International Journal of Computer Engineering in Research Trends
Multidisciplinary, Open Access, Peer-Reviewed and fully refereed
Research Paper Volume-9, Issue-10, 2022 Regular Edition E-ISSN: 2349-7084

Selection of MSER region based Ultrasound Doppler scan Image Big data classification using a faster RCNN network

S. Sandhya kumari[1], K.Sandhya Rani[2]

[1]*Research Scholar, Dept. of Computer Science, SPMVV,Tirupati,India.*
[2]*Professor, Dept. of Computer Science, SPMVV, Tirupati, India.*
Email ID: [1]sridharamsandhya@gmail.com, [2] sandhyaranikasireddy@yahoo.co.in
Corresponding author : K.Sandhya Rani
Available online at: http://www.ijcert.org

Received: 30/08/2022, Revised: 23/09/2022, Accepted: 06/10/2022, Published: 26/10/2022

Abstract: This paper proposes an ultrasound Doppler scan image big data classification approach that uses a selection process to estimate the best regions for extracting the feature of a faster region-based convolutional neural network (RCNN) network. This scheme initially pre-processes the Doppler scan images. From the pre-processed image, several maximally stable extremal regions (MSER) and residual regions are estimated. The residual region and a few of the regions selected from the stable regions are used to extract the features. A correlation-based approach is used to select the stable regions for extracting the features. The gradient values of selected regions are used to extract the triangular vertex transform-based features (TVT). The extracted TVT features are trained using the faster RCNN network to categorize the ultrasound Doppler scan image as the femur, brain, abdomen,cervix, thorax, and other regions. The evaluation metrics namely precision, recall, and F1-score are used to validate the algorithm. The proposed Doppler ultrasound classification approach provides a sensitivity, F1-score, precision, specificity, and accuracy of 96.13%, 94.74%, 94.26%, 98.82%, and 98.27% respectively.
Keywords: — Doppler scan image classification, MSER region, TVT feature extraction, faster RCNN, Image gradient

---

1. Introduction

During gestation, the maternal structure of the fetal growth can be examined using the ultrasound Doppler scan image. This ultrasound Doppler scan image is less expensive in examining the fetal weight, blood flow, internal structure of the mother, and the fetal organs. In magnetic resonance imaging (MRI) the moment of the fetus will degrade the MRI scan image quality and it is difficult to examine the internal structure of the mother and the fetal organs. The advantage of ultrasound Doppler scan image [1] over MRI scan image is that the visual quality is high in ultrasound Doppler scan image during the movement of the fetus. Even though examination like fetal electrocardiogram helps to understand the function of organs like the heart [2]. Examination of the internal structure of the fetus will help to predict fetal health and the complication during the pregnancy and delivery. In the manual classification of ultrasound Doppler scan images, the inter and intra-observer variation can lead to wrong classification results. A single ultrasound examination will generate more than 20 images during the second trimester. Classifying more than 20 images by the physician utilizes more time and also causes wrong classification results, gestation age estimation, and the position of the fetus. Even though the Doppler scan image has disadvantages like adipose tissue beam attenuation, low contrast, and a low field of view, it has a valuable advantage that it is less harmful to the mother and fetus.

Machine and deep learning algorithms play a crucial role in the classification of biomedical scan images like dermatology [3], Ultrasound scan [4], radiology [5], MRI imaging [6], fundus images[7], etc. Several feature extraction algorithms namely local binary pattern (LBP) [8], local directional pattern (LDP) [9], and histogram of oriented gradient (HOG) [10], etc. The extracted features are trained using the machine learning classifiers namely multilayer perceptron, multiclass support vector machine,

etc. The deep learning algorithm is preferred in classification applications where it is difficult to detect the region of interest or extraction of features. A semi-automatic classification was proposed by Ryou et al. [11] which also requiresthe position of the biometry plane for the detection of the body/head of the fetus. A convolutional neural network (CNN) with multiple posses was proposed by Li et al.[12] that uses an iterative approach in the detection of the fetal brain.Conditional random field (CRF) was applied to video frames by the author Maraci et al. [13] for heartbeat detection. The authors Sushma et al. [14] used the statistical features in the estimation of complex patterns for training in support vector machines. The authors Sibo et al. [15] used an explainable residual learning approach in the detection of congenital heart disease with 93% accuracy. Selvanthi et al. [16] used classifiers such as AlexNet, and CNN, where the classifier AlexNet provides maximum accuracy of 90.43% in the detection of abnormality. The scheme [17] categorized 14 different fetal structures using global and local features. Diagonal quadratic discriminate analysis along with Naïve Bayes classifier and k-nearest algorithm was proposed by Attallah et al. [18] that can categorize the brain scan images as normal and abnormal. The fetal tissue and amniotic fluid in fetal body detection were proposed by Li et al. [17] which provides 67% accuracy in classification.

