Beyond Core and Penumbra: Bi-Temporal Image-Driven Stroke Evolution Analysis

Md Sazidur Rahman[1] MD.S.RAHMAN@UIS.NO
Kjersti Engan[1] KJERSTI.ENGAN@UIS.NO
Kathinka Dæhli Kurz[1,2] KATHINKA.DEHLI.KURZ@SUS.NO
Mahdieh Khanmohammadi[1] MAHDIEH.KHANMOHAMMADI@UIS.NO
[1] *Department of Electrical Engineering and Computer Science, University of Stavanger, Norway*
[2] *Stavanger Medical Imaging Laboratory, Stavanger University, Hospital, Norway*

Editors: Under Review for MIDL 2026

Abstract

Computed tomography perfusion (CTP) at admission is routinely used to estimate ischemic core and penumbra, whereas follow-up diffusion-weighted MRI (DWI) obtained after treatment provides the definitive infarct outcome. However, single time-point segmentations do not capture the biological heterogeneity of stroke and ignore its continuous temporal evolution. We propose a bi-temporal analysis framework that characterizes ischemic tissue using statistical descriptors, radiomic texture features, and deep feature embeddings from two architectures (mJ-Net and nnU-Net). Bi-temporal refers to the admission (T_1) and the first post-treatment follow-up (T_2). All features are extracted at T_1 from CTP, and follow-up DWI is aligned with CTP to ensure spatial correspondence. Manually delineated masks at T_1 and T_2 are intersected to construct six regions of interest encoding both initial tissue state and final outcome. Extracted features were aggregated per region and analyzed in feature space. Evaluation on 18 patients with successful reperfusion demonstrated meaningful clustering of region-level representations. Regions classified as penumbra or healthy at T_1 that ultimately recovered exhibited feature similarity to preserved brain tissue, whereas infarct-bound regions formed distinct groupings.

Deep feature spaces, particularly mJ-Net, showed strong separation between salvageable and non-salvageable tissue, while nnU-Net further revealed a gradient-like organization with core and healthy brain at opposite ends and penumbra occupying transitional positions. These findings suggest that encoder-derived feature manifolds may reflect underlying tissue phenotypes and progression trajectories, providing an inshight toward imaging-based quantification of stroke evolution.

Keywords: Perfusion imaging, CTP, Diffusion MRI, DWI, Radiomic features, CNN embedding, Stroke progression analysis.

1. Introduction

Computed tomography perfusion (CTP) is widely used at the hyper-acute stage at hospital admission, referred to as T_1 in this paper, to quantify cerebral hemodynamics and estimate ischemic core and penumbra volumes. Ischemia is not clearly visible in raw CTP. However, from time-attenuation curves, perfusion parametric maps are derived to enable rapid assessment of tissue viability. The parametric maps include cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT), time to peak (TTP), and time-to-maximum (Tmax) (Wintermark et al., 2006). Using parametric maps, severely hypoperfused regions

are typically classified as ischemic core, whereas moderately hypoperfused regions represent penumbra, tissue that may still be salvageable with timely reperfusion.

Magnetic resonance diffusion-weighted imaging (DWI) is highly sensitive to irreversible cellular injury and is considered the definitive marker of infarcted brain tissue (Yang et al., 2022). Although DWI visualizes ischemia more clearly than CT, its longer acquisition time means it is typically obtained 24-72 hours post-stroke at the first follow-up time point (T_2). These images reveal the eventual infarct outcome, serving as ground truth for which threatened tissue ultimately died or survived (Yang et al., 2022).

Segmentation approaches on CTP and DWI provide operational definitions of core, penumbra and final infarct regions. Yet these methods often fail to capture the full biological heterogeneity, temporal dynamics of ischemic tissue injury, and how it evolves with treatment (Werdiger et al., 2023). Current stroke guidelines allow treatment in selected patients up to 24 hours after onset (Mainali et al., 2021). In practice, patients with small infarct cores and large regions of hypoperfused but salvageable tissue can benefit from reperfusion therapies well beyond the traditional time window of 4.5 or 6 hours (Mainali et al., 2021; Werdiger et al., 2023; Nogueira et al., 2018).

Imaging phenotyping involves extracting and defining meaningful patterns or subtypes from imaging data that reflect underlying tissue physiology or disease states (Van Timmeren et al., 2020). Population-scale imaging studies show that automated lesion segmentation yields robust imaging phenotypes (e.g., lesion volume and topography) that stratify patients and reflect stroke subtypes, highlighting the value of standardized quantitative descriptors over ad-hoc assessments (Wu et al., 2019). Trajectory-based analyses reveal that recovery patterns cluster into distinct, predictable subtypes, indicating that stroke phenotypes are dynamic processes rather than single time-point observations (Krishnagopal et al., 2022). The phenotyping in ischemic stroke means defining tissue-state subtypes that are not only relying on a single stroke time point but rather be defined using bi-temporal imaging signatures. It links stroke lesions admission-time perfusion characteristics to follow-up diffusion outcome lesion.

Recent advances in deep learning have significantly improved the estimation of ischemic core and penumbra from CTP (Tomasetti et al., 2022a,b). Instead of relying solely on derived perfusion maps, modern architectures leverage spatiotemporal CTP data to learn representations that more accurately segment at-risk tissue (Tomasetti et al., 2023). However, while deep learning enables high performance, the learned latent features that drive segmentation accuracy remain underexplored. We want to look at what characterizes an ischemic region at T_1 that ends up in a different region at T_2.

textbfContributions: In this paper we propose a framework, shown in figure 1, for quantitative bi-temporal analysis of ischemic tissue evolution. At admission (T_1), core, penumbra, and healthy regions are characterized using image-driven signatures comprising statistical descriptors, radiomic texture features, and deep convolutional neural network (CNN) embeddings. Deep representations are derived from two segmentation architectures, a 2D+time mJNet model (Tomasetti et al., 2020) and a 2D nnU-Net framework (Isensee et al., 2021), whose internal feature maps may provide multi-level descriptors of tissue state beyond segmentation outputs. These admission-time signatures are then tracked to follow-up imaging to determine which regions recover and which progress to infarction. Multiple

