CALHippo: Cell Segmentation for Neuronal Density Inference in the Human Hippocampus

Giovanni Casari*, Ettore Candeloro*, Daniela Gandolfi, Jonathan Mapelli, Federico Bolelli✉, and Costantino Grana

University of Modena and Reggio Emilia, Italy
{name.surname}@unimore.it

Abstract. Reliable estimates of cellular composition and anatomical distribution in the human brain are essential for biologically plausible circuit models. In the hippocampus, existing reconstructions rely on low-resolution (LR) data without explicit cell-type-resolved annotations, limiting quantitative maps of excitatory neurons, inhibitory interneurons, and glial cells. Using newly released 1 μm/px BigBrain sections of the right hippocampus, we present CALHippo, Cellular Annotation Library for the Hippocampus, a multiscale resource for cell-type-resolved reconstruction of the human CA complex. CALHippo includes the first expert-validated, cell-level annotated dataset spanning all Cornu Ammonis (CA1-CA4) subfields with explicit three-class labels, together with a lower-resolution mesoscale cellular point-cloud map. High-resolution (HR) cell instances are obtained through a human-in-the-loop pipeline combining foundation-model-based segmentation, iterative expert correction, and model ensembling, and are classified as excitatory neurons, inhibitory interneurons, or glial cells. To extend sparse HR annotations to the full volume, we project them into the 20 μm/px LR BigBrain space and use the resulting class-specific supervision maps to train a UNet-based density estimation model. The predicted density maps enable slice-by-slice inference across the full CA complex and are sampled to generate a class-resolved mesoscale cellular point cloud. Code[1] and dataset[2] are publicly released to support reproducibility.

Keywords: Human Brain · Cell Segmentation · Density Estimation

1 Introduction

Understanding how microscopic cellular organization gives rise to large-scale brain dynamics requires anatomically grounded multiscale models. While whole-brain simulation frameworks have advanced [18], biologically detailed human

[1] https://github.com/AImageLab-zip/CALHippo-Framework
[2] https://ditto.ing.unimore.it/calhippo
* Equal contribution. Authors are allowed to list their names first on their CVs.
✉ Corresponding author: federico.bolelli@unimore.it

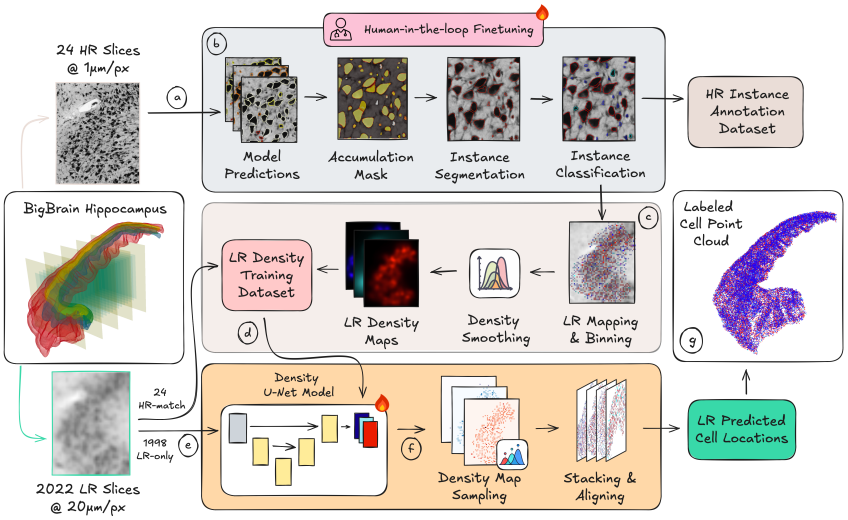

Fig. 1: CALHippo framework starts from high-resolution data extraction (a), instance segmentation with human-in-the-loop fine-tuning (b), and annotation dataset assembly (c). The density model is trained on matched high- and low-resolution slices (d), and subsequently applied to all low-resolution slices (e) to infer density maps (f), enabling the reconstruction of hippocampal volume (g).