In our previous work, an empirical mode decomposition (EMD) is used to decompose [19] the image to several intrinsic mode functions. Instead of using the entire frequency bands, a few essential frequency bands are selected for feature extraction. Instead of using a few frequency bands for feature extraction, a few essential regions are detected using the MSER algorithm. However, the extracted features on the essential regions are classified using a decision tree classifier. In order to improve the performance, a deep learning algorithm with faster RCNN [20] is used in the proposed work. The rest of the paper is arranged as follows. Section 2 shows the proposed doppler scan image classification system, Section 3 shows the experimental results, and finally, section 4 concludes the work.

2.Proposed ultrasound doppler scan image classification

The proposed ultrasound doppler scan image classification consists of processes namely (i) Pre-processing (ii) selection of unstable regions, (iii). Extraction of TVT features and (iv)Faster RCNN classifier with a dual sigmoid activation function.

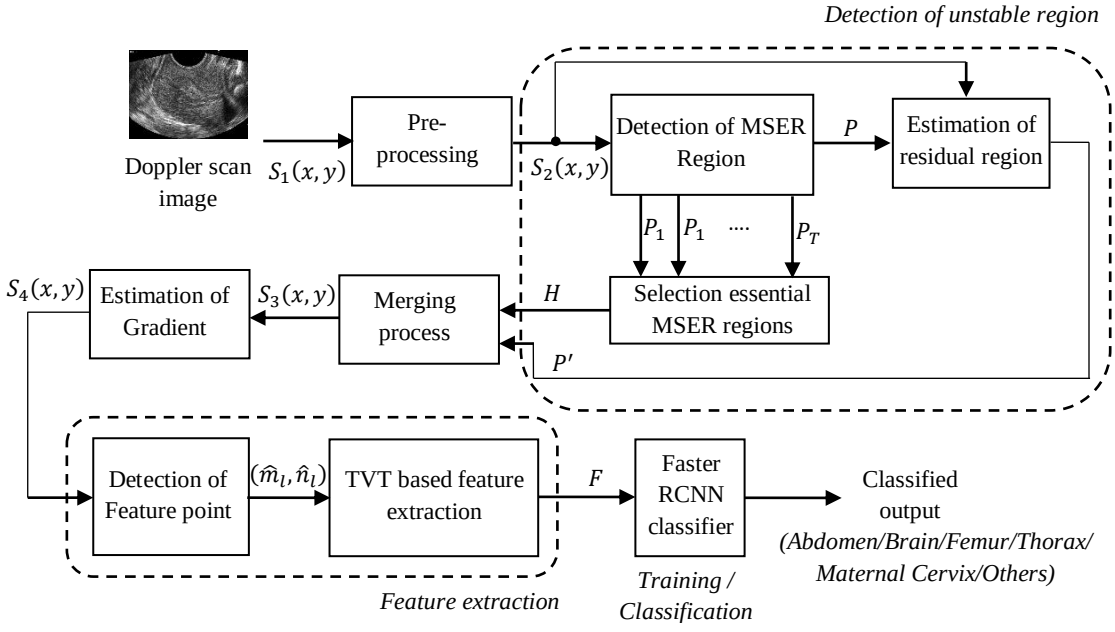


Fig 1: Block diagram representation of proposed ultrasound Doppler scan image classification

(a). Pre-processing

Let the input Doppler scan image be $S_1(x,y)$. The pre-processing includes the filtering of noise present in the scan image and resizing to a fixed size $S_x \times S_y$. The filtering is performed using a median filter in order to

remove the salt and pepper noise during the image acquisition. The pre-processed scan image be $S_2(x, y)$.

(b). Selection of unstable regions

Initially the image $S_2(x, y)$ is applied to MSER [21] algorithm. Let $P_1, P_2, \ldots . P_T$ represents T number of regions estimated by MSER algorithm, generally represented as P_t, $t = 1,2 \ldots . T$. The cumulative stable regions detected by the MSER algorithm is

$$P = P_1 \cup P_2 \cup P_3, \ldots . \cup P_T = \cup_{t=1}^{T} P_T \quad (1)$$

The unstable region represents the regions that are not covered by the MSER algorithm which is expressed as

$$P' = S_2 - P \quad (2)$$

Where $(-)$ and $(\cup)$ in equation (2) and (1) represents the set difference and union operator respectively. The region selection algorithm will select R number of regions fromthe region P_t.From T number of stable regions, R numbers of least stable regions are estimated. Let the least stable regions be represented as $\hat{P}_1$, $\hat{P}_2 \ldots\ldots \hat{P}_R$ generally expressed as $\hat{P}_r$. The MSER regions are estimated based on correlation and area of the regions. Let $\Delta_t = \{\Delta_1, \Delta_2, \ldots . \Delta_T\}$ represents the area of the stable regions. From the area Δ_t, the normalized area can be estimated as,

$$\hat{\Delta}_t = \frac{\Delta_t}{S_x \times S_y} \quad (3)$$

$$\hat{\Delta}_t = \{\hat{\Delta}_1, \hat{\Delta}_2, \ldots . \hat{\Delta}_T\} \quad (4)$$