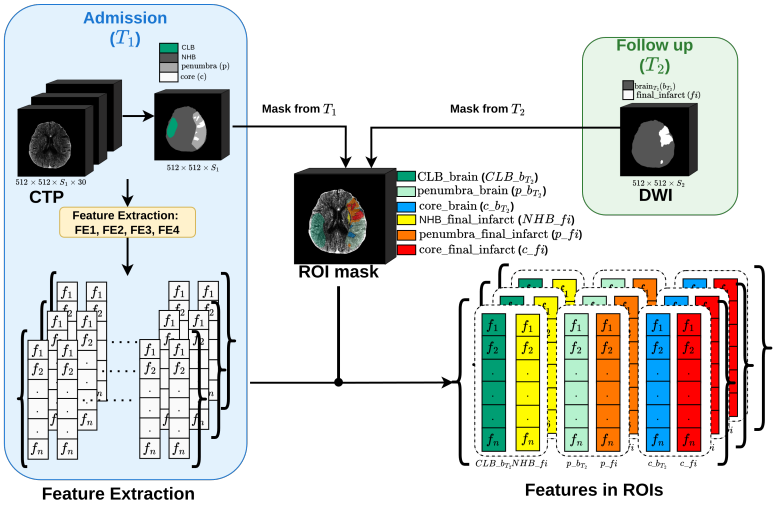


Figure 1: Framework of the proposed method. Acute ischemic regions from CTP at T_1 are combined with final tissue outcome from DWI at T_2 to form six ROI classes. Image-derived features from four extraction methods are computed on CTP and grouped according to ROI membership.

region of interest (ROI) are constructed by combining manually labeled regions in CTP taken at T_1 and in DWI obtained at T_2.

We hypothesize that the progression of ischemic tissue and its outcome can be described as trajectories or clustering patterns in feature space, where each coordinate represents a feature encoding intensity, texture, or latent CNN representation.

2. Dataset Description

The data was acquired at Stavanger University Hospital at two different time points: at admission (T_1) full CT protocol including non-contrast CT (NCCT), CT angiography and CT perfusion scans are obtained. Then after treatment at the first follow up (T_2) usually (24 to 72) hours later, magnetic resonance imaging (MRI) protocol is run and diffusion weighted images (DWI), are obtained. In this paper we use CTP from T_1 and DWI from T_2. CTP studies are acquired over a slab of the brain during the first pass of iodinated contrast (Omnipaque $350mg/ml$) followed by isotonic saline (total $80ml$; $6ml/s$, $4s$ delay). Each study comprises 512×512 images of 13-22 brain slices of resolution $0.425 \times 0.425 \times 5$ mm and a temporal sampling of 30 time points, the first 20 sampled with 1 sec interval, the last 10 at 2 sec. DWI scans are obtained as 3D volumes with typically 24–60 brain slices in the axial axis and in-plane resolution between (176×176) and (384×384).

The full dataset includes 152 acute ischemic stroke patients, where 109 have also gone through the MRI protocol. Among those 61 showed large vessel occlusion (LVO), where in 48 of them blood flow were preserved after treatment. We refer to them as recanalized patients in the reminder of this article. Among those 18 patients were selected for this analysis based on the availability of follow-up DWI imaging, successful spatial co-registration between CTP and DWI, and complete manual annotations of ischemic regions.

3. Preprocessing and region of interest construction

The CTP is 4-dimensional (4D) data we denote as $V_0(x, y, z, t) \in \mathbb{R}^{512\times512\times S_1\times30}$, were x and y are the spatial coordinates (resolution 512×512), S_1 corresponds to the number of brain slices taken at T_1,indexed by z and varies between (13-22). t is the time index ($t \in \{1 \cdots 30\}$). The CTP volumes were rigidly registered to the first time point using an intensity-based similarity transform to correct for motion within the perfusion sequence. An automatic brain extraction method (Najm et al., 2019) was used for skull stripping to remove bone, scalp, and other non-brain tissues. After skull stripping, Gamma correction ($\gamma = 0.5$) was applied to suppress low-intensity noise and enhance contrast in hypoperfused regions, followed by histogram equalization to normalize brightness variations across scans, described in (Tomasetti et al., 2023). We denote the preprosessed CTP volumes as $V(v, t)$ where we define $\Omega \subset \mathbb{Z}^3$ and $v \in \Omega$ as the volumetric index. Looking at CTP slice by slice but including all time points we denote this as $V(\bar{x}, z_i)$, $\bar{x} \in [x, y, t]$ for a specific slice z_i.

3.1. Region of interest construction

The registration and region of interest definitions are illustrated in figure 2.

T_1 and T_2 alignment: The DWI volume typically has lower in-plane resolution and covers a larger cranial extent than CTP. Therefore we need to register the DWI slices to the CTP brain slab. Thus, the DWI volumes $D_0(x, y, z)$ were first cropped in the z-direction to match the axial slices overlapping the CTP slab. The cropped DWI volume was then registered to the corresponding CTP using a similarity transform driven by an intensity-based similarity metric and resampled onto the CTP grid (in-plane 512×512). This procedure ensured spatial alignment between the follow-up infarct and the acute perfusion sequence. We denote the registered DWI-CTP pair as $(D(v), V(v, t))$, and for one slice $(D(\tilde{x}, z_1), V(\bar{x}, z_1))$, where $\tilde{x}$ corresponds to ($\bar{x}$ but exclusive the time axis).

Ground-truth labels were produced by two expert neuroradiologists through manual delineation using an in-house MATLAB-based software (Tomasetti et al., 2023, 2022b). For the T_1 time point, the full CT examination, including Non-Contrast CT, CT angiography and preprocessed CTP, and perfusion maps(CBV, CBF, TTP, and maximum intensity projection (MIP)) was used. Regions with increased TTP/Tmax and reduced CBF but preserved CBV were considered penumbra (p), whereas additional CBV reduction indicated ischemic core (c). DWI volume served as reference for generating the final infarct mask, $fi \in \Omega$, and the remaining brain voxels at T_2 were labeled as $b \in \Omega$. To account for small boundary inconsistencies, the ischemic core mask ($c \in \Omega$) was morphologically dilated by a small structure element, and in our work the penumbra ($p \in \Omega$) is defined as the marked penumbra region excluding all voxels belonging to the core. To avoid perfusion-related bias, a region from the contra-lateral hemisphere of the brain at T_1 was annotated and used as healthy brain tissue, called CLB $\in \Omega$. The remaining brain tissue at T_1 is denoted Not hypo-perfused brain: NHB $\in \Omega$.

