Supplementary Material
AmnioML: Amniotic Fluid Segmentation and Volume Prediction with Uncertainty Quantification

Daniel Csillag[1], Lucas Monteiro[2], Thiago Ramos[1], João Vitor Romano[1], Rodrigo Schuller[1], Roberto B. Seixas[1], Roberto I. Oliveira[1], and Paulo Orenstein[1]

[1]IMPA
[2]Harvard University

Abstract

Accurately predicting the volume of amniotic fluid is fundamental to assessing pregnancy risks, though the task usually requires many hours of laborious work by medical experts. In this paper, we present AmnioML, a machine learning solution that leverages deep learning and conformal prediction to output accurate volume estimates and segmentation masks from fetal MRIs in a few seconds, with Dice coefficient over 0.9, along with valid predictive intervals. To do so, we make available a novel, curated dataset for fetal MRIs with 853 exams, and benchmark the performance of many recent deep learning architectures. We also introduce modifications to conformal predictions tools that yield narrower predictive intervals with valid coverage, thus aiding doctors in quantifying pregnancy risks. A successful case study of AmnioML use in a medical setting is also reported. Real-world clinical benefits range from up to 20x segmentation time reduction, with over 60% of segmentations requiring no further human intervention. AmnioML's volume predictions were found to be highly accurate in practice, with mean absolute error below 55mL, and tight predictive intervals.

1 Notation

The set $\{(X_i, Y_i)\}_{i=1}^n$ will denote an independent and identically distributed sample of pairs of 3D exams X_i and their corresponding medical segmentation Y_i, with v voxels. Data indices will be partitioned into $I_{\text{train}} = \{1, \dots, n_{\text{train}}\}$, $I_{\text{val}} = \{n_{\text{train}} + 1, \dots, n_{\text{train}} + n_{\text{val}}\}$ and $I_{\text{test}} = \{n_{\text{train}} + n_{\text{val}} + 1, \dots, n_{\text{train}} + n_{\text{val}} + n_{\text{test}}\}$, with $n_{\text{train}} + n_{\text{val}} + n_{\text{test}} = n$ and $n_{\text{train}}, n_{\text{val}}, n_{\text{test}} \in \mathbb{N}_+$. Any model $\mathcal{M}$ will always be trained on $\{(X_i, Y_i)\}_{i \in I_{\text{train}}}$ and evaluated on $\{(X_i, Y_i)\}_{i \in I_{\text{test}}}$. Table 1 contains further notation employed throughout the text.

Symbol	Description	Domain
X	3D MRI exam	$[0,1]^v$
Y	AF segmentation mask	$\{0,1\}^v$
$\text{Vol}(Y)$	Exam volume	$\mathbb{R}_{\geq 0}$
$\mathcal{M}(X)$	Trained model output	$[0,1]^v$
$\mathcal{M}(X)_{\geq t}$	Model output thresholded at t	$\{0,1\}^v$
$A \odot B$	Element-wise multiplication	$\mathbb{R}^v \times \mathbb{R}^v \to \mathbb{R}^v$
$\overline{A}$	Mask negation: $\mathbf{1} - A$	$\{0,1\}^v \to \{0,1\}^v$

2 Algorithms

Algorithm 1: Thresholded Volume Prediction

```
Procedure calibrate(model M, validation set {(X_i, Y_i)}_{i∈I_val}, confidence 1 − α ∈ (0, 1))
    thresholds ← [ ]
    for i ∈ I_val do
        p ← proportion of zeros in Y_i
        best threshold ← p-quantile(M(X_i))
        append best threshold to thresholds
    end
    α̃ ← ⌈(n_val + 1)α⌉/n_val
    u ← (1 − α̃/2)-quantile of list thresholds
    l ← (α̃/2)-quantile of list thresholds
    return l, u
Procedure predict(prediction M(X), l, u)
    lower volume ← Vol(M(X)_{≥u})
    upper volume ← Vol(M(X)_{≥l})
    return [lower volume, upper volume]
```

Algorithm 2: Standard Volume Prediction

```
Procedure calibrate( model M, validation set {(X_i, Y_i)}_{i∈I_val}, confidence 1 − α ∈ (0, 1),
 normalizing function g : [0, 1]^v → R_+.)
    radii ← [ ]
    for i ∈ I_val do
        append (|Vol(M(X_i)_{≥.5}) − Vol(Y_i)| / g(M(X_i)_{≥.5})) to radii
    end
    α̃ ← ⌈(n_val + 1)α⌉/n_val
    r ← (1 − α̃)-quantile of radii
    return r
Procedure predict( prediction M(X), normalizing function g : [0, 1]^v → R_+, r)
    dv ← g(M(X)) · r
    lower volume ← Vol(M(X)_{≥.5}) − dv
    upper volume ← Vol(M(X)_{≥.5}) + dv
    return [lower volume, upper volume]
```