cellular scaffolds remain limited by scarce high-resolution anatomical data. To the best of our knowledge, the only BigBrain-based [1] cellular hippocampus scaffold to date is Gandolfi *et al.* [8], which targeted hippocampal CA1 and used 20 μm/px sections with intensity-based density inference, providing excitatory/inhibitory estimates without explicit glial modeling. Recently, selected BigBrain sections rescanned at 1 μm/px [19], previously leveraged for super resolution [5], enable direct soma visualization and, in principle, instance-level segmentation with class discrimination, but the lack of ground-truth annotations prevents supervised training and quantitative validation. To address this gap, we develop CALHippo, Cellular Annotation Library for the Hippocampus, a publicly available multiscale framework for cell-type-resolved reconstruction of the right human hippocampal CA complex. CALHippo combines human-in-the-loop instance segmentation and classification in 1 μm/px sections with supervised density inference in the 20 μm/px BigBrain space. Expert-validated high-resolution annotations are projected to low resolution to train a UNet-based model that predicts continuous, cell-type-specific density maps, from which we sample discrete spatial cell realizations to reconstruct a mesoscale, class-resolved point cloud of the full CA complex.

Key Contributions. We contribute CALHippo, comprising: (i) the first expert-validated, three-class (excitatory, inhibitory, glial) instance segmentation dataset for human hippocampal CA1-CA4 at 1 μm/px; (ii) a reusable HR-to-LR pipeline

that leverages HR cell annotations to supervise cell-type density prediction across LR sections; and (iii) a whole-CA, class-resolved 3D cellular point-cloud reconstruction of the mesoscale hippocampus.

2 Methods

2.1 High-Resolution Dataset Curation

Our work builds on two complementary BigBrain releases. The primary volumetric reference [1] comprises whole-brain coronal Merker silver-stained sections at 20 μm/px, forming an isotropic 3D reconstruction. From this volume, selected sections were recently re-scanned at 1 μm/px [19], providing a 20$\times$ in-plane resolution increase that enables direct visualization of cellular somata and supports instance-level, cell-type-aware segmentation. For both resolutions, per-section affine transforms are provided to align high- and low-resolution images. We restrict the analysis to the right hippocampal regions, using 2D ROI masks provided by [7] for CA1-CA4 regions, yielding 2,022 LR and 24 HR coronal sections (Fig. 1a). Because no ground-truth annotations exist, we develop a human-in-the-loop pipeline to produce validated instance masks and three-class labels. Two WSIs were independently segmented and classified by two expert annotators to quantify inter-annotator variability and were reserved as a held-out test set for the tasks of segmentation and classification (Fig. 1b).

Human-in-the-Loop Segmentation. Initial instance proposals were generated using a non-domain-adapted version of CellPose-SAM [15], selected after comparison with StarDist [20], HoverNet [11], and InstanSeg [10], for its zero-shot performance and human-in-the-loop dedicated fine-tuning pipeline. Expert annotators systematically corrected recurrent failure modes, including boundary inaccuracies, under-segmentation of adjacent cells, false positives in low-contrast regions, and misidentification of small glial cells. Corrected masks were iteratively incorporated into retraining cycles, progressively improving segmentation quality across architectures and confirming that supervision refinement, rather than model-specific tuning, was the primary driver of performance gains. The final expert-corrected annotation set was then used to train additional segmentation models, namely StarDist, HoverNet, and InstanSeg.

Model Ensembling & ATM. To increase robustness in morphologically heterogeneous regions, we ensembled independently trained instances of the above-mentioned models. The ensemble was complemented by a custom Adaptive Thresholding Method (ATM) that extracts candidate glial cells via adaptive thresholding and filters them using area (< 50 μm^2), eccentricity (near-circular morphology), and color thresholding, informed by prior knowledge [3] that glial cells are small and predominantly circular. Candidate instances are merged using a polygon-clustering strategy that groups overlapping detections (based on IoU) across models, aggregates spatial instance masks weighted by model scores, and extracts final boundaries via score thresholding, reducing missed detections and spurious instances relative to single-model predictions.