Bi-temporal ROI logic: To model tissue evolution over time, defined regions at T_1 were combined with regions at T_2 through region mapping and set intersections. Formally, each voxel v was assigned to one of the following ROIs based on the rule:

$$ROI_{R_{T1}}^{R_{T2}} := R_{T1} \cap R_{T2}$$

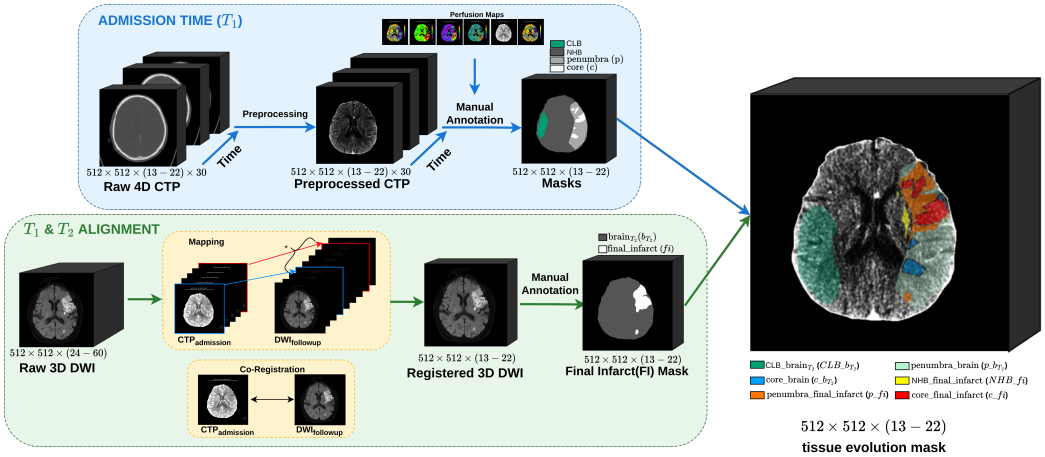


Figure 2: Construction of the tissue bi-temporal masks by combining manual delineation at admission (T_1) and follow-up (T_2). At T_1, preprocessed 4D CTP is annotated into brain, penumbra and core. At T_2, DWI is mapped to CTP slab and DWI co-registered to CTP is annotated for final infarct. The T_1 and T_2 masks are then merged into six-ROI mask.

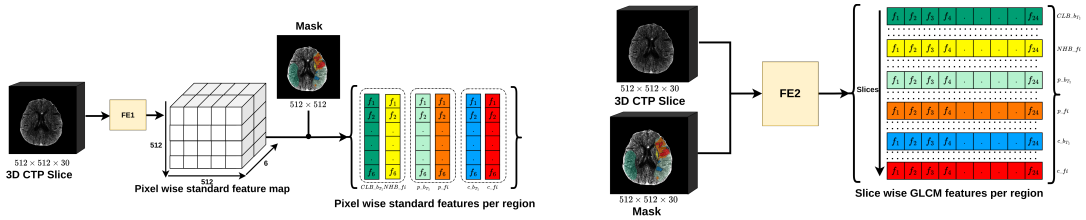
This resulted in a categorical bi-temporal tissue evolution maps $ROI_{R_{T1}}^{R_{T2}} \in \Omega$ that gives the regions for our further investigations:

ROI_{CLB}^{b}	healthy tissue from contralateral hemisphere at admission and remained healthy
ROI_{NHB}^{fi}	tissue initially appearing normal (not hypo-perfused(NHB)) but later infarcted
ROI_{p}^{b}	penumbral tissue that recovered fully
ROI_{p}^{fi}	penumbra that progressed to infarction
ROI_{c}^{b}	core tissue showing partial or unexpected recovery
ROI_{c}^{fi}	core that remained infarcted at follow-up

4. Methods

To characterize ischemic tissue, we do image feature extraction (FE) from preprocessed CTP and relate them to final outcome on follow-up DWI. We consider three feature families baseline statistical descriptors, radiomic texture metrics, and deep CNN embeddings which together capture complementary aspects of tissue heterogeneity and provide a compact representation of stroke evolution between the two time points.

(FE1) Baseline (BL) statistical features are directly derived from 3D CTP data $V(\bar{x}, z_i)$. To extract the features, we use a sliding window $\mathcal{W}(\bar{x}) \in \mathbb{R}^{3\times3\times30}$ centered around $\bar{x}$. From the intensities in $V(\mathcal{W}(\bar{x}), z_i)$, we compute six first order statistics: mean, standard deviation, skewness, kurtosis, minimum and maximum (see Appendix A.1 for formulas). Let $g : \mathbb{R}^{3\times3\times30} \to \mathbb{R}^6$ denote the feature operator. The baseline feature vector for each slice and location $\bar{x}$ is $\boldsymbol{f}_{z_i}^{\mathrm{BL}}(\bar{x}) = g\big(V(\mathcal{W}(\bar{x}), z_i)\big) \in \mathbb{R}^6$. Location wise feature vectors are mapped to bi-temporal tissue evolution masks (ROI_{RT1}^{RT2}). For each patient pt, slice z_i and bi–temporal ROI (Fig. 2), all locations in that ROI are aggregated by element–wise max pooling into a



(*a*) Baseline features (*b*) GLCM features

Figure 3: Baseline and GLCM feature extraction. (a) Baseline features are computed using a $3 \times 3 \times 30$ sliding window on CTP and aggregated into region-wise vectors via tissue evolution masks. (b) Radiomic-based 3D GLCM features are extracted per slice and assigned to each tissue evolution class.

single descriptor $\boldsymbol{F}^{\mathrm{BL}}_{pt,\mathrm{ROI},z_i} \in \mathbb{R}^6$, so each tissue–evolution class and slice is represented by one compact feature vector.

(FE2) Radiomic features capture higher order spatial texture beyond first order intensity statistics. We extract 3D gray level co–occurrence matrix (GLCM) features (Haralick et al., 1973) from $V(\bar{x}, z_i)$ using `PyRadiomics` (Van Griethuysen et al., 2017). For each patient pt and axial slice z_i, the volume $V(\bar{x}, z_i)$ is paired with a 3D tissue evolution mask of identical size, obtained by repeating the 2D bi–temporal ROI mask over all 30 time indexes. Symmetrical GLCMs are computed with bin width 8 and $\delta = 1$ (26 connectivity over 13 directions), yielding 24 GLCM features per slice (Fig. 3(*b*)). From these, we retain four descriptors: IMC1 and IMC2 (information measures of correlation 1 and 2), MCC (maximal correlation coefficient), and Correlation (formal definitions in Appendix A.2). IMC1/IMC2 quantify dependence between co–occurring gray levels, MCC reflects texture complexity, and Correlation measures linear dependence between neighbours. For each patient pt, slice z_i and bi–temporal ROI, the four values are concatenated into $\boldsymbol{F}^{\mathrm{glcm}}_{pt,\mathrm{ROI},z_i} \in \mathbb{R}^4$, yielding one radiomic feature vector per tissue ROI and slice.

