Partial Information Decomposition as a Multi-Contrast 3D MRI Selection Strategy for Resource-Constrained Deep Neural Network Training in Brain Tumor Segmentation

Agamdeep S. Chopra*[1] and Mehmet Kurt†[1]

[1]Department of Mechanical Engineering, University of Washington, Seattle, WA, USA
{achopra4,mkurt}@uw.edu

arXiv:2607.15396v1 [cs.CV] 16 Jul 2026

Abstract

Multi-contrast 3D MRI segmentation can be computationally demanding when all available sequences are used. We evaluate a pre-training Partial Information Decomposition framework that ranks input pairs according to their redundant, unique, and synergistic information about regional tumor burden and selects the highest-ranked pair for downstream training. Applied to T1n, T1c, T2w, and T2-FLAIR MRI, the framework selected T1c+T2-FLAIR. We then trained eleven architecturally identical lightweight 3D U-Nets using different input configurations. On an independent test cohort, T1c+T2-FLAIR was the strongest two-input configuration and ranked second overall in mean Dice (0.676 versus 0.687 for all four inputs). Independent Shapley analysis on the full-input model also identified T2-FLAIR and T1c as the most influential inputs and their pairwise interaction as the strongest. These findings demonstrate the practical value of PID based pre-training selection for identifying compact, informative MRI input sets before costly 3D model development.

Keywords: Partial information decomposition, MRI sequence selection, Brain tumor segmentation, Resource-constrained learning

1 Introduction

Multi-contrast MRI supports brain tumor segmentation because each sequence emphasizes complementary tissue characteristics. T1-weighted MRI depicts anatomy, contrast enhanced T1 highlights enhancing tumor, T2 MRI is sensitive to fluid, and T2-FLAIR suppresses cerebrospinal fluid to emphasize edema and infiltrative abnormalities. Consequently, brain tumor benchmarks and high-performing 3D segmentation frameworks commonly integrate all available contrasts for multi-region segmentation [1–4].

Although, this approach can be expensive in computationally constrained settings. Each additional sequence increases storage, preprocessing, data-transfer, memory, and training costs, which are amplified in 3D by large patch and volume tensors and by repeated experiments across folds, architectures, and hyperparameter settings. The practical question is therefore whether a smaller input subset can be selected before substantial resources are committed to downstream model development.

*ORCID: 0009-0003-4386-7894
†ORCID: 0000-0002-5618-0296

Standard subset selection relies on either exhaustive model training or single-source relevance measures. Exhaustive evaluation provides direct task performance estimates but scales poorly and must be repeated for each downstream architecture. Mutual information based methods offer a cheaper alternative by balancing target relevance against redundancy [5], but they may overlook synergy, whereby two sequences are jointly informative even when neither is individually dominant.

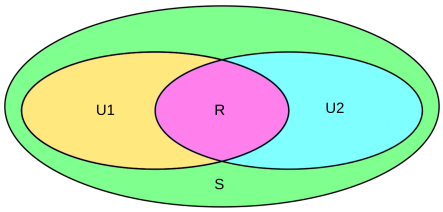

Figure 1: Diagram representing partial information decomposition for two sources, X_1 and X_2, relative to a target Y. The total information is partitioned into source-specific contributions, U_1 and U_2, redundant information R provided by both sources, and synergistic information S available only from their joint observation.

Partial information decomposition (PID) separates target information into redundant, unique, and synergistic components [6, 7]. It has been used to define feature relevance and guide feature selection [8, 9] , while recent works such as SFL-Net applied PID-guided latent factorization to multi-contrast MRI to PET synthesis [10] . Building on this work, prior studies of MRI sequence contributions in brain tumor segmentation [11, 12], and contrast-level Shapley attribution [13], we introduce a general n-to-2 pre-training framework for multi-contrast MRI tumor segmentation. A compact autoencoder is trained for each input sequence, the resulting embeddings are quantized, and all $\binom{n}{2}$ pairs are ranked using a region-aware minimum mutual information PID (MMI-PID) definition score [14]. The highest scoring pair is then selected for downstream model development. We evaluate the framework using four CoRe-BT MRI sequences [2], comparing the full four-input model, all six two-input combinations, and four single-input models with patient-level statistical testing and Shapley attribution under a restricted compute budget.

