

Additional File 4: Methods for Find-Confirm-Measure training, inference, and exploratory aggregation settings

This supplement consolidates the training settings, inference operating points, preprocessing constants, model-selection procedures, and exploratory aggregation parameters used in ATLAS.

Software environment	Training notebook: Ultralytics 8.4.90; Python 3.12.13; PyTorch 2.11.0+cu128. opencv-python-headless was installed, but its exact version is not retained in the final notebooks.
Runtime hardware	Training: NVIDIA A100-SXM4-80GB (Notebook 02). Separate manuscript inference profiling used RTX 2000 Ada Generation.

Terminology. Final configuration values are not automatically described as “tuned hyperparameters.” This supplement distinguishes fixed design settings, framework defaults, validation-selected settings, staged/manual development choices, exploratory sensitivity parameters, and evaluation-only thresholds.

Configuration taxonomy and provenance classes

Provenance class	Definition	ATLAS example
Fixed design setting	Specified as part of the implementation before locked-test/external evaluation.	Input size, preprocessing constants, crop padding, pixel-spacing guard.
Framework-selected/default	Chosen by the software framework or inherited default rather than manually searched.	Find optimizer recorded as optimizer=auto.
Validation-selected	Candidate model/checkpoint or setting chosen using development validation data only.	Confirm checkpoint selection; final source model selection where applicable.
Staged/manual development choice	Configuration altered across a predefined or iterative development curriculum without a formal search grid.	Confirm S0-S5 curriculum; 640/800 px stages.
Inference operating point	Fixed score threshold used to emit/withhold a computational output.	tau_F=0.25; tau_C=0.25.
Exploratory sensitivity parameter	Varied after the primary technical model specification for secondary/exploratory analysis.	TIES density 0.20 primary; 0.05-0.40 sensitivity.
Evaluation-only criterion	Used to score outputs rather than train or operate the model.	BPD tolerance 2.0 mm; HC tolerance 5.0%.

Locked preprocessing and shared pipeline constants

Setting	Value	Provenance	Interpretation
Bilateral filter diameter	5	Fixed preprocessing	Speckle attenuation while preserving boundaries.
Bilateral sigma_color	50	Fixed preprocessing	Applied uniformly under the locked preprocessing pipeline.
Bilateral sigma_space	50	Fixed preprocessing	Applied uniformly under the locked preprocessing pipeline.

Setting	Value	Provenance	Interpretation
CLAHE clip limit	2.0	Fixed preprocessing	Local contrast normalization.
CLAHE tile grid	8 x 8	Fixed preprocessing	Local contrast normalization.
Confirm crop padding	0.40 (40%)	Fixed pipeline geometry	Find crop expanded before Confirm processing.
Valid pixel-spacing guard	0.04-0.40 mm/px	Fixed data-quality rule	Spacing outside this range is not treated as valid image-specific metadata.
Pixel-spacing fallback	0.125 mm/px	Training-derived fixed fallback	Median valid spacing from the training partition only; fixed before validation, locked test, and external evaluation.
Random seed	42	Fixed reproducibility setting	Seed 42 was set globally and is recorded in the final Notebook 02 trainer logs for Find, Confirm stages, and Measure.

Find: final training and inference configuration

Parameter	Final value	Provenance class	Notes
Architecture / model family	YOLO26n detector	Fixed model choice	Cranial localization.
Input size	640 px	Fixed training/inference setting	Locked source detector input.
Batch size	16	Training setting	Reported against the prespecified BMC training contract and retained trainer log.
Maximum epochs	50	Training setting	Early stopping may terminate earlier.
Early-stopping patience	10	Training setting	Validation-driven stopping control.
Learning-rate schedule	Cosine decay; configured lr0=0.001; lrf=0.01	Training setting	optimizer=auto resolved to AdamW with lr=0.002 and momentum=0.9; weight decay=0.0005 and warmup=5 epochs. The configured lr0 was superseded by optimizer=auto.
Optimizer	optimizer=auto -> AdamW (resolved lr=0.002, momentum=0.9)	Framework-selected	Configured lr0=0.001 was ignored by Ultralytics because optimizer=auto; lrf=0.01, cosine schedule, weight decay=0.0005, warmup=5 epochs.
Mosaic augmentation	0.5	Training augmentation	Training only.
Flip augmentation	fliplr=0.5; flipud=0.5	Training augmentation	Training augmentation; selected BMC-locked Find run also used mosaic=0.5, scale=0.5 and translate=0.1.
Scale augmentation	+/-50%	Training augmentation	Training only.

