Article

Automated Ventricular and Midline Segmentation in Cranial Ultrasound with Metrology

Jaswant Vemulapalli * and Nicholus Vaughan

Posted Date: 31 March 2026

doi: 10.20944/preprints202603.2474.v1

Keywords: deep learning; cranial ultrasound; brain segmentation; sonolucent cranial implant; medical image segmentation; metrology

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2026 doi:10.20944/preprints202603.2474.v1

Article

Automated Ventricular and Midline Segmentation in Cranial Ultrasound with Metrology

Jaswant Vemulapalli * and Nicholus Vaughan

Longeviti Neuro Solutions
* Correspondence: jaswant.vemulapalli@longeviti.com

Abstract

Real-time ultrasound imaging through sonolucent cranial implants is an emerging modality for post-neurosurgical monitoring of the adult brain, but quantitative interpretation remains challenging due to speckle, attenuation, shadowing, and the difficulty of consistently delineating thin anatomical landmarks. We present a deep learning system developed at **Longeviti Neuro Solutions** for segmenting key intracranial structures–the ipsilateral and contralateral lateral ventricles and the cranial midline–in coronal-plane adult cranial ultrasound images from patients with **Longeviti ClearFit®** Acoustic Brain Interface (ABI)™ implants. Our dataset comprises 456 proprietary, de-identified ultrasound frames (JPEG with known pixel spacing) annotated in CVAT with ventricle and midline labels. We benchmark multiple encoder–decoder segmentation architectures and address severe class imbalance via class-weighted optimization, test-time augmentation (horizontal flip with left–right label swapping), and class-specific post-processing to reduce spurious components and improve mask coherence. The best-performing configuration achieves a foreground macro Dice of 0.869 on a held-out test set, with ventricle Dice values above 0.92 and midline Dice of approximately 0.75. Finally, we transform predicted masks into geometry-based metrology by estimating maximal perpendicular ventricle spans and ventricle-to-midline distances, producing standardized measurement overlays suitable for downstream review.

Keywords: deep learning; cranial ultrasound; brain segmentation; sonolucent cranial implant; medical image segmentation; metrology

---

1. Introduction

Ultrasound is an attractive modality for longitudinal neuro monitoring because it is portable, repeatable, and compatible with bedside acquisition. In adults, however, conventional transcranial ultrasound has historically been limited by the acoustic properties of the skull, which restrict acoustic penetration and degrade image quality. Sonolucent cranial implants provide a consistent acoustic window that can restore postoperative acoustic access to intracranial anatomy and enable repeated ultrasound imaging in adult patients [1,2].

Within the Longeviti ClearFit® Acoustic Brain Interface (ABI)™, implant-enabled cranial ultrasound creates an opportunity to move beyond qualitative review toward standardized, quantitative measurements that can be tracked over time. At the same time, achieving reliable quantitative assessment from adult cranial ultrasound remains challenging. Image quality varies substantially across frames due to speckle noise, attenuation, acoustic shadowing, and variability introduced by probe placement and insonation angle. These effects often produce incomplete or ambiguous boundaries, even when the acoustic window is consistent. Manual delineation of ventricular anatomy is therefore time intensive and difficult to scale across repeated examinations. Beyond ventricular boundaries, thin anatomical landmarks used as geometric references (e.g., the cranial midline) may be intermittently visible and sensitive to small discontinuities, which can destabilize downstream measurement overlays.

Automated segmentation offers a natural foundation for standardizing measurement workflows by producing consistent structure masks that can be transformed into calibrated geometric endpoints. Encoder–decoder convolutional networks, particularly the U-Net family, have become widely adopted in biomedical segmentation due to their ability to combine multi-scale context with boundary localization [3]. Nonetheless, adult cranial ultrasound introduces two properties that require particular attention in a deployment-oriented setting. First, the imaging domain is strongly artifact-driven: speckle patterns and shadowing can yield plausible-looking false edges or fragmented contours. Second, the learning problem is severely imbalanced, especially for thin reference structures that occupy a very small fraction of the image area. In such settings, predictions that appear reasonable under aggregate overlap metrics can still be topologically inconsistent or unstable across frames, undermining downstream measurement extraction.

This paper describes a deep learning system developed at Longeviti Neuro Solutions to segment three key structures from coronal-plane adult cranial ultrasound acquired through the ClearFit® implant window: the ipsilateral ventricle, the contralateral ventricle, and the cranial midline. The resulting masks are converted into geometry-based metrology, including ventricle spans and ventricle-to-midline distance measures, and rendered as standardized overlays for downstream review. The scope of this manuscript is segmentation and metrology overlay generation; it does not present clinical midline shift estimation as an endpoint and does not claim agreement of metrology against expert manual measurements, which is left for future clinical validation.

Contributions. (1) We introduce a production-oriented segmentation and metrology overlay pipeline for adult cranial ultrasound in the ClearFit implant setting, targeting bilateral ventricles and a midline reference structure. (2) We emphasize evaluation criteria aligned with thin-structure usability for downstream measurement extraction. (3) We provide an end-to-end workflow that produces standardized overlays suitable for downstream operational review and future clinical validation.

Figure 1 provides a high-level overview of the proposed workflow.

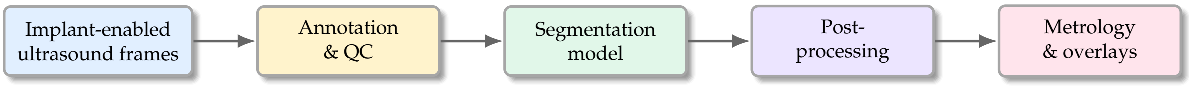


Figure 1. High-level pipeline: implant-enabled adult cranial ultrasound frames are annotated and reviewed, processed by a segmentation model, refined via deterministic post-processing, and converted into geometry-based metrology overlays for downstream review.

