.

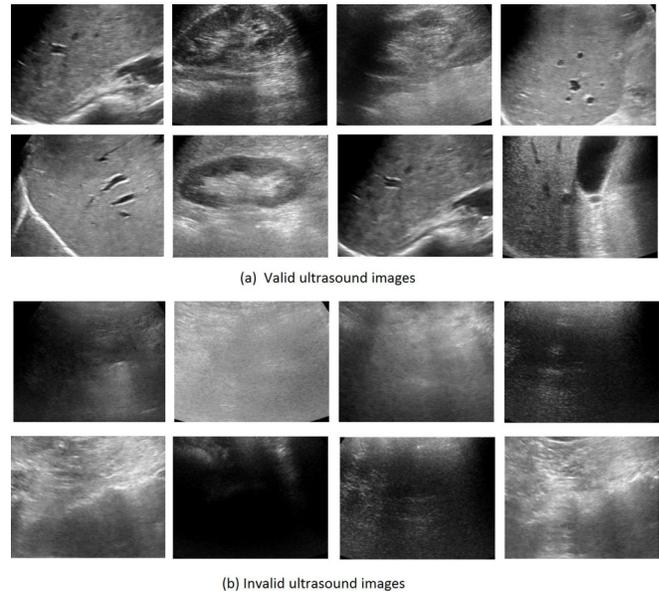


Fig. 2: (a) Valid ultrasound images, where structures of different organs can be seen. (b) Invalid ultrasound image, which appears homogeneous all over the image with low contrast.

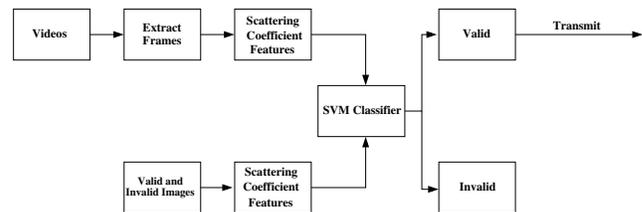


Fig. 3: Block diagram representation of the proposed ultrasound video validation algorithm.

different structures in an ultrasound image. The valid and invalid images representing organs and no organs is shown in Fig. 2. Video validation is formulated as detection of a number of valid frames present in the scanned video useful for diagnosis.

B. Proposed automated video validation algorithm

The block diagram representation of the proposed video validation algorithm is shown in Fig. 3. In training phase, a linear SVM classifier is trained with the scattering coefficient features of valid and invalid images. In the testing phase, every frame of video is classified as valid and invalid based on the scattering coefficient features of a frame. The valid frames are only transmitted to the sonographer while dropping the invalid frames.

C. Region of Interest (RoI) detection

The organ information embedded in the ultrasound image depends on the type of ultrasound probe and excitation used for scanning the organ. Linear probes with linear excitation

results in rectangular images, where organ information is present all over the image. Linear and curvilinear probe with phased array excitation results in sector images. In sector images, the organ information does not present all over the image, and it is constrained to center part of the image. Features are extracted only from the regions where most of the organ information is present. The ultrasound videos are of resolution 768×1024 . From ultrasound videos, an inference is made that a rectangular region of width 340×430 at the center of ultrasound video can cover a maximum region containing organ information as shown in Fig. 4. RoI is selected such that it does not include the outer portion of ultrasound image which does not represent organ information (pixels lying outside of the sector which has zero gray level).

D. Feature extraction

Scattering coefficients (SC) are used as features for classifying each frame of ultrasound video. The SC are translational and deformation invariant and provides robust representation for image classification. SC are computed via Invariant Scattering Convolution Network (ISCN) architecture. The detailed information regarding the ISCN architecture for computing the SC can be found in [14], [17]. The efficiency of SC in classification lies in building invariant representation for images; it builds the invariant features by progressively cascading wavelet transform with modulus and averaging operators. Since high-frequency components are the main source of variability, ISCN maps high-frequency wavelet coefficients to lower frequencies and averages the lower frequencies to get translational invariant representation. Deformation invariance to the features is obtained with the wavelet transform since wavelets are localized waveforms stable to deformations. The layer-wise mathematical operations involved in ISCN are explained below.