2 Method

2.1 PID Pair Selection

Figure 2 summarizes the selection process. For each training subject, we sampled 128 spatially corresponding $32 \times 32 \times 32$ patches from all MRI sequences and the tumor mask. 70% of patch centers were drawn from the tumor foreground. For each tumor region $c \in \{\text{edema}, \text{enhancing tumor}, \text{necrotic}\}$, tumor burden was defined as the fraction of patch voxels assigned to that region. Zero-burden patches formed class 0, while positive burdens were divided into three training-set quantile classes.

A separate shallow 3D autoencoder was trained for each MRI sequence to reconstruct its input patches using mean-squared error. The encoder comprised three $3 \times 3 \times 3$ convolutional blocks with stride two and 8, 16, and 32 output channels respectively. Each convolution was followed by instance normalization and a ReLU activation. For a $32 \times 32 \times 32$ input patch, the encoder produced a $32 \times 4 \times 4 \times 4$ feature map, which was flattened and linearly projected to a 32-dimensional latent vector. After training, the encoders were frozen and the decoders were discarded. For each sequence m, a separate K-means model with $K = 16$ was fitted to the training latent vectors, and each cluster assignment defined the discrete source code q_m [15, 16].

For every input pair (a, b) and tumor region c, mutual information was estimated from the

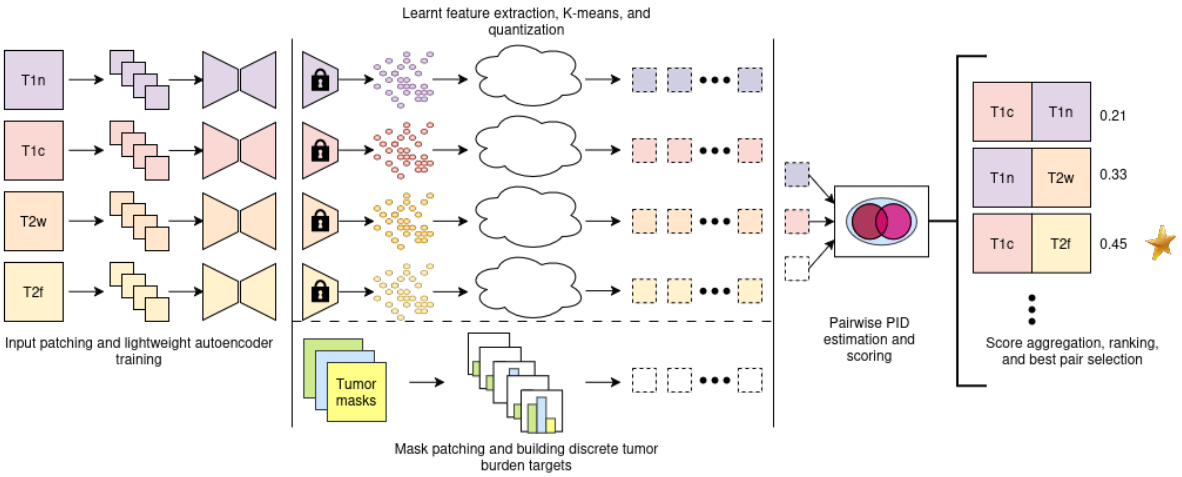

Figure 2: PID pair-selection pipeline. Spatially aligned patches from each MRI sequence are encoded using sequence-specific autoencoders. The frozen latent representations are quantized into discrete source codes, while aligned tumor masks provide discretized regional tumor-burden targets. All two-input combinations are evaluated using region-aware PID, ranked by their aggregate scores, and the highest-ranked pair is retained for downstream model training.