2. Related Work

2.1. *Implant-Enabled Adult Cranial Ultrasound Monitoring*

Transcranial ultrasound in adults is constrained by the strong acoustic impedance mismatch and attenuation introduced by the intact skull, which limits image quality and restricts the use of high-frequency probes in many clinical settings. Recent work on sonolucent cranial implants has revisited ultrasound as a modality for post-neurosurgical monitoring by restoring an acoustic window and enabling repeated bedside imaging through the implant. Prior feasibility studies have demonstrated the potential of implant-enabled ultrasound for neuroimaging in controlled phantom and cadaveric settings, establishing the practical basis for adult intracranial visualization through a cranial window [1,2]. Within this emerging context, reliable segmentation of clinically relevant intracranial structures becomes a key step toward standardized, quantitative interpretation.

2.2. *Ultrasound Ventricle Segmentation and Quantitative Measurement*

Most prior work on brain ultrasound segmentation focuses on neonatal or fetal imaging, where the open fontanelle provides a natural acoustic window and large datasets are more commonly available. In neonatal cranial ultrasound, segmentation of the lateral ventricles has been explored for downstream tasks such as monitoring ventricular dilation and supporting risk stratification; for example, Tabrizi *et al.* propose an automated pipeline for ventricle delineation in 2D cranial ultrasound to support outcome prediction in intraventricular hemorrhage [4]. Deep learning methods have also

doi:10.20944/preprints202603.2474.v1

been investigated for 3D ultrasound ventricle segmentation, including 3D U-Net based approaches for neonatal cerebral ventricles [5]. In parallel, ultrasound-based measurement extraction has been studied in other anatomical settings such as fetal ventriculomegaly screening, where automated ventricle measurements from ultrasound are used to reduce operator variability [6]. Compared to these neonatal and fetal settings, adult cranial ultrasound through a cranial implant presents distinct challenges (different anatomy, acoustic characteristics, and acquisition constraints) and remains relatively underexplored, motivating segmentation systems tailored to this modality.

2.3. Architectural Motifs for Medical Image Segmentation

Fully convolutional encoder-decoder designs remain the dominant paradigm for biomedical image segmentation, with U-Net providing a widely adopted baseline due to its multiscale feature aggregation and skip connections [3]. Many subsequent variants incorporate architectural components that improve optimization and representation capacity. Residual connections, introduced in ResNet, enable deeper networks by stabilizing gradient flow and are frequently used in segmentation encoders and residual U-Net variants [7]. Channel attention mechanisms such as squeeze-and-excitation (SE) blocks recalibrate feature channels adaptively and are commonly integrated into medical segmentation backbones to improve sensitivity to subtle structures [8]. For transfer learning, efficient classification backbones such as EfficientNet offer favorable accuracy-efficiency tradeoffs and have been adopted as encoders in segmentation pipelines, particularly when training data is limited or domain shift is a concern [9]. These design patterns motivate the family of architectures evaluated in this work (baseline U-Net, residual/SE-augmented variants, and an EfficientNet-based encoder) while keeping the overall pipeline compatible with dense pixel-wise prediction.

2.4. Thin-Structure Segmentation and Topology-Aware Evaluation

Accurate segmentation of thin, low-prevalence structures remains challenging because overlap-based metrics can be insensitive to connectivity errors, and class imbalance can dominate optimization dynamics. Topology-aware metrics have therefore been proposed to complement region overlap measures by explicitly evaluating structural continuity. In particular, clDice was introduced as a connectivity-sensitive similarity measure based on centerline overlap and has been shown to better reflect topological correctness for curvilinear structures than Dice alone [10]. For applications where continuity is clinically meaningful, reporting both overlap (e.g., Dice) and a connectivity-aware metric provides a more complete assessment of segmentation quality. This motivates our inclusion of a connectivity-aware evaluation alongside standard segmentation metrics for thin anatomical targets.

3. Materials and Methods

3.1. Dataset Provenance and Governance

We use a proprietary, de-identified dataset of adult cranial ultrasound frames acquired in the coronal plane through a sonolucent cranial implant window (Longeviti ClearFit®). Images and labels were generated within Longeviti Neuro Solutions for internal research and development; therefore, the dataset and training code cannot be released publicly. Per-frame pixel spacing metadata is available and is used to convert pixel distances into physical units for metrology (Section 4).

3.2. Frame Export, JPEG Representation, and Pixel Calibration

Ultrasound frames were exported for annotation and model development as RGB JPEG images derived from de-identified DICOM studies. For downstream physical measurements (metrology), per-frame pixel calibration was computed from ultrasound region calibration metadata in the DICOM headers, specifically *Physical Delta X* and *Physical Delta Y* (DICOM tags (0018,602C) and (0018,602E)), together with the associated *Physical Units* tag. In the exported studies used here, the delta values were expressed in centimeters and converted to millimeters as:

$$\Delta x_{\mathrm{mm}} = 10\,\Delta x_{\mathrm{cm}}, \qquad \Delta y_{\mathrm{mm}} = 10\,\Delta y_{\mathrm{cm}}. \tag{1}$$

The resulting calibration values were stored alongside image identifiers and used to convert pixel distances to millimeters during metrology computation (Section 4).

3.3. Annotation Protocol, Quality Control, and Dataset Split

Ground-truth labels target three intracranial structures: contralateral lateral ventricle, ipsilateral lateral ventricle, and intracranial midline; all remaining pixels are assigned to background. Masks were created in CVAT by trained internal employees following a standardized labeling guide for ventricle contours and midline tracing. Quality control was performed by two expert sonographers who reviewed all annotations; identified issues were returned for correction prior to model training and evaluation. Final masks were exported as RGB label images using a fixed color map, enabling deterministic LUT-based conversion to class-index masks during preprocessing.