(FE3 & FE4) Deep CNN-based embeddings are capturing hierarchical, data–driven representations of ischemic tissue. We derive embeddings from two CTP segmentation encoders trained on our cohort (excluding the 18 recanalized LVO patients): (i) 2D+time mJ-Net (Tomasetti et al., 2020) that segments core and penumbra from CTP, and (ii) 2D nnU-Net (Isensee et al., 2021), also a segmentation model where each axial slice is a 512×512 image with 30 time points stacked as channels (fig. 4). In both cases, the trained networks are used in inference mode purely as feature extractors, without further fine–tuning.

For mJ-Net, spatio–temporal patches of size $16 \times 16 \times 30$ are densely sampled over $V(\bar{x}, z_i)$ (stride 1), passed through the encoder, and activations from the last convolutional block ($4 \times 4 \times 1 \times 256$) are globally average–pooled. The resulting 256D vector is assigned to the patch centre, yielding a dense map of location–wise deep features. For each patient pt, slice z_i and bi–temporal ROI, all vectors within the ROI are aggregated by element–wise max pooling to form a slice–level mJ-Net embedding $\boldsymbol{F}^{\mathrm{mJNet}}_{pt,\mathrm{ROI},z_i} \in \mathbb{R}^{256}$.

For nnU-Net, intermediate feature maps of size $256 \times 64 \times 64$ (channels $\times$ height $\times$ width) are extracted from $V(\bar{x}, z_i)$. Bi–temporal ROI masks are downsampled to 64×64

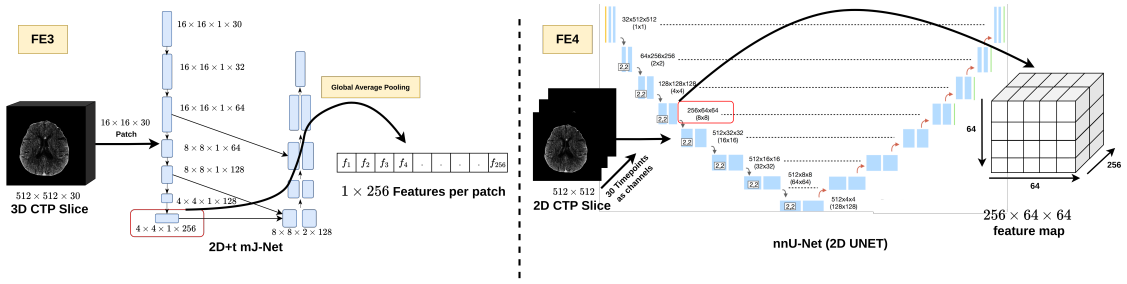


Figure 4: CNN-based deep feature extraction. Left: the 2D+time mJ-Net encoder processes $16 \times 16 \times 30$ patches and applies global average pooling on the last convolutional block to yield a 256D feature vector per patch. Right: the nnU-Net encoder produces a $256 \times 64 \times 64$ feature map, providing a 256D feature vector at each position of a downsampled 64×64 grid.

using nearest–neighbor interpolation so that each grid location inherits a tissue label. For each patient pt, slice z_i and ROI, all 256D feature vectors whose grid locations fall inside the downsampled ROI are max-pooled to obtain a slice–level nnU-Net embedding $\boldsymbol{F}^{\text{nnUNet}}_{pt,\text{ROI},z_i} \in \mathbb{R}^{256}$. These embeddings act as compact CNN-based descriptors of local CTP dynamics for downstream similarity analysis, clustering and visualization (see Appendix A.3 for details).

4.1. Assessment techniques

Box plots summarize the distribution of baseline statistical and GLCM radiomic features across bi–temporal ROIs, providing a compact view of central tendency, variability and class–level overlap to qualitatively assess feature separability.

t-SNE visualization qualitatively assess the geometric structure of feature representations. The high-dimensional feature vectors were projected into two-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) (Maaten and Hinton, 2008). The resulting embeddings visualize whether samples belonging to different tissue evolution ROI exhibit separable structure in feature space.

Cosine similarities quantify how similarly the CNN encoders represent tissue evolution ROIs and their value can vary from -1 to 1. In cosine space, "direction" reflects whether tissue regions share a common latent evolution pattern. Therefore, cosine similarities close to zero imply representational differences and in case of negative values imply opposing representational trajectory. We consider 1 as high and -1 as low for cosine similarity interpretation. For each patient pt, ROI and slice z_i, let $\boldsymbol{F}^{\text{cnn}}_{pt,ROI,z_i} \in \mathbb{R}^{256}$ denote the ℓ_2–normalized feature vector extracted from mJ-Net or nnU-Net. Slice–level vectors within an ROI are aggregated by element–wise max pooling, $\boldsymbol{\psi}_{pt,ROI} = \max_{z_i} \boldsymbol{F}^{\text{cnn}}_{pt,ROI,z_i}$. We define $S_{pt}(ROI_1, ROI_2) = \cos\big(\boldsymbol{\psi}_{pt,ROI_1}, \boldsymbol{\psi}_{pt,ROI_2}\big)$ as *per–patient* cosine similarity and $\bar{S}(ROI_1, ROI_2) = \frac{1}{|\mathcal{P}|} \sum_{pt \in \mathcal{P}} S_{pt}(ROI_1, ROI_2)$ as *group–level* where $\mathcal{P}$ is the set of patients for which both ROIs are present.

5. Results and discussion

The proposed framework was evaluated on 18 recanalized LVO patients. We examined 4 categories of features including BL, GLCM, and deep embeddings from mJ-Net and nnU-

Net across six bi-temporal ROIs defined by combining CTP and DWI labels. The CNN models were trained on all 152 patients excluding the 18 LVO patients. Figure 5 shows box plots of BL and GLCM features, where colors match ROI colors in bi-temporal mask.