Let ρ_t be the correlation coefficient estimated between the region P_t and residual region P'

$$\rho_t = \{\rho_1, \rho_2, \ldots . \rho_T\} \quad (5)$$

A region P_t is selected as an essential region if the product of the normalized area of $P_t(\hat{\Delta}_t)$ and the cumulative correlation coefficient of other regions $(\hat{\tau}_t)$ is less than the product of the cumulative normalized area of the remaining regions of $P_t(\delta_t)$ and correlation coefficient of the region P_t (ρ_t)

$$\hat{\Delta}_t \hat{\tau}_t < \delta_t \rho_t \quad (6)$$

$$\delta_t = \sum_{k \neq t} \hat{\Delta}_t \quad (7)$$

$\hat{\tau}_t$ is the correlation between the residual region P' and $\cup_{k \neq t} P_t$. The least stable regions are the regions that satisfy the criteria provided in equation (6). Therefore the cumulative least stable region is expressed as

$$H = \cup_{r=1}^{R} \hat{P}_r \quad (8)$$

The least stable region H and residual region P' are merged to obtain the unstable regions

$$S_3(x, y) = H \cup P' \quad (9)$$

Let $S_3(x, y)$ be the merged regions from which the gradient is estimated. Let $S_4(x, y)$ be the gradient image from which feature points and TVT features are estimated.

(c) Extraction of features

SURF [22] feature points are then detected on the gradient image. Let $(\hat{m}_l, \hat{n}_l)$ represents the SURF feature points

$$\hat{m}_l, \hat{n}_l = \{(\hat{m}_1, \hat{n}_1), (\hat{m}_2, \hat{n}_2), \ldots . (\hat{m}_L, \hat{n}_L)\} \quad (10)$$

From the 3×3neighborhood of the SURF feature point, the TVT features [23] are extracted. Let the 3×3 neighborhood of the SURF feature point $(\hat{m}_l, \hat{n}_l)$ be represented as

$$G(\hat{m}_l, \hat{n}_l) = \begin{bmatrix} g(\hat{m}_t - 1, \hat{n}_t - 1) & g(\hat{m}_t - 1, \hat{n}_t) & g(\hat{m}_t - 1, \hat{n}_t + 1) \\ g(\hat{m}_t, \hat{n}_t - 1) & g(\hat{m}_t, \hat{n}_t) & g(\hat{m}_t, \hat{n}_t + 1) \\ g(\hat{m}_t + 1, \hat{n}_t - 1) & g(\hat{m}_t + 1, \hat{n}_t) & g(\hat{m}_t + 1, \hat{n}_t + 1) \end{bmatrix} \quad (11)$$

In the 3×3 neighborhood the bottom left, bottom right, top right, top left, and center pixels are represented by G_5, G_4, G_3, G_2 and G_1 respectively, which is expressed as

