SCIENTIFIC CULTURE
SCIENTIFIC CULTURE, Vol. 12, No.5.1, (2026), pp. 160-170
Open Access. Online & Print

www.sci-cult.com

DOI: 10.5281/zenodo.12511014

MATHEMATICAL MODELING AND IMAGE PROCESSING TECHNIQUES FOR EARLY DETECTION OF TUMOR HETEROGENEITY IN MEDICAL IMAGING

[1*]Dr. C. Balaji, [2]Er. Rishabh Aryan, [3]Rachit Ahuja, [4]Dr. Prabhat Bansal, [5]Dr. Krishnan Bandyopadhyay, [6]Prem Kumar

[1]Assistant Professor, Department of Computer Applications, SRM Institute of Science and Technology Tiruchirappalli, Email ID: balajicj80@gmail.com Orcid ID: 0009-0002-5892-5498*
[2]M.Tech (Artificial Intelligence and Data Science), Department of Computer Science and Engineering,, Indian Institute of Information Technology, Bhagalpur (Bihar), E-mail: rishabh.250201011@iiitbh.ac.in, Orcid ID: 0009-0004-7595-9440
[3]Assistant Professor, Department of Radiation Oncology, Shri Guru Ram Rai Institute of Medical and Health Sciences, Email ID: rachitahuja.dr@gmail.com, Orcid ID: 0000-0002-7946-0987
[4] Assistant Professor, Institute of Applied Sciences, Mangalayatan University, Aligarh, Uttar Pradesh, Email ID: prabhat.bansal@mangalayatan.edu.in, Orcid ID: 0000-0001-9380-8700
[5]Assistant Professor, Department of Medical Electronics Engineering, Dayananda Sagar College of Engineering, Kumaraswamy Layout, Bangalore, Karnataka 560111, Orcid ID: 0000-0003-0147-936X Email ID: krishnan-ml@dayanandasagar.edu
[6]Assistant Professor Computer Science & Engineering (CS&DS) Brainware University, Kolkata, Email ID: premashokcv@gmail.com Orcid ID: 0009000032282821

Abstract
Tumor heterogeneity is a significant cause of treatment resistance and poor clinical outcomes, and the medical images used to diagnose and treat patients are routinely interpreted in a largely qualitative manner and may not reflect subtle patterns of heterogeneity early in disease progression. This study is aimed at proposing an integrated mathematical modeling and image processing framework for quantification of intra-tumoral heterogeneity and early detection using medical imaging. Paired magnetic resonance and computed tomography brain tumor images along with the expert-annotated segmentation masks were analyzed. Following the data ingestion and quality control process, images were standardized and modality specific preprocessing pipelines were followed. Tumor regions of interest were extracted in order to calculate morphological descriptors, first order radiomic features, texture measures, and multi-scale entropy maps. A composite heterogeneity index was developed from the combination of entropy statistics, dispersion measures, and spatial autocorrelation to describe both the variation of intensity and the organization of space in tumors. Slices-level heterogeneity measures were aggregated to patients level and an interpretable L1 regularized logistic regression model was applied for the early vs. advanced heterogeneity discrimination under patient-wise cross-validation. The proposed heterogeneity-informed model showed stable performance and obvious contributions of spatial heterogeneity components. The results show that mathematically derived heterogeneity indices can be robust and interpretable imaging biomarkers to aid in the early characterization of tumor complexity and provide motivation for their validation in larger, clinically annotated cohorts.

Keywords: Tumor heterogeneity; Radiomics; Mathematical modeling; Entropy mapping; Medical imaging

1.Introduction

Tumor heterogeneity is a basic property of a cancer due to spatial and temporal differences in the cell composition, molecular expression, and physiological behavior in a single tumor mass. Such heterogeneity has important implications on the disease course, therapeutic resistance, and patient prognosis, making it a major problem in modern oncology. Medical imaging modalities such as magnetic resonance imaging (MRI) and computed tomography (CT) are noninvasive methods of gaining access to tumor structure and tissue characteristics and thus can assess heterogeneity at clinically relevant spatial scales. However, the clinical interpretation of medical images is still mostly done qualitatively by observers, limiting the scope of detecting subtle intra-tumoral variations that may represent early-stage heterogeneity.

Radiomics has become a revolutionary paradigm with a novel approach that redefines medical images as quantitative data and not simply visual data. Through the systematic consolidation of high dimensional features that describe intensity distributions, texture patterns and spatial relationships, radiomics allows for a more complete characterization of tumor phenotype in addition to visual inspection (Gillies et al., 2016). Seminal work has been done to show the noninvasive decoding of tumour phenotype by radiomic features to uncover meaningful relations with underlying biological processes and clinical outcome (Aerts et al., 2014). Despite these advances, several radiomics-based investigations have relied on rather empirically derived sets of features and predictive modelling approaches that are not explicitly based on a mathematical theory and which can limit interpretability and reproducibility.