The dataset contains 456 labeled frames and is partitioned into fixed training, validation, and holdout test sets containing 365, 68, and 24 frames respectively. The split is defined by text files listing image identifiers and is used consistently across training, checkpoint selection, post-processing tuning, and evaluation.

3.4. Preprocessing, Label Encoding, and Training-Time Augmentation

Input frames are represented as 8-bit RGB JPEG images and normalized to $[0, 1]$ by dividing intensities by 255. All models are trained and evaluated at a fixed spatial resolution of 768×1024 (height × width) with three input channels.

Segmentation masks are exported as RGB label images and converted into dense class-index maps $M \in \{0, 1, 2, 3\}^{H \times W}$ using a deterministic lookup-table (LUT) derived from a project labelmap file. Following the evaluation pipeline, each RGB pixel (r, g, b) is mapped to an integer key

$$k = (r \ll 16) \mid (g \ll 8) \mid b, \tag{2}$$

and a precomputed LUT maps $k \mapsto c$, where $c \in \{0, 1, 2, 3\}$ denotes the semantic class. The resulting index map uses the fixed class ordering: background (0), contralateral ventricle (1), ipsilateral ventricle (2), and midline (3). This LUT-based conversion guarantees consistent label decoding across training, checkpoint evaluation, and post-processing.

To improve robustness to acquisition variability and ultrasound-specific artifacts, training employs stochastic data augmentation applied consistently to both images and masks. Augmentations include mild geometric perturbations (small rotations, translations, and isotropic scaling), horizontal flips, and intensity-domain perturbations (brightness/contrast and gamma adjustments), together with blur and noise operations to emulate speckle-like degradations. Augmentation is implemented as a per-sample stochastic transform pipeline, ensuring that the mask undergoes the identical spatial transform as the corresponding image.

3.5. Class Imbalance and Frequency-Derived Weights

Pixel-level class frequencies are highly imbalanced, with background dominating most frames and the midline occupying a small fraction of pixels. To improve optimization stability under this imbalance, we apply frequency-derived weights in the cross-entropy component of the training objective (the Dice term remains unweighted).

Let n_c denote the total number of training pixels belonging to class $c \in \{0, \ldots, C-1\}$, and let λ denote a Laplace smoothing constant. We compute smoothed frequencies

$$\tilde{n}_c = n_c + \lambda, \tag{3}$$

$$f_c = \frac{\tilde{n}_c}{\sum_{j=0}^{C-1} \tilde{n}_j}. \tag{4}$$

We then form inverse-frequency weights with a gentle exponent p,

$$u_c = \left(\frac{1}{f_c + \epsilon}\right)^p, \tag{5}$$

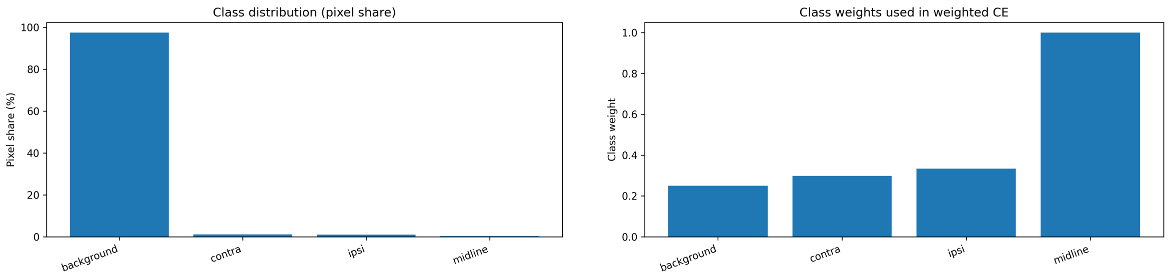
optionally apply modest foreground boost factors b_c (used only to avoid under-emphasizing thin structures), and normalize to unit mean:

$$\hat{w}_c = \frac{u_c\, b_c}{\frac{1}{C}\sum_{j=0}^{C-1} u_j\, b_j}. \tag{6}$$

Finally, weights are clamped to prevent extreme ratios:

$$w_c = \mathrm{clip}(\hat{w}_c,\, w_{\min},\, w_{\max}). \tag{7}$$

In the reported experiments, the effective weights used for weighted cross-entropy were $w = \{0.25, 0.298, 0.334, 1.0\}$ for {background, contralateral ventricle, ipsilateral ventricle, midline}, respectively (Figure 2). This assigns the largest penalty to midline misclassification while keeping background gradients bounded.



(a) Pixel-level class distribution.

(b) Effective class weights used in weighted cross-entropy.

Figure 2. Class imbalance and frequency-derived weighting. The midline is severely underrepresented, motivating increased emphasis relative to background and ventricles.

3.6. Model Families

We evaluate three encoder–decoder model families for four-class semantic segmentation (background, contralateral ventricle, ipsilateral ventricle, and intracranial midline). All models take $768 \times 1024 \times 3$ RGB inputs and output per-pixel class probabilities via a 1×1 convolution followed by a four-way softmax. These families were selected to test whether architectural capacity (residual learning), channel attention (SE), and pretrained representations (EfficientNet) improve delineation of the thin midline and low-contrast ventricle boundaries under severe pixel-level class imbalance.

Baseline U-Net (optimized).

Our primary baseline is a U-Net style architecture with four resolution levels and symmetric decoding. The encoder follows a doubling schedule (64–128–256–512 channels) with a 1024-channel bottleneck, and skip connections concatenate encoder features to the corresponding decoder stage. Upsampling is performed using learned transpose convolutions, followed by convolutional refinement at each scale (Figure 3).