empirical joint distribution of the two source codes and the discretized tumor-burden target T_c [17]. Following the bivariate PID framework [6], we decomposed the information that each pair of quantized source codes provided about the regional tumor-burden target. We used the MMI formulation as a computationally simple proxy for bivariate PID [14]. For input pair (a, b) and tumor region c, the PID terms were

$$\begin{aligned} R^c_{ab} &= \min\{I(q_a; T_c), I(q_b; T_c)\}, \\ U^c_{a|b} &= I(q_a; T_c) - R^c_{ab}, \qquad U^c_{b|a} = I(q_b; T_c) - R^c_{ab}, \\ S^c_{ab} &= I((q_a, q_b); T_c) - R^c_{ab} - U^c_{a|b} - U^c_{b|a}. \end{aligned} \tag{1}$$

To rank the candidate pairs, we defined an aggregate score inspired by PID-based feature selection criteria that favor informative and non-redundant feature sets [8, 9]:

$$\text{Score}(a, b) = \sum_c w_c \left[I((q_a, q_b); T_c) + \lambda \left(U^c_{a|b} + U^c_{b|a} + S^c_{ab} \right) \right], \qquad \lambda = 0.5, \tag{2}$$

where w_c was the normalized inverse frequency of positive patches for region c. Thus, joint target information determined the baseline pair score, while unique and synergistic information received an additional weight relative to redundant information. The selected pair was

$$\mathcal{M}^* = \arg \max_{\{a,b\} \subset \mathcal{M}} \text{Score}(a, b), \qquad |\mathcal{M}^*| = 2. \tag{3}$$

2.2 Lightweight 3D U-Net

The downstream segmentation model was a lightweight 3D U-Net with three resolution levels [3]. Each encoder and decoder block contained two $3 \times 3 \times 3$ convolutions followed by ReLU activations and Instance Normalization. Max pooling was used for downsampling, while nearest-neighbor interpolation followed by $1 \times 1 \times 1$ convolution was used for upsampling. A final $1 \times 1 \times 1$ convolution produced the output logits.

Models were trained using $64 \times 64 \times 64$ patches, a batch size of one, and 256 patches per epoch. The training objective combined Dice loss and binary cross-entropy, balancing region overlap with voxelwise supervision. Optimization used AdamW with a weight decay of 10^{-4} [18]. The learning rate was reduced from 10^{-3} to 10^{-5} using a cosine schedule over 200 epochs [19]. A 30 subject validation cohort was used for development stage benchmarking, and final evaluation was performed on an independent held-out 28 subject test cohort.

3 Experiments

We used 132 training, 30 validation, and 28 test subjects from the MRI and tumor-mask component of CoRe-BT dataset [2]. Subjects missing any required sequence were excluded. Inputs were T1n, T1c, T2w, and T2-FLAIR. Reported targets were edema, enhancing tumor, and necrotic/non-enhancing tumor. Volumes were cropped, clipped to the 0.5th–99.5th percentile range, and normalized for training.

PID selected T1c+T2-FLAIR as the highest-ranked input pair for this task. We then trained 11 architecturally identical 3D U-Net models (one four-input model, all six pairwise models, and four single-input models). This design tested whether the PID ranking predicted downstream segmentation performance.

The models were evaluated on the held-out set using patient-level macro-averaged Dice, HD95, sensitivity, and precision across tumor regions. Statistical analyses included 10,000 bootstrap resamples, Spearman correlation, paired Wilcoxon tests, Friedman tests with Holm correction, and a -0.03 Dice non-inferiority margin [20]. Shapley values over all 16 input coalitions were computed for the four-input model, replacing omitted inputs with training set channel means inside the brain mask and zeros outside [13, 21].

The complete experiment used one RTX 3090 with GPU memory capped at 2 GB, four logical CPU cores, and 4 GB of system memory for training and feature extraction.

4 Analysis and Results

4.1 PID Selection and Segmentation

PID ranked T1c+T2-FLAIR first with a score of 0.963. Across the six pair models, the PID score was positively associated with test-set Dice ($\rho = 0.600$, $p = 0.208$) and negative HD95 ($\rho = 0.600$, $p = 0.208$). Both associations were directionally consistent with downstream performance, but neither was statistically significant, and the analysis had limited power because only six pairs were available.