Mathematical modeling allows the development of a principled system [framework] for representing complexity, disorder and spatial organization in imaging data. Concepts drawn from information theory, multiscale signal processing and spatial statistics make it possible to express tumor heterogeneity in terms of interpretable mathematical descriptors. Wavelet-based representations allow for the multiresolution analysis of imaging signals which allow for capturing of localized variations in texture across different spatial scales and allows for the rigorous mathematical characterization of image structure (Daubechies, 1992). When combined with image processing and extraction of radiomic features, such mathematical techniques provide a systematic method of how to quantify heterogeneity in addition to descriptive statistics.

Nevertheless, there are still a number of methodological challenges. Many of the current methods highlight global or averaged features which may mask heterogeneity early in cancer development occurring in small tumor regions. Furthermore, the heterogeneity does not have any uniform mathematical formulations for comparability across studies and difficulty in biological interpretation. Prior investigations by MRI and PET imaging have shown the feasibility of the quantitative measurement of intratumoral heterogeneity, but these studies were mostly performed in advanced or well-established tumors where the signals of heterogeneity are more prominent (Asselin et al. 2012). Due to this, however, the detection of heterogeneity at earlier stages - when there is less tumour burden and the variation is subtle in space - is relatively underexplored.

The issue of robustness and reproducibility is also a persistent problem in the field of radiomics research. Features based on imaging may be sensitive to acquisition schemes, preprocessing decisions and segmentation variability and this brings up the issue of generalizability across datasets. The most recent reviews and methodological guidelines put a strong focus on the need to develop standardized, transparent, and robust radiomics pipelines in order to obtain reliable and reproducible results (Avanzo et al., 2025). Moreover, predictive modeling in radiomics can have complicated learning methods, which can be hard to interpret and can be overfitted, especially in small-sample studies. These aspects explain why interpretable modeling methods should be used with stringent validation procedures.

The current research is dedicated to the quantitative description of the tumor heterogeneity based on the two-dimensional MR and CT imaging data and the relative segmentation masks. Heterogeneity is considered at the imaging level, which focuses on spatial and textural heterogeneity in tumor regions as opposed to molecular or genomic heterogeneity. The absence of clinical staging or longitudinal outcome data would operationally define tumor burden proxies (imaging-derived) that were used to detect cancer early. Although this method allows conducting the systematic methodological study, it has limitations connected with the use of retrospective data, the limited sample size, and the impossibility to make the direct causal or clinical conclusion.

In this respect, the research would add to the literature by incorporating image processing methods with mathematically based heterogeneity metrics in a clear and reproducible framework. Heterogeneity quantification can be based on the use of entropy-based measures, multiscale texture analysis, and spatial statistical measures, which are interpretable. The receiver operating characteristic analysis is used to measure model performance, and it is a threshold-independent measure of classification ability (Hanley and McNeil, 1982). Precisionrecall analysis is also used to assess predictive performance when there exists a class imbalance, which is usual in the early detection case (Davis and Goadrich, 2006). The study contributes to the creation of credible imaging-based measures of tumor heterogeneity by ensuring that methodological decisions are made in accordance with the existing standards of robustness and interpretability.

The research objectives of this study are specific and they are:

- To create a mathematical-based model of quantifying intratumoral heterogeneity of MR and

CT images based on image processing and radiomic feature extraction methods.

- To build interpretable heterogeneity indices, which are multiscale texture responsive and spatially statistic responsive to characterize heterogeneity on early stages.
- To assess the usefulness of the proposed heterogeneity-based features in supporting early detection with patient-level predictive model.

2.Literature Review

Tumor heterogeneity is commonly known as a core property of cancer and it is defined as both spatial and temporal variations in cellular structure, genetic changes, and phenotypic behavior of a single tumor. This inherent variability facilitates treatment resistance, disease progression, and unpredictable clinical outcomes hence making it difficult to effectively manage cancer (Marusyk and Polyak, 2010). Consequently, there has been increased interest in coming up with quantitative techniques that are able to measure and describe heterogeneity in a noninvasive and clinically viable fashion.

