# Asymmetric Contour Uncertainty Estimation for Medical Image Segmentation

Thierry Judge[1⋆], Olivier Bernard[3], Woo-Jin Cho Kim[2], Alberto Gomez[2], Agisilaos Chartsias[2], and Pierre-Marc Jodoin[1]

[1] Department of Computer Science, University of Sherbrooke, Canada
[2] Ultromics Ltd., Oxford, OX4 2SU, UK
[3] University of Lyon, CREATIS, CNRS UMR5220, Inserm U1294, INSA-Lyon, University of Lyon 1, Villeurbanne, France

**Abstract.** Aleatoric uncertainty estimation is a critical step in medical image segmentation. Most techniques for estimating aleatoric uncertainty for segmentation purposes assume a Gaussian distribution over the neural network's logit value modeling the uncertainty in the presence of a class at that location. However, in many cases of segmentation, such as heart ultrasound or chest X-ray segmentation, there is no uncertainty about the presence of a specific structure but rather about the precise outline of that structure. For this reason, we propose to explicitly model the location uncertainty by reframing the usual pixel-by-pixel segmentation task into a contour regression problem. This allows for modeling the uncertainty of contour points using a more appropriate multivariate distribution. Also, since countour uncertainly is often anisotropic, we use a multivariate skewed Gaussian distribution. In addition to being directly interpretable, our uncertainty estimation method outperforms previous methods on three datasets using two different image modalities.

**Keywords:** Uncertainty estimation · Medical image segmentation · Deep learning.

## 1 Introduction

Segmentation is a key task in medical image analysis. Most state-of-the-art methods tackle segmentation with pixel-wise encoder-decoder type neural networks and achieve results within inter-expert variability for multiple tasks [?,?,?]. But recent work has shown the advantage of tackling the task of segmenting organs (e.g. heart, lungs) using a point based contouring approach [?,?]. Moreover, this approach is more interpretable as it is closer to how humans label data.

While point based approaches have improved various aspects of the segmentation pipeline, the crucial aspect of uncertainty has not yet been addressed. Uncertainty estimation methods can help identify out of distribution input data and/or inaccurate results.It is widely documented that uncertainty can be of epistemic or aleatoric nature [?]. While epistemic uncertainty is the uncertainty

⋆ Corresponding author

in the model [?], aleatoric uncertainty is the uncertainty in the data [?]. Most pixel-wise segmentation methods estimate aleatoric uncertainty per pixel by modeling a normal distribution over each output logit [?]. When segmenting organs of regular shape and location, there is no uncertainty in the presence of the organ but rather its precise delineation.

In this work, we therefore propose a novel method to estimate aleatoric uncertainty of contouring regression for segmentation by modeling each contour points with an interpretable bivariate normal distribution. As will be shown, our method also benefits from extending the standard normal distribution to a skew-normal distribution to allow the model to predict asymmetrical uncertainty. In turn, we show that the advantages of point-based segmentation approaches extent to uncertainty estimation.

Our contributions summarize as follows:

- we model assymetric uncertainties through outlines in segmentation tasks by model contour uncertainty using skew-normal distributions.
- We show how to convert regression uncertainty into a pixel-wise uncertainty map to allow a fair comparison with state-of-the-art methods.
- Our method outperforms classical segmentation uncertainty estimations in a rigorous evaluation using three datasets and two modalities.

## 2 Related work

### 2.1 Uncertainty Estimation

Both epistemic and aleatoric uncerainty are important to predict and decorrelate to assess the model performance and data quality, both of which are required to use image segmentation models in clinical settings. As previously mentioned epistemic uncertainty refers to the uncertainty in the model itself and is expressed by posing weights as a distribution instead of a point estimate. Many methods can be used to estimate epistemic uncertainty such as Bayesian neural networks [?], MC dropout [?,?] and ensembles [?]. In most cases, estimating epistemic uncertainty is task agnostic as it can be estimated by computing multiple forward passes.

