



HAL
open science

Combining Bayesian and Deep Learning Methods for the Delineation of the Fan in Ultrasound Images

Hind Dadoun, Hervé Delingette, Anne-Laure Rousseau, Eric de Kerviler,
Nicholas Ayache

► **To cite this version:**

Hind Dadoun, Hervé Delingette, Anne-Laure Rousseau, Eric de Kerviler, Nicholas Ayache. Combining Bayesian and Deep Learning Methods for the Delineation of the Fan in Ultrasound Images. ISBI 2021 - 18th IEEE International Symposium on Biomedical Imaging, Apr 2021, Nice, France. 10.1109/ISBI48211.2021.9434112 . hal-03127809

HAL Id: hal-03127809

<https://inria.hal.science/hal-03127809>

Submitted on 2 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COMBINING BAYESIAN AND DEEP LEARNING METHODS FOR THE DELINEATION OF THE FAN IN ULTRASOUND IMAGES

Hind Dadoun* Hervé Delingette* Anne-Laure Rousseau† Eric de Kerviler‡ Nicholas Ayache*

* Université Côte d’Azur, Inria Epione Team, Sophia Antipolis, France

† Hôpital Européen Georges Pompidou, NHance, Nabla, Paris, France

‡ Radiology, Saint Louis Hospital, AP-HP, Paris, France

ABSTRACT

Ultrasound (US) images usually contain identifying information outside the ultrasound fan area and manual annotations placed by the sonographers during exams. For those images to be exploitable in a Deep Learning framework, one needs to first delineate the border of the fan which delimits the ultrasound fan area and then remove other annotations inside. We propose a parametric probabilistic approach for the first task. We make use of this method to generate a training data set with segmentation masks of the region of interest (ROI) and train a U-Net to perform the same task in a supervised way, thus considerably reducing computational time of the method, one hundred and sixty times faster. These images are then processed with existing inpainting methods to remove annotations present inside the fan area. To the best of our knowledge, this is the first parametric approach to quickly detect the fan in an ultrasound image without any other information than the image itself.

Index Terms— Ultrasound imaging, Deep Learning, Ultrasound fan area detection, pre-processing

1. INTRODUCTION

Ultrasound (US) imaging is one of the most common techniques for medical diagnosis. According to the WHO, two thirds of the world’s population do not have access to medical imaging, and ultrasound associated with X-ray could cover 90 % of these needs. The decrease in ultrasound hardware prices allows its diffusion, but limitations persist because acquiring and interpreting an ultrasound image is a difficult examiner-dependent task with few trained operators [1, 2, 3]. Hence the importance of developing the research around the entire processing chain in US imaging [4, 5, 6]. Interested readers may refer to [7] for an overview of Deep Learning strategies for ultrasound-specific processing methods.

Ultrasound imaging provides real-time anatomical information and allows the measurement of functional parameters. These measurements along with other annotations are manually added to the image during an exam to provide a complete report. In order to make them available to the research



Fig. 1: Left picture shows the original image. One can see that the ultrasound fan area is limited to a conic section, and that a number of text and graphic elements are present. Right picture shows the result of our pre-processing. The white lines delimiting the ultrasound fan area are automatically detected and all graphic and text elements are removed and replaced by a plausible intensity value.

community following regulatory guidelines, they are usually converted to JPEG/PNG format, and processed to remove all metadata, including acquisition parameters. The presence of biometric measurements and other machine dependent characteristics may be challenging for the task of automated image analysis. First, because there is no guarantee that these annotations do not include sensitive information. Second, because we would like to make sure that they do not induce a bias during the training of a neural network for the task of classification, segmentation or detection.

In [8], the authors “blacked out” the identifying patient information on videos by setting the corresponding intensities of pixels that remained static throughout the entire clip to minimal intensity. Unfortunately, this method only works if we have access to the video sequence of the exam. As for the biometric measurements, [9] removed all the annotations using the inpainting algorithm proposed in [10], but no information was given on how to create the inpainting masks. In this work we propose a pre-processing pipeline for ultrasound images to detect the ultrasound fan area combining Bayesian and Deep Learning methods and show how to apply existing inpainting methods to remove biometric measurements as shown in Fig 1.