Cell Classification. Each segmented instance is assigned to one of three classes (excitatory, inhibitory, glial) using a hybrid feature-based classifier. First, 30 handcrafted morphological and spatial descriptors are computed, informed by known cytoarchitectural differences between cell types [16,26]—e.g., area, eccentricity, and orientation—and neighborhood-aware features computed over a 15-nearest-neighbor graph, including local density and orientation deviation. Deep embeddings are then extracted using the UNI2-h model [6] from both the isolated cell crop and its surrounding context, and reduced to 20 principal components via PCA, heuristically preventing them from dominating other descriptors. The resulting normalized and concatenated 50-dimensional feature vectors are classified using logistic regression trained on human-in-the-loop annotations.

2.2 Density Estimation

Because the released 1 μm/px sections sparsely sample the hippocampal volume, microscopic counts alone cannot produce complete volumetric density maps. Inspired by crowd-density estimation works [9], we cast cellular estimation as supervised localized regression in the 20 μm/px domain: given a low-resolution patch, the model predicts a continuous, cell-type-specific density field encoding the expected number of cells per unit area (from which discrete cell locations can be sampled). Supervision is obtained by projecting and registering matching high-resolution annotations into the low-resolution coordinate space. Finally, we reconstruct the mesoscale hippocampal CA point-cloud volume by sampling each coronal-slice density prediction and stacking them slice by slice in the global low-resolution whole brain space.

Low-Resolution Density Map Dataset. Cell centroids annotated at 1 μm/px are mapped to the 20 μm/px space using the provided affine transforms (Sec. 2.1) and refined via deformable registration [22] to correct residual HR-LR tissue mismatch and prevent misaligned density targets. We then build continuous ground-truth density maps for three classes by replacing each cell with a geometry-aware Gaussian kernel [27]. Kernel variance adapts to local cytoarchitecture: for each cell we compute the mean same-class k-NN distance $\overline{d}$ and set the variance proportional to $\overline{d}$, applying coefficient β, while normalizing each kernel to preserve total counts (allowing per-pixel densities > 1). This adaptive smoothing reduces sparse-region discontinuities and yields more learnable supervision signals; empirically, $\beta = 0.5$ and $k = 5$ performed best. From the 24 matched HR/LR WSI pairs, we extract 128×128 patches with 96-px overlap to mitigate data scarcity, yielding 505 patches, and divide them into an 80/20 train/test split, stratified by coronal slice coordinates (WSI ID) to avoid data leakage (Fig. 1c).

Architecture & Losses. We use a modernized UNet backbone [13,17] for low-resolution density estimation, leveraging multi-scale context and skip connections to preserve spatial localization. The network outputs three continuous density channels (one per cell class). To match receptive field to histological context while limiting overfitting under scarce data, we employ a compact 4-stage UNet with 32 initial convolution channels, along with batch normalization

that improves feature stability in cell-dense patches, and a final ReLU activation, enforcing non-negative density predictions. Regarding the loss, the standard Mean Squared Error (MSE) used in classic density estimation imposes a quadratic pixel-wise penalty that strongly over-penalizes small spatial shifts, encouraging blurred, low-amplitude maps that reduce squared loss while distorting the spatial structure and total count. The Mean Absolute Error (MAE) better preserves global mass but still tends to diffuse peaks and underestimate high-density regions. In addition, hippocampal data exhibit strong cell-class imbalance and substantial variability in cell density across patches. These factors motivate the use of a bounded objective designed to remain stable across density scales while preserving localized peaks. We therefore optimize the hybrid loss $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{NGAME}} + \mathcal{L}_{\text{AClsNL1}}$, where:

$$\mathcal{L}_{\text{NGAME}} = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{N_c}\sum_{i=1}^{4^l}\left|\hat{N}_{c,i} - N_{c,i}\right|,$$

$$\mathcal{L}_{\text{AClsNL1}} = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{N_c}\sum_{h,w}\lambda_{c,h,w}\left|\hat{Y}_{c,h,w} - Y_{c,h,w}\right|,\ \lambda_{c,h,w} = \begin{cases}\alpha & \text{if } \hat{Y}_{c,h,w} < Y_{c,h,w}\\ 1 & \text{otherwise}\end{cases}$$