Residual U-Net variants.

To improve optimization and feature reuse under limited data, we evaluate residual U-Net variants that replace the U-Net convolutional blocks with residual blocks. We additionally evaluate an SE-augmented variant that performs channel recalibration after residual fusion.

Transfer-learning encoder (EfficientNet-B4).

Finally, we evaluate a transfer-learning configuration that replaces the encoder with an ImageNet-pretrained EfficientNet-B4 backbone and uses intermediate encoder activations as multiscale skip features for a U-Net style decoder. The decoder progressively upsamples and fuses skip features to recover full-resolution predictions.

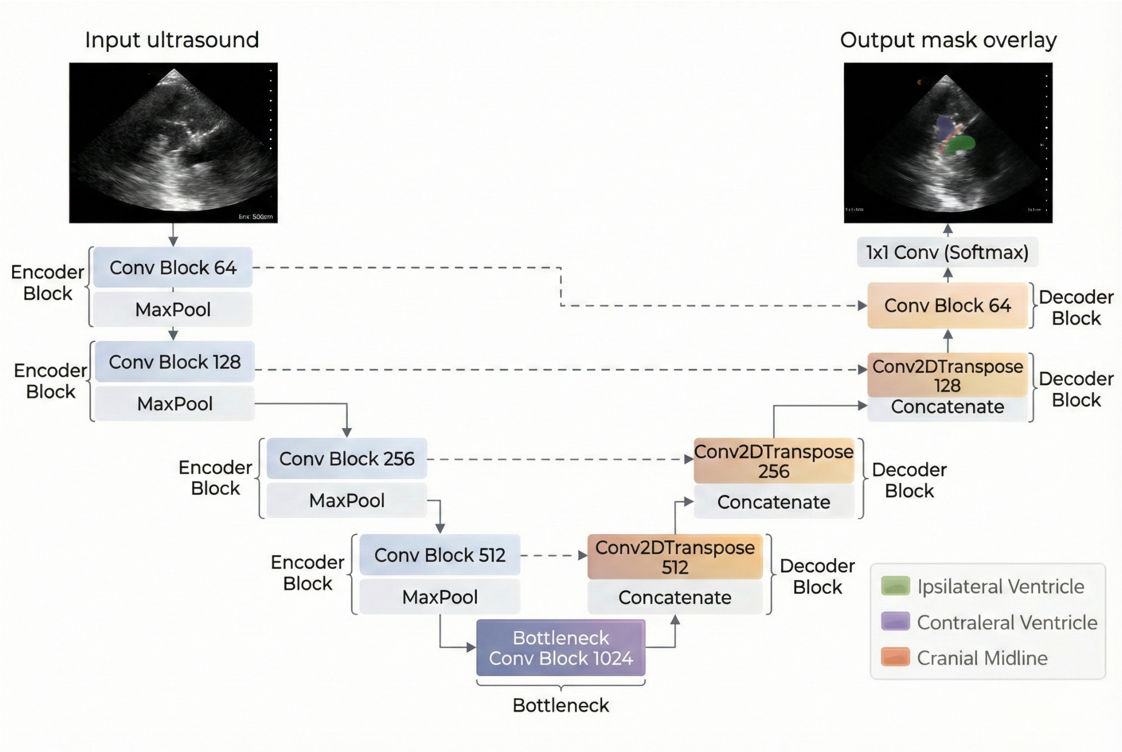


Figure 3. U-Net style encoder–decoder used in this work.

3.7. Composite Loss with Label Smoothing and Midline Focal-Tversky

Let $y(x) \in \{0,1\}^C$ denote the one-hot class label at pixel x (with $C = 4$), and $p(x) \in [0,1]^C$ the corresponding softmax probabilities. To jointly address extreme pixel imbalance (midline) and preserve overlap quality for larger structures (ventricles), training uses a composite objective that combines: (i) class-weighted categorical cross-entropy with label smoothing, (ii) an unweighted multi-class soft Dice loss, and (iii) a small midline-only focal-Tversky term that specifically targets thin-structure errors while remaining a secondary regularizer.

Label smoothing.

We apply label smoothing to reduce overconfidence and improve generalization [11]:

$$y^{\text{smooth}}(x) = (1-\varepsilon)\, y(x) + \frac{\varepsilon}{C}\mathbf{1}, \qquad \varepsilon = 0.03. \tag{8}$$

Class-Weighted Cross-Entropy.

Using per-class weights $\{w_c\}$ (Eq. (7)), per-pixel weights are computed from the smoothed label:

$$w(x) = \sum_{c=0}^{C-1} y_c^{\text{smooth}}(x)\, w_c, \tag{9}$$

and the weighted mean categorical cross-entropy is

$$\mathcal{L}_{\text{WCE}} = \frac{\sum_x w(x)\, \text{CE}\Big(y^{\text{smooth}}(x), p(x)\Big)}{\sum_x w(x) + \epsilon}. \tag{10}$$

Weights are applied only to the cross-entropy component to emphasize minority classes (notably the midline) while avoiding instability from heavily weighted overlap terms.

Multi-class soft Dice loss.

We include an unweighted Dice-based overlap loss [12], computed per class and averaged:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{1}{C}\sum_{c=0}^{C-1} \frac{2\sum_x y_c(x) p_c(x) + s}{\sum_x y_c(x) + \sum_x p_c(x) + s}, \tag{11}$$

where s is a smoothing constant (set to $s{=}1.0$ in our implementation).

Midline focal-Tversky.