$$G_5 = g(\hat{m}_l + 1, \hat{n}_l + 1) \quad (12)$$

$$G_4 = g(\hat{m}_l + 1, \hat{n}_l - 1) \quad (13)$$

$$G_3 = g(\hat{m}_l - 1, \hat{n}_l + 1) \quad (14)$$

$$G_2 = g(\hat{m}_l - 1, \hat{n}_l - 1) \quad (15)$$

$$G_1 = g(\hat{m}_l, \hat{n}_l) \quad (16)$$

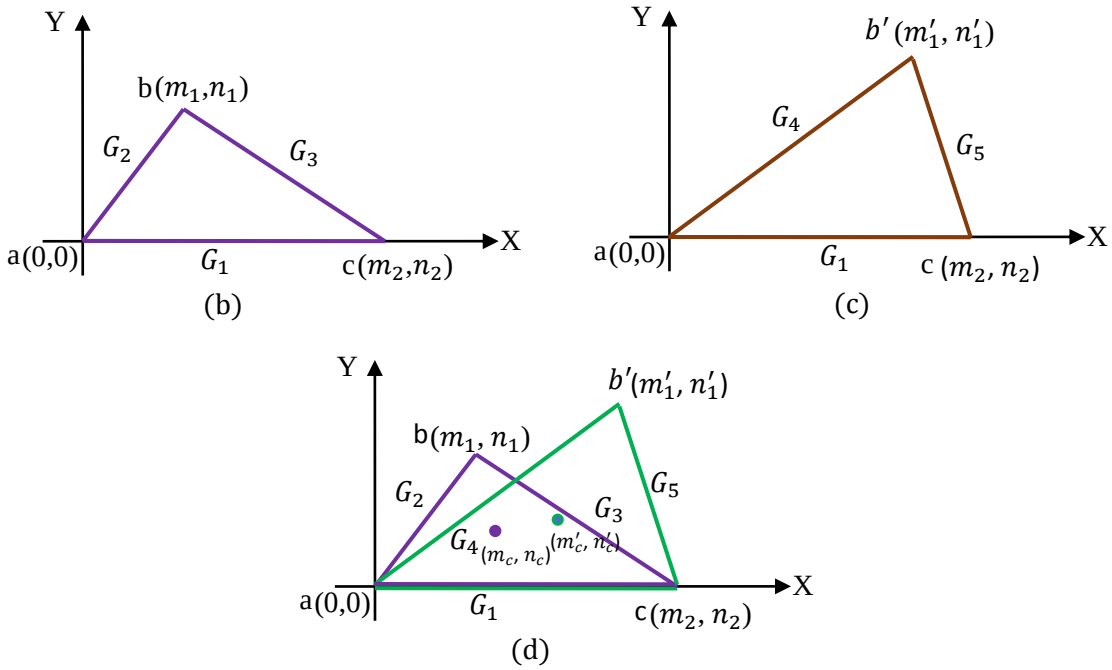


Fig 2: Extraction of feature using TVT transform (a) Construction of first TVT triangle (b) Construction of second TVT triangle (c) Extraction of TVT feature from two triangles

Using the gradient pixels G_5, G_4, G_3, G_2 and G_1 two triangular sections are constructed.G_1, G_2, and G_3 are used to construct the first triangular section, and G_4, G_5, and G_6 are used to construct the second triangular section as depicted in Fig. 2. The vertices of the first TVT transform $a(0,0)$, $b(m_1,m_1)$, $c(m_2,m_2)$ can be estimated from the relations

$$m_1 = \left|\frac{G_2^2-G_3^2+G_1^2}{2G_1}\right| \quad (17)$$

$$n_1 = \left|\sqrt{\frac{G_2^2-G_2^2+G_1^2-G_3^2}{2G_1}}\right| \quad (18)$$

$$m_2 = G_1, n_2 = 0 \quad (19)$$

Similarly, the vertices of the second TVT section $a(0,0)$, $b'(m_1',n_1')$, $c(m_2,m_2)$ can be estimated as,

$$m_1' = \left|\frac{G_4^2-G_5^2+G_1^2}{2G_1}\right| \quad (20)$$

$$n_1' = \left|\sqrt{\frac{G_4^2-G_4^2+G_5^2+G_1^2}{2G_1}}\right| \quad (21)$$

From the two triangular sections, the centroids are estimated. The centroid of the first triangular sections abc is given by

$$(m_c,n_c) = \left(\frac{m_1+G_1}{3},\frac{n_1}{3}\right) \quad (22)$$

The centroid of the second triangular section $ab'c$ is given by

$$(m_c',n_c') = \left(\frac{m_1'+G_1}{3},\frac{n_1'}{3}\right) \quad (23)$$

The distance between the centroids of two triangular sections is given by

$$\omega = \sqrt{(m_c-m_c')^2+(n_c-n_c')^2} \quad (24)$$

Let $\omega_{k,l}$ represent the feature extracted from the image k having L number of feature $l = [1,2,\dots\dots L]$. Therefore the extracted features from K training images are

$$\hat{F}_{train} = [\omega_{k,l}] = [\omega_{1,l},\omega_{2,l},\omega_{3,l}\dots\dots\omega_{K,l}],\ wherel = [1,2,\dots..L] \quad (25)$$

(d) Faster RCNN classifier with dual sigmoid activation function

The region proposal is generated by a convolutional network named as regional proposal network (RPN). The RPN and fast R-CNN are combined to form a faster R-CNN network [20, 24] as depicted in Fig.3. Dual sigmoid-based activation is used instead of using a traditional sigmoid activation function. Let z be the input to dual sigmoid activation where the output is expressed as

$$D(z) = \frac{1}{1+e^{-z}} \quad (26)$$

The output of the dual sigmoid activation function is given by

$$D(z) = D_2(z)+D_1(z) \quad (27)$$

$$D(z) = \frac{1}{1+e^{-\mu(z-\alpha)}}+\frac{1}{1+e^{-\mu(z+\alpha)}}-1 \quad (28)$$

Where $D(z)$ lies between -1 and 1 as depicted in Fig. 4. It also depicts the dual sigmoid activation function with $\alpha = 100$ and different values of μ. For higher values of μ, the activation function provides an S-shaped curve function, while it performs the linear function for the lower value of μ.