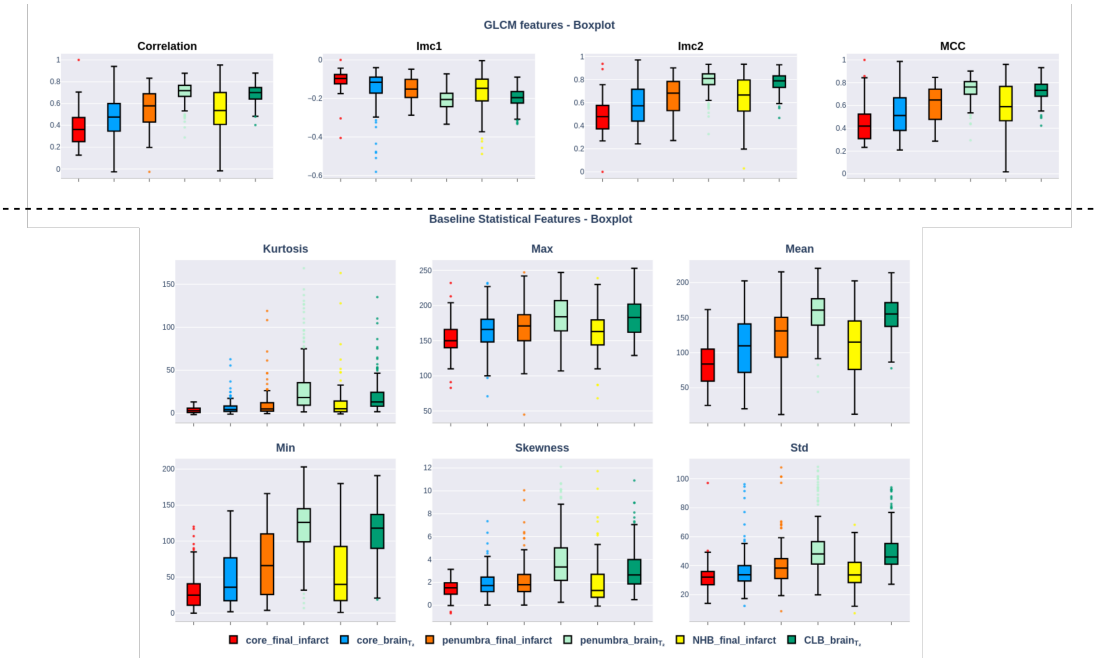


Figure 5: Box plots of GLCM and statistical features over the six ROIs on slice level for 18 recanalized LVO patients.

Across both GLCM and BL features a consistent trend emerges. Healthy tissue that remains viable (ROI^{b}_{CLB}) exhibits high median values, narrow interquartile ranges, and low variability, reflecting a stable structural phenotype. In contrast, core tissue evolving into final infarct (ROI^{fi}_{c}) shows lower values and greater dispersion, consistent with a disrupted or degenerating phenotype. Salvaged penumbra (ROI^{b}_{p}) generally lies between these extremes and often more closely resembles healthy tissue, suggesting a reversible phenotype. Among GLCM descriptors Imc2 provides the clearest separation between healthy/penumbra and infarcted tissue. ROI^{fi}_{c} and ROI^{b}_{c} display similar distributions across both GLCM and baseline features, suggesting that early core tissue at T_1 already shares textural properties with its eventual infarct outcome. Overall, while baseline intensity features show the same directional trend as GLCM, they offer weaker separation across regions, supporting the observation that texture-based metrics capture ischemic degradation more effectively than raw intensity values.

Figure 6 shows the t-SNE embeddings of FE1-FE4. Symbols encode anatomical region (circle: brain, square: penumbra, cross: core), while colour represents tissue outcome from T_1 to T_2 (green: salvaged to brain, red: progressing to final infarct). BL and GLCM feature spaces show partial overlap between tissue types, but a visible tendency for salvaged regions

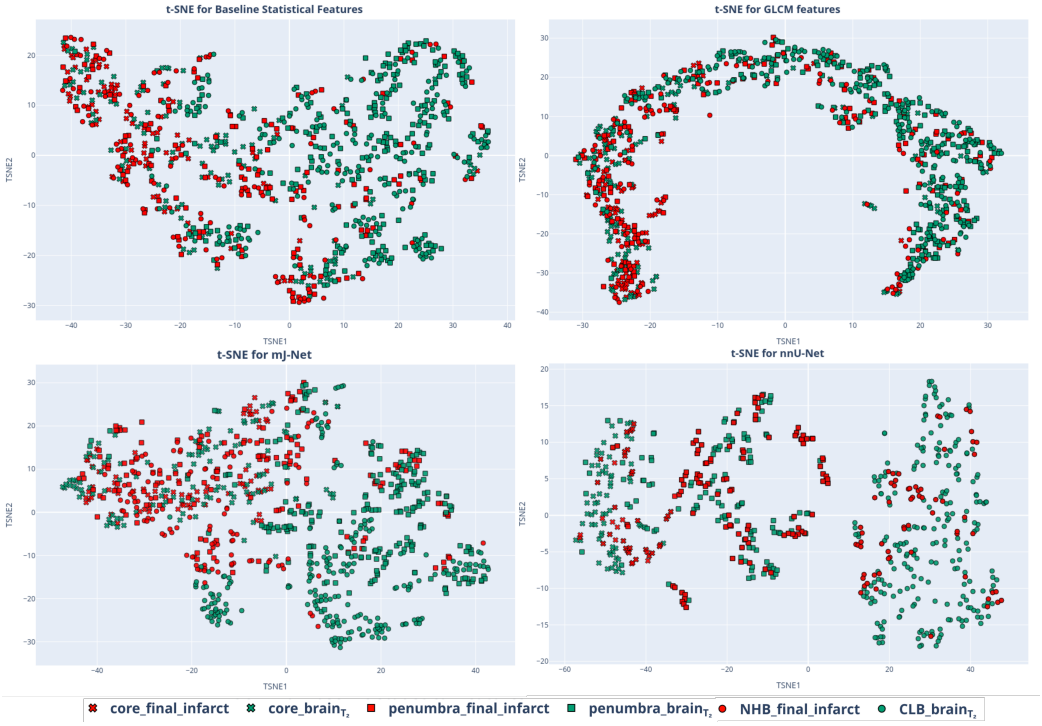


Figure 6: t-SNE for mapping distribution of features in six ROIs, a) BL, b) GLCM, c) mJ-Net and d) nnU-Net. Each point represents one ROI in one brain slice.

(ROI^b_{CLB}, ROI^b_p, ROI^b_c) to cluster toward one side and infarct-evolving regions (ROI^{fi}_{NHB}, ROI^{fi}_p, ROI^{fi}_c) toward the other, with GLCM features forming slightly tighter clusters.

The mJ-Net, exhibits more distinct partitioning between salvaged and infarct-prone tissue. The nnU-Net embedding shows an apparent gradient structure, with core tissue positioned at one extreme, healthy brain at the opposite end, and penumbra forming a transitional band between them. Penumbra (ROI^b_p and ROI^{fi}_p) occupy intermediate regions, reflecting their mixed outcome trajectories. Regions destined for infarction form compact clusters, whereas partially or fully salvaged regions display broader spread, consistent with greater biological heterogeneity.

Table 1: Cosine similarity of CNN features (a) mJ-Net, (b) nnU-Net.