Here, $Y, \hat{Y} \in R^{C\times H\times W}$ are the ground-truth and predicted density maps, h, w index pixels, and $c \in \{1, \ldots, C\}$ indexes classes. N_c and $\hat{N}_c$ are ground-truth and predicted class counts computed over Y and $\hat{Y}$, respectively. $\mathcal{L}_{\text{NGAME}}$, Normalized Grid Average Mean Absolute Error [12], measures absolute count errors over 4^l sub-regions, where $N_{c,i}$ and $\hat{N}_{c,i}$ denote the corresponding regional counts. $\mathcal{L}_{\text{AClsNL1}}$ is a class-normalized ℓ_1 loss with asymmetric weighting $\alpha > 1$ applied only over under-prediction to promote peak recovery.

Model Training. We train and evaluate using stratified five-fold cross-validation on the low-resolution training set previously described, stratifying by coronal slice coordinates associated with each WSI identifier (Fig. 1d). Optimization uses AdamW [14] with gradient clipping to mitigate gradient spikes from highly dense cellular regions. Training runs for 100 epochs with batch size 32 and learning rate 1×10^{-4}. Data augmentation applies horizontal/vertical flips and 90° rotations to both images and masks to mitigate data scarcity; inputs are normalized using the training-set mean and std. We set $\alpha = 2$ for the $\mathcal{L}_{\text{AClsNL1}}$ and $l = 2$ for the $\mathcal{L}_{\text{NGAME}}$ based on best-performing validation-set experiments.

WSIs and Volume Inference. We perform LR WSI inference with a sliding-window strategy: for each slice, 128×128 patches with 50% overlap are extracted, ensuring each pixel is observed 4 times, and the predictions are merged via Gaussian blending. Predicted multi-class density maps are constrained to the CA-area ROIs. To obtain discrete points, each density map is sampled N times using a multinomial distribution, where N is the predicted total cell count. We repeat this procedure for all LR coronal slices (Fig. 1e-f).

Finally, the points can be mapped to the global BigBrain coordinate system and stacked across slices in isotropic space to obtain the final 3D mesoscale point cloud of the right hippocampus (Fig. 1g). This reconstruction is anatomically

Table 1: Segmentation performance across fine-tuned ensembled models on held-out test set. Best results per section are in **bold**, second best are underlined.

	Global					Excitatory		Inhibitory		Glial	
Model	Dice	HD95	F1	IoU	PQ	Recall	HD95	Recall	HD95	Recall	HD95
CellPose-SAM	**0.79**	2.44	0.71	**0.60**	0.42	**0.93**	**2.47**	0.82	2.10	0.40	2.45
StarDist	**0.79**	2.65	**0.85**	0.52	0.45	0.92	3.17	0.87	**1.88**	**0.70**	2.38
HoverNet	0.77	2.99	0.77	0.57	0.44	0.92	3.71	**0.90**	2.82	0.53	2.40
InstanSeg	0.74	2.67	0.83	0.56	**0.46**	**0.93**	3.57	0.86	2.37	0.64	**2.06**
ATM	0.14	**2.42**	0.53	0.39	0.20	0.03	3.88	0.62	2.09	0.52	2.09
Merge (no ATM)	0.80	2.40	0.86	**0.55**	**0.47**	**0.91**	**2.63**	**0.95**	2.10	0.73	**2.28**
Merge (All)	**0.81**	**2.39**	**0.87**	0.54	**0.47**	**0.91**	2.63	0.93	**1.99**	**0.77**	2.29

plausible because hippocampal cell somata are generally on the order of, and often smaller than, 20 μm in diameter, so the 20 μm coronal step rarely bisects them, and the isotropic-space assembly remains realistic.

