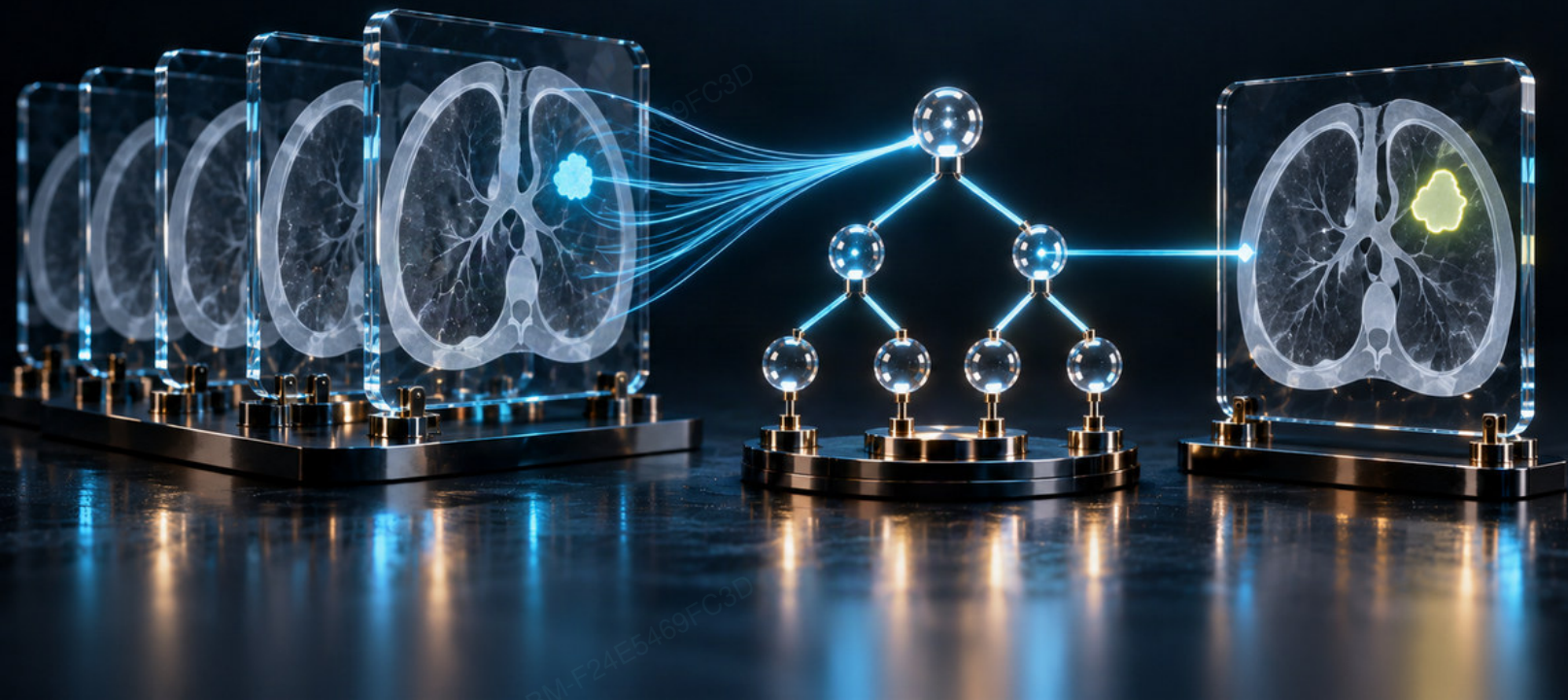




RESEARCH MANUSCRIPT | OCTOBER 8, 2026

Compact Decision Programs over Frozen Neural Features for Lesion Segmentation

Brian Mingus

Readable decision programs over frozen neural features: patient-grouped lesion segmentation, compact XGBoost conversion, and complete binary trees with every decision exposed.

Collection cover added October 8, 2026. All 3 manuscript pages follow unchanged.

AI-generated conceptual artwork; separate from the scientific figures.

brianmingus.com

Compact Decision Programs over Frozen Neural Features for Lesion Segmentation

Brian Mingus¹

Frozen neural features support compact decision programs for location-assisted lesion segmentation. On 32 RIDER patients and 62 scans, patient-grouped evaluation yields mean Dice of **0.765**, versus **0.734** for the pretrained mask and **0.758** for an earlier hybrid procedure. Current and earlier all-patient refits yield complete binary trees with 181 and 21 nodes, respectively. Both trees are shown in full. The neural feature pipeline remains necessary; these exploratory results demonstrate adaptable, inspectable downstream decisions.