Table 1: Region-specific PID scores for all input-pair combinations. Nec., Ed., and Enh. denote necrotic tumor, edema, and enhancing tumor, respectively.

Input pair	Nec.	Ed.	Enh.	Aggregate	Rank
T1c+T2-FLAIR	**0.789**	**1.106**	**1.032**	**0.963**	**1**
T1c+T2w	0.785	0.777	0.947	0.841	2
T1n+T2-FLAIR	0.745	0.932	0.736	0.793	3
T2w+T2-FLAIR	0.669	0.849	0.651	0.712	4
T1n+T2w	0.662	0.680	0.590	0.642	5
T1n+T1c	0.455	0.409	0.728	0.540	6

On the test cohort, the full model achieved average Dice 0.687 and HD95 16.09. T1c+T2-FLAIR achieved Dice 0.676 and HD95 16.77, making it the strongest pair and second overall (Table 2). The selected pair retained 98.5% of the full-model mean Dice while using half the input channels.

Table 2: Test-set performance for the top input configurations. Ed, En, and Ne denote edema, enhancing tumor, and necrotic tumor, respectively. HD95, average Dice, sensitivity, and precision are patient-level macro averages across the three regions.

Inputs	Ch.	HD95	Avg. Dice	Sens.	Prec.	Dice Ed	Dice En	Dice Ne
All four	4	**16.09**	**0.687**	**0.709**	**0.714**	0.550	**0.800**	**0.710**
T1c+T2-FLAIR	2	16.77	0.676	**0.709**	0.699	**0.570**	0.767	0.692
T1c+T2w	2	18.06	0.650	0.670	0.683	0.520	0.721	0.707
T1c	1	21.71	0.589	0.606	0.622	0.534	0.574	0.658
T1n+T1c	2	18.21	0.587	0.594	0.612	0.488	0.613	0.661

The selected vs. full Dice difference was -0.010 (95% CI $[-0.052, 0.034]$; paired Wilcoxon $p = 0.0505$). HD95 increased by 0.68 (95% CI $[-4.54, 5.09]$; $p = 0.174$). Sensitivity was nearly identical (difference < 0.001; $p = 0.991$), and precision differed by -0.015 ($p = 0.099$). Dice non-inferiority was not established given our choice of margin.

Model configuration affected Dice ($\chi^2_{10} = 164.63$, $p < 10^{-29}$, Kendall's $W = 0.588$) and HD95 ($\chi^2_{10} = 79.07$, $p < 10^{-12}$, $W = 0.282$). The selected pair significantly outperformed most lower-performing pairwise and single-input models after Holm correction. However, differences between the selected pair and either the full model or T1c+T2w were not significant after correction.