Algorithm 3: Segmentation Prediction

Procedure `calibrate`(*model* $\mathcal{M}$*, validation set* $\{(X_i, Y_i)\}_{i \in I_{\text{val}}}$*, leniency* $\lambda \in [0,1]$ *and confidence* $1 - \alpha \in (0,1)$)

  upper thresholds $\leftarrow$ [ ]
  lower thresholds $\leftarrow$ [ ]
  for $i \in I_{\text{val}}$ **do**
    $\lambda_{\text{upper}} \leftarrow \lambda$
    $t_u \leftarrow \lambda_{\text{upper}}$-quantile$(\mathcal{M}(X_i)|Y_i[v] = 1)$
    append $\min(t_u, 0.5)$ to upper thresholds
    $\lambda_{\text{lower}} \leftarrow 1 - \lambda \cdot \text{Vol}\,(Y_i)\,/\text{Vol}\,(1 - Y_i)$
    $t_l \leftarrow \lambda_{\text{lower}}$-quantile$(\mathcal{M}(X_i)|Y_i[v] = 0)$
    append $\max(t_l, 0.5)$ to lower thresholds
  end
  $\tilde{\alpha} \leftarrow \lceil (n_{\text{val}} + 1)\alpha \rceil / n_{\text{val}}$
  $l \leftarrow (1 - \tilde{\alpha}/2)$-quantile of lower thresholds
  $u \leftarrow (\tilde{\alpha}/2)$-quantile of upper thresholds
  return l, u

Procedure `predict`(*prediction* $\mathcal{M}(X)$, l, u)

  $\mathcal{U} \leftarrow \mathcal{M}(X)_{\geq u}$
  $\mathcal{L} \leftarrow \mathcal{M}(X)_{> l}$
  return $[\mathcal{L}, \mathcal{U}]$

3 Main Theorems and Proofs

Proofs of our main theorems relies on [AB21, Theorem A.1] which is an adaptation of the classical Conformal Prediction result [PPVG02, Propostion 1].

Proposition 1. *Let $\{Z_i\}_{i=1}^n$ be an iid sample. Given a confidence $1 - \alpha \in (0,1)$ define $\widehat{q}_{1-\tilde{\alpha}}$ as*

$$\widehat{q}_{1-\tilde{\alpha}} = Z_{(\lceil (n+1)(1-\alpha) \rceil)},$$

where $Z_{(k)}$ indicates the kth order statistics of the list of random variables $\{Z_1, \ldots, Z_n\}$. Then, for a new independent identically distributed random variable Z

$$\mathbb{P}[Z \leq \widehat{q}_{1-\tilde{\alpha}}] \geq 1 - \alpha. \tag{1}$$

Theorem 2. *For a confidence level $1 - \alpha$, take $\mathcal{I}_\alpha^{\text{tvp}}$ as in Algorithm 1, using $\{(X_i, Y_i)\}_{i \in \text{val}}$ as validation set. Then, for any $j \in I_{\text{test}}$,*

$$\mathbb{P}[\text{Vol}\,(Y_j) \in \mathcal{I}_\alpha^{\text{tvp}}(X_j)] \geq 1 - \alpha.$$

Proof. Given any segmentation image Y, note that $1 - \frac{\text{Vol}(Y)}{v}$ represents its proportion of 0's. For $i \in I_{\text{val}}$, define $\beta_i = 1 - \frac{\text{Vol}(Y_i)}{v}$ and $Z_i = \mathcal{M}(X_i)[(\lceil v\beta_i \rceil)]$, where $\mathcal{M}(X_i)[(k)]$ indicates the kth order statistic of the set of random variables $\{\mathcal{M}(X_i)[1], \ldots, \mathcal{M}(X_i)[v]\}$.

