
Automatic Shadow Detection in 2D Ultrasound

Qingjie Meng¹, Christian Baumgartner², Matthew Sinclair¹, James Housden³,
Martin Rajchl¹, Alberto Gomez³, Benjamin Hou¹, Nicolas Toussaint³, Jeremy Tan¹,
Jacqueline Matthew³, Daniel Rueckert¹, Julia Schnabel³, Bernhard Kainz¹

¹Imperial College London, ²ETH Zürich, ³Kings College London

q.meng16@imperial.ac.uk

Abstract

Automatically detecting acoustic shadows is of great importance for automatic 2D ultrasound analysis ranging from anatomy segmentation to landmark detection. However, variation in shape and similarity in intensity to other structures in the image make shadow detection a very challenging task. In this paper, we propose an automatic shadow detection method to generate a pixel-wise shadow confidence map from weakly labelled annotations. Our method jointly uses; (1) a feature attribution map from a Wasserstein GAN and (2) an intensity saliency map from a graph cut model. The proposed method accurately highlights the shadow areas in two 2D ultrasound datasets comprising standard view planes as acquired during fetal screening. Moreover, the proposed method outperforms the state-of-the-art quantitatively and improves failure cases for automatic biometric measurement.

1 Introduction

2D ultrasound (US) is a popular medical imaging modality, well known for its portability, low cost, and high temporal resolution. However, acoustic shadows that caused by sound-opaque occluders potentially omit vital anatomical information in 2D US and thus can be a big burden for automatic medical image analysis. Although, these shadow artefacts have been well studied in the clinical literature [1, 2], automatic shadow detection has rarely been focused. [3] and [4] have developed different shadow awareness techniques but either of these techniques are application specific or sensitive to US transducer settings. Recently, deep learning methods have paved the way to fully automatic semantic real-time image understanding. Inspired by [5], we propose a novel shadow detection model to predict a dense, anatomically agnostic shadow-confidence map using only weak annotations. In this model, we jointly use a feature attribution map from a Wasserstein GAN (WGAN) model and an intensity confidence map from a graph cut model. To our knowledge, this is the first shadow detection model for ultrasound images that generates a dense, shadow-focused confidence map. Fig. 1 shows an overview over our method.

2 Method

(1) Shadow image discrimination: We use a fully convolutional neural network (FCN) to classify shadow images ($l = 1$) and clear images ($l = 0$). The classifier we use has a similar architecture to SonoNet-32 [6], which performs well for 2D ultrasound fetal standard view classification and this classifier provides soft predictions $p(x_i|l = 1)$ for image x_i during testing.

(2) Saliency map generation: In the shadow class images, shadows have features such as a typical direction and relatively low intensity. These features are highlighted in saliency map s_{m_i} which is generated by guided back-propagation [7] so that shadow pixel candidates are obtained.

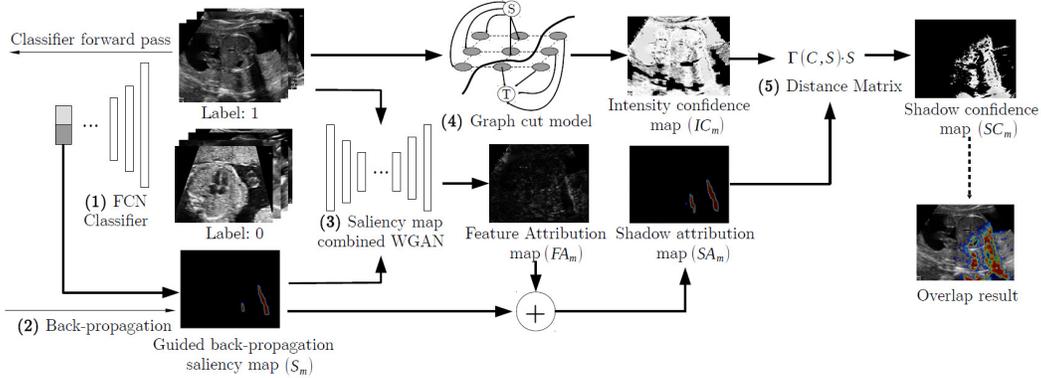


Figure 1: Overview of our anatomy agnostic shadow detection approach.

