

# Mitigating Data Scarcity in the Classification of Glioma Molecular Subtypes: The Power of Generative Imaging

Nghi C. D. Truong<sup>1,\*</sup>, Chandan Ganesh Bangalore Yogananda<sup>1</sup>, Benjamin C. Wagner<sup>1</sup>, Niloufar Saadat<sup>1</sup>, James M. Holcomb<sup>1</sup>, Divya Reddy<sup>1</sup>, Sadeem Lodhi<sup>1</sup>, Jason Bowerman<sup>1</sup>, Kimmo J. Hatanpaa<sup>2</sup>, Toral R. Patel<sup>3</sup>, Baowei Fei<sup>1,4</sup>, Matthew D. Lee<sup>5</sup>, Rajan Jain<sup>5,6</sup>, Richard J. Bruce<sup>7</sup>, Marco C. Pinho<sup>1</sup>, Ananth J. Madhuranthakam<sup>8</sup>, and Joseph A. Maldjian<sup>1</sup>

<sup>1</sup>Department of Radiology, UT Southwestern Medical Center, Texas, USA

<sup>2</sup>Department of Pathology, UT Southwestern Medical Center, Texas, USA

<sup>3</sup>Department of Neurological Surgery, UT Southwestern Medical Center, Texas, USA

<sup>4</sup>Department of Bioengineering, University of Texas at Dallas, Texas, USA

<sup>5</sup>Department of Radiology, NYU Grossman School of Medicine, New York, USA

<sup>6</sup>Department of Neurosurgery, NYU Grossman School of Medicine, New York, USA

<sup>7</sup>Department of Radiology, University of Wisconsin-Madison, Wisconsin, USA

<sup>8</sup>Department of Radiology, Mayo Clinic, Rochester, MN, USA

## ABSTRACT

Isocitrate dehydrogenase (IDH) mutation status is a critical prognostic indicator in glioma patients. Numerous studies have focused on developing non-invasive methodologies to classify IDH status using pre-operative MRI scans. However, the challenge lies in data scarcity and class imbalance in IDH mutations. This study explores generative AI methods to augment training data and enhance IDH classification accuracy. We developed a 3D conditional latent diffusion model (LDM) for generating 3D multi-contrast brain tumor MRI data ( $128 \times 128 \times 64$  with a voxel spacing of  $1.5 \times 1.5 \times 2.0$  mm) with whole tumor mask and IDH mutation status as conditions. The LDM comprises a 3D autoencoder for perceptual compression and a conditional 3D diffusion model (DM) for generating multi-contrast synthetic samples guided by tumor masks and the IDH mutation status. We incorporated two types of attention modules within the denoising UNet of the LDM to capture the semantic class-dependent data distribution driven by the provided whole tumor mask and IDH status. The LDM was trained using two brain tumor datasets: The Cancer Genome Atlas dataset and an internal dataset from the University of Texas Southwestern Medical Center. The synthetic images generated by the LDM were then used to train IDH classification models, which were subsequently tested on real brain tumor data comprising 327 mutated and 1,394 wild-type cases from the University of California San Francisco Preoperative Diffuse Glioma MRI dataset, the Erasmus Glioma Database, the University of Pennsylvania glioblastoma, and two held-out internal datasets. The IDH classification models, trained on synthetic images and tested on real data, achieved an excellent overall classification accuracy of 94.02%. This approach has the potential to be extended to other molecular markers where data scarcity presents a challenge.

**Keywords:** Latent diffusion model, Generative models, Brain tumor imaging, Synthetic data

## 1. INTRODUCTION

Data plays a crucial role in the development and training of machine learning models, as the accuracy of these models is highly dependent on the quality and quantity of the training data. In the domain of medical imaging, however, acquiring large and diverse data is particularly challenging due to stringent privacy regulations for protecting patient confidentiality and the substantial costs associated with data collection and annotation. These challenges result in a limited availability of medical imaging data, creating a bottleneck in the training process.

---

Further author information: (Send correspondence to Nghi C. D. Truong)

Nghi C. D. Truong: E-mail: nghi.truong@utsouthwestern.edu

Recent advancements in image generative models,<sup>1,2</sup> particularly the emergence of diffusion models (DM),<sup>2-5</sup> offer promising potential in addressing the data scarcity in medical imaging. Diffusion models can generate high-quality synthetic images that closely resemble real medical images, thereby expanding the pool of available data. This expanded dataset can enhance the development of more robust machine-learning models for various tasks, including image segmentation and classification.