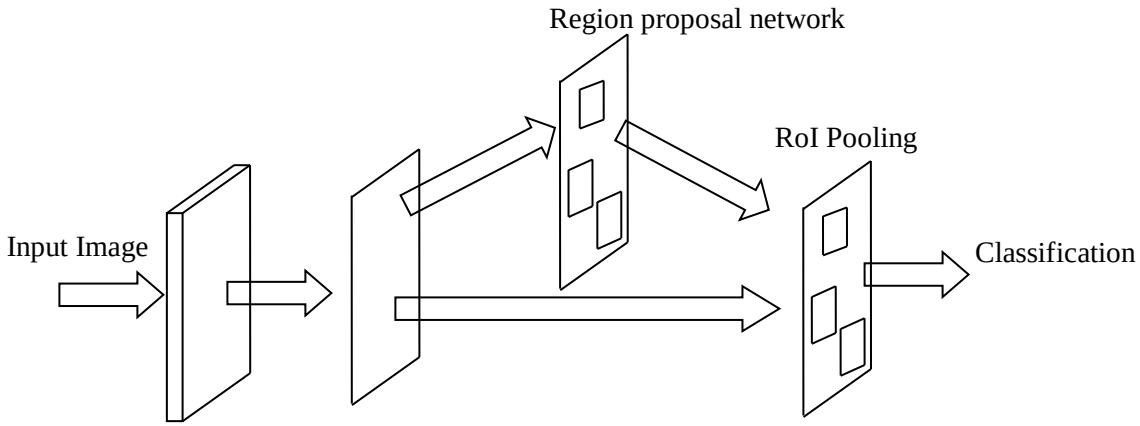


Fig 3: Representation of Faster RCNN network

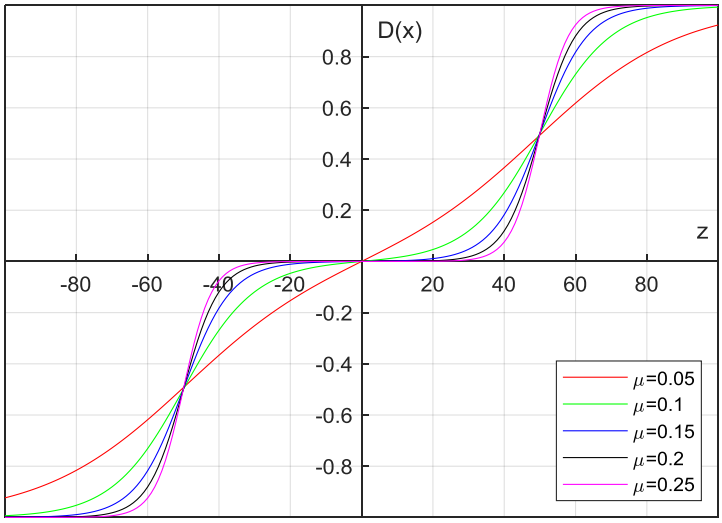


Fig 4: Dual sigmoid activation function

The algorithm for the proposed Doppler scan image classification can be summarized as follows:

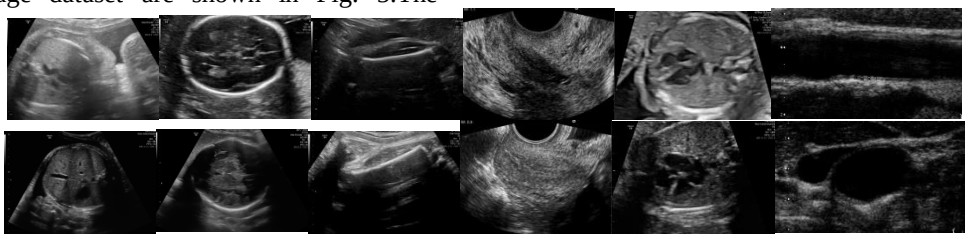
Algorithm:
Input: Test image
Output: Classified Doppler scan image result
Step 1: Pre-process the ultrasound Doppler scan image.
Step 2: Apply the MSER algorithm to obtain a maximally stable extremal region.
Step 3: Apply the region selection process to estimate the regions that are required for feature extraction.
Step 4: Estimate the residual region.
Step 5: Merge the regions obtained in steps 3 and 4, and find the gradient magnitude.
Step 6: Apply the SURF feature point detection algorithm to estimate the location of efficient features.
Step 7: Construct a 3×3 neighborhood of each feature point and estimate the TVT feature from the neighborhood.
Step 8: Classify the extracted features on the trained faster RCNN network to obtain the classified result.