(3) Shadow attribution map: Inspired by VA-GAN [5], we develop a novel WGAN model in which the generator directly produces a fake clear image of the original shadow image to learn the visual feature attribution FA_m between classes. The generator has a U-Net structure with all convolution layers being replaced by residual-units while the discriminator is a FCN without dense layers. The WGAN model is optimised with a combined L1-loss and L2-loss. $\hat{G} = \arg \min_G \max_D (\mathbf{E}_{\phi(t_i) \sim p(\phi(t_i)|l=0)} [D(x_i)] - \mathbf{E}_{\phi(t_i) \sim p(\phi(t_i)|l=1)} [D(G(\phi(t_i)))] + \lambda_1 L_1 + \lambda_2 L_2)$. L1-loss is defined as $L_1 = \|G(\phi(t_i)) - \phi(t_i)\|_1$ to guarantee small changes of the feature attribution map while L2-loss is defined as $L_2 = \|G(\phi(t_i)_B - \phi(t_i)_B\|_2$ to encourage changes to happen only in potential shadow areas. Here, $G(\phi(t_i))$ is the fake clear image produced by the generator for each tuple $t_i = (x_i | l_i = 1, s_{mi})$. $\phi(t_i) = \psi(x_i | l_i = 1, T(s_{mi}))$ with $T(\cdot)$ being a threshold operation producing a binary mask. The threshold is $s_{mi} \cdot (0.02 \leq s_{mi} \leq 0.98) = 0$ and is decided according to the histogram of s_{mi} . ψ replaces pixels in $x_i (T(s_{mi}) = 1)$ with the mean intensity value of $x_i (T(s_{mi}) = 0)$. Combining the difference between fake and clear images and the saliency map from (1) yields an shadow attribution map $SA_m = |G(\phi(t_i)) - x_i| + s_{mi}$.

(4) Graph cut model: To integrate the intensity feature of the shadows, we build a graph cut model using intensity information as weights to connect each pixel in the image to shadow class and background class. The weights that connect pixels to the shadow class give an intensity saliency map IC_m . For a pixel x_{ij} with intensity I_{ij} in the image x , the score of being a shadow pixel F_{ij} is given by $F_{ij} = -\frac{|I_{ij} - I_S|}{|I_{ij} - I_S| + |I_{ij} - I_B|}$ while the score of being a background pixel B_{ij} is given by $B_{ij} = -\frac{|I_{ij} - I_B|}{|I_{ij} - I_S| + |I_{ij} - I_B|}$. I_S and I_B are shadow mean intensity and background mean intensity respectively. The weight from x_{ij} to source (shadow class) is set as $W_{F_{ij}} = \frac{F_{ij}}{F_{ij} + B_{ij}}$ and the weight from x_{ij} to sink (background) is $W_{B_{ij}} = \frac{B_{ij}}{F_{ij} + B_{ij}}$. We use a 4-connected neighbourhood to set weights between pixels and all the weights between neighbourhood pixels are set to 0.5.

(5) Distance matrix: We propose a distance matrix \mathbf{D} combining SA_m from (3) with IC_m from (4) to produce a shadow confidence map SC_m . The distance matrix $\mathbf{D} = \Gamma(IC_m, SA_m) \cdot IC_m$. $\Gamma(IC_m, SA_m) = 1 - \frac{Dis}{\max(Dis)}$ computes the distance score between pixel IC_{mij} in $IC_m (IC_{mij} \neq 0)$ to potential shadow areas in SA_m . Here, $Dis_{ij} = \min_{1 \leq v \leq t} \delta(IC_{mij}, SA_{mv})$, SA_{mv} is the center of

the v th connected component of all t connected components in SA_m and $\delta(\cdot)$ computes the distance between two positions. When multiplying $\Gamma(IC_m, SA_m)$ with IC_m map, pixels with similar shadow area intensity but far away from potential shadow areas achieve a lower score in SC_m .

3 Evaluation and Results

We test the proposed model on two data sets. Data set **A** consists of 993 2D ultrasound images sampled from 14 different anatomical standard planes and data set **B** comprises of 643 brain images. 48 non-brain images in data set **A** and whole data set **B** have been accurately manually segmented. We have shown shadow detection in Fig.2 and evaluated DICE overlap in Table 1. Because shadows

Table 1: Threshold ranges and DICE scores of different shadow detection methods: RW [4] vs. intermediate results from our approach and the final shadow confidence map.