Parameter	Final value	Provenance class	Notes
Inference confidence threshold, tau_F	0.25	Fixed inference operating point	Computational detection filter; not a clinical plane-adequacy threshold and not a calibrated probability.
Checkpoint selection metric	Best validation-loss checkpoint within the prespecified BMC-locked run	Validation checkpointing	A separate Find ablation ranked OPT_A/OPT_B/OPT_C by validation mAP50(B), then mAP50-95(B), then fitness; these optimized artifacts were not substituted for the manuscript BMC-locked model.
Selected checkpoint identity	BrainLocator_YOLO26n_MuSGD/weights/best.pt	Model artifact	SHA-256: bcc55db31feaf617d1643185c89c743b1450ef5de4b9dff2e422702a66ff14cd

Confirm: staged anatomical-evidence model configuration

Parameter	Final value	Provenance class
Task	CSP/LV evidence detection	Fixed task definition
Input region	Find crop + 40% padding	Fixed pipeline setting
Training curriculum	S0-S5 staged curriculum	Staged/manual development choice
Training resolutions	S0 640; S1_CSP 640; S1_LV 640; S2 640; S3_CSP 800; S3_LV 800; S4 640; S5 800 px	Staged/manual development choice
Pseudo-reference assistance	Offline SAM2-Small during data construction	Fixed development method
Stage-specific epochs	S0 20; S1_CSP 15; S1_LV 15; S2 40; S3_CSP 30; S3_LV 30; S4 20; S5 15 epochs	Training setting
Stage-specific batch sizes	S0 32; S1_CSP 32; S1_LV 32; S2 32; S3_CSP 16; S3_LV 16; S4 24; S5 12	Training setting
Loss terms / weights	YOLO box/cls/df1 weights: S0 7.5/0.5/3.0; S1 6.0/4.0/1.5; S2 7.5/0.5/1.5; S3 9.5/0.5/2.0; S4 7.5/0.5/1.5; S5 10.0/0.5/2.5	Training setting
Optimizer / learning-rate schedule	optimizer=auto -> AdamW (logged lr=0.001667, momentum=0.9); cos_lr=True; lrf=0.01; warmup=5 epochs	Training setting
Inference confidence threshold, tau_C	0.25	Fixed inference operating point

Parameter	Final value	Provenance class
Checkpoint selection	Final S5 framework-generated best.pt checkpoint	Validation checkpointing within staged curriculum
Selected checkpoint identity / class map	UniDet_GrandMaster_YOLO26_S5_PixelPerfect/weights/best.pt; class map 0=CSP, 1=LV	Model artifact

Measure: segmentation-derived geometry and deterministic comparator

Parameter / setting	Final value	Provenance class	Notes
Primary learned Measure pathway	Segmentation-derived fetal-head geometry	Model choice	Used to derive BPD/OFD/HC from head geometry plus physical scale.
Comparator	Deterministic Find-box geometry	Prespecified comparator	No learned regression; retained as a separate measurement route.
Segmentation architecture	YOLO26n-seg (single class: HeadEllipse)	Model specification	Run: MeasureHeadEllipse_YOLO26n_Seg; derived ellipse-polygon masks contain 72 polygon points.
Segmentation input size	640 px	Training/inference setting	Reported from the final Measure training call and retained trainer log.
Batch size / epochs / patience	Batch 16; 30 epochs; patience 100	Training setting	Reported from the final Measure trainer log; no separate formal hyperparameter search was documented.
Optimizer / LR schedule	optimizer=auto -> AdamW (lr=0.002, momentum=0.9); cosine schedule; lrf=0.01; warmup=3 epochs	Training setting	Weight decay=0.0005. Configured lr0=0.01 was ignored because optimizer=auto resolved optimizer/LR.
Loss function(s) / weights	Logged framework weights: box=7.5, cls=0.5, dfl=1.5; training reports box, segmentation, classification, DFL and semantic loss terms	Training setting	No custom segmentation-specific loss weights were supplied in the Measure training call; additional framework-internal segmentation/semantic weights are not exposed in the retained notebook log and are not inferred here.
Augmentation	mosaic=1.0; flipplr=0.5; flipud=0.0; scale=0.5; translate=0.1; HSV h/s/v=0.015/0.7/0.4; erasing=0.4	Training augmentation	degrees=0, shear=0, perspective=0, mixup=0 and copy-paste=0 in the retained run.
Geometry conversion	Ellipse/head geometry -> BPD/OFD/HC	Deterministic post-processing	Not a tuned model hyperparameter; mathematical derivation is fixed.