3.Experimental results

The evaluation of the proposed algorithm was performed using the metrics namely Sensitivity, F1-score, Precision, Specificity, and accuracy with the Ultrasound Doppler scan image big data dataset [25]. The dataset consists of 6 different classes of Doppler scan images namely fetal abdomen, fetal brain, fetal femur, maternal

 https://doi.org/10.22362/ijcert/2022/v9/i10/v9i1001

cervix, thorax, and other regions. Thus the dataset totally has 12,400 images collected from 1792 patients. The number of images used in training and testing is7129 and 5271 respectively. Few of the sample images from the doppler scan image dataset are shown in Fig. 5.The evaluation metrics are estimated using the following relations



(a) (b) (c) (d) (e) (f)

Fig 5: Sample images from the dataset (a)Abdomen (b) Brain (c) Femur (d) maternal cervix (e) thorax (f) other regions

$$Sensitivity = \frac{T_{pos}}{(T_{pos}+F_{neg})} \quad (29)$$

$$F1\ Score = \frac{2\times(precision\times sensitivity)}{precision+sensitivity} \quad (30)$$

$$Precision = \frac{T_{pos}}{(T_{pos}+F_{pos})} \quad (31)$$

$$Specificity = \frac{T_{neg}}{(T_{neg}+F_{pos})} \quad (32)$$

$$Accuracy = \frac{T_{pos}+T_{neg}}{(T_{pos}+T_{neg}+F_{pos}+F_{neg})} \quad (33)$$

Table 1: Comparison of performance metrics for the proposed method with traditional schemes

Schemes	Sensitivity (%)	F1-Score (%)	Precision (%)	Specificity (%)	Accuracy (%)
MobileNet	90.45	90.87	89.07	92.01	87.5
VGG	89.16	90.16	88.64	92.48	92.1
SE-ResNeXt-101	93.01	91.89	92.23	96.11	92.7
ResNet-152	90.32	93.71	91.99	95.05	92.8
Inception-v3	91.11	93.2	90.47	93.69	93.5
		8			
DenseNet-169	93.66	92.45	93.01	97.56	93.6
ResneXt-101	92.89	91.73	90.56	94.21	94
EMD-FRCNN	94.73	93.96	93.32	98.94	94.78
MSER-DT	95.86	94.12	93.57	98.73	97.96
Proposed	96.13	94.74	94.26	98.82	98.27

For the experimental comparison of the proposed doppler scan classification, we have used the traditional classifiers such as MobileNet [26], VGG [27], SE-ResNeXt-101 [28], ResNet-152 [29], Inception-152 [30], DenseNet-169 [31], ResneXt-101 [32]. We have also used our recent works namely EMD-FRCNN [19] and MSER-DT for comparison. The proposed Doppler ultrasound classification approach provides a Sensitivity, F1-score, Precision, Specificity, and accuracy of 96.13%, 94.74%, 94.26%, 98.82%, and 98.27% respectively as shown in Table 1. The proposed approach provides 0.27%, 0.62%, 0.94%, 0.09%, 0.31% higher sensitivity, F1-score, Precision, Specificity, and accuracy than the MSER-DT approach. The graphical comparison is illustrated in Fig. 6.

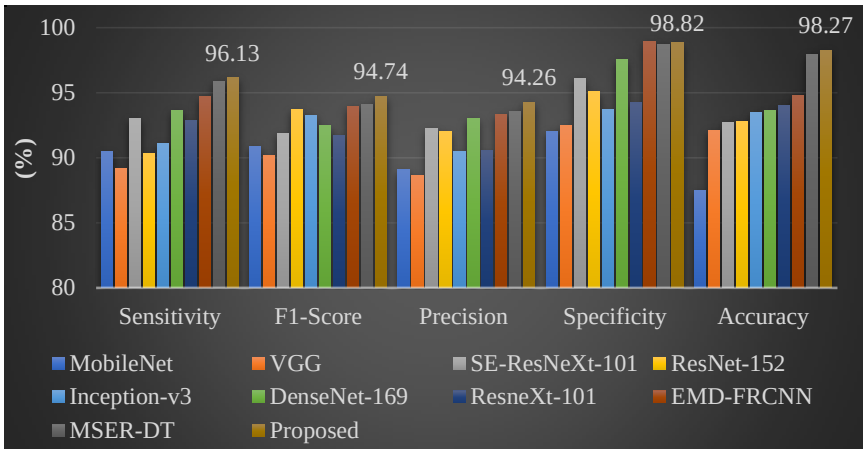


Fig 6: Graphical comparison of metrics Sensitivity, F1-score, Precision, Specificity, and accuracy for the proposed approach with other recent classifier schemes.