Since each Z_i only depends on the pair (X_i, Y_i), we can apply Proposition 1 to show that for $j \in I_{\text{test}}$,

$$\mathbb{P}\left[\mathcal{M}(X_j)[(\lceil v\beta_j \rceil)] \leq u\right] \geq 1 - \alpha/2,$$

where u is as defined in Algorithm 1. But then,

$$\begin{aligned} 1 - \alpha/2 &\leq \mathbb{P}\left[\mathcal{M}(X_j)[(\lceil v\beta_j \rceil)] \leq u\right] \\ &\leq \mathbb{P}\left[\text{Vol}(\mathcal{M}(X_j)_{\geq \mathcal{M}(X_j)[(\lceil v\beta_j \rceil)]}) \geq \text{Vol}(\mathcal{M}(X_j)_{\geq u}\right] \\ &= \mathbb{P}\left[\text{Vol}(Y_i) \geq \text{Vol}(\mathcal{M}(X_j)_{\geq u}\right] \end{aligned}$$

where the last inequality comes from the fact that $\mathrm{Vol}(\mathcal{M}(X_j)_{\geq \mathcal{M}(X_j)[(\lceil v\beta_j \rceil)]}$ counts how many elements of the vector $\mathcal{M}(X_j)$ are larger than its $(\lceil v\beta_j \rceil)$th order static element, that is,

$$\begin{aligned}\mathrm{Vol}(\mathcal{M}(X_j)_{\geq L_j}) &= v - \lceil v\beta_j \rceil \\ &= v - \left\lceil v\left(1 - \frac{\mathrm{Vol}(Y_j)}{v}\right)\right\rceil \\ &= \mathrm{Vol}(Y_j).\end{aligned}$$

Again by Proposition 1, we have that for $j \in I_{\text{test}}$,

$$\mathbb{P}\left[\mathcal{M}(X_j)[(\lceil v\beta_j \rceil)] \geq l)\right] \geq 1 - \alpha/2,$$

where l is as defined in Algorithm 1. As before, we have that

$$\begin{aligned}1 - \alpha/2 &\leq \mathbb{P}\left[\mathcal{M}(X_j)[(\lceil v\beta_j \rceil)] \geq l)\right] \\ &\leq \mathbb{P}\left[\mathrm{Vol}(\mathcal{M}(X_j)_{\geq \mathcal{M}(X_j)[(\lceil v\beta_j \rceil)]}) \leq \mathrm{Vol}(\mathcal{M}(X_j)_{\geq l})\right] \\ &\leq \mathbb{P}\left[\mathrm{Vol}(Y_i) \leq \mathrm{Vol}(\mathcal{M}(X_j)_{\geq l})\right].\end{aligned}$$

Therefore, for $j \in I_{\text{test}}$ and the union bound,

$$\begin{aligned}\mathbb{P}[\mathrm{Vol}(Y_j) \in \mathcal{I}_\alpha^{\text{tvp}}(X_j)] &= 1 - \mathbb{P}[\mathrm{Vol}(Y_j) \notin \mathcal{I}_\alpha^{\text{tvp}}(X_j)] \\ &\geq 1 - \mathbb{P}\left[\mathrm{Vol}(Y_i) > \mathrm{Vol}(\mathcal{M}(X_j)_{\geq l})\right] - \mathbb{P}\left[\mathrm{Vol}(Y_i) < \mathrm{Vol}(\mathcal{M}(X_j)_{\geq u})\right] \\ &\geq 1 - \alpha/2 - \alpha/2 \\ &= 1 - \alpha.\end{aligned}$$

□

Theorem 3. *Given a normalizing function $g : [0,1]^v \to \mathbb{R}_+$ and a confidence level $1 - \alpha$, take $\mathcal{I}_\alpha^{\text{svp}}$ as Algorithm 2. Then, for $j \in I_{\text{test}}$*

$$\mathbb{P}[\mathrm{Vol}(Y_j) \in \mathcal{I}_\alpha^{\text{svp}}(X_j)] \geq 1 - \alpha.$$