(a) mJ-Net

	ROI^{b}_{CLB}	ROI^{fi}_{NHB}	ROI^{b}_{p}	ROI^{fi}_{p}	ROI^{b}_{c}	ROI^{fi}_{c}
ROI^{b}_{CLB}	**1.00**	0.11	0.46	0.25	0.07	-0.13
ROI^{fi}_{NHB}	0.11	**1.00**	-0.12	0.41	0.35	0.47
ROI^{b}_{p}	0.46	-0.12	**1.00**	0.25	0.02	-0.28
ROI^{fi}_{p}	0.25	0.41	0.25	**1.00**	0.36	0.43
ROI^{b}_{c}	0.07	0.35	0.02	0.36	**1.00**	0.65
ROI^{fi}_{c}	-0.13	0.47	-0.28	0.43	0.65	**1.00**

(b) nnU-Net

	ROI^{b}_{CLB}	ROI^{fi}_{NHB}	ROI^{b}_{p}	ROI^{fi}_{p}	ROI^{b}_{c}	ROI^{fi}_{c}
ROI^{b}_{CLB}	**1.00**	-0.78	0.59	-0.44	-0.52	-0.67
ROI^{fi}_{NHB}	-0.78	**1.00**	-0.40	0.47	0.56	0.64
ROI^{b}_{p}	0.59	-0.40	**1.00**	-0.15	-0.10	-0.21
ROI^{fi}_{p}	-0.44	0.47	-0.15	**1.00**	0.53	0.56
ROI^{b}_{c}	-0.52	0.56	-0.10	0.53	**1.00**	0.57
ROI^{fi}_{c}	-0.67	0.64	-0.21	0.56	0.57	**1.00**

The cosine similarity values are shown in Table 1. Both for mJ-Net and nnU-Net high alignment is observed for core turning to brain (ROI_c^b) or ending up as final infarct(ROI_c^{fi}), 0.65 in mJ-Net and 0.57 in nnU-Net. None hypo-perfused region at T_1 turning to final infarct, ROI_{NBH}^{fi}, shows high similarities to core and penumbra turning to infarct (ROI_c^{fi}, ROI_p^{fi}), 0.47, 0.41 in mJ-Net and 0.64, 0.47 in nnU-Net, confirming the observation in t-SNE plots. Features obtained from nnU-Net show strong polarity between salvageable and infarct-bound tissue (e.g. ROI_{CLB}^b, ROI_c^{fi} (-0.67) and ROI_{NBH}^{fi}, ROI_{CLB}^b (0.78)), similar values for mJ-Net are (-0.13) and (0.11) respectively.

The nnU-Net encodes recovered tissue at one extreme and infarct tissue at the opposite extreme. These findings indicate that nnU-Net representations more clearly separate ROIs and can encode outcome-related phenotypes better than mJ-Net. In t-SNE plots for nnU-Net we saw less distinction between salvaged and infarct-prone tissue however this is observed in cosine similarities. This can be because cosine similarities are calculated on class-level aggregated embeddings and while local samples may overlap (seen in t-SNE plots), their pooled ROI-level centroids can point in opposite directions (negative values in cosine similarities).

Interestingly, the NHB region, which was manually annotated as none-hypoperfused, frequently aligns with tissue evolving into final infarct, suggesting that the embedded feature spaces do not fully agree with initial expert labels. Another limitation arises from co-registration between CTP and DWI. The two modalities differ in acquisition geometry, and CTP is frequently acquired with a head tilt to increase brain coverage under dose constraints. This introduces an axial angulation mismatch that standard 2D registration cannot fully compensate for, potentially affecting voxel-level correspondence between tissue states at T_1 and T_2.

6. Conclusion

We presented an image-driven framework for stroke tissue evolution characterization and used statistical, textural and deep embedding features. We observed that healthy tissue from contralateral hemisphere showed a different characteristic compared to ischemic core at admission regardless of its outcome and salvaged penumbra lies between them but more similar to CLB. None-hypoperfused at admission but infarct-evolving regions clustering with core/infarct suggests that learned representations may capture latent vulnerability not apparent in initial expert labels. Clear trajectory from core, penumbra to healthy tissue was seen specially in nnU-Net. Overall, our findings support the feasibility of using acute CTP-derived feature spaces as a basis for imaging phenotyping of stroke tissue evolution. While the study is limited by cohort size and reliance on a single centre, the proposed framework opens avenues for larger validation and for introducing stroke tissue phenotypes into stroke analysis.

7. Compliance with Ethical Standards

This study was conducted in accordance with institutional guidelines and approved by the regional ethics committee and all data were pseudonymised prior to analysis.

References

Robert M. Haralick, K. Sam Shanmugam, and Its'hak Dinstein. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.*, 3:610–621, 1973. URL https://api.semanticscholar.org/CorpusID:206786900.

Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211, 2021.

Sanjukta Krishnagopal, Keith Lohse, and Robynne Braun. Stroke recovery phenotyping through network trajectory approaches and graph neural networks. *Brain Inf.*, 9(1):13, June 2022. ISSN 2198-4026. doi: 10.1186/s40708-022-00160-w. URL https://doi.org/10.1186/s40708-022-00160-w.

Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.

Shraddha Mainali, Marin E. Darsie, and Keaton S. Smetana. Machine learning in action: Stroke diagnosis and outcome prediction. *Frontiers in Neurology*, Volume 12 - 2021, 2021. ISSN 1664-2295. doi: 10.3389/fneur.2021.734345. URL https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2021.734345.

Mohamed Najm, Hulin Kuang, Alyssa Federico, Uzair Jogiat, Mayank Goyal, Michael D Hill, Andrew Demchuk, Bijoy K Menon, and Wu Qiu. Automated brain extraction from head ct and cta images using convex optimization with shape propagation. *Computer Methods and Programs in Biomedicine*, 176:1–8, 2019.

Raul G Nogueira, Ashutosh P Jadhav, Diogo C Haussen, Alain Bonafe, Ronald F Budzik, Parita Bhuva, Dileep R Yavagal, Marc Ribo, Christophe Cognard, Ricardo A Hanel, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *New England Journal of Medicine*, 378(1):11–21, 2018.

Luca Tomasetti, Kjersti Engan, Mahdieh Khanmohammadi, and Kathinka Dæhli Kurz. Cnn based segmentation of infarcted regions in acute cerebral stroke patients from computed tomography perfusion imaging. In *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*, pages 1–8, 2020.

Luca Tomasetti, Liv Jorunn Høllesli, Kjersti Engan, Kathinka Dæhli Kurz, Martin Wilhelm Kurz, and Mahdieh Khanmohammadi. Machine Learning Algorithms Versus Thresholding to Segment Ischemic Regions in Patients With Acute Ischemic Stroke. *IEEE Journal of Biomedical and Health Informatics*, 26(2):660–672, February 2022a. ISSN 2168-2208. doi: 10.1109/JBHI.2021.3097591. URL https://ieeexplore.ieee.org/document/9488246/?arnumber=9488246. Conference Name: IEEE Journal of Biomedical and Health Informatics.