2. DATASET

Data for this pilot study were obtained in collaboration with the Clinical Data Warehouse of Greater Paris University Hospitals. Data were collected prospectively in Saint Louis Hospital and we worked closely with volunteer physicians from the NHANCE NGO to construct an heterogeneous database that includes images from various manufacturers. The database was composed of 1280 images from which 1150 were used to train and validate the method. The 130 remaining images were used to evaluate the method by a trained engineer.

3. DETECTION OF THE ULTRASOUND FAN AREA

We present a fully parametric method to detect the ultrasound fan area in the image and thus remove most of the annotations present outside. This is achieved by generating segmentation masks of the US region using a probabilistic approach.

3.1. Probabilistic model of US fan area

Let I be our image of size $n \cdot m$ where n is the width and m the height of the image. We define $\theta = \{\epsilon_1, \dots, \epsilon_{10}\}$ as the set of parameters describing the truncated cone Ω_θ modeling the US fan area (see Fig. 2). We seek to optimize θ for every input US image via a probabilistic formulation. More precisely, we introduce the hidden binary variable $Z_i \in [0, 1]$, indicating whether voxel i belongs to the fan area Ω_θ . The θ parameters are equipped with a uniform prior and we propose to maximize the marginal log likelihood $p(\theta|I)$ as follows:

$$\begin{aligned} \log(p(\theta|I)) &\propto \log(p(\theta)) + \log(p(I|\theta)) \\ \log(p(I|\theta)) &= \sum_{i=0}^{n \cdot m} \log(p(I_i|Z_i = 1) \cdot p(Z_i = 1|\theta) \\ &\quad + (p(I_i|Z_i = 0)) \cdot (1 - p(Z_i = 1|\theta))) \end{aligned}$$

To model the distributions, we make use of intensity values extracted from bounding boxes of all training images. The distribution of a pixel intensity in the US fan $f_i = P(I_i|Z_i = 1)$ is captured by a mixture of two Gaussians (see Fig 3.b) whose parameters are estimated using bounding boxes of size 5×5 located in the center of images. As for the background distribution $b_i = P(I_i|Z_i = 0)$ it is modeled as a uniform distribution whose parameters are estimated using bounding boxes of size 2×2 located in the top left, top right, bottom left and bottom right corners of the image.

$P(Z_i = 1|\theta)$ is the probability of voxel i to be inside the fan area Ω_θ which is defined in a closed form manner. Indeed, the truncated cone region Ω_θ is defined analytically as the set of points $\mathbf{x} = (x, y)$ for which $f_i(\mathbf{x}, \theta) \geq 0$

where f_1, f_2 are respectively the equations of the two straight lines AD, CF and f_3, f_4 are respectively the equations of the two parabolas through ABC, DEF as shown in

Figure 2. From this piecewise analytic description, we derive a soft implicit definition of the fan shape by summing up the four implicit functions $f_i(\mathbf{x}, \theta)$. The prior label probability is defined as Bernoulli distribution whose parameter depends on the sigmoid of the regularized implicit function :

$$p(Z_i = 1|\theta) = \sigma\left[\sum_{i=1}^4 f_i(\mathbf{x}_i, \theta)\right]$$