Prior studies have utilized image-generative models to create high-quality, realistic medical imaging data.<sup>6</sup> Specifically, several works have focused on developing models for brain MRI data.<sup>7-12</sup> Some studies have aimed to enhance diagnostic tasks such as tumor segmentation<sup>13,14</sup> or mutation prediction.<sup>15</sup> For instance, Dorjsembe et al.<sup>14</sup> introduced semantic tumor mask conditioning to control the synthesis process and then combined the synthetic samples with the real ones to enhance the segmentation performance. Moon et al.<sup>15</sup> trained 2D score-based diffusion models separately for IDH mutated and wild-type classes to generate 2D IDH-specific samples, which were subsequently evaluated using Turing’s test and employed to train an IDH classification model.

This study leverages the power of latent diffusion models (LDM)<sup>2</sup> to generate IDH-specific brain tumor MRI samples. IDH mutations are a critical glioma biomarker, influencing prognosis and treatment strategies.<sup>16</sup> We developed a diffusion model capable of generating synthetic MR images that accurately reflect the characteristics of IDH-specific brain tumors. The generation process is guided by a conditional whole tumor mask and IDH mutation status, ensuring the synthetic data closely mirrors the characteristics of the tumors. The synthetic samples were employed to train IDH classification models, which subsequently underwent rigorous testing on real IDH brain tumor data acquired from diverse datasets to validate the effectiveness of the synthetic data.

## 2. METHODOLOGY

### 2.1 Datasets and Image Preprocessing

This study utilized retrospective MRI scans from four publicly available datasets and three internal datasets. The public datasets included: The Cancer Genome Atlas (TCGA),<sup>17</sup> the University of California San Francisco Preoperative Diffuse Glioma MRI dataset (UCSF),<sup>18</sup> the Erasmus Glioma Database (EGD),<sup>19</sup> and the University of Pennsylvania glioblastoma (UPENN).<sup>20</sup> The internal datasets were collected from three geographically distinct institutions: UT Southwestern Medical Center (UTSW), New York University (NYU), and the University of Wisconsin-Madison (UWM). Data included in this study was required to have the preoperative MRI scans of all four sequences: T1-weighted (T1), post-contrast T1-weighted (T1C), T2-weighted (T2), and T2-weighted fluid-attenuated inversion recovery (FLAIR). The TCGA and UTSW datasets were used to train the latent diffusion model (LDM) (Table 1(a)), while the remaining five datasets (Table 1(b)) were reserved for testing the IDH classification models. A summary of the IDH mutations available in each dataset is provided in Table 1.

Table 1: (a) Training data for the diffusion model, and (b) Testing data for IDH classification models.

(a)	TCGA	UTSW	Total	(b)	UCSF	NYU	UWM	EGD	UPENN	Total
<b>MT</b>	91	104	195	<b>MT</b>	103	47	16	150	11	327
<b>WT</b>	107	256	363	<b>WT</b>	392	129	180	306	387	1,394
<b>Total</b>	198	360	558	<b>Total</b>	495	176	196	456	398	1,721

Note.—Data represent the number of patients. TCGA = The Cancer Genome Atlas, UCSF = University of California San Francisco Preoperative Diffuse Glioma MRI dataset, EGD = Erasmus Glioma Database, UPENN = The University of Pennsylvania glioblastoma, UTSW = UT Southwestern Medical Center, NYU = New York University, UWM = the University of Wisconsin-Madison; MT = IDH-mutant, WT = IDH- wild-type.

For the UCSF and EGD datasets, we utilized the preprocessed skull-stripped images provided by the original data sources. For the TCGA, UPENN, and the three internal datasets, MRI scans underwent preprocessing using the federated tumor segmentation (FeTS) tool.<sup>21</sup> The preprocessing steps included co-registering MRI scans to a template atlas, correcting for bias field distortion, and skull stripping. Additionally, the FeTS tool was employed for tumor segmentation in the TCGA and UTSW datasets. The resulting tumor masks from these datasets were subsequently used as conditions for training the diffusion models. To train the diffusion model, the images were further resampled and cropped to the dimensions of  $128 \times 128 \times 64$  with a voxel spacing of  $1.5 \times 1.5 \times 2.0$  mm.