Table 2 shows the performance comparison of the proposed approach with and without using the region selection process. In the 'without region selection' evaluation process the feature is extracted from the entire image without detecting any regions. In the 'with region selection' evaluation process the feature is extracted from the region selected by the region selection process. The time of training the train images is 6,137s and the time of classification for one image is 1.06s. Using the region selection process the accuracy improves by 2.89% in the proposed approach.

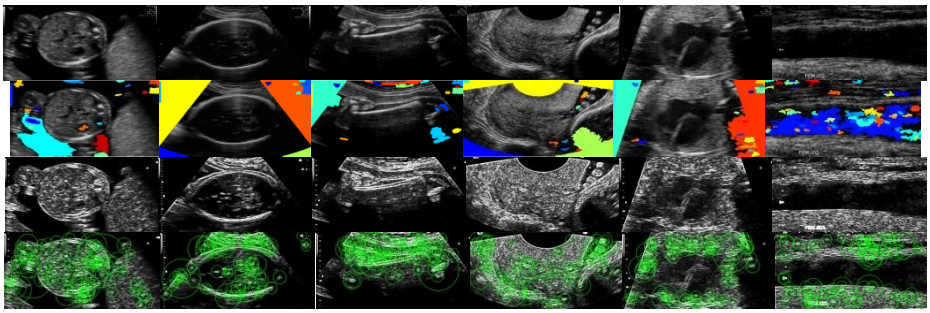
Table 2: Comparison of classification performance with and without region selection

Scheme	Method	Accuracy (%)	Time of training (s)	Time of classification (s)
MSER-DT	Without region selection	94.21	***5,821***	***0.97***
	With region selection	***97.96***	6,453	1.13
Proposed	Without region selection	95.38	5,426	0.95
	With region selection	98.27	6,137	1.06

Fig. 7 shows the confusion matrix obtained on the test images during the classification process.

Fig 7: Confusion matrix chart obtained during testing

The experimental results obtained during region selection, gradient estimation, and SURF feature point detection process is illustrated in Fig. 8. Here row1 represents the input Doppler scan image, row2 represents the regions detected by the MSER algorithm, row3 represents the Gradient estimated on the regions estimated by the region selection process and row4 represent the feature points detected by the SURF feature point detection algorithm.



(a) (b) (c) (d) (e) (f)

Fig 8: Experimental result of the proposed approach (a) Abdomen (b) Brain (c) Femur (d) maternal cervix (e) thorax (f) other regions

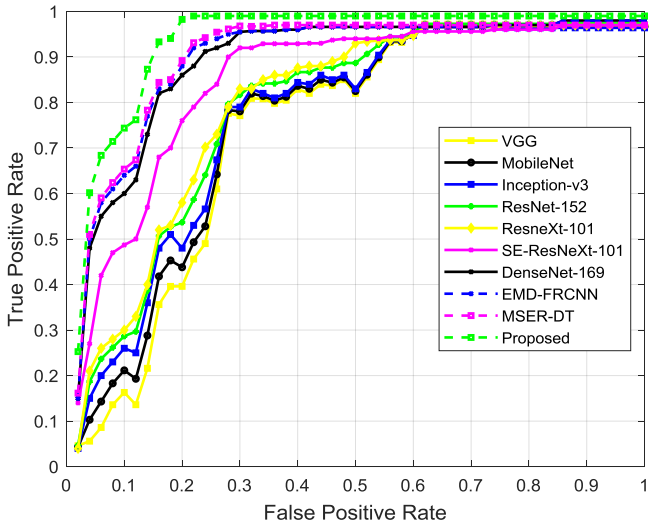


Fig 9: Comparison of ROC curve for different schemes

Fig.9 shows the comparison of the ROC curve for different classifiers. The proposed algorithm shows a higher AUC of 0.9392 which is higher than the traditional approaches.

4. Conclusion

An algorithm for ultrasound doppler scan image classification system is proposed in this paper. Initially, the algorithm decomposes the image into a maximally stable extremal region using the MSER algorithm. The residual regions along with the selected least stable regions are used to estimate the essential regions for estimating the features. The gradient image is then estimated on the essential regions. From the gradient image, the SURF feature points are detected and the feature value is estimated by the TVT transform around the neighborhood of the feature point. Finally, the faster RCNN network with the dual sigmoid activation function is used to classify the doppler scan image as Femur, cervix, abdomen, thorax, brain, and other regions. The evaluation was done using the metrics such as precision, recall, sensitivity, specificity, and accuracy. The proposed Doppler ultrasound classification approach provides a sensitivity, F1-score, precision, specificity, and accuracy of 96.13%, 94.74%, 94.26%, 98.82%, and 98.27% respectively which is higher than other classifiers used in Doppler scan image classification.

References

[1]. Pellicer, B., et al. "Ultrasound bioeffects in rats: quantification of cellular damage in the fetal liver after pulsed Doppler imaging." *Ultrasound in obstetrics & gynecology* 37.6 (2011): 643-648.