To explicitly improve performance on the thin midline, we apply a focal-Tversky loss only to the midline channel [13,14]. Let $y_m(x)$ and $p_m(x)$ denote the midline ground truth and probability. Define

$$\text{TP} = \sum_x y_m(x) p_m(x), \tag{12}$$

$$\text{FP} = \sum_x (1 - y_m(x)) p_m(x), \tag{13}$$

$$\text{FN} = \sum_x y_m(x)(1 - p_m(x)), \tag{14}$$

and the Tversky index

$$T = \frac{\text{TP} + s}{\text{TP} + \alpha\,\text{FP} + \beta\,\text{FN} + s}. \tag{15}$$

The midline focal-Tversky term is

$$\mathcal{L}_{\text{FT}}^{\text{mid}} = (1 - T)^{\gamma}, \qquad \alpha = 0.7,\ \beta = 0.3,\ \gamma = 0.75, \tag{16}$$

with $s{=}1.0$ matching the implementation.

Total loss.

The final training objective is

$$\mathcal{L} = 0.7\,\mathcal{L}_{\text{WCE}} + 0.3\,\mathcal{L}_{\text{Dice}} + 0.05\,\mathcal{L}_{\text{FT}}^{\text{mid}}. \tag{17}$$

3.8. Optimization and Training Protocol

All models are implemented in TensorFlow/Keras and trained with the composite loss described above. For the baseline optimized U-Net runs, we use a batch size of 4 and optimize with AdamW (learning rate 10^{-4}, weight decay 10^{-5}). Training is run for up to 120 epochs, and we select the final checkpoint using the validation foreground macro Dice, computed over the three foreground classes (contralateral ventricle, ipsil ateral ventricle, and midline).

To improve convergence stability, we apply ReduceLROnPlateau on validation macro dice with no background and early stopping on the same metric. ModelCheckpoint saves the best-performing weights according to validation macro dice with no background.

For the optimized U-Net baseline, we additionally perform a short continuation phase: after the initial 120-epoch run, training is resumed for 30 more epochs (epochs 121–150) while preserving the optimizer state, again selecting the best checkpoint by validation foreground macro Dice.

3.9. Inference with Test-Time Augmentation (TTA)

At inference time, we apply a lightweight test-time augmentation consisting of a horizontal flip. For each batch, we concatenate the original inputs with a horizontally flipped copy, run a single forward pass, invert the flip on the second half of the predicted probability maps, and then average the two predictions:

$$\hat{p}(x) = \tfrac{1}{2}\Big(p(x) + \mathrm{flip}^{-1}(p(\mathrm{flip}(x)))\Big). \tag{18}$$

This two-view averaging reduces prediction variance and produces smoother, more stable probability maps, which is particularly helpful for the thin midline channel. This inference procedure matches the evaluation pipeline used to generate the model comparison table across checkpoints.

3.10. Post-Processing and Validation-Tuned Selection

Raw softmax outputs can contain small spurious foreground islands and fragmented midline responses. We therefore apply a lightweight, class-specific post-processing procedure to the TTA-averaged probability maps prior to evaluation. This is particularly important for the midline, where thin-structure predictions can fragment into discontinuous segments that degrade downstream metrology.

For each foreground class (contralateral ventricle, ipsilateral ventricle, and midline), we (i) threshold the class probability map using a class-specific threshold, (ii) apply morphological closing with a 3×3 structuring element (with class-specific iteration counts), and (iii) remove connected components below a class-specific minimum area. Outside the retained components, the corresponding foreground probabilities are suppressed to zero, while the background probability is left unchanged. Denoting the resulting score maps by $S_c(x)$, the final label map is obtained by

$$\hat{M}(x) = \arg\max_{c \in \{0,1,2,3\}} S_c(x), \tag{19}$$

where $c = 0$ is background and $c \in \{1, 2, 3\}$ are the three foreground classes.

Post-processing hyperparameters are tuned by grid search on the validation set using the same TTA probability outputs used at test time (validation labels only). We use a reduced grid of threshold triples, minimum-area triples, and closing-iteration triples; opening iterations are disabled in all configurations. The selected configuration maximizes validation foreground macro Dice (averaged over the three foreground classes), with midline Dice used as a secondary tie-breaker and foreground macro IoU as a tertiary. Only this validation-selected configuration is applied when reporting test-set results, ensuring that tuning does not use test labels.

4. Metrology Module

Segmentation masks are converted into clinically interpretable distances in millimeters using per-frame pixel spacing (Section 3). The metrology module operates on post-processed discrete masks and estimates three measurements: the maximum perpendicular extent from the midline to the ipsilateral ventricle, the analogous contralateral extent, and the maximum ventricle-to-ventricle span along a line normal to the fitted midline.

4.1. Midline Line Fit

Given a binary midline mask, we extract its foreground contour points and fit a straight line by least squares. The fitted line is represented parametrically by a point (x_0, y_0) and a direction vector (v_x, v_y). We define the corresponding unit normal $\mathbf{n} = (-v_y, v_x)/\|(-v_y, v_x)\|_2$. To restrict evaluation to the portion of the line supported by the midline mask, we compute scalar projection values $t = (x - x_0)v_x + (y - y_0)v_y$ over all contour points and retain the range $[t_{\min}, t_{\max}]$.

4.2. Perpendicular Span Search by Discrete Ray Sampling

For a set of uniformly sampled parameters t over $[t_{\min}, t_{\max}]$, we form midline points $\mathbf{p}(t) = (x_0, y_0) + t(v_x, v_y)$. At each $\mathbf{p}(t)$, we evaluate the line through $\mathbf{p}(t)$ in the normal direction $\mathbf{n}$ by discrete sampling at 1-pixel steps. For a given ventricle mask, the farthest in-mask sample along the appropriate side of the normal line is treated as the outermost extent at that t. The maximum extent across all sampled t values is reported. For the ventricle-to-ventricle span, we require valid hits in both ipsilateral and contralateral masks at the same t and take the distance between the two farthest hits.