Aleatoric uncertainty on the other hand pertains to the uncertainty in the data. This type of uncertainty must be modeled by assuming a distribution from which the data is drawn from. For regression, a common practice is to assume the data is iid from a normal distribution. To model aleatoric uncertainty, the model parameterized by $\theta$ must output the distribution parameters: $f_\theta(x) = (\mu(x), \sigma(x))$. The resulting loss function, derived by maximizing the log-likelihood for both $\mu$ and $\sigma$, is given by [?]

$$\mathcal{L}_{\mathcal{N}_1} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\log\sigma(x_i)^2 + \frac{1}{2\sigma(x_i)}||\mu(x_i) - y_i||^2, \tag{1}$$

where $y_i$ is the target value of $x_i$ (for simplicity $\mu(x_i) = \mu_i$).

Other methods attempt to model the aleatoric uncertainty by adding test-time augmentation [?,?]. The model is given multiple examples of the same data, and an uncertainty distribution can be estimated from the different outputs.

Some methods estimate uncertainty without explicitly modeling uncertainty (neither epistemic nor aleatoric). Some methods include training with custom losses [?] or adding an auxiliary confidence network [?]. Other methods, learn representations to predict uncertainty based on encoded prior [?] or to sample the representation space [?,?]. The latter however requires a dataset containing multiple annotations per image to obtain optimal results.

### 2.2 Segmentation as Regression

Posing segmentation as a regression problem provides multiple advantages. This task involves regressing a fixed number [say $K$] 2D points along the contour. This task is quite common in human pose estimation [?] and medical landmark regression [?]. The simplest way to do this is to regress $2K$ values with a standard CNN. This, however, achieves sub-optimal results due to the loss of spatiality in the flatten layer. A better approach involve a convolutional encoder and graph decoder [?] and heatmap regression [?]. In this work, we focus on the latter as it offers a direct way of estimating spatial uncertainty. One effective approach is that by Nibali *et al.* in which they propose the *differentiable spatial to numerical transform* (DSNT) to extract numerical coordinates from a heatmap in a differentiable operation [?].

Given a heatmap $Z \in R^{m\times n}$ of the size of the input image, two coordinate maps are defined $\mathbf{I} \in R^{m\times n}$ and $\mathbf{J} \in R^{m\times n}$ where $\mathbf{I}_{i,j} = \frac{2j-(n+1)}{n}$ and $\mathbf{J}_{i,j} = \frac{2i-(m+1)}{m}$. The heatmap is normalized to obtain $\hat{Z} = \phi(Z)$ ($\phi(\cdot)$ is a 2D flat softmax) so it can represent a probability mass function for each coordinate $c$:

$$P(c = [\mathbf{I}_{i,j}, \mathbf{J}_{i,j}]) = \hat{Z}_{i,j}. \tag{2}$$

Then, the mean of the heatmap can be found as the expected value of $c$,

$$DSNT(\hat{Z}) = E[c_x] = \mu = \left[\langle \hat{Z}, \mathbf{I}\rangle_F, \langle \hat{Z}, \mathbf{J}\rangle_F\right] \tag{3}$$

where $\langle \cdot, \cdot \rangle_F$ is the Frobenius inner product. The operation is differentiable and allows for sub-pixel accuracy. The variance of $\hat{Z}$ in x direction can be computed as follows (the $y$ variance is found similarly):

$$\begin{aligned} Var[c_x] &= E[(c_x - E[c_x])^2) \\ &= \langle \hat{Z}, (\mathbf{I} - \mu_x) \odot (\mathbf{I} - \mu_x)\rangle_F \end{aligned} \tag{4}$$

where $\odot$ is the Hadamar product. Nibali et al.'s work uses the variance to regularize the heatmaps with a loss between the output variance and a target variance set as a hyperparamters. The following section will explain how we use the DSNT layer to learn the variance to express uncertainty.