[2]. Dhas, Edwin, and M. Suchetha. "Extraction of Fetal ECG From Abdominal and Thorax ECG Using a Non-Causal Adaptive Filter Architecture." IEEE Transactions on Biomedical Circuits and Systems (2022).

[3]. Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. Nat. 542, 115 (2017).

[4]. Sobhaninia, Zahra, et al. "Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning." 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, 2019.

[5]. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. &Aerts, H. J. Artificial intelligence in radiology. Nat. Rev. Cancer 18, 500 (2018).

[6]. Litjens, G. et al. A survey on deep learning in medical image analysis. Med. image analysis 42, 60–88 (2017).

[7]. Das, Sraddha, et al. "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy." Biomedical Signal Processing and Control 68 (2021): 102600.

[8]. Guo, Z., Zhang, L. and Zhang, D., 2010. A completed modeling of local binary pattern operator for texture classification. IEEE transactions on image processing, 19(6), pp.1657-1663.

[9]. Jabid, Taskeed, Md Hasanul Kabir, and Oksam Chae. "Local directional pattern (LDP)–A robust image descriptor for object recognition." 2010 7th IEEE international conference on advanced video and signal based surveillance. IEEE, 2010.
[10]. Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). Vol. 1. Ieee, 2005.
[11]. Ryou, H. et al. Automated 3d ultrasound biometry planes extraction for first trimester fetal assessment. In Machine Learning in Medical Imaging, 196–204 (2016).
[12] Li, Y. et al. Standard plane detection in 3d fetal ultrasound using an iterative transformation network. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, 392–400 (2018).
[13]. Maraci, M., Bridge, C., Napolitano, R., Papageorghiou, A. & Noble, J. A framework for analysis of linear ultrasound videos to detect fetal presentation and heartbeat. Med. Image Analysis 37, 22–36 (2017).
[14] Sushma, T. V., et al. "Classification of Fetal Heart Ultrasound Images for the Detection of CHD." Innovative Data Communication Technologies and Application. Springer, Singapore, 2021. 489-505.
[15] Qiao, Sibo, et al. "RLDS: An explainable residual learning diagnosis system for fetal congenital heart disease." Future Generation Computer Systems 128 (2022): 205-218.
[16] Xie, H. N., et al. "Using deep-learning algorithms to classify fetal brain ultrasound images as normal or abnormal." Ultrasound in Obstetrics &Gynecology 56.4 (2020): 579-587.
[17]. Lee, Lok Hin, Yuan Gao, and J. Alison Noble. "Principled Ultrasound Data Augmentation for Classification of Standard Planes." International Conference on Information Processing in Medical Imaging. Springer, Cham, 2021.
[18]. Attallah, Omneya, Maha A. Sharkas, and Heba Gadelkarim. "Fetal brain abnormality classification from MRI images of different gestational age." Brain sciences 9.9 (2019): 231.
[19]. Kumari, S. Sandhya, and K. Sandhya Rani. "Empirical mode Decomposition and Dual Sigmoid Activation Function-Based Faster RCNN for Big Data Doppler Scan Image Classification." measurement 8.9 (2021): 151-165.
[20]. Suchetha, M., et al. "Region of interest-based predictive algorithm for subretinal hemorrhage detection using faster R-CNN." Soft Computing 25.24 (2021): 15255-15268.
[21]. Donoser, Michael, and Horst Bischof. "Efficient maximally stable extremal region (MSER) tracking." *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*. Vol. 1. Ieee, 2006.
[22]. Bay, Herbert, et al. "Speeded-up robust features (SURF)." Computer vision and image understanding 110.3 (2008): 346-359.
[23]. Prabha, K., and I. Shatheesh Sam. "A novel blind color image watermarking based on Walsh Hadamard Transform." *Multimedia Tools and Applications* 79.9 (2020): 6845-6869.
[24]. Jiang, Huaizu, and Erik Learned-Miller. "Face detection with the faster R-CNN." *2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017)*. IEEE, 2017.
[25]. Burgos-Artizzu, Xavier P., et al. "Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes." Scientific Reports 10.1 (2020): 1-12.
[26]. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. & Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. CVPR(2018).
[27]. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556 (2014).
[28]. Hu, J., Shen, L. & Sun, G. Squeeze-and-excitation networks. CoRR abs/1709.01507 (2017).
[29]. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. CoRR abs/1512.03385 (2015).
[30]. Szegedy, C. et al. Going deeper with convolutions. CoRR abs/1409.4842 (2014).
[31]. Burgos-Artizzu, Xavier P., et al. "Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes." Scientific Reports 10.1 (2020): 1-12.
[32]. Xie, S., Girshick, R., Dollar, P., Tu, Z. & He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1492–1500 (2017).