Pixel distances are converted to millimeters using the per-frame spacing s (mm/pixel). If the midline fit fails or a required ventricle intersection is absent, the corresponding measurement is not reported for that frame.

4.3. Algorithm

Algorithm 1 summarizes the computation of the maximum ventricle-to-ventricle span; the midline-to-ventricle extents are computed analogously using a single ventricle mask.

Algorithm 1 Maximum perpendicular ventricle-to-ventricle span (discrete sampling)

```
Require: ipsi mask $M_i$, contra mask $M_c$, midline fit $(x_0, y_0, v_x, v_y, t_{\min}, t_{\max})$, pixel spacing $s$ (mm/pixel)
Ensure: $d_{\text{global}}$ (mm) and endpoints $(\mathbf{a}, \mathbf{b})$
$\mathbf{n} \leftarrow (-v_y, v_x)/\|(-v_y, v_x)\|_2$
Sample $t_1, \ldots, t_N$ uniformly in $[t_{\min}, t_{\max}]$ (e.g., $N$=121)
$d_{\max} \leftarrow 0$, endpoints $\leftarrow$ None
for $k = 1$ to $N$ do
    $\mathbf{p} \leftarrow (x_0, y_0) + t_k(v_x, v_y)$
    Sample pixels along $\mathbf{p} + u\mathbf{n}$ at 1-pixel steps for $u \in [-U, U]$
    Find farthest in-mask hit $u_i$ in $M_i$ on the ipsi side; farthest hit $u_c$ in $M_c$ on the contra side
    if both hits exist then
        $d \leftarrow |u_i - u_c|$
        if $d > d_{\max}$ then
            $d_{\max} \leftarrow d$ and store endpoints $(\mathbf{a}, \mathbf{b})$
        end if
    end if
end for
if endpoints is None then
    return None
end if
return $d_{\text{global}} = s \cdot d_{\max}$ and stored endpoints
```

4.4. Visualization Outputs

For qualitative review, the module generates an overlay on the original ultrasound frame showing the fitted midline, ventricle contours, and the selected measurement segments, along with a compact text panel reporting the pixel spacing and computed distances. Representative examples are presented with the metrology results in Section 5.

5. Experiments and Results

5.1. Evaluation Metrics

We report Dice similarity coefficient for each foreground class (contralateral ventricle, ipsilateral ventricle, and midline), as well as the foreground macro Dice defined as the mean Dice across these three classes. Validation foreground macro Dice is used for checkpoint selection during training. Unless otherwise stated, results are computed using the full inference pipeline described in Methods, including test-time augmentation and validation-tuned post-processing.

5.2. Segmentation Performance

Table 1 summarizes representative models from each architectural family. Overall, the optimized U-Net yields the strongest test foreground macro Dice in this dataset setting. Residual and attention-augmented variants remain competitive on ventricle segmentation but show lower midline Dice, reflecting the difficulty of thin-structure delineation under extreme class imbalance. The EfficientNet-

doi:10.20944/preprints202603.2474.v1

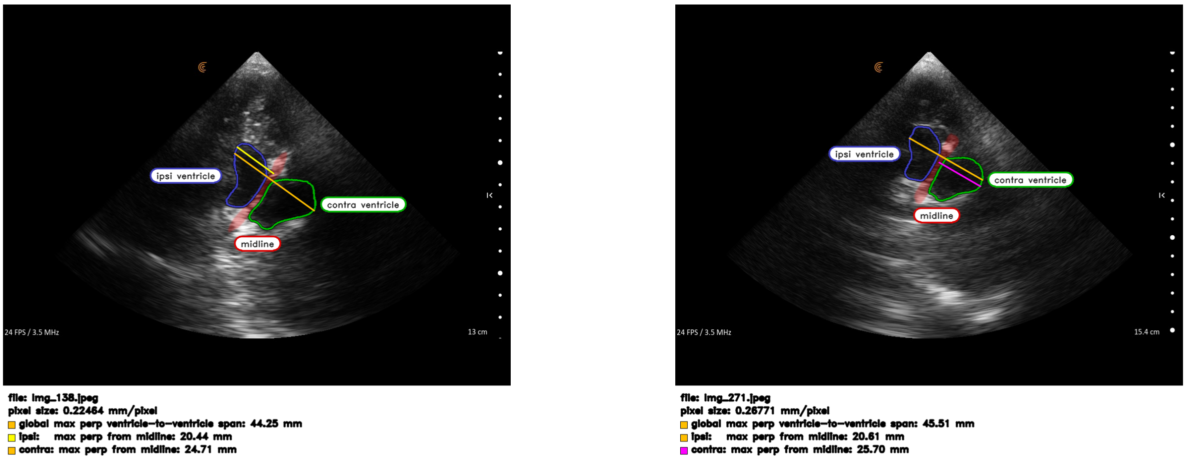
B4 encoder configuration provides reasonable ventricle performance but does not surpass the simpler optimized U-Net in overall foreground accuracy.

Table 1. Representative test performance (with test-time augmentation and validation-tuned post-processing).