where \mathbf{x}_i is the position of voxel i in the image.-

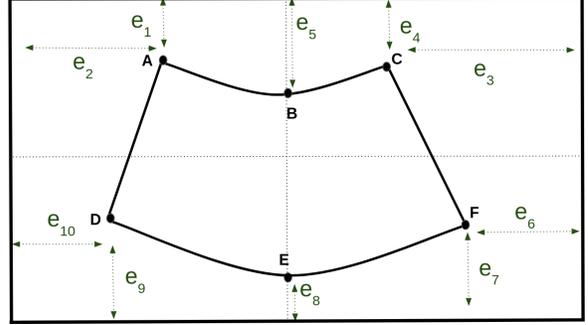


Fig. 2: Parameterisation of the region of interest

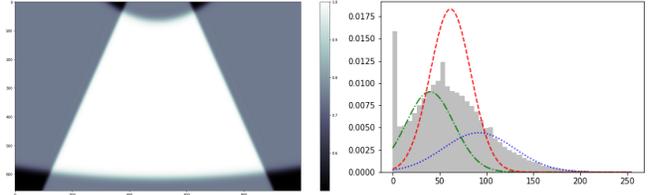


Fig. 3: (Left) Prior label probability $p(Z_i = 1|\theta)$ parameterized by θ ; (Right) Normalized histograms of the ROI distribution. In green/blue lines the two Gaussians and in red the mixture of the Gaussians.

3.2. Expectation-Maximization (E-M) algorithm

Maximizing the log joint probability is not easy since we have the log of sums. Instead, we use a lower bound which is much simpler to compute. More precisely, we replace this maximization :

$$\log p(I, \theta) = \log(p(\theta) + \log(p(I|\theta)))$$

With this one :

$$\log p(I, \theta) - D_{\text{KL}}(U||p(Z|I))$$

Where $U = \{U_n\}$ is a surrogate function for the posterior label probability $p(Z|I)$ and D_{KL} is the Kullback-Leibler divergence. This can easily be solved using the E-M algorithm.

- **E-Step.** Compute the posterior label probability for the current value of θ :

$$\begin{aligned} U_i &= p(Z_i = 1|I_i) \\ &= \frac{r_i p(Z_i = 1|\theta)}{r_i p(Z_i = 1|\theta) + (1 - r_i)(1 - p(Z_i = 1|\theta))} \end{aligned}$$

where $r_i = p(I_i|Z_i = 1)/p(I_i|Z_i = 1) + p(I_i|Z_i = 0)$.

- **M-Step.** Find θ by maximizing the variational lower bound which is equivalent to minimizing the following expression with the current value of U :

$$\begin{aligned} \mathcal{L}(\theta) &= -D_{\text{KL}}(U||p(Z|I)) + \log(p(\theta)) \\ &= \sum_{i=0}^{n \cdot m} -U_i \log(p(Z_i = 1|\theta)) \\ &\quad + (1 - U_i) \log(1 - p(Z_i = 1|\theta)) \\ &\quad + \log(p(\theta)) \end{aligned}$$

During the M-step we use an optimization algorithm, Limited-memory BFGS algorithm (L-BFGS). This algorithm approximates the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) using a limited amount of computer memory [11, 12].

3.3. Reducing the computational time of the method using Deep Learning

We use the segmentation masks generated by the method to train a neural network for the same task. This is done by training on CPU a simple U-Net on 70 % of our dataset. We validate the method on 20 % of our dataset and test it on the remaining 10 % . We use the Binary Cross Entropy (BCE) with logits loss. In inference, the processing of one frame takes approximately 0.45 seconds which is 160 times faster than the Bayesian method.

4. IMPLEMENTATION AND RESULTS

In this section we explicit the implementation details of our method and show the quantitative and qualitative results.

4.1. Implementation of the E-M algorithm

We only optimize the Kullback-Leibler term during this step since we use a uniform prior as $p(\theta)$. Indeed, we do not have access to a preferred range of values for θ . We show in Figure 4 the optimization curve of the log-likelihood during the E-M algorithm when the prior on θ is far from the ground truth. We can see that the algorithm is robust to the change of shape in the ultrasound fan area.

4.2. Training details of the U-Net

We use Adam Optimizer [13] with learning rate 0.001 and a batch size of 1. After 10 epochs, we achieve a validation loss of 0.062 and Dice loss of 0.017.