The first layer of the ISCN starts with low pass filtering of an image x .

$$S_0 = x \star \phi_{2^J}, \quad (1)$$

where, ϕ denotes Gaussian low-pass filter, \star denotes the convolution operation and J corresponds to the scale space

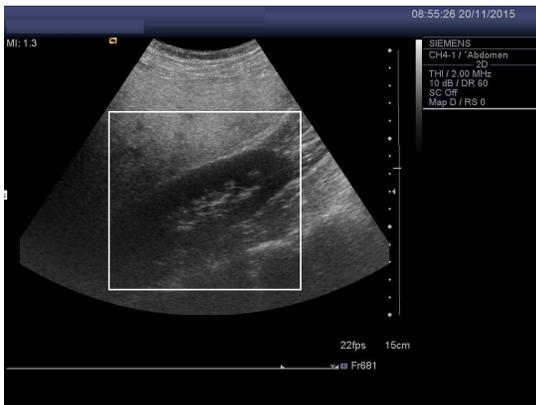


Fig. 4: Ultrasound image scanned with curvilinear probe, rectangular box indicates the RoI used to extract features.

variable. In the second layer, modulus of the complex wavelet transform of an image is convolved with low pass filter.

$$S_{\lambda_1} = |x \star \psi_{\lambda_1}| \star \phi_{2^J}. \quad (2)$$

ψ is a complex wavelet filter represented as

$$\psi(u) = \psi_a(u) + i\psi_b(u); u = [u_1, u_2], \quad (3)$$

u represent the spatial location of the pixel in the image. The term $\psi_\lambda(u) = 2^{-2j}\psi(2^{-j}r_\theta u)$ represents all the rotated and dilated versions of the wavelet with $\lambda = (2^j, \theta)$, $0 \leq j < J$, j is a scale space variable, θ is given by $2\pi l/L$, where $0 \leq l < L$ and r represent group of rotations θ . The third layer of the network is obtained as

$$S_{\lambda_1, \lambda_2} = ||x \star \psi_{\lambda_1}| \star \psi_{\lambda_2}| \star \phi_{2^J}. \quad (4)$$

The coefficients are computed only for the scales $2^{j_2} < 2^{j_1}$, since coefficients of S_{λ_1, λ_2} becomes negligible at $2^{j_2} \geq 2^{j_1}$ [18]. Similarly, the network can be extended to deeper layers in the following way.

$$S_{\lambda_1, \lambda_2, \dots, \lambda_m} = ||\dots|x \star \psi_{\lambda_1}| \star \psi_{\lambda_2}| \dots \psi_{\lambda_m}| \star \phi_{2^J}. \quad (5)$$

After convolving the modulus of complex wavelet transform with ϕ_{2^J} , the image is subsampled with an interval 2^J . The coefficients obtained after subsampling are the SC which is going to be translational and deformation invariant. The complex wavelet ψ is chosen to be Morlet wavelet given by

$$\psi(u) = e^{-|u|^2/(2\sigma^2)} \alpha(e^{i \cdot u \cdot \zeta} - \beta). \quad (6)$$

The Morlet wavelet is obtained by multiplying the Gaussian window with complex exponential. Here β is adjusted such that the area under the wavelet becomes zero. For experimentation analysis, $\sigma = 0.8$, $\zeta = 3\pi/4$ is chosen.

SC computed from all the layers are used as features for classification. As the depth of the network increases,

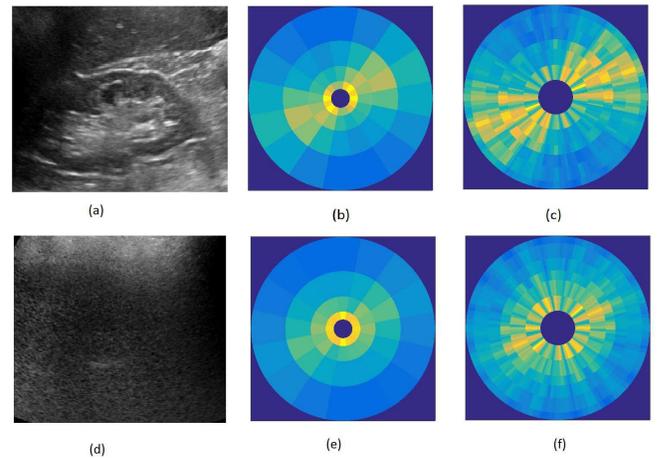


Fig. 5: Perceptual difference observed in SC of valid and invalid images. (a), (d). Valid and Invalid image. (b), (e). First order SC of valid and invalid image. (c), (f). Second order SC of valid and invalid image respectively.

