SLIMBRAIN database

https://slimbrain.citsem.upm.es/

Hyperspectral imaging (HSI) and machine learning (ML) have been employed in the medical field for classifying highly infiltrative brain tumors. Although existing HSI databases of in vivo human brains are available, they present two main deficiencies. First, the amount of labelled data is scarce, and second, 3D-tissue information is unavailable. To address both issues, Martín-Pérez et al. from the Research Center on Software Technologies and Multimedia Systems, Universidad Politécnica and Neurosurgery Department, Hospital Universitario 12 de Octubre, Madrid Spain, present the SLIMBRAIN database, a multimodal image database of in vivo human brains that provides HS brain tissue data within the 400-1000 nm spectra, as well as RGB, depth, and multiview images. Two HS, depth, and RGB sensors were used to capture images and videos from 193 patients. All the data in the SLIMBRAIN database can be used in a variety of ways, for example, to train machine learning models with more than 1 million HS pixels available and labelled by neurosurgeons, to reconstruct 3D scenes or to visualize RGB brain images with different pathologies, offering unprecedented flexibility for both the medical and engineering communities [1)].

The application of hyperspectral imaging (HSI) in neuro-oncology is gaining traction due to its potential to provide real-time, non-contact, and label-free differentiation of brain tissue based on spectral signatures. However, most existing studies have been hampered by two significant limitations: limited labelled datasets and the lack of 3D tissue context. These deficiencies have constrained the training and generalizability of machine learning models in surgical settings.

The SLIMBRAIN database is presented as a direct response to these issues, positioning itself as the most comprehensive multimodal in vivo brain imaging dataset currently available.

2. Strengths of the Study Multimodality: SLIMBRAIN includes hyperspectral data (400–1000 nm), RGB images, depth maps, and multiview recordings. This rich combination enables not only tumor classification but also 3D reconstruction, tissue segmentation, and data fusion studies.

Scale and Clinical Validity: The inclusion of 193 patients with intraoperative data and labelled pixels by neurosurgeons significantly enhances the clinical credibility and utility of the dataset. Over 1 million labelled pixels provide an unprecedented foundation for robust model training and validation.

Engineering and Clinical Interoperability: By bridging raw HSI data with depth and RGB, the dataset is valuable both for computer vision researchers and clinical teams, enabling translational research across disciplines.

3. Methodological Considerations and Limitations

Lack of Standardization Across Surgeries: While the scale is impressive, variability in lighting conditions, sensor calibration, or inter-patient differences could introduce noise or domain shift, affecting ML model generalizability. The authors do not explicitly describe normalization protocols for inter-surgical data variability.

Labeling Strategy: Although pixel-level labeling by neurosurgeons is a major strength, the study lacks detail on how consistency was ensured across surgeons, especially for infiltrative margins where subjective interpretation may vary. This is crucial for downstream model performance.

Missing Histopathological Correlation: The labels are intraoperative (based on neurosurgical judgment), but there is no systematic histopathological validation of HSI pixel classifications. This limits ground-truth fidelity, especially in ambiguous zones such as tumor margins.

Absence of Functional or Dynamic Data: While structural and spectral data are well represented, the absence of functional imaging modalities (e.g., perfusion, fluorescence, or electrocorticography) restricts the dataset's utility for comprehensive intraoperative decision-making.

4. Broader Implications and Future Directions

The SLIMBRAIN database fills a crucial gap in multimodal intraoperative imaging research and sets a new benchmark for open-access datasets in neurosurgery. It is well-positioned to:

Foster the development of real-time ML-based surgical guidance tools.

Enable transfer learning and domain adaptation studies using rich spectral and geometric features.

Serve as a validation benchmark for new HSI sensors and fusion algorithms.

Future work should emphasize:

Cross-validation with postoperative histopathology.

Standardization of acquisition protocols and interoperability with hospital systems.

Expansion to include longitudinal follow-up and outcome-based annotations.

Conclusion

The SLIMBRAIN database is a landmark contribution to the field of surgical imaging and AI in neuro-oncology. Despite certain methodological limitations—chiefly regarding labeling consistency and histopathological correlation—its scale, multimodal depth, and open availability make it a transformative resource. Its full potential will be realized when integrated into prospective, outcome-linked ML workflows and subjected to external clinical validation.

[1)]
Martín-Pérez A, Villa M, Rosa Olmeda G, Sancho J, Vazquez G, Urbanos G, Martinez de Ternero A, Chavarrías M, Jimenez-Roldan L, Perez-Nuñez A, Lagares A, Juarez E, Sanz C. SLIMBRAIN database: A multimodal image database of in vivo human brains for tumour detection. Sci Data. 2025 May 21;12(1):836. doi: 10.1038/s41597-025-04993-y. PMID: 40399336.