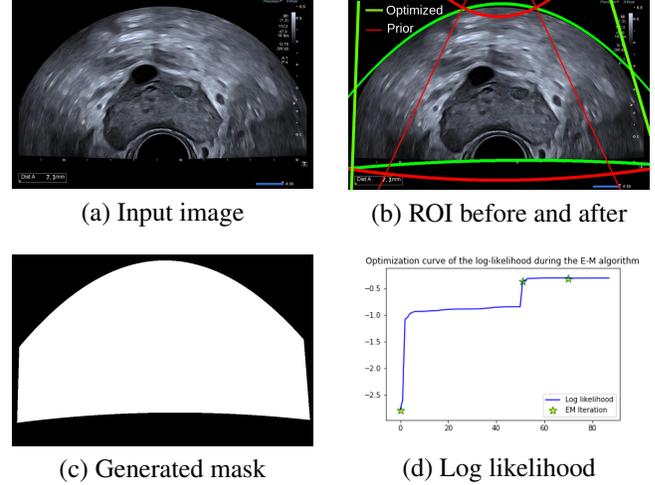


Fig. 4: Log-likelihood optimization during the EM algorithm.

4.3. Bayesian method compared to the U-Net

The Bayesian method developed in Section 3 is constrained by our piecewise analytic description. It provides a good approximation of the ground truth mask, but results in a rigid delineation of the ultrasound fan area. Whereas the masks generated by the U-Net are less regularized and can therefore capture more information on the ultrasound fan area. An example of the two masks is shown in Figure 5.

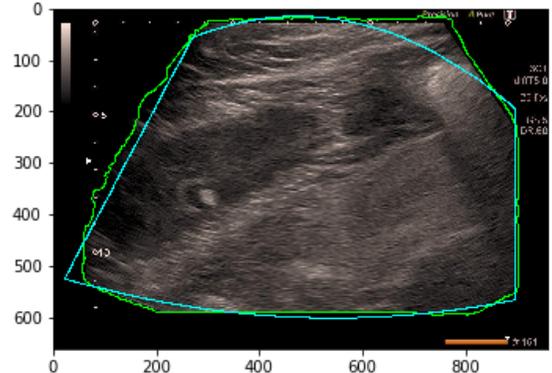


Fig. 5: (Blue) Mask generated by the Bayesian method. (Green) Mask generated by the U-Net.

4.4. Evaluation of the method

For the remaining 10 % of the dataset, which we had never used before, we asked a trained engineer to evaluate the results of the method using 3 labels:

- Perfect match between the ultrasound fan area and the detected area.
- Good detection when the area of the missing part is less than 1% of the image.

Table 1: Evaluation of the detection method on 130 images

Label	# Images	Mean mismatch area
Perfect detection	90 (69.2 %)	0.00%
Good detection	37 (28.5 %)	0.15%
Total	127 (97.7 %)	0.05%
Poor detection	3 (2.3 %)	5.0%

- Poor detection when the area of the missing part is more than 1% of the image.

We see in Table 1 that 90 images were labeled as a perfect match, 37 images were labeled as good detection, with mean area of mismatch $<0.15\%$. This part corresponds to the corners of the fan that were slightly cropped due to the parametric definition of the fan area. Yet those tiny errors on the fan margin have no impact on the interpretation of the image content. Finally 3 images were labeled as missing a relatively large part of the detected ultrasound fan area. This happens when a part of the ultrasound fan area is totally dark, the method then mixes up the background with the foreground. Examples are shown in Fig 6.

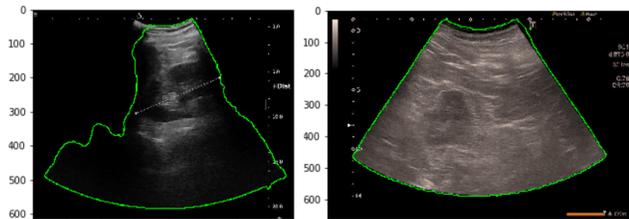


Fig. 6: (Left) Example of the label 'Poor detection', a part of the fan is not detected because it is filled with low intensity pixels. (Right) Example of the label 'Good detection', the missing area corresponding to the cropped corners is barely visible to the naked eye.