more invariant and discriminative features are obtained. To visualize the discriminative nature of SC, the disk covering the entire frequency support of image is displayed as sectors as shown in Fig. 5. If $\psi(w)$ is the Fourier transform centered at frequency η then $\psi(w)_{2^{-j}r}$ has a support centered at $2^{-j}r$ with bandwidth proportional to 2^{-j} . Each sector in the disk corresponds to local Fourier transform energy of an image over support of $\psi_\lambda(w)$. From Fig. 5, we can infer that SC gives discriminative features for valid and invalid images both in the second and third layer. The SC obtained in each index is of dimension $[M/2^{J-1}, N/2^{J-1}]$, where M, N are the dimensions of the image. The feature representation for each index is computed by summing all the SC in each index resulting in one feature. To reduce the number of computations in computing the SC, the convolution operations are computed in the frequency domain. The operations involved in computing the SC is shown in Fig. 6.

III. FEATURES USED IN THE LITERATURE FOR ULTRASOUND IMAGE CLASSIFICATION

Some of the popular features used to classify the ultrasound images are discussed below.

A. GLCM features

The Gray Level Co-occurrence Matrix (GLCM) captures the spatial relationship between the pixels present in an image. The spatial relationship between the pixels is characterized regarding how often two pixels with intensities i and j occur in specific direction and distance [19]. The texture features of four GLCM's corresponding to directions $0^\circ, 45^\circ, 90^\circ, 135^\circ$ with distance between two pixels being one unit is computed [20]. From each GLCM, 13 features were computed constituting a total of 52 features for four GLCM.

B. GLRLM features

The Gray Level Run Length Matrix (GLRLM) features capture the texture information by computing the run-length of a pixel with specific gray value occurring in a specific

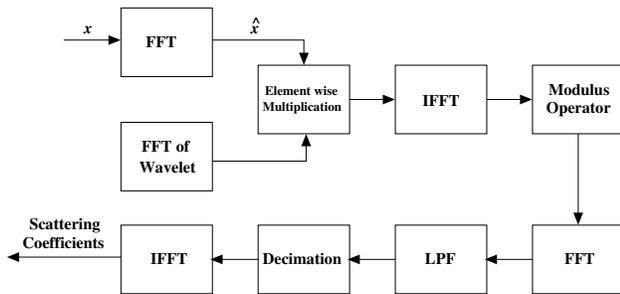


Fig. 6: Computational blocks involved in computing SC. FFT and IFFT represents 2D fast Fourier transform and inverse fast Fourier transform respectively. LPF denotes low pass filtering.

direction [21]. Eleven GLRLM's corresponding to each direction $0^\circ, 45^\circ, 90^\circ, 135^\circ$ are computed, constituting a total of 44 features [22].

C. Multiresolution features

The multiresolution framework proposed in [20] has been used in this paper for comparison. The images are decomposed using M-band wavelet and Gabor filter bank. From each decomposed sub-image, features like energy and energy deviation are computed. Using M-band wavelet, the image is decomposed into 45 sub-images, resulting in 45 M-band wavelet magnitude (Wav_Mag) and wavelet energy (Wav_Eng) features. Gabor filter bank with five radial frequencies ($\sqrt{2}/2^5, \sqrt{2}/2^4, \sqrt{2}/2^3, \sqrt{2}/2^2$ and $\sqrt{2}/2^1$), six orientations ($0^\circ, 30^\circ, 60^\circ, 90^\circ$ and 150°) have been used for obtaining a total of 31 sub-images. Energy in Gabor sub-images (Gabor_Eng) and Energy deviation in Gabor sub-images (Gabor_Dev) are used as the features for representing the image.