Luca Tomasetti, Mahdieh Khanmohammadi, Kjersti Engan, Liv Jorunn Høllesli, and Kathinka Dæhli Kurz. Multi-input segmentation of damaged brain in acute ischemic stroke patients using slow fusion with skip connection. *nldl*, 3, March 2022b. ISSN 2703-6928.

doi: 10.7557/18.6223. URL https://septentrio.uit.no/index.php/nldl/article/view/6223.

Luca Tomasetti, Kjersti Engan, Liv Jorunn Høllesli, Kathinka Dæhli Kurz, and Mahdieh Khanmohammadi. Ct perfusion is all we need: 4d cnn segmentation of penumbra and core in patients with suspected acute ischemic stroke. *IEEE Access*, 11:138936–138953, 2023. ISSN 2169-3536.

Joost JM Van Griethuysen, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina GH Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo JWL Aerts. Computational radiomics system to decode the radiographic phenotype. *Cancer research*, 77(21):e104–e107, 2017.

Janita E Van Timmeren, Davide Cester, Stephanie Tanadini-Lang, Hatem Alkadhi, and Bettina Baessler. Radiomics in medical imaging—"how-to" guide and critical reflection. *Insights into imaging*, 11(1):91, 2020.

Freda Werdiger, Mark W. Parsons, Milanka Visser, Christopher Levi, Neil Spratt, Tim Kleinig, Longting Lin, and Andrew Bivard. Machine learning segmentation of core and penumbra from acute stroke ct perfusion data. *Frontiers in Neurology*, Volume 14 - 2023, 2023. ISSN 1664-2295. doi: 10.3389/fneur.2023.1098562. URL https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2023.1098562.

Max Wintermark, Adam E. Flanders, Birgitta Velthuis, Reto Meuli, Maarten van Leeuwen, Dorit Goldsher, Carissa Pineda, Joaquin Serena, Irene van der Schaaf, Annet Waaijer, James Anderson, Gary Nesbit, Igal Gabriely, Victoria Medina, Ana Quiles, Scott Pohlman, Marcel Quist, Pierre Schnyder, Julien Bogousslavsky, William P. Dillon, and Salvador Pedraza. Perfusion-ct assessment of infarct core and penumbra. *Stroke*, 37 (4):979–985, 2006. doi: 10.1161/01.STR.0000209238.61459.39. URL https://www.ahajournals.org/doi/abs/10.1161/01.STR.0000209238.61459.39.

Ona Wu, Stefan Winzeck, Anne-Katrin Giese, Brandon L. Hancock, Mark R. Etherton, Mark J.R.J. Bouts, Kathleen Donahue, Markus D. Schirmer, Robert E. Irie, Steven J.T. Mocking, Elissa C. McIntosh, Raquel Bezerra, Konstantinos Kamnitsas, Petrea Frid, Johan Wasselius, John W. Cole, Huichun Xu, Lukas Holmegaard, Jordi Jiménez-Conde, Robin Lemmens, Eric Lorentzen, Patrick F. McArdle, James F. Meschia, Jaume Roquer, Tatjana Rundek, Ralph L. Sacco, Reinhold Schmidt, Pankaj Sharma, Agnieszka Slowik, Tara M. Stanne, Vincent Thijs, Achala Vagal, Daniel Woo, Stephen Bevan, Steven J. Kittner, Braxton D. Mitchell, Jonathan Rosand, Bradford B. Worrall, Christina Jern, Arne G. Lindgren, Jane Maguire, Natalia S. Rost, on behalf of the MRI-GENIE, and GISCOME Investigators. Big data approaches to phenotyping acute ischemic stroke using automated lesion segmentation of multi-center magnetic resonance imaging data. *Stroke*, 50(7):1734–1741, 2019. doi: 10.1161/STROKEAHA.119.025373. URL https://www.ahajournals.org/doi/abs/10.1161/STROKEAHA.119.025373.

Wenjin Yang, Jan W. Hoving, Miou S. Koopman, Manon L. Tolhuisen, Henk van Voorst, Olvert A. Berkheme, Jonathan M. Coutinho, Ludo F. M. Beenen, and Bart J. Emmer.

Agreement between estimated computed tomography perfusion ischemic core and follow-up infarct on diffusion-weighted imaging. *Insights into Imaging*, 13(1):191, December 2022. ISSN 1869-4101. doi: 10.1186/s13244-022-01334-0. URL https://doi.org/10.1186/s13244-022-01334-0.

Appendix A. Supplementary Materials

A.1. FE1: Baseline (BL) statistical features

We provide the formulas for the baseline statistical features used in Sec. 4. Consider the sliding window $\mathcal{W}(\bar{x}) \in \mathbb{R}^{3\times3\times30}$ denote the local spatio–temporal window centred at $\bar{x}$ as defined in Sec. 4. The intensities in $V(\mathcal{W}(\bar{x}), z_i)$ are collected into a vector $\{\zeta_j\}_{j=1}^{N}$ with $N = 3 \times 3 \times 30 = 270$.

$$\mu(\bar{x}) = \frac{1}{N}\sum_{j=1}^{N}\zeta_j \qquad \sigma(\bar{x}) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\zeta_j - \mu(\bar{x}))^2}$$
$$\mathrm{skew}(\bar{x}) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\zeta_j - \mu(\bar{x})}{\sigma(\bar{x})}\right)^3 \qquad \mathrm{kurt}(\bar{x}) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\zeta_j - \mu(\bar{x})}{\sigma(\bar{x})}\right)^4 - 3$$
$$v_{\min}(\bar{x}) = \min_{1\le j\le N}\zeta_j \qquad v_{\max}(\bar{x}) = \max_{1\le j\le N}\zeta_j$$

Here $\mu(\bar{x})$ and $\sigma(\bar{x})$ denote the local mean and standard deviation, $\mathrm{skew}(\bar{x})$ and $\mathrm{kurt}(\bar{x})$ are the skewness and kurtosis of the local intensity distribution, and $v_{\min}(\bar{x})$, $v_{\max}(\bar{x})$ are the minimum and maximum attenuation values within $\mathcal{W}(\bar{x})$. The "-3" in $\mathrm{kurt}(\bar{x})$ sets the kurtosis of a normal distribution to zero for easier interpretation.