## 2.2 Generative Model

Our latent diffusion model (LDM) for generating IDH-specific samples built upon previous work,<sup>12</sup> with notable enhancements to the conditioning mechanisms. The autoencoder used for perceptual compression was retained from the previous work, as it effectively reduces the dimensionality of the input data while preserving critical features for the LDM. In the second stage of the LDM, two types of conditioning were introduced: the whole tumor mask and text describing the IDH mutation status. The tumor mask was first resampled to match the latent space dimensions and then concatenated with the input of the denoising UNet in the LDM, providing spatial context that guides the denoising process. The IDH mutation status was embedded using a pre-trained CLIP model.<sup>22</sup> Within the denoising UNet, two types of attention mechanisms were employed. The self-attention module helps the model focus on the semantic-class-dependent features by attending to specific regions indicated by the conditional tumor mask. Moreover, the cross-attention modules reinforce the learning of class-dependent data distributions by integrating the IDH mutation status. These two attention modules allow the model to leverage both visual and textual information simultaneously, ensuring that the generated samples not only conform to the spatial characteristics dictated by the tumor mask but also accurately reflect the molecular subtype of the tumor, as determined by the IDH mutation status.

## 2.3 IDH Classification Model

To assess the effectiveness of these synthetic samples, the synthetic IDH-specific samples generated by the LDM were utilized to train IDH classification models. The classification model was developed using the powerful UNet architecture,<sup>23</sup> designed to classify tumor voxels as either IDH mutated or wild-type. During training, the tumor mask, which serves as the conditioning input for generating the synthetic sample, is labeled according to the IDH status of the sample. This labeling allows the network to learn relevant features across the whole tumor region. The final classification is determined by averaging the probabilities across all tumor voxels, with the class having the higher average probability being assigned as the final classification.

# 3. RESULTS

## 3.1 Image Generation with 3D Multi-contrast Conditional LDM

Figure 1 shows representative axial slices from two multi-contrast MRI samples generated by the LDM using the whole tumor mask and the IDH mutation status. The conditional tumor mask (red contour) was the same to generate the mutated (first row) and wild-type (second row) samples. The LDM successfully generated realistic images with distinct brain parenchymal features and expected IDH mutation status tumor characteristics (e.g., wild-type demonstrates more enhancement than mutated).

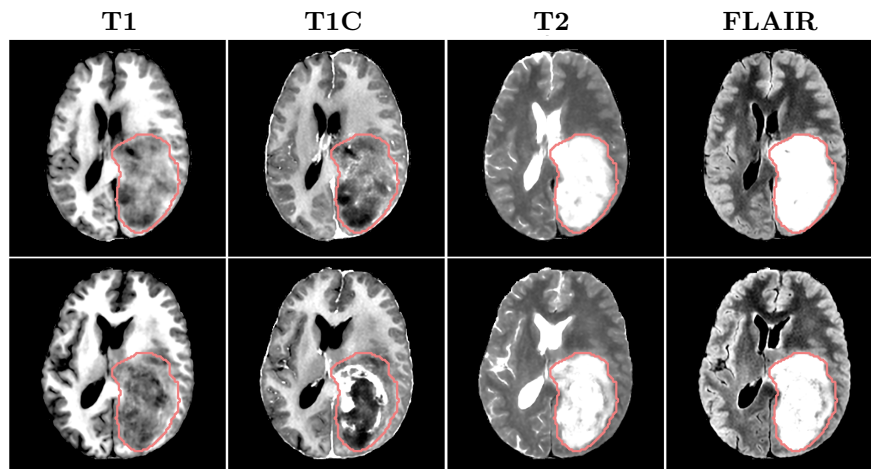


Figure 1: Multi-contrast MRI samples generated by the LDM using the same tumor mask (red contour) and the IDH mutation status. First row: IDH mutated example. Second row: IDH wild-type example.