4.2 Post-hoc Attribution

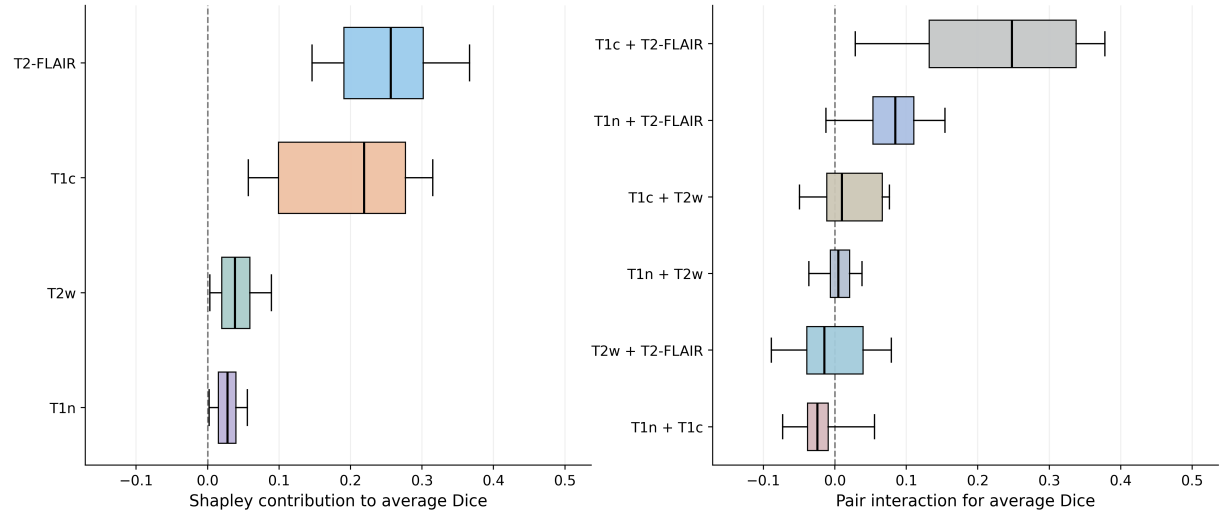

Figure 3: Test-set Shapley analysis of the four-input model. T2-FLAIR and T1c had the largest input contributions, and their pairwise interaction was the strongest.

Shapley analysis ranked T2-FLAIR first (0.336; 95% CI 0.298–0.374) and T1c second (0.258; 95% CI 0.203–0.311). Their interaction was also largest (0.307; 95% CI 0.241–0.373; Fig. 3). Pair interactions differed across the six combinations ($\chi^2_5 = 71.22$, $p < 10^{-13}$, Kendall's $W = 0.509$). This independent analysis agreed with the PID-selected input pair.

5 Discussion

5.1 PID Identified the Strongest Input Pair

The central result is that the MMI-PID score selected the strongest two-input configuration before segmentation training. T1c+T2-FLAIR ranked first by the information score and was independently identified as the best pair in the 11-model evaluation. This agreement supports the use of the proposed score as a pre-training input pair selector.

The selected model used half the input channels and retained 98.5% of the full input model mean Dice. Paired tests did not provide clear evidence of differences in Dice, HD95, sensitivity, a larger study is required to reliably draw conclusions. However, our work only serves as a proof of concept and is limited by only 4 input contrasts and small dataset splits.

5.2 The Selected Pair Matches the Imaging Task

The selection of T1c and T2-FLAIR is consistent with the reported regions. T1c directly emphasizes enhancing tumor, while T2-FLAIR is sensitive to edema and surrounding tissue abnormality. Their combination therefore covers two distinct components of the segmentation target. The region-specific PID scores reflected this pattern as (Table 1). T1c+T2-FLAIR achieved the highest score for all three evaluated regions, including edema, enhancing tumor, and necrotic tumor.

Shapley attribution provided a separate model-based audit. T2-FLAIR and T1c had the two largest individual contributions, and their interaction was larger than every other pair interaction. However, PID and Shapley address different questions. PID ranks inputs before downstream training, whereas Shapley measures how a specific trained full model uses its inputs under a defined missing-input baseline. Their agreement strengthens the interpretation that T1c+T2-FLAIR captures both strong individual signal and useful cross-input dependence.

5.3 Implications for Resource-Constrained Training

Selection occurs before costly segmentation training. For n candidate inputs, the method trains n shallow autoencoders and evaluates $\binom{n}{2}$ pair scores, avoiding a separate downstream model for every pair.

These pre-training steps are comparatively lightweight and reusable across architectures. Reducing the number of inputs also lowers data-transfer, preprocessing, latency, augmentation, memory, and computation costs.

6 Conclusion

In this work, our framework using shallow autoencoder encoding, latent quantization, and a region-aware MMI-PID score identified the strongest two-input configuration before segmentation training. The framework selected T1c+T2-FLAIR, which halved the number of input channels, retained 98.5% of the full-model test Dice, and was independently supported by Shapley attribution. This study provides a focused proof of concept for PID based input selection as a practical and interpretable pre-training strategy for reducing multi-contrast 3D input complexity. Future work should evaluate its robustness in larger cohorts and settings with more than four candidate inputs.