3 Results

3.1 CALHippo Instance Annotation Dataset Curation

Instance Segmentation Metrics. Instance predictions are matched to ground-truth polygons using a centroid-based protocol. Let $\mathcal{S}$ and $\hat{\mathcal{S}}$ denote the sets of ground-truth and predicted instances. A prediction $\hat{s} \in \hat{\mathcal{S}}$ is considered a true positive (TP) if the Euclidean distance between its centroid and that of a ground-truth instance $s \in \mathcal{S}$ is below 3 pixels, following the methodology of [21,23], with each ground-truth instance matched to at most one prediction. Unmatched predictions and ground-truth instances are counted as false positives (FP) and false negatives (FN). Segmentation accuracy is evaluated using Dice and the 95th percentile Hausdorff distance (HD95) on matched instances. Overall instance quality is summarized by Panoptic Quality (PQ), defined as $\mathrm{PQ} = \mathrm{F1} \times \mathrm{IoU}$ [11], where F1 (Detection Quality) is computed over matched instances and IoU (Segmentation Quality) is the mean intersection-over-union value across TP pairs. As segmentation predictions are class-agnostic, only ground-truth-derived metrics Recall (detection) and HD95 (boundary accuracy) are reported per cell type.

Segmentation & Classification Results. As shown in Table 1, CellPose-SAM shows stronger boundary adherence (higher Dice, lower HD95) but reduced glial Recall, whereas alternative models better capture small glial instances at the expense of contour precision in dense regions. To mitigate these trade-offs, predictions are fused via the clustering-and-confidence merging strategy described in Sec. 2.1: polygons with pairwise IoU ≥ 0.3 are clustered, and a normalized per-pixel confidence map is thresholded (> 0.3) to obtain final masks. The merged output, compared with single-model metrics, consistently improves Dice, PQ,

Table 2: Density estimation model ablation studies. Results correspond to the best epoch, averaged first over cell classes and then across folds (± std), and are ordered by the risk-adjusted confidence bound (rank) U.

Loss	Mean NAE % (↓)	Mean SSIM % (↑)	Rank U (↑)
MSE	36.37 ± 9.72	34.07 ± 2.83	0.40
MSE + NormL1	<u>23.34</u> ± 5.25	22.10 ± 3.00	0.48
NGAME + NormL1	**23.22** ± 5.12	35.48 ± 3.10	0.82
NGAME + ClassNormL1	24.63 ± 5.01	<u>40.62</u> ± 3.28	<u>0.91</u>
NGAME + AsymClassNormL1	23.77 ± 4.03	**41.41** ± 2.15	**1.00**

and glial Recall while preserving boundary accuracy. Detected instances were then classified into excitatory, inhibitory, or glial cells using the method described in Sec. 2.1. Under leave-one-out cross-validation, the classifier achieved 95.08% mean accuracy and 80.67% mean F1-score.

Annotation Variability. To assess reliability, we compute the Annotation Error Rate (AER), defined as $(FP + FN)/(TP + FN)$ [15], on the two fully annotated test WSIs, using annotator A as a reference. Annotator B reaches an inter-annotator AER of 0.35, while repeated annotation by the same experts yields an intra-annotator AER of 0.30. Disagreement is mainly driven by glial cells, whose weak boundaries in Merker silver-stained tissue make them difficult to distinguish from artifacts and debris. The predicted annotations achieve an AER of 0.25, below inter-annotator disagreement and within expert variability. All predictions were expert-validated through manual review, ensuring anatomically reliable masks and labels across all CA subfields.