### 3.2 IDH Classification Results

To assess the fidelity of the synthetic MRI samples, we generated a substantial number of IDH-specific samples using synthetic tumor masks. These 4-contrast synthetic MRI samples were then employed to train IDH classification models. In Experiments 1-4, the number of synthetic samples increased progressively from 200 to 3,000, with a balanced distribution of IDH mutated and wild-type cases. Table 2 presents the IDH classification results on various test sets corresponding to different training data sizes (Experiments 1-4). Additionally, we conducted two further experiments in which we evaluated the similarity among synthetic samples by extracting radiomics features from the whole tumor.<sup>24</sup> Samples exhibiting high similarity were excluded based on a predefined threshold. This process resulted in training sample sizes of 496 mutated and 614 wild-type cases in Experiment 5 and 872 mutated and 1,329 wild-type cases in Experiment 6.

Table 2: Performance of IDH mutation classification models trained on synthetic samples with different training sizes and tested on real data. The number in each experiment indicates the number of mutated and wild-type samples used for training.

		Exp. 1 100/100	Exp. 2 400/400	Exp. 3 800/800	Exp. 4 1,500/1,500	Exp. 5 496/614	Exp. 6 872/1,329
UCSF	ACC	84.24	84.65	90.30	91.31	91.31	92.73
	SEN (MT)	93.20	92.23	87.38	88.35	84.47	86.41
	SPE (WT)	81.89	82.65	91.07	92.09	93.11	94.39
	AUC	96.31	96.56	95.83	96.63	96.17	96.65
NYU	ACC	84.66	86.36	86.93	87.50	90.91	91.48
	SEN (MT)	87.23	87.23	82.98	82.98	80.85	87.23
	SPE (WT)	83.72	86.05	88.37	89.15	94.57	93.02
	AUC	92.63	92.34	93.13	94.05	93.76	91.32
UWM	ACC	81.63	87.76	88.27	87.76	90.31	95.92
	SEN (MT)	81.25	75.00	75.00	75.00	81.25	81.25
	SPE (WT)	81.67	88.89	89.44	88.89	91.11	97.22
	AUC	88.35	89.83	90.16	90.30	95.07	95.21
EGD	ACC	87.72	89.69	87.72	87.72	92.76	92.32
	SEN (MT)	81.33	78.67	75.33	75.33	86.67	90.67
	SPE (WT)	90.85	95.10	93.79	93.79	95.75	93.14
	AUC	93.77	95.12	95.20	95.64	96.56	95.75
UPENN	ACC	94.47	95.73	96.98	96.98	97.49	97.74
	SEN (MT)	63.64	63.64	63.64	72.73	63.64	72.73
	SPE (WT)	95.35	96.64	97.93	97.67	98.45	98.45
	AUC	90.25	89.41	90.60	90.73	89.73	93.75
Overall	ACC	87.27	89.08	90.59	90.88	92.97	94.02
	SEN (MT)	85.32	83.49	79.82	80.43	84.10	87.77
	SPE (WT)	87.73	90.39	93.11	93.33	95.05	95.48
	AUC	92.95	93.33	93.53	94.02	94.41	95.03

Note.— ACC = accuracy, SEN = sensitivity, SPE = specificity, and AUC = area under the receiver operating characteristic curve. SEN and SPE correspond to the accuracy of the mutated (MT) and wild-type (WT) classes, respectively. UCSF = University of California San Francisco Preoperative Diffuse Glioma MRI dataset, NYU = New York University, UWM = the University of Wisconsin-Madison, EGD = Erasmus Glioma Database, UPENN = The University of Pennsylvania glioblastoma.

The IDH classification results demonstrated that increasing the training size for the classification model enhanced the overall classification accuracy and the AUC. Additionally, by introducing a step of excluding similar samples from the synthetic brain tumor data, the performance of the IDH classification models was further enhanced, as evidenced by the results of experiments 5 and 6.

## 4. CONCLUSIONS

This study presents the development of a conditional LDM designed to generate 3D multi-contrast IDH-specific brain tumor MRI data. The models successfully produced high-quality synthetic MRI samples that accurately

reflected the expected IDH mutation tumor characteristics. When tested on real datasets, the IDH classification models trained on these synthetic images achieved a notable overall classification accuracy of 94.02%. This approach holds significant potential for application to other molecular markers, particularly in cases where data scarcity presents a challenge.

## ACKNOWLEDGMENTS

This research was supported by the NIH/NCI U01CA207091 (AJM, JAM) and R01CA260705 (JAM).