Proof. Taking $Z_i = \frac{\left|\mathrm{Vol}\left(\mathcal{M}(X_i)_{\geq .5}\right) - \mathrm{Vol}(Y_i)\right|}{g(\mathcal{M}(X_i)_{\geq .5})}$ with $i \in I_{\text{val}}$ in Proposition 1 yields for any $j \in I_{\text{test}}$,

$$\mathbb{P}\left[\frac{\left|\mathrm{Vol}\left(\mathcal{M}(X_j)_{\geq .5}\right) - \mathrm{Vol}\left(Y_j\right)\right|}{g(\mathcal{M}(X_j)_{\geq .5})} \leq r\right] \geq 1 - \alpha,$$

where r is as defined in Algorithm 2. That is, if $dv_j = g(\mathcal{M}(X_j)_{\geq .5}) \cdot r$

$$\begin{aligned}\mathbb{P}\left[-dv_j + \mathrm{Vol}\left(\mathcal{M}(X_j)_{\geq .5}\right) \leq \mathrm{Vol}\left(Y_j\right) \leq dv_j + \mathrm{Vol}\left(\mathcal{M}(X_j)_{\geq .5}\right)\right] &= \mathbb{P}[\mathrm{Vol}(Y_j) \in \mathcal{I}_\alpha^{\text{svp}}(X_j)] \\ &\geq 1 - \alpha.\end{aligned}$$

□

Theorem 4. *For confidence level $1-\alpha$ and lenience value $\lambda \in [0,1]$, consider the upper and lower thresholds $u, l \in [0,1]$ returned by Algorithm 3. Then, for $j \in I_{\text{test}}$, with probability at least $1 - \alpha$*

$$\max\left\{\mathrm{Vol}\left(Y_j \odot \overline{\mathcal{M}(X_j)_{\geq u}}\right), \mathrm{Vol}\left(\mathcal{M}(X_j)_{>l} \odot \overline{Y_j}\right)\right\} \leq \lambda \mathrm{Vol}(Y_j).$$

Proof. For $i \in I_{\text{val + test}}$, define $Z_i = \inf\left\{t \ : \ \frac{\sum_{k=1}^{v} \mathbb{I}_{[\mathcal{M}(X_i)[k] \leq t]}\mathbb{I}_{[Y_i[k]=1]}}{\sum_{k=1}^{v} \mathbb{I}_{[Y_i[k]=1]}} \geq \lambda\right\}$. If we apply Proposition 1 with $\{Z_i\}_{i \in I_{\text{val}}}$, we have that for $j \in I_{\text{test}}$,

$$\mathbb{P}\left(Z_j \geq u\right) \geq 1 - \alpha/2,$$

where u is as defined in Algorithm 3. But then,

$$
\begin{aligned}
1-\alpha/2 &\le \mathbb{P}\left[Z_j \ge u\right] \\
&\le \mathbb{P}\left[\mathrm{Vol}(Y_j \odot \overline{\mathcal{M}(X_j)_{\ge Z_j}}) \ge \mathrm{Vol}(Y_j \odot \overline{\mathcal{M}(X_j)_{\ge u}})\right] \\
&= \mathbb{P}\left[\mathrm{Vol}(Y_j \odot \overline{\mathcal{M}(X_j)_{\ge Z_j}}) \ge \mathrm{Vol}(Y_j \odot \overline{\mathcal{U}(\mathcal{M}(X_j))})\right].
\end{aligned}
$$

But note that by the definition of Z_j,

$$
\begin{aligned}
\mathrm{Vol}(Y_j \odot \overline{\mathcal{M}(X_j)_{\ge Z_j}}) &= \sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=1]}\mathbb{I}_{[\mathcal{M}(X_j)[k]<Z_j]} \\
&\le \lambda \sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=1]} \\
&= \lambda \mathrm{Vol}(Y_j),
\end{aligned}
$$

where the inequality comes from the fact that for any $t < Z_j$, by the definition of Z_j

$$
\sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=1]}\mathbb{I}_{[\mathcal{M}(X_j)[k]<t]} < \lambda \sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=1]},
$$

thus,

$$
\begin{aligned}
\sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=1]}\mathbb{I}_{[\mathcal{M}(X_j)[k]<Z_j]} &= \lim_{t \nearrow Z_j} \sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=1]}\mathbb{I}_{[\mathcal{M}(X_j)[k]<t]} \\
&\le \lambda \sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=1]}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
1-\alpha/2 &\le \mathbb{P}\left[\mathrm{Vol}(Y_j \odot \overline{\mathcal{M}(X_j)_{\ge Z_j}}) \ge \mathrm{Vol}(Y_j \odot \overline{\mathcal{U}(\mathcal{M}(X_j))})\right] \\
&\le \mathbb{P}\left[\lambda \mathrm{Vol}(Y_j) \ge \mathrm{Vol}(Y_j \odot \overline{\mathcal{U}(\mathcal{M}(X_j))})\right].
\end{aligned}
$$