3.2 Density Estimation Results

Metrics and Ablation. Following [9], we evaluate both counting and localization. Counting is measured by Normalized Absolute Error (NAE), which accounts for strong density (count) variation and class imbalance, while localization fidelity is assessed with SSIM [24]. We study loss effectiveness through stratified five-fold ablations on the training set and rank the resulting models by a risk-adjusted score U that penalizes fold variability. Specifically, we compute worst-case bounds $B_{\mathrm{NAE}} = \mu_{\mathrm{NAE}} + \sigma_{\mathrm{NAE}}$ and $B_{\mathrm{SSIM}} = \mu_{\mathrm{SSIM}} - \sigma_{\mathrm{SSIM}}$, min-max normalize them, and combine them as $U = 0.5(1-\tilde{B}_{\mathrm{NAE}})+0.5\tilde{B}_{\mathrm{SSIM}}$. Table 2 follows this incremental ranking and shows consistent gains: jointly using $\mathcal{L}_{\mathrm{NGAME}}$ and $\mathcal{L}_{\mathrm{AClsNL1}}$ maximizes U, improving count accuracy and peak alignment in high-density regions compared with standard MSE and L_1 losses.

Experimental Results. Table 3 reports performance on a held-out test set of 5 WSIs. We generate full-slide density maps with the inference pipeline in Sec. 2.2 and evaluate predictions within CA ROIs. For fair comparison with Gandolfi *et al.* [8], which is limited to CA1 and discrete predictions, we convert their point maps into density maps within matched coronal regions using

Table 3: Held-out density estimation performance. Gandolfi et al. baseline is reported only for CA1 because it is the closest available BigBrain hippocampus baseline; All-CA reports our full CA1-CA4 evaluation. Results are mean ± std.

		Excitatory		Inhibitory		Glial	
Model	Area	NAE % (↓)	SSIM % (↑)	NAE % (↓)	SSIM % (↑)	NAE % (↓)	SSIM % (↑)
Gandolfi [8]	CA1	82.8 ± 24.6	7.2 ± 7.5	90.9 ± 4.4	**58.4** ± 18.7	-	-
Ours	CA1	**22.8** ± 38.8	**49.4** ± 13.7	**79.2** ± 59.5	47.7 ± 11.5	25.3 ± 10.6	33.0 ± 11.9
Ours	All CA	29.5 ± 35.0	50.7 ± 12.8	21.6 ± 21.6	48.2 ± 9.9	20.9 ± 17.8	28.9 ± 8.4

the same geometry-aware Gaussian kernel described in Sec. 2.2. The baseline substantially underestimates excitatory-cell density, whereas our model reduces excitatory count error and improves spatial agreement. For inhibitory cells, results are mixed: our model achieves lower NAE but lower SSIM, consistent with sparse interneuron targets and the difficulty of resolving inhibitory-enriched patterns at 20 μm/px. The All-CA evaluation extends the analysis to the full CA1-CA4 complex, reducing data sparsity and yielding more accurate and stable inhibitory-cell density estimates than the CA1-only setting.

Hippocampal Mesoscale Modeling. We processed the remaining 1,998 low-resolution BigBrain slices to reconstruct the mesoscale hippocampal volume as described in Sec. 2.2. Across CA1-CA4, we predict $>$ 35M cells (18.1M excitatory neurons, 15.4M glia, 2.19M inhibitory interneurons), yielding an inhibitory-to-excitatory ratio of $\sim$12%, consistent with prior neuroanatomical estimates [4,16]. Region-wise neuron-like counts (excitatory+inhibitory) closely match the estimates of West *et al.* [25]: CA1 15.2M vs. 16M, CA2+CA3 2.8M vs. 2.7M, and CA4 2.3M vs. 2.0M, further supporting biological plausibility. For the additional glial class, we observe an approximate 1:1 glia-to-neuron ratio, consistent with Azevedo *et al.* [2]. Gandolfi *et al.* [8], the only directly comparable prior BigBrain work, reports CA1-only estimates of 4.80M excitatory and 480k inhibitory cells. For CA1 alone, our framework estimates 13.5M excitatory and 1.6M inhibitory cells, in closer agreement with stereological estimates. In addition, our reconstruction includes glia and extends the analysis from CA1 to the full CA1-CA4 complex. To our knowledge, this is the first large-scale, cell-type-resolved reconstruction of all human BigBrain CA subfields at this resolution.