## DISCLOSURES

The authors declare no conflict of interest.

## REFERENCES

- [1] Abdal, R., Qin, Y., and Wonka, P., “Image2StyleGAN: How to Embed Images Into the StyleGAN Latent Space?,” in [*Proceedings of the IEEE/CVF International Conference on Computer Vision*], 4432–4441 (2019).
- [2] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B., “High-Resolution Image Synthesis With Latent Diffusion Models,” in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 10684–10695 (2022).
- [3] Dhariwal, P. and Nichol, A., “Diffusion Models Beat GANs on Image Synthesis,” in [*Advances in Neural Information Processing Systems*], **34**, 8780–8794, Curran Associates, Inc. (2021).
- [4] Nichol, A. Q. and Dhariwal, P., “Improved Denoising Diffusion Probabilistic Models,” in [*Proceedings of the 38th International Conference on Machine Learning*], 8162–8171, PMLR (July 2021).
- [5] Zhang, L., Rao, A., and Agrawala, M., “Adding Conditional Control to Text-to-Image Diffusion Models,” in [*Proceedings of the IEEE/CVF International Conference on Computer Vision*], 3836–3847 (2023).
- [6] Kazerouni, A., Aghdam, E. K., Heidari, M., Azad, R., Fayyaz, M., Hacihaliloglu, I., and Merhof, D., “Diffusion models in medical imaging: A comprehensive survey,” *Medical Image Analysis* **88**, 102846 (Aug. 2023).
- [7] Na, Y., Kim, K., Ye, S.-J., Kim, H., and Lee, J., “Generation of Multi-modal Brain Tumor MRIs with Disentangled Latent Diffusion Model,” in [*Medical Imaging with Deep Learning, Short Paper Track*], (Apr. 2023).
- [8] Khader, F., Müller-Franzes, G., Tayebi Arasteh, S., Han, T., Haarburger, C., Schulze-Hagen, M., Schad, P., Engelhardt, S., Baeßler, B., Foersch, S., Stegmaier, J., Kuhl, C., Nebelung, S., Kather, J. N., and Truhn, D., “Denoising diffusion probabilistic models for 3D medical image generation,” *Sci Rep* **13**, 7303 (May 2023).
- [9] Pinaya, W. H. L., Tudosiu, P.-D., Dafflon, J., Da Costa, P. F., Fernandez, V., Nachev, P., Ourselin, S., and Cardoso, M. J., “Brain Imaging Generation with Latent Diffusion Models,” in [*Deep Generative Models: Second MICCAI Workshop, DGM4MICCAI 2022, Held in Conjunction with MICCAI 2022, Singapore, September 22, 2022, Proceedings*], 117–126, Springer-Verlag, Berlin, Heidelberg (Sept. 2022).
- [10] Pinaya, W. H. L., Graham, M. S., Kerfoot, E., Tudosiu, P.-D., Dafflon, J., Fernandez, V., Sanchez, P., Wolleb, J., da Costa, P. F., Patel, A., Chung, H., Zhao, C., Peng, W., Liu, Z., Mei, X., Lucena, O., Ye, J. C., Tsaftaris, S. A., Dogra, P., Feng, A., Modat, M., Nachev, P., Ourselin, S., and Cardoso, M. J., “Generative AI for Medical Imaging: Extending the MONAI Framework,” (July 2023).
- [11] Peng, W., Adeli, E., Bosschieter, T., Park, S. H., Zhao, Q., and Pohl, K. M., “Generating Realistic Brain MRIs via a Conditional Diffusion Probabilistic Model,” (Sept. 2023).
- [12] Truong, N. C. D., Yogananda, C. G. B., Wagner, B. C., Holcomb, J. M., Reddy, D., Saadat, N., Hatanpaa, K. J., Patel, T. R., Fei, B., Lee, M. D., Jain, R., Bruce, R. J., Pinho, M. C., Madhuranthakam, A. J., and Maldjian, J. A., “Synthesizing 3D multicontrast brain tumor MRIs using tumor mask conditioning,” in [*Medical Imaging 2024: Imaging Informatics for Healthcare, Research, and Applications*], **12931**, 116–120, SPIE (Apr. 2024).
- [13] Kebaili, A., Lapuyade-Lahorgue, J., Vera, P., and Ruan, S., “3D MRI Synthesis with Slice-Based Latent Diffusion Models: Improving Tumor Segmentation Tasks in Data-Scarce Regimes,” (June 2024).