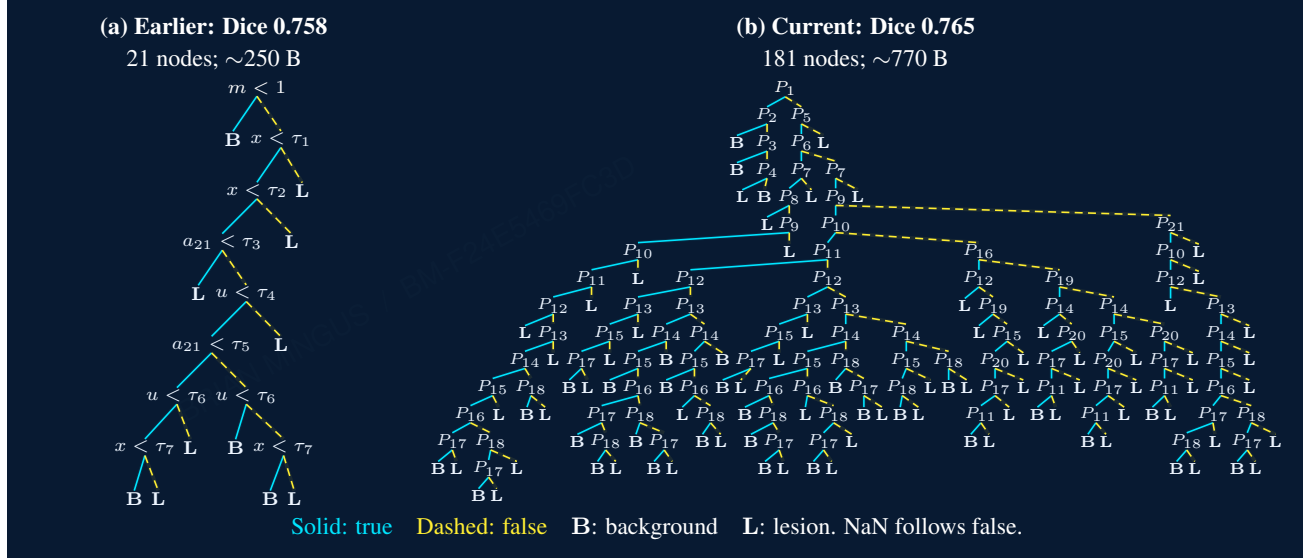


Figure 1. Both complete binary trees, with no collapsed branches. The labels give each procedure’s patient-grouped score; these separately selected all-patient refits have no additional held-out score. Repeated P_j labels denote the same predicate at distinct positions. Every test and terminal decision is shown. The exact predicates are listed on page 2. Training weights and cutoff searches differ between procedures, so the accuracy gain is not attributed to tree size alone.

1. Accuracy and Complete Decision Trees

We pool each patient’s scan counts, calculate $D_g = 2TP_g / (2TP_g + FP_g + FN_g)$, and average over patients. The current procedure improves 28 patients, ties one, and lowers three versus the fixed mask: mean gain 0.031, with descriptive 95% patient-bootstrap interval [0.010, 0.052] (10,000 paired-patient resamples of saved predictions; unadjusted for repeated selection).

The 0.007 gain over the earlier procedure accompanies larger final decision logic (Fig. 1). Fresh all-patient selection chooses depth 3, 16 rounds and FP32 cutoff 0.325. Its 32 source trees contain 480 nodes; conversion merges

Table 1. Cross-validated procedure scores and sizes of separately selected all-patient refits. The pictured trees have no additional held-out score. Storage excludes the CNN.

Procedure	Dice	Tree nodes	Storage
Fixed CNN mask	0.734	—	—
Earlier XGBoost	0.758	21	~250 B
Current XGBoost	0.765	181	~770 B

identical continuations into 63 shared nodes and 768 compact bytes in about half a second. The complete expanded tree has 90 tests and 91 terminal nodes. Construction checks pass; optional independent exhaustive qualification was not run. Native XGBoost 3.4.1 and C++23/CUDA complete evaluation and refitting in about 12 minutes after feature preparation.

¹Cognatory, LLC, Denver, Colorado, USA. Correspondence to: Brian Mingus <brian@cognatory.com>.

2. Patient-Grouped Learning

We combine frozen LesionLocator features (Rokuss et al., 2025) with XGBoost (Chen & Guestrin, 2016). RIDER Lung CT version 3 (Zhao et al., 2015; 2024) supplies 32 patients: 30 repeat pairs and two single scans. Independent spreadsheet locations define an 80^3 grid at 2-mm spacing. Two intended acquisitions lack usable location mapping and are withheld without contour-based correction. Missed foreground outside the region counts as false negatives; one patient’s zero score is retained.