# 3 Method

figures/figure.png

Fig. 1: a) Example of uncertainty over-estimation due to symmetric gaussian. b) Example of skewed distributions and skewed uncertainty map.

## 3.1 Contouring Uncertainty

Segmentation datasets typically comes with $N$ pairs of data each containing an image $x_i \in R^{H,W}$ (we assume a one-channel input for simplicity) with its associated segmentation mask $y_i$. In this work, $y_i$ is a polygon $Q$ made of $K_q$ contour vertices. Note that $y_i$ can comprise several polygons $Q$ associated with several regions of different classes (i.e the two lungs, right and left ventricles, etc.). The procedure for extracting the vertices from the segmentation masks will be explained in section **??**.

To model the uncertainty of a contour $Q$, the uncertainty for each 2D vertex point must be modeled. This can be achieved using equation 1. This, however, assumes no covariance between the uncertainty along the horizontal and vertical axis. Instead, one should model the uncertainty of the entire contour using a $2K_q$ dimensional multivariate Gaussian [?]. This would consider the covariance between all points. This, however, is ineffective in practice as it requires predicting a $2K_q \times 2K_q$ covariance matrix which is unstable for neural networks. A good compromise is to model each contour point with a bivariate Gaussian. The networks must therefore take the following from $f_\theta(x) = (\mu, \Sigma)$ with $\mu \in R^{K_q \times 2}$ and $\Sigma \in R^{K_q \times 2 \times 2}$. The resulting loss function is

$$\mathcal{L}_{\mathcal{N}_2} = \frac{1}{NK_q} \sum_{i=1}^{N} \sum_{k=1}^{K_q} \log |\Sigma_i^{(k)}| + (\mu_i^{(k)} - y_i^{(k)})^T (\Sigma_i^{(k)})^{-1} (\mu_i^{(k)} - y_i^{(k)}). \quad (5)$$

The DSNT layer already offers the possibility of obtaining coordinates and variance from heatmaps. We add covariance output to the DSNT layer as follows:

$$\begin{aligned} cov[c] &= E[(c_x - E[c_x])(c_y - E[c_y]) \\ &= \langle \hat{Z}, (\mathbf{I} - \mu_x) \odot (\mathbf{J} - \mu_y) \rangle_F . \end{aligned} \quad (6)$$

While modeling the uncertainty of 2D contour points with a bivariate Gaussian distribution appears logical, it does have some drawbacks. For instance, this requires the uncertainty to be isotropic at each vertex, which is a strong assumption. To relax that constrain, we propose extending the modeling distribution to a bivariate *skew normal* distribution whose PDF is :

$$\mathcal{SN}_n(y|\zeta, \Sigma, \alpha) = 2\phi_n(y|\mu, \Sigma)\Phi_1(\alpha^T \omega^{-1}(y - \zeta)) \quad (7)$$

where $\phi_m$ is a multivariate normal, $\Phi_1$ is the cumulative distribution function of unit normal, $\Sigma = \omega\bar{\Sigma}\omega$ and $\alpha \in R^2$ is the skewness parameter. Note that we now use $\zeta$ to denote the location parameter that is different from the mean.

From that point on, the model predicts the skewness parameter for each contour point: $f_\theta(x) = (\zeta, \Sigma, \alpha)$ with $\zeta \in R^{K \times 2}$, $\Sigma \in R^{K \times 2 \times 2}$ and $\alpha \in R^{K \times 2}$. Deriving the maximum likelihood estimate for the bivariate skew-normal distribution yield the following loss function:

$$\begin{aligned} \mathcal{L}_{\mathcal{SN}_2} = \frac{1}{NK_q} \sum_{i=1}^{N} \sum_{k=1}^{K_q} & \log |\Sigma_i^{(k)}| + (\zeta_i^{(k)} - y_i^{(k)})^T (\Sigma_i^{(k)})^{-1} (\zeta_i^{(k)} - y_i^{(k)}) \\ & + \log \Phi_1((\alpha_i^{(k)})^T (\omega_i^{(k)})^{-1} (y_i^{(k)} - \zeta_i^{(k)})). \end{aligned} \quad (8)$$

### 3.2 Uncertainty Map

While the estimation of uncertainty provides numerous advantages such as the quantification of uncertainty per region and of specific anatomical landmark, there are still situations in which the uncertainty is required to be expressed in pixels. Indeed, an uncertainty map, in which each pixels represents the uncertainty can be an excellent tool to visualize and compare uncertainty estimation methods. We therefore need to construct an uncertainty map.