- [14] Dorjsembe, Z., Pao, H.-K., Odonchimed, S., and Xiao, F., “Conditional Diffusion Models for Semantic 3D Brain MRI Synthesis,” *IEEE J. Biomed. Health Inform.* **28**, 4084–4093 (July 2024).
- [15] Moon, H. H., Jeong, J., Park, J. E., Kim, N., Choi, C., Kim, Y.-H., Song, S. W., Hong, C.-K., Kim, J. H., and Kim, H. S., “Generative AI in glioma: Ensuring diversity in training image phenotypes to improve diagnostic performance for IDH mutation prediction,” *Neuro Oncol* **26**, 1124–1135 (June 2024).
- [16] Louis, D. N., Perry, A., Wesseling, P., Brat, D. J., Cree, I. A., Figarella-Branger, D., Hawkins, C., Ng, H. K., Pfister, S. M., Reifenberger, G., Soffiatti, R., von Deimling, A., and Ellison, D. W., “The 2021 WHO Classification of Tumors of the Central Nervous System: A summary,” *Neuro-Oncology* **23**, 1231–1251 (Aug. 2021).
- [17] Ceccarelli, M., Barthel, F. P., Malta, T. M., Sabedot, T. S., Salama, S. R., Murray, B. A., Morozova, O., Newton, Y., Radenbaugh, A., Pagnotta, S. M., Anjum, S., Wang, J., Manyam, G., Zoppoli, P., Ling, S., Rao, A. A., Grifford, M., Cherniack, A. D., Zhang, H., Poisson, L., Carlotti, C. G., Tirapelli, D. P. d. C., Rao, A., Mikkelsen, T., Lau, C. C., Yung, W. K. A., Rabadan, R., Huse, J., Brat, D. J., Lehman, N. L., Barnholtz-Sloan, J. S., Zheng, S., Hess, K., Rao, G., Meyerson, M., Beroukhi, R., Cooper, L., Akbani, R., Wensch, M., Haussler, D., Aldape, K. D., Laird, P. W., Gutmann, D. H., Anjum, S., Arachchi, H., Auman, J. T., Balasundaram, M., Balu, S., Barnett, G., Baylin, S., Bell, S., Benz, C., Bir, N., Black, K. L., Bodenheimer, T., Boice, L., Bootwalla, M. S., Bowen, J., Bristow, C. A., Butterfield, Y. S. N., Chen, Q.-R., Chin, L., Cho, J., Chuah, E., Chudamani, S., Coetzee, S. G., Cohen, M. L., Colman, H., Couce, M., D’Angelo, F., Davidsen, T., Davis, A., Demchok, J. A., Devine, K., Ding, L., Duell, R., Elder, J. B., Eschbacher, J. M., Fehrenbach, A., Ferguson, M., Frazer, S., Fuller, G., Fulop, J., Gabriel, S. B., Garofano, L., Gastier-Foster, J. M., Gehlenborg, N., Gerken, M., Getz, G., Giannini, C., Gibson, W. J., Hadjipanayis, A., Hayes, D. N., Heiman, D. I., Hermes, B., Hilty, J., Hoadley, K. A., Hoyle, A. P., Huang, M., Jefferys, S. R., Jones, C. D., Jones, S. J. M., Ju, Z., Kastl, A., Kessler, A., Kim, J., Kucherlapati, R., Lai, P. H., Lawrence, M. S., Lee, S., Leraas, K. M., Lichtenberg, T. M., Lin, P., Liu, Y., Liu, J., Ljubimova, J. Y., Lu, Y., Ma, Y., Maglinte, D. T., Mahadeshwar, H. S., Marra, M. A., McGraw, M., McPherson, C., Meng, S., Mieczkowski, P. A., Miller, C. R., Mills, G. B., Moore, R. A., Mose, L. E., Mungall, A. J., Naresh, R., Naska, T., Neder, L., Noble, M. S., Noss, A., O’Neill, B. P., Ostrom, Q. T., Palmer, C., Pantazi, A., Parfenov, M., Park, P. J., Parker, J. S., Perou, C. M., Pierson, C. R., Pihl, T., Protopopov, A., Radenbaugh, A., Ramirez, N. C., Rathmell, W. K., Ren, X., Roach, J., Robertson, A. G., Saksena, G., Schein, J. E., Schumacher, S. E., Seidman, J., Senecal, K., Seth, S., Shen, H., Shi, Y., Shih, J., Shimmel, K., Sicotte, H., Sifri, S., Silva, T., Simons, J. V., Singh, R., Skelly, T., Sloan, A. E., Sofia, H. J., Soloway, M. G., Song, X., Sougnez, C., Souza, C., Staugaitis, S. M., Sun, H., Sun, C., Tan, D., Tang, J., Tang, Y., Thorne, L., Trevisan, F. A., Triche, T., Van