For the second part, for $i \in I_{\mathrm{val+test}}$, define $W_i = \inf\left\{t \;:\; \frac{\sum_{k=1}^{v} \mathbb{I}_{[\mathcal{M}(X_i)[k]\le t]}\mathbb{I}_{[Y_i[k]=0]}}{\sum_{k=1}^{v} \mathbb{I}_{[Y_i[k]=0]}} \ge 1-\lambda \frac{\mathrm{Vol}(Y_i)}{\mathrm{Vol}(1-Y_i)}\right\}$, where $1-Y_i = (1-Y_i[1], \ldots, 1-Y_i[v])$. If we apply Proposition 1 with $\{W_i\}_{i\in I_{\mathrm{val}}}$, we have that for $j \in I_{\mathrm{test}}$,

$$
\mathbb{P}\left[W_j \le l\right] \ge 1-\alpha/2,
$$

where l is as defined in Algorithm 3. But then,

$$
\begin{aligned}
1-\alpha/2 &\le \mathbb{P}\left[W_j \le l\right] \\
&\le \mathbb{P}\left[\mathrm{Vol}(\mathcal{M}(X_j)_{>W_j} \odot \overline{Y_j}) \ge \mathrm{Vol}(\mathcal{M}(X_j)_{>l} \odot \overline{Y_j})\right] \\
&= \mathbb{P}\left[\mathrm{Vol}(\mathcal{M}(X_j)_{>W_j} \odot \overline{Y_j}) \ge \mathrm{Vol}(\mathcal{L}(\mathcal{M}(X_j)) \odot \overline{Y_j})\right]
\end{aligned}
$$

But note that by the definition of W_j,

$$\begin{aligned}
\mathrm{Vol}(\mathcal{M}(X_j)_{>W_j} \odot \overline{Y_j}) &= \sum_{k=1}^{v} \mathbb{I}_{[\mathcal{M}(X_j)[k]>W_j]} \mathbb{I}_{[Y_j[k]=0]} \\
&= \sum_{k=1}^{v} \mathbb{I}_{[Y_j[k]=0]} - \sum_{k=1}^{v} \mathbb{I}_{[\mathcal{M}(X_j)[k]\leq W_j]} \mathbb{I}_{[Y_j[k]=0]} \\
&\leq \mathrm{Vol}(1-Y_j) - \left(1 - \lambda \frac{\mathrm{Vol}(Y_j)}{\mathrm{Vol}(1-Y_j)}\right) \mathrm{Vol}(1-Y_j) \\
&= \lambda \mathrm{Vol}(Y_j).
\end{aligned}$$

Hence,

$$\begin{aligned}
1-\alpha/2 &\leq \mathbb{P}\left(\mathrm{Vol}(\mathcal{M}(X_j)_{>W_j} \odot \overline{Y_j}) \geq \mathrm{Vol}(\mathcal{L}(\mathcal{M}(X_j)) \odot \overline{Y_j})\right) \\
&\leq \mathbb{P}\left(\lambda \mathrm{Vol}(Y_j) \geq \mathrm{Vol}(\mathcal{L}(\mathcal{M}(X_j)) \odot \overline{Y_j})\right).
\end{aligned}$$

By the union bound,

$$\begin{aligned}
&\mathbb{P}\left[\lambda \mathrm{Vol}(Y_j) \geq \max\{\mathrm{Vol}(\mathcal{L}(\mathcal{M}(X_j)) \odot \overline{Y_j}), \mathrm{Vol}(Y_j \odot \overline{\mathcal{U}(\mathcal{M}(X_j))})\}\right] \\
&\geq 1 - \mathbb{P}\left[\lambda \mathrm{Vol}(Y_j) < \mathrm{Vol}(\mathcal{L}(\mathcal{M}(X_j)) \odot \overline{Y_j})\right] - \mathbb{P}\left[\lambda \mathrm{Vol}(Y_j) < \mathrm{Vol}(Y_j \odot \overline{\mathcal{U}(\mathcal{M}(X_j))})\right] \\
&\geq 1-\alpha.
\end{aligned}$$

□

References

[AB21] Anastasios Nikolas Angelopoulos and Stephen Bates. A gentle introduction to conformal prediction and distribution-free uncertainty quantification, 2021.

[PPVG02] Harris Papadopoulos, Kostas Proedrou, Volodya Vovk, and Alex Gammerman. Inductive confidence machines for regression. In Tapio Elomaa, Heikki Mannila, and Hannu Toivonen, editors, *Machine Learning: ECML 2002*, pages 345–356, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.