IV. WEBRTC IMPLEMENTATION

WebRTC [16] provides browsers real-time communication capabilities with a set of application programming interfaces (APIs) and communication protocols. WebRTC is an open source project supported by Google, Mozilla and Opera. The standardization of WebRTC is jointly developed by World Wide Web Consortium (W3C) and the Internet Engineering Task Force (IETF). The set of real-time communication protocols is standardized by IETF, while W3C standardizes APIs. Before WebRTC, to enable voice and video services within the browser, the user has to install that particular service real-time communication stack as plug-in. The multimedia communication stack is built into the web browser internals with WebRTC. Hence the developers can make use of HTML5 APIs for developing multimedia services. WebRTC uses real-time congestion control algorithm proposed by Google [23]. The complete receiver side congestion of WebRTC is reported in [24]. WebRTC did not standardize the signaling protocol and the users can choose any preexisting signaling protocol like session initiation protocol, jingle, etc. WebRTC eliminates the need for installing third-party plugins like flash players, customized multimedia stacks for playing multimedia content making it widely adaptable in developing videophone services [25].

WebRTC provides a peer-to-peer communication by extending client-server semantics. In general, WebRTC uses Trapezoidal or Triangle architectural web server models. In WebRTC Trapezoid model, both browsers are running a web application downloaded from a different server, while in Triangle model both browsers are running the same web application downloaded from the same webpage. Here, we employed the Triangle model web architecture as shown in Fig. 7. The web application is implemented on a dedicated server using nodejs [26]. To traverse firewalls and NATs, we run multiple TURN servers on dedicated machines. Websockets is used for signaling between browsers and server. WebRTC supported browsers Google Chrome and

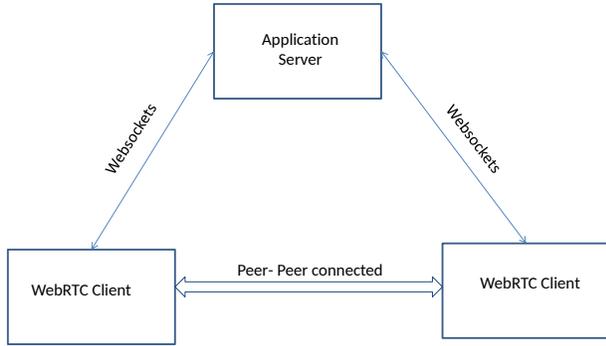


Fig. 7: WebRTC network architecture.

Firefox are used to establish a connection with a web server and to download the WebRTC javascript to configure the browser internals during call establishment. If the end points are behind NATs and firewalls, the media data flows through TURN servers else it flows directly between endpoints. The server application is run on a Intel core i7 processor with 16 GB RAM running on a 2.8 GHz clock cycle.

V. RESULTS

The images for this study is acquired using a Siemens Acuson S1000 ultrasound scanner with a phased array transducer from Asian Institute of Gastroenterology, Hyderabad, India. The database consists of 409 valid, 192 invalid and 20 ultrasound videos. The ground truth for the image is jointly annotated by the two sonographers, who are vastly experienced in sonography for more than 20 years. The valid ultrasound images include 160 kidneys, 125 liver, 80 spleen and 44 cardiac ultrasound images. Out of 20 ultrasound videos, ten correspond to the kidney, six correspond to the liver and four videos correspond to the spleen. The invalid images are captured corresponding to improper settings of gain knobs, focusing depth and the images appeared homogeneous resembling no information etc.

The results of the classification algorithm depend on the number of scales and orientations used to generate the SC. Table. I refers the accuracy of the proposed algorithm referring to different scales and orientations. SVM with a linear kernel is used to classify the extracted features, and ten-fold cross-validation is employed to validate the proposed algorithm. The accuracy of the algorithm increased with increase in scales and orientations. The proposed algorithm resulted with a maximum accuracy of 96.5% for four scales and six orientations. Further increase in scales and orientations did not increase the accuracy of the algorithm. Hence, the efficiency of the ISCN is studied with four scales and six orientations.

The classification accuracy of the SC features with respect to network depth is shown in Table. II. Highest accuracy is achieved for a network of depth 2. The classification accuracy reduced as further increase in the depth of the network. This is resulted due to the propagation of redundant information from parent layers to the child layers.

The confusion matrix of the proposed algorithm in classifying the valid and invalid images is shown in Table. III. The algorithm classified valid images with a sensitivity of 98% (401 out of 409) and invalid images with a specificity of 93.2% (179 out of 192). High sensitivity is preferred for the proposed application since we do not want valid images to classify as invalid images, while low specificity is allowed since the cost paid in transmitting invalid images is less.