4 Conclusion

We presented CALHippo, a multiscale resource for cell-type-resolved reconstruction of the right human hippocampal CA complex. Using 1 μm/px HR BigBrain sections, we produced the first expert-validated, three-class (excitatory, inhibitory, glial) instance-level dataset spanning CA1-CA4, via a human-in-the-loop pipeline with iterative expert correction and multi-model ensembling. To generalize beyond sparsely sampled HR sections, we projected and registered

annotations into the LR 20 μm/px space to derive class-specific density supervision, and trained a compact UNet with a hybrid loss that enforces regional count consistency and improves peak recovery under strong class imbalance. On held-out WSIs, the model improved CA1 excitatory density estimation relative to the baseline [8], while enabling inhibitory and glial density inference across the full CA1-CA4 complex. Applied to all available LR slices, the resulting density-derived volume produced biologically plausible neuron-like counts. The main limitations remain the sparse availability of HR sections and the difficulty of distinguishing glial cells in Merker-stained tissue. Additional high-resolution data would improve density supervision and reduce uncertainty in the inferred volume. More broadly, the same HR-to-LR annotation and density-estimation pipeline could potentially be applied to other anatomical regions of the BigBrain dataset. The released CALHippo annotations, models, code, and sampled 3D point cloud provide a reproducible scaffold for cytoarchitectonic analysis and large-scale hippocampal circuit modeling.

Acknowledgments. This project has received funding from Fondazione di Modena, FAR 2024 (CUP E93C24002080007) and PRJ-0408-FAR2024-UNIMORE-EPICURE (CUP E93C24002030005). Project IR0000011, CUP B51E22000150006, "EBRAINS-Italy".

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Amunts, K., Lepage, C., Borgeat, L., Mohlberg, H., Dickscheid, T., Rousseau, M.É., Bludau, S., Bazin, P.L., Lewis, L.B., Oros-Peusquens, A.M., et al.: BigBrain: An Ultrahigh-Resolution 3D Human Brain Model. Science (2013)
2. Azevedo, F.A., Carvalho, L.R., Grinberg, L.T., Farfel, J.M., Ferretti, R.E., Leite, R.E., Filho, W.J., Lent, R., Herculano-Houzel, S.: Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. Journal of Comparative Neurology (2009)
3. Bedner, P., Jabs, R., Steinhäuser, C.: Properties of human astrocytes and NG2 glia. Glia (2020)
4. Bezaire, M.J., Soltesz, I.: Quantitative Assessment of CA1 Local Circuits: Knowledge Base for Interneuron-Pyramidal Cell Connectivity. Hippocampus (2013)
5. Casari, G., Bolelli, F., Grana, C.: Histological Brain Imaging Super-resolution with Frequency-guided Diffusion Models. In: IEEE International Symposium on Biomedical Imaging (ISBI) (2026)
6. Chen, R.J., Ding, T., Lu, M.Y., Williamson, D.F.K., Jaume, G., Song, A.H., Chen, B., Zhang, A., Shao, D., Shaban, M., Williams, M., Oldenburg, L., Weishaupt, L.L., Wang, J.J., Vaidya, A., Le, L.P., Gerber, G., Sahai, S., Williams, W., Mahmood, F.: Towards a general-purpose foundation model for computational pathology. Nature Medicine (2024)
7. DeKraker, J., Lau, J.C., Ferko, K.M., Khan, A.R., Köhler, S.: Hippocampal subfields revealed through unfolding and unsupervised clustering of laminar and morphological features in 3D BigBrain. NeuroImage (2020)
8. Gandolfi, D., Mapelli, J., Solinas, S.M., Triebkorn, P., D'Angelo, E., Jirsa, V., Migliore, M.: Full-scale scaffold model of the human hippocampus CA1 area. Nature Computational Science (2023)
9. Gao, G., Gao, J., Liu, Q., Wang, Q., Wang, Y.: A survey of deep learning methods for density estimation and crowd counting. Vicinagearth (2025)
10. Goldsborough, T., Philps, B., O'Callaghan, A., Inglis, F., Leplat, L., Filby, A., Bilen, H., Bankhead, P.: InstanSeg: An embedding-based instance segmentation algorithm optimized for accurate, efficient and portable cell segmentation. arXiv:2408.15954 (2024)
11. Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., Rajpoot, N.: Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Medical Image Analysis (2019)
12. Guerrero-Gómez-Olmedo, R., Torre-Jiménez, B., López-Sastre, R., Maldonado-Bascón, S., Onoro-Rubio, D.: Extremely Overlapping Vehicle Counting. In: Iberian Conference on Pattern Recognition and Image Analysis (2015)
13. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods (2021)
14. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. In: International Conference on Learning Representations (2017)
15. Pachitariu, M., Rariden, M., Stringer, C.: Cellpose-SAM: superhuman generalization for cellular segmentation. bioRxiv (2025)
16. Pelkey, K.A., Chittajallu, R., Craig, M.T., Tricoire, L., Wester, J.C., McBain, C.J.: Hippocampal GABAergic Inhibitory Interneurons. Physiological Reviews (2017)
17. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer Assisted Intervention (2015)