Den Berg, D. J., Veluvolu, U., Voet, D., Wan, Y., Wang, Z., Warnick, R., Weinstein, J. N., Weisenberger, D. J., Wilkerson, M. D., Williams, F., Wise, L., Wolinsky, Y., Wu, J., Xu, A. W., Yang, L., Yang, L., Zack, T. I., Zenklusen, J. C., Zhang, J., Zhang, W., Zhang, J., Zmuda, E., Noushmehr, H., Iavarone, A., and Verhaak, R. G. W., “Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma,” *Cell* **164**, 550–563 (Jan. 2016).
- [18] Calabrese, E., Villanueva-Meyer, J. E., Rudie, J. D., Rauschecker, A. M., Baid, U., Bakas, S., Cha, S., Mongan, J. T., and Hess, C. P., “The University of California San Francisco Preoperative Diffuse Glioma MRI Dataset,” *Radiology: Artificial Intelligence* **4**, e220058 (Nov. 2022).
- [19] van der Voort, S. R., Incekara, F., Wijnenga, M. M. J., Kapsas, G., Gahrman, R., Schouten, J. W., Dubbink, H. J., Vincent, A. J. P. E., van den Bent, M. J., French, P. J., Klein, S., and Smits, M., “The Erasmus Glioma Database (EGD): Structural MRI scans, WHO 2016 subtypes, and segmentations of 774 patients with glioma,” *Data in Brief* **37**, 107191 (Aug. 2021).
- [20] Bakas, S., Sako, C., Akbari, H., Bilello, M., Sotiras, A., Shukla, G., Rudie, J. D., Santamaría, N. F., Kazerooni, A. F., Pati, S., Rathore, S., Mamourian, E., Ha, S. M., Parker, W., Doshi, J., Baid, U., Bergman, M., Binder, Z. A., Verma, R., Lustig, R. A., Desai, A. S., Bagley, S. J., Mourelatos, Z., Morrisette, J., Watt, C. D., Brem, S., Wolf, R. L., Melhem, E. R., Nasrallah, M. P., Mohan, S., O’Rourke, D. M., and Davatzikos, C., “The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: Advanced MRI, clinical, genomics, & radiomics,” *Sci Data* **9**, 453 (July 2022).

- [21] Pati, S., Baid, U., Edwards, B., Sheller, M. J., Foley, P., Anthony Reina, G., Thakur, S., Sako, C., Bilello, M., Davatzikos, C., Martin, J., Shah, P., Menze, B., and Bakas, S., “The federated tumor segmentation (FeTS) tool: An open-source solution to further solid tumor research,” *Phys Med Biol* **67** (Oct. 2022).
- [22] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I., “Learning Transferable Visual Models From Natural Language Supervision,” (Feb. 2021).
- [23] Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., and Maier-Hein, K. H., “nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation,” *Nature methods* **18**(2), 203–211 (2021).
- [24] Truong, N. C. D., Bangalore Yogananda, C. G., Wagner, B. C., Holcomb, J. M., Reddy, D., Saadat, N., Hatanpaa, K. J., Patel, T. R., Fei, B., Lee, M. D., Jain, R., Bruce, R. J., Pinho, M. C., Madhuranthakam, A. J., and Maldjian, J. A., “Two-Stage Training Framework Using Multicontrast MRI Radiomics for IDH Mutation Status Prediction in Glioma,” *Radiology: Artificial Intelligence* **6**, e230218 (July 2024).