Parameter / setting	Final value	Provenance class	Notes
Physical scale	Metadata when valid; otherwise 0.125 mm/px training-derived fallback	Fixed measurement rule	Scale provenance retained per frame.
Selected checkpoint identity	MeasureHeadEllipse_YOLO26n_Seg/weights/best.pt	Model artifact	SHA-256: 3b6d2932d9c6370f979ca4bfd1d88cffe2b39662c8fb1be89b0023d0e21c3e
BPD tolerance for joint endpoint	2.0 mm	Evaluation-only criterion	Not used for training, inference gating, diagnosis, or calibration.
HC tolerance for joint endpoint	5.0%	Evaluation-only criterion	Not used for training, inference gating, diagnosis, or calibration.

Site-specific adaptation and data-free parameter aggregation

Setting	Value	Provenance class
Site-specific models	Separate FP- and UCL-specialist Find models	Secondary technical analysis
Raw-data pooling	Not used for parameter aggregation step	Design property of aggregation analysis
Aggregation method	TIES parameter merging	Exploratory method
Primary merge density	0.20	Exploratory operating setting
Sensitivity range	0.05-0.40	Exploratory sensitivity parameter
Other TIES settings	Trim by per-domain absolute-magnitude quantile at density d; elect consensus sign as $\text{sign}(\text{sum}(\text{sign}(\text{trimmed})))$; discard sign-disagreeing updates; merge surviving updates by elementwise mean; add merged task vector to base. No extra scaling coefficient configured. Primary d=0.20; sensitivity d=0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.40.	Exploratory setting
Selection/use of aggregated model	Exploratory only; did not replace locked primary source model	Analysis boundary

Model selection, threshold locking, and test-set separation

Control	ATLAS specification	Status / timing
Development/validation split	Patient-exclusive training and validation partitions	Model-development control
Locked internal test	586 frames / 279 patients	Evaluation-only
External FP/UCL	Subject-exclusive external evaluation	Evaluation-only
Find threshold τ_F	0.25	Locked before test evaluation
Confirm threshold τ_C	0.25	Locked computational evidence threshold

Control	ATLAS specification	Status / timing
Reference information	Contours/landmarks/biometric reference values unavailable during locked inference	Evaluation separation
Model updating	Site adaptation/TIES analysed as secondary models	Secondary/exploratory

Hyperparameter-search and optimization

Component	Search/tuning strategy	Optimization / selection metric	Reported candidate values
Find	Prespecified BMC-locked run; separate three-candidate validation ablation (OPT_A/OPT_B/OPT_C) reported as secondary	Primary run: best validation-loss checkpoint; ablation ranking: mAP50(B) -> mAP50-95(B) -> fitness	640 px, batch 16, 50 epochs, patience 10, mosaic/flip/scale settings, tau_F=0.25; secondary ablation varied resolution, epochs/patience, mosaic, flipud, scale and configured lr0
Confirm	Predefined/manual S0-S5 staged curriculum; no grid/random/Bayesian search	Stage-wise framework best.pt validation checkpoint; no separate custom selection metric retained in Notebook 02	Exact S0-S5 epochs, batch sizes, resolutions, loss weights and optimizer schedule reported in S4; tau_C=0.25
Measure	Single fixed YOLO26n-seg training configuration; no formal hyperparameter search documented	Framework-generated best.pt from validation during the fixed training run; no separate ATLAS selection metric retained	640 px; batch 16; 30 epochs; patience 100; optimizer=auto -> AdamW lr=0.002; cosine schedule; augmentation as reported in S5
TIES	Exploratory sensitivity analysis	Downstream technical performance	Density 0.20 primary; 0.05-0.40 sensitivity