18. Sanz Leon, P., Knock, S.A., Woodman, M.M., Domide, L., Mersmann, J., McIntosh, A.R., Jirsa, V.: The Virtual Brain: a simulator of primate brain network dynamics. Frontiers in Neuroinformatics (2013)
19. Schiffer, C., Lepage, C., Omidyeganeh, M., Mohlberg, H., Brandstetter, A., Bludau, S., Heuer, K., Toussaint, P.J., Wenzel, S., Dickscheid, T., Evans, A.C., Amunts, K.: Selected 1 micron scans of BigBrain histological sections (v1.0). EBRAINS (2022)
20. Schmidt, U., Weigert, M., Broaddus, C., Myers, G.: Cell Detection with Star-convex Polygons. In: Medical Image Computing and Computer Assisted Intervention (2018)
21. Sirinukunwattana, K., Raza, S.E.A., Tsang, Y.W., Snead, D.R., Cree, I.A., Rajpoot, N.M.: Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. IEEE Transactions on Medical Imaging (2016)
22. Tustison, N.J., Cook, P.A., Holbrook, A.J., Johnson, H.J., Muschelli, J., Devenyi, G.A., Duda, J.T., Das, S.R., Cullen, N.C., Gillen, D.L., et al.: The ANTsX ecosystem for quantitative biological and medical imaging. Scientific Reports (2021)
23. Vadori, V., Graïc, J.M., Peruffo, A., Vadori, G., Finos, L., Grisan, E.: CISCA and CytoDArk0: A cell instance segmentation and classification method for histo(patho)logical image Analyses and a new, open, Nissl-stained dataset for brain cytoarchitecture studies. Computers in Biology and Medicine (2025)
24. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing (2004)
25. West, M.J., Gundersen, H.: Unbiased stereological estimation of the number of neurons in the human hippocampus. Journal of Comparative Neurology (1990)
26. Zhang, T., et al.: Neuron type classification in rat brain based on integrative convolutional and tree-based recurrent neural networks. Scientific Reports (2021)
27. Zhang, Y., Zhou, D., Chen, S., Gao, S., Ma, Y.: Single-Image Crowd Counting via Multi-Column Convolutional Neural Network. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2016)