The accuracy of SC compared to other textural features is shown in Table. IV. Linear SVM classifier is used to evaluate the performance of all the features. The proposed SC features performed with the highest accuracy of 96.5% and the next best accuracy 92.5% is achieved with Wavelet magnitude and standard deviation of the Gabor coefficients, while GLRLM features performed with the lowest accuracy of 87.7%.

In general, compression depends on the number of valid and invalid frames present in the scanned video. To quantify the amount of compression achieved, the proposed algorithm is tested on five ultrasound videos, which includes two liver, two kidneys, and one spleen ultrasound videos. The number of valid and invalid frames present in each video is shown in Table. V. The original videos are in avi format, and the valid frames extracted from ultrasound video is stored in jpg format. By transmitting only valid images of five ultrasound videos to the cloud, 66% of the reduction in overall data is achieved.

Since the smartphone-based ultrasound scanners are presently available for scanning, which inherently comes with Internet compatibility, the proposed WebRTC framework can provide easy access to connect any expertise with high security. The expert can be located anywhere and need to have a multimedia device with Internet connectivity for accessing the ultrasound data for analysis. Additionally, the proposed video validation algorithm filters the insignificant data getting into the cloud thus helping the experts to diagnose more patients in less time. By knowing the

TABLE I: Accuracy(%) of the SC features for different scales and orientations with linear SVM classifier, depth of the layer=2. Features from all the layers are concatenated as a single feature vector to represent each image.

Scales (j)	Orientations (θ)							
	1	2	3	4	5	6	7	8
2	88.5	91.5	92	92.5	92.5	93.2	93.5	93.5
3	91.3	92.5	94.0	94.2	94.5	94.8	94.5	95.7
4	92.2	95.5	96.2	96.5	96.2	96.5	96.0	95.8
5	95.7	96.0	95.0	96.0	95.8	96.0	95.5	95.3
6	95.7	96.2	96.3	96.3	96.3	96.0	96.0	96.0

TABLE II: Efficiency of the SC features in classification with respect to network depth with $j=4$ and $\theta=6$.

Network depth	Feature Dimension	Accuracy(%)
1	17	93.8
2	113	96.5
3	369	95.3

TABLE III: Confusion matrix of the proposed algorithm tested on the ultrasound images.

True class	Predicted class	
	Valid	Invalid
Valid (409)	401	8
Invalid (192)	13	179

TABLE IV: Performance comparison of features with linear SVM classifier.

Features	Accuracy(%)
GLCM	91.2
GLRLM	87.7
Wavelet_Eng	89.4
Wavelet_Mag	92.5
Gabor_Eng	91.8
Gabor_Dev	92.5
Proposed Method	96.5

TABLE V: Reduction in size of ultrasound data after video validation.

Video	Total Frames	Valid Frames	Invalid Frames	Video size(MB)	Data after validation(MB)
1	78	45	33	7.9	3.7
2	47	10	37	4.9	0.8
3	47	23	24	4.6	1.8
4	75	37	38	7.8	1.9
5	74	24	50	4.5	1.8

number of representative frames present in scanned video, the technician can scan better by adjusting the scanning settings or can go for rescanning to get more representative data for diagnosis. The proposed framework will unleash the potentiality of ultrasound scanning to widely use in PoC diagnostics, thus bringing a significant impact in IoT driven healthcare.

ACKNOWLEDGMENT

The authors are thankful to Dr. Mateen Mohammad, radiologist, Hyderabad, India and his team for providing the ultrasound database, and for their consistent support and feedback for developing the framework. This work was partially supported and funded by Department of Science and Technology (DST) India under the project "Data Science-based Farming Support System For Sustainable Crop Production Under Climatic Changes (DSFS)" project no: MST/IBCD/EE/F066/2016-17G48.