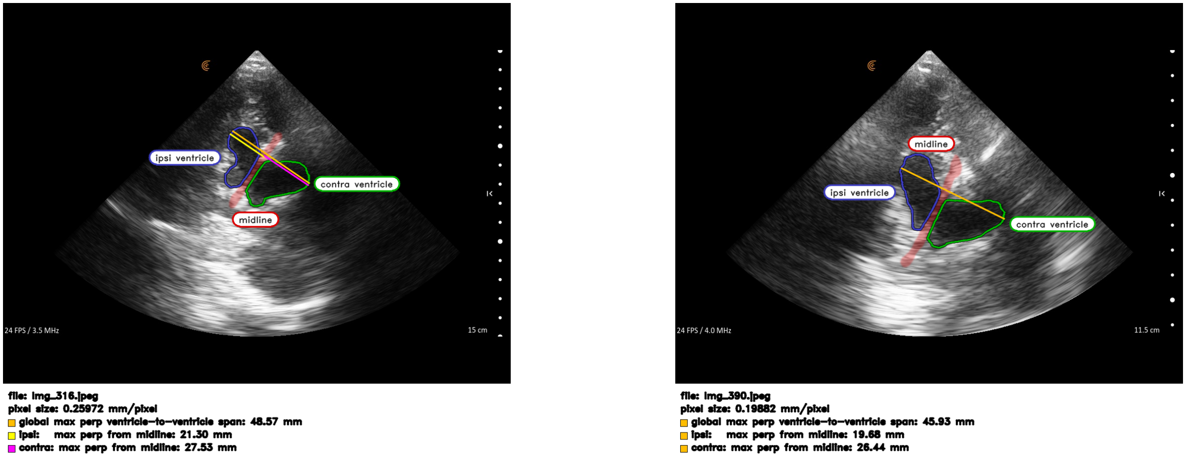
Model family	Contra Dice	Ipsi Dice	Midline Dice	FG macro Dice
Optimized U-Net	**0.928**	**0.924**	**0.756**	**0.869**
SE-ResUNet	0.918	0.904	0.724	0.848
SE-ResUNet + thin-structure loss	0.909	0.884	0.675	0.823
EfficientNet-B4 U-Net	0.895	0.877	0.659	0.810

5.3. Qualitative Results

Figure 4 shows representative overlays produced by the segmentation-to-metrology pipeline. The overlays visualize the fitted midline, ventricle contours, and the selected measurement segments rendered on the original ultrasound frame. These examples illustrate typical behavior across cases, including scenarios where the midline mask exhibits minor fragmentation but the fitted line and perpendicular span estimation remain visually stable.



(a) Representative case. **(b)** Representative case.



(c) Representative case. **(d)** Representative case.

Figure 4. Representative overlays produced by the segmentation-to-metrology pipeline. Colored masks indicate the predicted ventricles and midline; line segments depict the selected maximum perpendicular extents and ventricle-to-ventricle span, converted to millimeters using per-frame pixel spacing.

6. Discussion

This study shows that implant-enabled adult cranial ultrasound can support reliable, fully automated segmentation of the lateral ventricles and the cranial midline in coronal-plane frames when training is explicitly designed around ultrasound noise characteristics and extreme pixel imbalance. Across the evaluated model families, ventricle segmentation was consistently strong, while the midline remained the most demanding target due to its thin morphology and low pixel prevalence. These findings are consistent with the core challenge of the problem: accurately preserving a narrow, low-contrast linear landmark under speckle and acoustic artifacts, without trading off ventricle boundary fidelity.

A key observation from the model comparison is that the optimized U-Net configuration achieved the best overall foreground macro Dice in this dataset setting. This suggests that, for the current scale of labeled data, architectural simplicity paired with careful optimization choices (composite loss design, class weighting, regularization, and validation-driven model selection) can be more effective than increasing encoder complexity. The transfer-learning configuration with an EfficientNet-family encoder did not surpass the optimized U-Net, which may reflect the limited benefit of natural-image pretraining when the input domain exhibits ultrasound-specific texture and speckle statistics.

Beyond segmentation, the metrology module provides a practical bridge from pixel-level predictions to standardized, reviewable measurements. The central geometric design choice is to measure ventricle extents along lines normal to a fitted midline, which yields a consistent definition even when ventricular boundaries are irregular or when the midline mask exhibits small local discontinuities. The resulting overlays and numeric outputs offer an interpretable presentation layer that aligns with how measurements are typically reviewed in practice, while remaining fully derived from model predictions and per-frame pixel calibration.

6.1. Limitations

This work is constrained by the limited size of the labeled dataset and by the fact that neither the data nor broad public benchmarking are available at this stage. The midline label is intrinsically difficult to annotate and learn: it is both thin and visually ambiguous in some frames, and small boundary inconsistencies can disproportionately affect overlap metrics. Additionally, while the metrology module produces consistent geometric outputs from masks and known pixel spacing, a formal accuracy study against an independent measurement reference is not included here; such validation is necessary to quantify clinical measurement error and to characterize failure modes under challenging image conditions (e.g., low contrast or shadowing).

6.2. Future Work

This study should be viewed as a baseline for an emerging imaging setting where relatively little prior segmentation work exists for adult implant-enabled cranial ultrasound. The most direct next step is to expand the labeled dataset to increase diversity in image appearance and to improve robustness, particularly for midline continuity. A second direction is to broaden the pipeline beyond coronal frames to additional acquisition planes commonly used in cranial ultrasound practice (e.g., sagittal and transverse/axial views), which would require plane-aware training and evaluation. Third, the segmentation targets can be extended to additional clinically relevant structures and landmarks within the same framework, enabling richer quantitative summaries from a single scan.