References

[1] Bjoern H. Menze et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.

[2] Juampablo E. Heras Rivera, Daniel K. Low, Xavier Xiong, Jacob J. Ruzevick, Daniel D. Child, Wen-wai Yim, Mehmet Kurt, and Asma Ben Abacha. CoRe-BT: A multimodal radiology-pathology-text benchmark for robust brain tumor typing. *arXiv preprint arXiv:2603.03618*, 2026.

[3] Özgün Çiçek, Ahmed Abdulkadir, Sören S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d U-Net: Learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, volume 9901 of *Lecture Notes in Computer Science*, pages 424–432. Springer, 2016.

[4] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18:203–211, 2021.

[5] Hanchuan Peng, Fuhui Long, and Chris Ding. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, 2005.

[6] Paul L. Williams and Randall D. Beer. Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515*, 2010.

[7] Nils Bertschinger, Johannes Rauh, Eckehard Olbrich, Jürgen Jost, and Nihat Ay. Quantifying unique information. *Entropy*, 16(4):2161–2183, 2014.

[8] Patricia Wollstadt, Sebastian Schmitt, and Michael Wibral. A rigorous information-theoretic definition of redundancy and relevancy in feature selection based on (partial) information decomposition. *Journal of Machine Learning Research*, 24(131):1–44, 2023.

[9] Charles Westphal, Stephen Hailes, and Mirco Musolesi. Partial information decomposition for data interpretability and feature selection. In Yingzhen Li, Stephan Mandt, Shipra Agrawal, and Emtiyaz Khan, editors, *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pages 1873–1881. PMLR, 2025.

[10] Agamdeep S. Chopra, Caitlin Neher, Tianyi Ren, Juampablo E. Heras Rivera, Hesamoddin Jahanian, and Mehmet Kurt. SFL-Net: Source-factorized latent representation learning for multi-contrast MRI to tau-PET synthesis. *arXiv preprint arXiv:2602.22545*, 2026.

[11] Tongxue Zhou, Stéphane Canu, Pierre Vera, and Su Ruan. Feature-enhanced generation and multi-modality fusion based deep neural network for brain tumor segmentation with missing MR modalities. *Neurocomputing*, 466:102–112, 2021.

[12] Yousef Sadegheih and Dorit Merhof. Segmentation of brain metastases in MRI: A two-stage deep learning approach with modality impact study. In Islem Rekik, Ehsan Adeli, Sang Hyun Park, and Celia Cintas, editors, *Predictive Intelligence in Medicine*, volume 15155 of *Lecture Notes in Computer Science*, pages 196–206. Springer, 2024.

[13] Tianyi Ren, Juampablo E. Heras Rivera, Hitender Oswal, Yutong Pan, Agamdeep S. Chopra, Jacob J. Ruzevick, and Mehmet Kurt. Here comes the explanation: A shapley perspective on multi-contrast medical image segmentation. In *Joint Proceedings of the xAI 2025 Late-Breaking Work, Demos and Doctoral Consortium*, volume 4017 of *CEUR Workshop Proceedings*, pages 185–192, 2025.

[14] Adam B. Barrett. Exploration of synergistic and redundant information sharing in static and dynamical gaussian systems. *Physical Review E*, 91(5):052802, 2015.

[15] Geoffrey E. Hinton and Ruslan R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[16] Stuart Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

[17] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Hoboken, NJ, 2 edition, 2006.

[18] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

[19] Ilya Loshchilov and Frank Hutter. SGDR: Stochastic gradient descent with warm restarts. In *International Conference on Learning Representations*, 2017.

[20] Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

[21] Lloyd S. Shapley. A value for n-person games. In Harold W. Kuhn and Albert W. Tucker, editors, *Contributions to the Theory of Games II*, number 28 in Annals of Mathematics Studies, pages 307–317. Princeton University Press, Princeton, NJ, 1953.