An observation that can be made about the estimated uncertainty is that when considering the contour as a whole, not all components of the uncertainty express in 2D are relevant. As there are a finite number of points along the contour, a specific point can be moved by a small distance along the contour without influencing the overall shape of the contour. One the contrary, moving the pixel along the direction perpendicular to the contour will have a maximal effect on the shape. This extends to the uncertainty, as the uncertainty expressed along the contour does not reflect the uncertainty of the shape. However, the uncertainty perpendicular to the contour better represents the uncertainty of the shape itself.

Uncertainty perpendicular is the relevant uncertainty

Given a vector $v$ with angle $\theta$ that is perpendicular to the contour. We want to find the standard deviation

This is done by computing the conditional of the rotated covariance matrix.

$$\Sigma^{'} = R \Sigma R^T \tag{9}$$

where R is a rotation matrix with angle $-\theta$. Conditional of $\Sigma^{'}$ is given by [?]

$$\begin{aligned} \sigma &= \Sigma^{'}_{xx} - \Sigma^{'}_{xy} (\Sigma^{'}_{yy})^{-1} \Sigma^{'}_{xy} \\ &= \frac{\Sigma^{'}_{xx} \Sigma^{'}_{yy} - \Sigma^{'}_{xy} \Sigma^{'}_{xy}}{\Sigma^{'}_{yy}} \\ &= \frac{|\Sigma^{'}|}{\Sigma^{'}_{yy}} \\ &= \frac{|R||\Sigma||R^T|}{\Sigma_{xx} \sin^2 \theta + \Sigma_{yy} \cos^2 \theta - 2\Sigma_{xy} \sin \theta \cos \theta} \\ &= \frac{|\Sigma|}{\Sigma_{xx} \sin^2 \theta + \Sigma_{yy} \cos^2 \theta - 2\Sigma_{xy} \sin \theta \cos \theta} \end{aligned} \tag{10}$$

Given the perpendicular distribution, we can construct contours that follow equal probability.

Bivariate normal

Bivaraite skew normal marginal

$$\hat{\alpha}_1 = \frac{\alpha_1 + \bar{\Omega}_{11}^{-1} \bar{\Omega}_{12} \alpha_2}{(1 + \alpha_2^T \bar{\Omega}_{22 \cdot 1} \alpha_2)^{1/2}} \tag{11}$$

$$\bar{\Omega}_{22 \cdot 1} = \bar{\Omega}_{22} - \bar{\Omega}_{22} \bar{\Omega}_{11}^{-1} \bar{\Omega}_{12} \tag{12}$$

## 4 Experimental Setup

### 4.1 Data

Our evaluation was conducted on two heart ultrasound datasets and one chest X-Ray dataset.

**CAMUS.** We used the CAMUS dataset [?] which contains cardiac ultrasounds from 500 patients. Two-chamber and four-chamber sequences were acquired for each patient. Manual annotations for the endocardium and epicardium borders of the left ventricle (LV) and the left atrium were obtained from a cardiologist for the end-diastolic (ED) and end-systolic (ES) frames. The dataset was split into 400 training patients, 50 validation patients and 50 testing patients. Contour points where extracted by finding the basal points of the endocardium and epicardium and subsequently the apex as the farthest points along the edge of the mask. Nine points where equally spaced between each basal point and apex for a total of 42 points for both contours.