5. INPAINTING IMAGES WITH ANNOTATIONS INSIDE THE ULTRASOUND FAN AREA

Here we show how to use open CV's module inpaint to replace segments and annotations present on the cone by pixels of the background. More precisely we use in-paint Telea which is based on [10]. Values of pixels of the region to be inpainted are replaced by a weighted sum of neighboring pixels starting from the boundary. The challenge is to generate in a fully unsupervised way a mask of the region to be inpainted. This is done by maximizing the contrast of the image and masking all pixels below a threshold value. We also replace all colored pixels in the image with random shades of gray so that the inpainting algorithm doesn't use colored pixels present in the boundary. Finally we denoise the resulting image using non-local-means filtering [14]. The method uses small patches centered on pixels. The patch of interest is compared to other patches. Then it replaces the value of the

pixel by the average intensity of pixels that have patches close to the current patch. An example is shown in Fig 7.

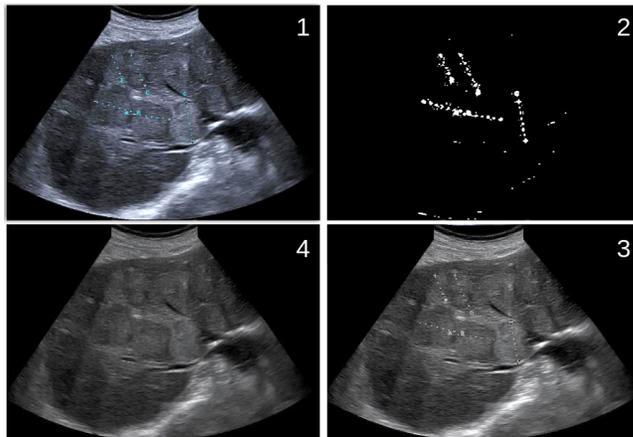


Fig. 7: Pipeline to generate masks for the inpainting algorithm. We maximize the contrast of the input image (1) and mask all pixels below a threshold value (2). We also replace all colored pixels in the image with random shades of gray (3) so that the inpainting algorithm doesn't use colored pixels present in the boundary. Finally we apply the inpainting algorithm and denoise the resulting image using non-local-means filtering(4).

6. CONCLUSION

We achieved our primary objective, namely the construction of an automated pipeline for ultrasound fan area detection. We have shown that this method is scalable by evaluating it on 130 varied images obtained from different machines and various shapes of the ultrasound fan area. The next step is to work closely with the Clinical Data Warehouse of Greater Paris University Hospitals, to assess the method on a larger database. The primary novelty of this method is the use of Bayesian statistics to generate training data for a Deep Learning application. We believe that this work is an important step for a larger and better exploitation of ultrasound images in the area of medical image analysis using Deep Learning. A possible improvement of the method is to replace the uninformative uniform $p(\theta)$ with a distribution estimated on the training set and include it in the M-step of the EM algorithm. The method could be further improved by adding a regularization term in the U-Net loss function. This would allow for more regular approximations of the ultrasound fan area. We intend to release the code publicly.

7. REFERENCES

- [1] Christopher L Moore and Joshua A Copel, "Point-of-care ultrasonography," *New England Journal of Medicine*, vol. 364, no. 8, pp. 749–757, 2011.