Let $g : \mathbb{R}^{3\times3\times30} \to \mathbb{R}^6$ denote the baseline feature operator that maps a window $\mathcal{W}(\bar{x})$ to the six statistics above. The baseline feature vector at slice z_i and location $\bar{x}$ is then

$$\boldsymbol{f}_{z_i}^{\mathrm{BL}}(\bar{x}) = g\big(\mathcal{W}(\bar{x})\big) = \big[\mu(\bar{x}),\, \sigma(\bar{x}),\, \mathrm{skew}(\bar{x}),\, \mathrm{kurt}(\bar{x}),\, v_{\min}(\bar{x}),\, v_{\max}(\bar{x})\big]^\top \in \mathbb{R}^6.$$

For each patient pt, slice z_i and bi–temporal region of interest ROI, we define the slice–level baseline descriptor by element–wise max pooling,

$$\boldsymbol{F}_{pt,\mathrm{ROI},z_i}^{\mathrm{BL}} = \max_{\bar{x}\in\mathrm{ROI}_{pt,z_i}} \boldsymbol{f}_{z_i}^{\mathrm{BL}}(\bar{x}) \in \mathbb{R}^6,$$

where ROI_{pt,z_i} denotes the set of locations $\bar{x}$ on slice z_i assigned to that bi–temporal tissue–evolution class.

A.2. FE2: Radiomic features

This appendix gives the formal definitions (Haralick et al., 1973) of the four GLCM–based radiomic descriptors used in Sec. 4. All notation and preprocessing (3D GLCMs computed from $V(\bar{x}, z_i)$ and the corresponding 3D tissue–evolution masks, bin width 8, $\delta = 1$, 26–connectivity) follow Sec. 4.

Let ϵ be an arbitrary small positive number, $p(i, j)$ denote the normalized gray–level co–occurrence matrix (GLCM) for a fixed distance/orientation pair, and let N_g be the number of gray levels. The marginal distributions are

$$p_x(i) = \sum_{j=1}^{N_g} p(i,j), \qquad p_y(j) = \sum_{i=1}^{N_g} p(i,j),$$

with means μ_x, μ_y and standard deviations σ_x, σ_y. The entropies of p_x, p_y and $p(i,j)$ are

$$HX = -\sum_{i=1}^{N_g} p_x(i)\log_2\big(p_x(i)+\epsilon\big), \qquad HY = -\sum_{j=1}^{N_g} p_y(j)\log_2\big(p_y(j)+\epsilon\big),$$
$$HXY = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log_2\big(p(i,j)+\epsilon\big).$$

The Haralick cross–entropy terms are

$$HXY1 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log_2\big(p_x(i)p_y(j)+\epsilon\big),$$

$$HXY2 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_x(i)p_y(j)\log_2\big(p_x(i)p_y(j)+\epsilon\big).$$

Using these quantities, the four radiomic descriptors are defined as

$$\mathrm{Imc1} = \frac{HXY - HXY1}{\max\{HX, HY\}}, \quad \mathrm{Imc2} = \sqrt{1-\exp\big(-2\,(HXY2 - HXY)\big)},$$
$$\mathrm{MCC} = \sqrt{\lambda_2(Q)}, \qquad \mathrm{corr} = \frac{\sum_{i,j=1}^{N_g} i\,j\,p(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y},$$

where $\lambda_2(Q)$ is the second–largest eigenvalue of

$$Q(i,j) = \sum_{k=1}^{N_g} \frac{p(i,k)\,p(j,k)}{p_x(i)\,p_y(k)}.$$

IMC1 and IMC2 quantify the statistical dependence between co–occurring gray levels (IMC1 in $[-1,0]$, IMC2 in $[0,1)$), MCC measures texture complexity (typically in $[0,1]$), and corr captures the linear dependence of neighbouring gray levels. For each patient pt, slice z_i and bi–temporal ROI, these four quantities are computed on the corresponding 3D tissue–evolution mask and concatenated into the radiomic feature vector

$$\boldsymbol{F}^{\mathrm{glcm}}_{pt,\mathrm{ROI},z_i} = \big[\mathrm{Imc1},\ \mathrm{Imc2},\ \mathrm{MCC},\ \mathrm{corr}\big] \in \mathbb{R}^4,$$

A.3. FE3 & FE4

(FE3 & FE4) Deep CNN-based embeddings provide slice-wise representations derived from two pretrained CTP segmentation encoders: a 2D + time mJ-Net variant and a 2D nnU-Net (see main text and Fig. 4). In both cases, the networks are used in inference mode as frozen feature extractors; no additional fine-tuning is applied when generating embeddings.

mJ-Net encoder (FE3). From each 3D CTP slice $V(\bar{x}, z_i)$, we extract spatio–temporal patches $16 \times 16 \times 30$ on a dense grid with unit in–plane stride. We use last encoder block to get features of size $4 \times 4 \times \ 1 \times 256$. Global average pooling over the spatial and depth axes yields a 256-dimensional descriptor, which we associate with the central in–plane location, giving a dense map of location-wise features $\boldsymbol{f}_{z_i}^{\mathrm{mJNet}}(\bar{x}) \in \mathbb{R}^{256}$. These maps are aligned with the bi-temporal tissue–evolution masks. For each patient pt, slice z_i and ROI, all $\boldsymbol{f}_{z_i}^{\mathrm{mJNet}}(\bar{x})$ whose locations fall inside the corresponding region are aggregated by element-wise max pooling,

$$\boldsymbol{F}_{pt,\mathrm{ROI},z_i}^{\mathrm{mJNet}} = \max_{\bar{x} \in \mathrm{ROI}_{pt,z_i}} \boldsymbol{f}_{z_i}^{\mathrm{mJNet}}(\bar{x}) \in \mathbb{R}^{256},$$

yielding one 256D mJ-Net embedding per slice and tissue–evolution class.

nnU-Net encoder (FE4). For the nnU-Net backbone, we extract intermediate encoder feature maps of size $256 \times 64 \times 64$ (channels $\times$ height $\times$ width) for each slice z_i. Let $\boldsymbol{f}_{z_i}^{\mathrm{nnUNet}}(u, v) \in \mathbb{R}^{256}$ denote the feature vector at grid location (u, v) in this 64×64 map. The bi-temporal ROI masks defined at 512×512 resolution are downsampled to 64×64 using nearest–neighbour interpolation so that each (u, v) inherits a unique tissue label. For each patient pt, slice z_i and ROI, the nnU-Net features whose grid locations lie within the downsampled region are aggregated by element-wise max pooling,

$$\boldsymbol{F}_{pt,\mathrm{ROI},z_i}^{\mathrm{nnUNet}} = \max_{(u,v) \in \mathrm{ROI}_{pt,z_i}^{64\times 64}} \boldsymbol{f}_{z_i}^{\mathrm{nnUNet}}(u, v) \in \mathbb{R}^{256}.$$

Thus, both mJ-Net (FE3) and nnU-Net (FE4) provide one compact 256D deep embedding per patient, slice and bi-temporal ROI, which is used for similarity analysis, clustering and visualization in the main experiments.