Finally, the pipeline can be advanced toward a unified real-time framework that provides on-device or near-real-time segmentation overlays during acquisition, alongside automated measurement readouts. This would require system-level engineering (latency profiling on target hardware, model compression where appropriate) and careful validation to ensure stable behavior across a wide range of acquisition conditions.

doi:10.20944/preprints202603.2474.v1

7. Conclusion

We developed a proprietary deep learning pipeline at **Longeviti Neuro Solutions** for adult implant-enabled cranial ultrasound that segments the ipsilateral ventricle, contralateral ventricle, and cranial midline in coronal-plane frames and translates these masks into calibrated, millimeter-scale measurements. On a held-out test set, the best-performing optimized U-Net configuration achieved a foreground macro Dice of 0.869, with strong ventricle Dice values and midline Dice around 0.75 despite severe class imbalance and thin-structure challenges. The accompanying metrology module renders interpretable overlays and standardized distance estimates directly on the ultrasound image, supporting quantitative review and providing a practical path from segmentation to clinically meaningful outputs. This work establishes a foundation for future expansion to larger datasets, additional planes, and real-time integrated inference overlays for implant-enabled adult cranial ultrasound.

Author Contributions: Jaswant Vemulapalli: model development, training and evaluation, metrology implementation, analysis, and manuscript drafting. Nicholus Vaughan: dataset curation, coordination of the annotation workflow, annotation dataset export and organization, and manuscript review.

Institutional Review Board Statement: Images were exported in de-identified form for research and engineering purposes. No personally identifiable information was included in the working dataset.

Data Availability Statement: The ultrasound dataset used in this study is proprietary to the data owner and contains de-identified clinical images; it cannot be released publicly. The codebase used for model training, inference, and metrology was developed within the company environment and is not publicly available. Representative de-identified examples are included for illustration only.

Acknowledgments: The authors acknowledge Neil Christopher, Longeviti VP of Operations and Advanced Programs, for conceiving and sponsoring this initiative; the Longeviti annotation team for dataset labeling; and the expert clinical reviewers (sonographers) for their systematic quality review and refinement of the annotations. We thank our clinical collaborators, including neurosurgeons and staff, for their domain guidance on relevant anatomical landmarks and clinically interpretable presentation of measurements.

Conflicts of Interest: The authors were employed by Longeviti Neuro Solutions during the development of this work.

References

1. Belzberg, M.; Ben Shalom, N.; Lu, A.; Yuhanna, E.; Manbachi, A.; Tekes, A.; Huang, J.; Brem, H.; Gordon, C. Transcranioplasty Ultrasound Through a Sonolucent Cranial Implant Made of Polymethyl Methacrylate: Phantom Study Comparing Ultrasound, Computed Tomography, and Magnetic Resonance Imaging. *The Journal of Craniofacial Surgery* **2019**, *30*, e626–e629. https://doi.org/10.1097/SCS.0000000000005651.
2. Belzberg, M.; Ben Shalom, N.; Yuhanna, E.; Manbachi, A.; Tekes, A.; Huang, J.; Brem, H.; Gordon, C.R. Sonolucent Cranial Implants: Cadaveric Study and Clinical Findings Supporting Diagnostic and Therapeutic Transcranioplasty Ultrasound. *The Journal of Craniofacial Surgery* **2019**, *30*, 1456–1461. https://doi.org/10.1097/SCS.0000000000005454.
3. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer, 2015, Vol. 9351, *Lecture Notes in Computer Science*, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
4. Tabrizi, P.R.; Obeid, R.; Cerrolaza, J.J.; Penn, A.; Mansoor, A.; Linguraru, M.G. Automatic Segmentation of Neonatal Ventricles from Cranial Ultrasound for Prediction of Intraventricular Hemorrhage Outcome. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, July 2018, pp. 3136–3139. https://doi.org/10.1109/EMBC.2018.8513097.
5. Szentimrey, Z.; de Ribaupierre, S.; Fenster, A.; Ukwatta, E. Automated 3D U-Net based segmentation of neonatal cerebral ventricles from 3D ultrasound images. *Medical Physics* **2022**, *49*, 1034–1046. https://doi.org/10.1002/mp.15432.

6. Chen, H.; Wu, L.; Dou, Q.; Qin, J.; Li, S.; Cheng, J.Z.; Ni, D.; Heng, P.A. Automatic Measurements of Fetal Lateral Ventricles in Ultrasound Images for Ventriculomegaly Screening. *Frontiers in Physiology* **2020**, *11*, 794. https://doi.org/10.3389/fphys.2020.00794.
7. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2016, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.
8. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2018, pp. 7132–7141. https://doi.org/10.1109/CVPR.2018.00745.
9. Tan, M.; Le, Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In Proceedings of the Proceedings of the 36th International Conference on Machine Learning; Chaudhuri, K.; Salakhutdinov, R., Eds. PMLR, June 2019, Vol. 97, *Proceedings of Machine Learning Research*, pp. 6105–6114.
10. Shit, S.; Paetzold, J.C.; Sekuboyina, A.; Ezhov, I.; Unger, A.; Zhylka, A.; Pluim, J.P.W.; Bauer, U.; Menze, B.H. clDice – A Novel Topology-Preserving Loss Function for Tubular Structure Segmentation. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2021, pp. 16560–16569.
11. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2016, pp. 2818–2826. https://doi.org/10.1109/CVPR.2016.308.
12. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV). IEEE, Oct 2016, pp. 565–571. https://doi.org/10.1109/3DV.2016.79.
13. Salehi, S.S.M.; Erdogmus, D.; Gholipour, A. Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks. In *Machine Learning in Medical Imaging*; Springer: Cham, 2017; Vol. 10541, *Lecture Notes in Computer Science*, pp. 379–387. https://doi.org/10.1007/978-3-319-67389-9_44.
14. Abraham, N.; Khan, N.M. A Novel Focal Tversky Loss Function With Improved Attention U-Net for Lesion Segmentation. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, April 2019, pp. 683–687. https://doi.org/10.1109/ISBI.2019.8759329.