**Proprietary US.**

**JSRT.**The Japanese Society of Radiological Technology (JSRT) dataset consists of 247 chest X-Rays [?]. We used the 120 landmakrs for the lungs and heart annotation made available in [?]. The set of landmarks contain specific anatomical points for each structure (4 for the right lung, 5 for the left lung and 4 for the heart) and equally spaced points between each anatomical point. We reconstructed the segmentation map with 3 classes (background, lungs, heart) with these landmarks and used the same train-val-test split of 70%-10%-20% as [?].

### 4.2 Implementation Details

1. enet / deeplab
2. B-Spline for contours / derivatives
3. US / CHEST reconstructions
4. best validation loss
5. Adam, lr=1e-3, weight decay = 1e-4
6. data augmentation
7. all images to 256x256

### 4.3 Evaluation Metrics

As well as average class Dice, the following metrics were used to asses the quality of the uncertainty estimates at both a image and pixel-wise level.

**Correlation.** The correlation between the the image uncertainty and the Dice was computed using the absolute value of the Peasion correlation score. We obtained image uncertainty be taking the sum of the uncertainty map and dividing it by the number of foregroud pixels. Higher correlation values indicated better uncertainty estimation.

**Expected Calibration Error (ECE)** Expected calibration error measures if a classifier's (by extension segmentation method) confidence represents its probability of being correct [?].

**Uncertinaty error mutual-information**. As proposed in [?], uncertainty error mutual-information measures the degree of overlap between the uncertainty map and the pixel-wise error map without requiring the uncertainty map to be thresholded. We report the average uncertainty error mutual information weighted by the number of erroneous pixels in each image.

| Data | CAMUS | | | | Proprietary | | | | JSRT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Method | Dice | Corr. | ECE | MI | Dice | Corr. | ECE | MI | Dice | Corr. | ECE | MI |
| Aleatoric | | | | | | | | | | | | |
| Epistemic | | | | | | | | | | | | |
| TTA | | | | | | | | | | | | |
| Edge [**?**] | | | | | | | | | | | | |
| Epistemic | | | | | | | | | | | | |
| TTA | | | | | | | | | | | | |
| $\mathcal{N}_1$ | | | | | | | | | | | | |
| $\mathcal{N}_2$ | | | | | | | | | | | | |
| $\mathcal{SN}_2$ | | | | | | | | | | | | |

Table 1: Uncertainty estimation results for segmentation (top rows) and regression (bottom rows) methods. Bold values indicate best results.

## 5 Results

Due its similar construction of the uncertainty, we also compare with the EDGE method presented in [**?**]. This method adds decreasing uncertainty around the contour trough successive dilations and errosions. Note however, that the EDGE method will always predict equal uncertainty around the contour which is not the case with our method.

1. Segmentation
   (a) aleatoric
   (b) epistemic
   (c) TTA
   (d) EDGE
2. Regression
   (a) Epistemic
   (b) TTA
   (c) Univariate normal 1
   (d) Bivarite normal
   (e) Bivariate skew-normal

## 6 Discussion and Conclusion

Although segmentation aleatoric uncertainty modeling with pixel-wise uncertainty is unconstrained, that is each pixel can produce uncertainty independently, practical results ins previous work and this work have shown that the results are sub-optimal. This method was originally proposed for multiclass segmentation of natural images where the uncertainty should reflect the probability of being in each class. In many medical image segmentation applications, however, the

Our results show that modeling the uncertainty associated with the location of the segmentation structure rather than the presence yields better results. Furthermore, by allowing the distribution to

While this method cannot be applied to all medical image segmentation tasks, this method is still broad enough to cover many applications, especially related to segmentation that is later used for downstream tasks such as clinical metric estimation.

Future work will look to expand this method to more general variations of elliptical distributions and combine the aleatoric and epistemic uncertainty to obtain the full predictive uncertainty.

**Acknowledgements** Please place your acknowledgments at the end of the paper, preceded by an unnumbered run-in heading (i.e. 3rd-level heading).