REFERENCES

- [1] B. Xu *et al.* "Ubiquitous Data Accessing Method in IoT-Based Information System for Emergency Medical Services," in *IEEE Trans. on Industrial Informatics*, vol. 10, no. 2, pp. 1578-1586, May 2014.
- [2] A. Zanella, *et al.* "Internet of Things for Smart Cities," in *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22-32, Feb. 2014.
- [3] Business Insider "BI Intelligence projects 34 billion devices will be connected by 2020" Nov.6, 2015.
- [4] R. Bharath, and Pachamuthu Rajalakshmi. "Automatic organ validation of b-mode ultrasound images for transmission to cloud." In *Advances in Computing, Communications and Informatics (ICACCI)*, International Conference on, pp. 479-482. IEEE, 2014.
- [5] R. Bharath, *et al.* "Implementation of diagnostically driven compression algorithms via WebRTC for IoT enabled tele-sonography." *Biomedical Engineering and Sciences (IECBES)*, *IEEE*, 2016.
- [6] Kang, Jeeun, *et al.* "A System-on-Chip Solution for Point-of-Care Ultrasound Imaging Systems: Architecture and ASIC Implementation." *IEEE Trans. on biomedical circuits and systems* vol. 10, no. 2, pp. 412-423, 2016.
- [7] Tang, *et al.* "Miniaturizing Ultrasonic System for Portable Health Care and Fitness." *IEEE Trans. on biomedical circuits and systems* vol. 9, no. 6, 767-776, 2015.
- [8] Stawicki, Stanislaw Peter, and David Paul Bahner. "Modern sonology and the bedside practitioner: evolution of ultrasound from curious novelty to essential clinical tool;" *European Journal of Trauma and Emergency Surgery* 41.5, 457-460, 2015.
- [9] Bharath, R., *et al.* "Portable ultrasound scanner for remote diagnosis." E-health Networking, Application & Services (HealthCom), 2015 17th International Conference on. *IEEE*, 2015.
- [10] C. Christodoulou and E. Rajo-Iglesias, "A tutorial for emerging wireless medical video transmission systems, *IEEE Antennas Propag. Mag.*, vol. 53, no. 2, pp. 202-213, Apr. 2011.
- [11] C.S. Pattichis, *et al.* "Wireless telemedicine systems: an overview." *IEEE Antennas and Propagation Magazine* 44, 2, pp. 143-153, 2002.
- [12] Panayides, Andreas S., *et al.* "Mobile-health systems use diagnostically driven medical video technologies." *IEEE Signal Proc. Mag.*, vol. 30, no. 6, pp. 163-172, yr. 2013.
- [13] R Bharath, P Rajalakshmi, "Subjective Liver Ultrasound Video Quality Assessment of Internet based Videophone Services for Real-Time Telesonography" Accepted for publication in *IEEE Healthcom'*, yr. 2017.
- [14] Bruna, Joan, and Sthpane Mallat. "Invariant scattering convolution networks." *IEEE transactions on pattern analysis and machine intelligence* 35.8, 1872-1886, 2013.
- [15] Chang, Chih-Chung, and Chih-Jen Lin. "LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2.3, 27, 2011.
- [16] WebRTC, "<https://webrtc.org/>"
- [17] Mallat, Sthpane. "Understanding deep convolutional networks. *Phil.Trans. R. Soc. A* 374.2065, 20150203, 2016.
- [18] Mallat, Sthpane. "Group invariant scattering." *Communications on Pure and Applied Mathematics* 65.10, 1331-1398, 2012.
- [19] Haralick, Robert M., and Karthikeyan Shanmugam. "Textural features for image classification." *IEEE Transactions on systems, man, and cybernetics* 3.6, 610-621, 1973.
- [20] Wu, Cheng-Chi, *et al.* "Evolution-based hierarchical feature fusion for ultrasonic liver tissue characterization." *IEEE journal of biomedical and health informatics*, 17.5, 967-976, 2013.
- [21] Galloway, Mary M. "Texture analysis using gray level run lengths." *Computer graphics and image processing* 4.2, 172-179, 1975.
- [22] Tang, Xiaou. "Texture information in run-length matrices." *IEEE transactions on image processing*, 7.11, 1602-1609, 1998.
- [23] H. Alvestrand, S. Holmer, and H. Lundin, "A Google Congestion Control Algorithm for Real-Time Communication on the World Wide Web," 2013, IETF Internet Draft.
- [24] Varun Singh, *et al.* "Performance analysis of receive-side real-time congestion control for WebRTC." In 2013 20th International Packet Video Workshop, pp. 1-8. *IEEE*, 2013.
- [25] Google: What's next for WebRTC? [online] <https://www.youtube.com/watch?v=HCE3S1E5UwY>
- [26] nodejs "<https://nodejs.org/en/>"