	RW	S_m	FA_m	SA_m	SC_m
Dataset B	$T < 0.3$ 0.06	$T < 0.01 \cup T > 0.99$ 0.25	$T < 0.01 \cup T > 0.85$ 0.06	$T < 0.01 \cup T > 0.96$ 0.27	$T > 0.80$ 0.55
Dataset C	$T < 0.3$ 0.11	$T < 0.01 \cup T > 0.99$ 0.28	$T < 0.01 \cup T > 0.80$ 0.08	$T > 0.90$ 0.31	$T > 0.70$ 0.36

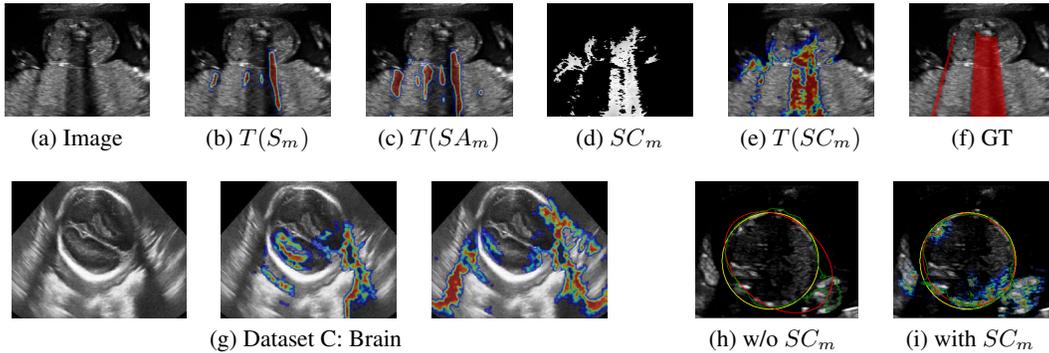


Figure 2: Rows 1 shows an example for shadow detection on kidney. The key steps from Fig. 1 are illustrated from (a), the input image to (f), the ground truth (GT) from manual segmentation. (g) illustrates the importance of the WGAN model (input image – w/o WGAN – with WGAN). (h-i) Improving automatic biometric measurements through applying SC_m as additional channel to a FCN [8] (yellow=GT, red=prediction, green=segmentation boundary).

can be harder to consistently annotate than anatomy, the DICE evaluation may seem low. However, our final aim is to provide a confidence map, which cannot be compared quantitatively to a ground truth. As such, we show the advantage of confidence maps via integration into an automatic method to measure biometrics using a FCN (Fig.2).

Conclusion: We have presented an accurate method to generate shadow-focused, pixel-wise confidence maps for ultrasound imaging. Such confidence maps can be used to identify less certain regions in the images, which is important for fully automatic segmentation tasks or automatic, image-based biometric measurements.

References

- [1] F. W. Kremkau and K.J. Taylor. Artifacts in ultrasound imaging. *J Ultrasound Med*, 5(4):227–237, 1986.
- [2] B. Bouhemad and et al. Clinical review: bedside lung ultrasound in critical care practice. *Critical Care*, 11(1):205, 2007.
- [3] P. Hellier and et al. An automatic geometrical and statistical method to detect acoustic shadows in intraoperative ultrasound brain images. *Med Image Anal*, 14(2):195–204, 2010.
- [4] A. Karamalis and et al. Ultrasound confidence maps using random walks. *Medical Image Analysis*, 16(6):1101–1112, 2012.
- [5] C. F. Baumgartner. and et al. Visual feature attribution using wasserstein gans. *CoRR*, abs/1711.08998, 2017.
- [6] C. F. Baumgartner and et al. SonoNet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE Trans. Med. Imaging*, 36(11):2204–2215, 2017.
- [7] J. T. Springenberg and et al. Striving for simplicity: The all convolutional net. *CoRR*, abs/1412.6806, 2014.
- [8] M. Sinclair and et al. Human-level performance on automatic head biometrics in fetal ultrasound using fully convolutional neural networks. In *EMBC’18*, 2018.