Each voxel has 54 inputs: 20 CT/location features, the pretrained mask, foreground probability, and 32 decoder activations. Probability uses the released five-model/eight-view ensemble; activations use fold zero’s unmirrored pass. No neural weights are fitted locally. Unavailable continuous neural support is zero-filled.

Five outer folds hold out patients; five inner folds select the model and decision rule. Repeat scans always stay together. Eight weighted-cross-entropy candidates cross depths $\{1, 2, 3, 5\}$ with $\{16, 64\}$ rounds ($\eta = 0.1$, $L_2 = 1$, 128 bins). Rules include argmax and 161 FP32 probability cutoffs. A second XGBoost stage offers four combiners, trained on patient-excluded predictions using three further folds.

Fitting retains all foreground and samples background at stride 32. The current procedure weights retained class- c rows from patient g by

$$w_{gc} = \frac{n}{G} \frac{N_{gc}}{N_g n_{gc}}, \quad (1)$$

where G, n count fitting patients and retained rows, and N_g, N_{gc}, n_{gc} are full patient, full class, and retained class counts. This preserves patient prevalence with equal patient mass. The earlier procedure uses global inverse-retention weights and 19 cutoffs. Within outer evaluation, selection and fitting weights exclude held-out patients.

3. Reading the Complete Trees

Each solid branch means the test is true; each dashed branch means false. Terminal B and L denote background and lesion. Repeated P_j labels identify the same test at different positions, not a collapsed branch.

Here x is normalized CT intensity; u, s are local mean (2-mm radius) and standard deviation (4-mm radius); v is intensity at the supplied point; m, p are CNN mask and foreground probability; a_k is decoder activation k .

The constants below round-trip to stored FP32 thresholds. NaN follows the dashed branch. The node labels and line styles preserve meaning independently of color.

Current predicates

$P_1 : m < 1$ $P_2 : p < 0.0192318726$
 $P_3 : x < -0.537887037$ $P_4 : v < 0.147721395$
 $P_5 : x < -0.266500354$ $P_6 : a_{28} < -0.0118003832$
 $P_7 : x < -0.419816613$ $P_8 : a_{28} < -0.012050854$
 $P_9 : x < -1.07635784$ $P_{10} : u < -0.660179913$
 $P_{11} : x < -1.4028399$ $P_{12} : a_{21} < -0.0135068875$
 $P_{13} : s < 1.26291883$ $P_{14} : a_{21} < -0.0102521647$
 $P_{15} : v < 0.207296506$ $P_{16} : a_{21} < -0.0117333885$
 $P_{17} : s < 1.18441689$ $P_{18} : u < -0.969308496$
 $P_{19} : a_{22} < -0.0138405655$ $P_{20} : u < 0.195440516$
 $P_{21} : x < -0.76002878$

Earlier thresholds

$\tau_1 = -0.407344371$ $\tau_2 = -0.711924314$
 $\tau_3 = -0.0117178345$ $\tau_4 = -0.635912418$
 $\tau_5 = -0.0102527272$ $\tau_6 = -0.785341203$
 $\tau_7 = -1.02144098$

4. Scope and Availability

The cohort informed earlier development; nested grouping does not make repeated use confirmatory (Cawley & Talbot, 2010). External checkpoint training overlap remains unverified. Independent patient evaluation is needed. The complete trees expose both procedures’ decision logic while retaining the upstream CNN. Code, models and paper: github.com/mingusb/rider-lesion-segmentation.

References

- Cawley, G. C. and Talbot, N. L. C. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11:2079–2107, 2010. <https://www.jmlr.org/papers/v11/cawley10a.html>.
- Chen, T. and Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016. doi: 10.1145/2939672.2939785.
- Rokuss, M., Kirchhoff, Y., Akbal, S., Kovacs, B., Roy, S., Ulrich, C., Wald, T., Rotkopf, L. T., Schlemmer, H.-P., and Maier-Hein, K. LesionLocator: Zero-shot universal tumor segmentation and tracking in 3D whole-body imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 30872–30885, 2025.
- Zhao, B., Schwartz, L. H., Kris, M. G., and Riely, G. J. Coffee-break lung CT collection with scan images reconstructed at multiple imaging parameters. *The Cancer Imaging Archive*, 2015. Dataset, version 3, updated June 25, 2024. doi: 10.7937/K9/TCIA.2015.U1X8A5NR.
- Zhao, B., Dercle, L., Yang, H., Riely, G. J., Kris, M. G., and Schwartz, L. H. Annotated test-retest dataset of lung cancer CT scan images reconstructed at multiple imaging parameters. *Scientific Data*, 11:1259, 2024. doi: 10.1038/s41597-024-04085-3.