- [2] Max J Liebo, Rachel L Israel, Elizabeth O Lillie, Michael R Smith, David S Rubenson, and Eric J Topol, “Is pocket mobile echocardiography the next-generation stethoscope? a cross-sectional comparison of rapidly acquired images with standard transthoracic echocardiography,” *Annals of internal medicine*, vol. 155, no. 1, pp. 33–38, 2011.
- [3] Brian G Choi, Monica Mukherjee, Praveen Dala, Heather A Young, Cynthia M Tracy, Richard J Katz, and Jannet F Lewis, “Interpretation of remotely downloaded pocket-size cardiac ultrasound images on a web-enabled smartphone: validation against workstation evaluation,” *Journal of the American Society of Echocardiography*, vol. 24, no. 12, pp. 1325–1330, 2011.
- [4] Richard Droste, Lior Drukker, Aris T Papageorgiou, and J Alison Noble, “Automatic probe movement guidance for freehand obstetric ultrasound,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 583–592.
- [5] Thomas LA van den Heuvel, Hezekiel Petros, Stefano Santini, Chris L de Korte, and Bram van Ginneken, “Automated fetal head detection and circumference estimation from free-hand ultrasound sweeps using deep learning in resource-limited countries,” *Ultrasound in medicine & biology*, vol. 45, no. 3, pp. 773–785, 2019.
- [6] Moi Hoon Yap, Gerard Pons, Joan Martí, Sergi Ganau, Melcior Sentís, Reyer Zwiggelaar, Adrian K Davison, and Robert Martí, “Automated breast ultrasound lesions detection using convolutional neural networks,” *IEEE journal of biomedical and health informatics*, vol. 22, no. 4, pp. 1218–1226, 2017.
- [7] Ruud JG van Sloun, Regev Cohen, and Yonina C Eldar, “Deep learning in ultrasound imaging,” *Proceedings of the IEEE*, vol. 108, no. 1, pp. 11–29, 2019.
- [8] Jeffrey Zhang, Sravani Gajjala, Pulkit Agrawal, Geoffrey H Tison, Laura A Hallock, Lauren Beussink-Nelson, Eugene Fan, Mandar A Aras, ChaRandle Jordan, Kirsten E Fleischmann, et al., “A computer vision pipeline for automated determination of cardiac structure and function and detection of disease by two-dimensional echocardiography,” *arXiv preprint arXiv:1706.07342*, 2017.
- [9] Christian F Baumgartner, Konstantinos Kamnitsas, Jacqueline Matthew, Tara P Fletcher, Sandra Smith, Lisa M Koch, Bernhard Kainz, and Daniel Rueckert, “Sononet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound,” *IEEE transactions on medical imaging*, vol. 36, no. 11, pp. 2204–2215, 2017.
- [10] Alexandru Telea, “An image inpainting technique based on the fast marching method,” *Journal of graphics tools*, vol. 9, no. 1, pp. 23–34, 2004.
- [11] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu, “A limited memory algorithm for bound constrained optimization,” *SIAM Journal on scientific computing*, vol. 16, no. 5, pp. 1190–1208, 1995.
- [12] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal, “Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization,” *ACM Transactions on Mathematical Software (TOMS)*, vol. 23, no. 4, pp. 550–560, 1997.
- [13] Diederik P Kingma and Jimmy Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [14] Antoni Buades, Bartomeu Coll, and J-M Morel, “A non-local algorithm for image denoising,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*. IEEE, 2005, vol. 2, pp. 60–65.

8. COMPLIANCE WITH ETHICAL STANDARDS

IRB approval (IRB00011591) was obtained for this multi-center study and informed consent was waived. Data for this pilot study were obtained in collaboration with the Clinical Data Warehouse of Greater Paris University Hospitals registered at the National Commission for Data Protection and Liberties (CNIL-France) under the number 1980120.

9. ACKNOWLEDGMENTS

This work has been supported by the French government, through the 3IA Côte d’Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.

The authors are grateful to the OPAL infrastructure from Université Côte d’Azur, the Radiology team from Saint Louis Hospital of Greater Paris University Hospitals and the French Health Data Hub for providing resources and support.

We thank Guillaume Oules, engineer at GoPro, for his help in the design of the NHance project. We thank the Clinical Data Warehouse of Greater Paris University Hospitals and especially the imaging team. The authors declare that they have no conflicts of interest.