Medical imaging has become one of the most effective methods of monitoring tumor heterogeneity because it can be used to visualize structural and functional differences throughout the entire tumor volume. Radiomics is a continuation of traditional imaging analysis that identifies high-dimensional quantitative characteristics that characterize intensity patterns, texture, shape, and spatial organization of medical images. The initial radiomics research has shown that latent information in imaging is not seen in visual evaluation but that features can be extracted systematically to provide biologically relevant tumor features (Lambin et al., 2012). Later studies further defined radiomics as a transition point between medical imaging and personalized medicine, where radiomics can be used to stratify risks and plan treatment on an individual basis (Lambin et al., 2017).

The key element of heterogeneity analysis using radiomics is the extraction of texture features, which is meant to measure the spatial associations among pixel intensities. One of the most popular texture descriptors is gray-level co-occurrence matrix (GLCM) features that were first introduced to describe image patterns to classify them (Haralick et al., 2007). These characteristics correspond to second-order statistical characteristics of contrast, homogeneity, and correlation allowing the quantitative evaluation of spatial complexity in tumor areas. Micro-texture Complementary texture measures, which are defined on the basis of local binary patterns, represent an alternative description of micro-texture, and encode local intensity differences in a rotationally invariant form, with the advantage that they are resistant to changes in illumination (Ojala et al., 1996).

In addition to the texture descriptors, mathematical concepts are important towards formalizing tumor heterogeneity. Information theory gives a natural methodology of measuring uncertainty and disorder in imaging data. The use of entropy as a quantitative measure of information content has been extensively used to describe variability and randomness of intensity within regions of tumors, which is why it is especially appropriate to characterize heterogeneity (Shannon, 1948). Entropy-based metrics are particularly useful in the initial analysis, when the slight differences are not presented as significant structural changes.

Spatial statistics also provide a better modeling of heterogeneity by considering the spatial dependence of imaging features. I by Moran is a classical spatial autocorrelation measure that determines the extent to which similar values are clustered in space, thus, providing an opportunity to evaluate spatial organization in tumors (Moran, 1950). The analysis of heterogeneity can be improved by considering spatial autocorrelation to distinguish noise at random and biologically significant spatial patterns and enhance the interpretability of the metrics derived using imaging.

Proper heterogeneity analysis requires proper tumor segmentation and definition of region of interest as well. Overlap measures have also been extensively applied to determine similarity between segmentation masks to guarantee consistency in region delineation e.g. the Sorensen coefficient (Sorensen, 1948). Meaningful feature extraction is not possible in the absence of reliable segmentation since the errors in region definition can spread in the radiomics pipeline and corrupt the estimates of heterogeneity.

The fact that the dimensionality of a radiomic feature set is growing is an issue that requires effective feature selection and regularization bars to curb overfitting with the aim of improving interpretations. The least absolute shrinkage and selection operator (LASSO) offers an objective method of feature selection that puts sparsity restrictions on the predictors to keep only the most informative ones (Tibshirani, 1996). The LASSO-based modeling is especially appropriate to radiomics studies that have small sample size because it strikes a balance between predictive power and simplicity of models.

In order to make radiomics analysis reproducible and scalable, computational models have been created that standardize the process of feature extraction and modeling. Open-source systems allow the open implementation of radiomics pipelines and allow the same features to be computed across studies, which leads to methodological rigor and reproducibility (Van Griethuysen et al., 2017). Nonetheless, inconsistency in purchasing procedures and processing measures is still an issue, which leads to the establishment of official standardization programs.

Image Biomarker Standardization Initiative (IBSI) is one of the largest attempts to standardize radiomic features definitions and computational practices throughout the research community. IBSI is expected to enhance comparability and reliability of radiomics studies and hasten clinical practice adoption through the provision of standardized nomenclature and reference values and validation guidelines (Zwanenburg et al., 2020). Such standards are being considered to be a necessary requirement of the credibility of radiomics-based heterogeneity research.

Nevertheless, despite the significant improvements, there are still significant limitations. Radiomics characteristics are likely to be sensitive to image acquisition parameters, variability in segmentation, and preprocessing options that can have an impact on robustness and generalizability. Critical assessment of radiomics applications has indicated the following limitations and the necessity to carefully validate and report (Yip & Aerts, 2016). To overcome these difficulties, it is necessary to have integrated frameworks, which entail strong image processing, mathematically based heterogeneity measures, and interpretable modeling strategies.

Taken together, the available literature highlights the significance of image processing, radiomics, and mathematical modeling in the quest to develop the quantitative description of tumor heterogeneity. Nevertheless, there is a limited use of such integrated solutions in early detection situations. This is where a gap is created to foster frameworks that are focused on spatially resolved heterogeneity modeling, interpretability, and robustness, especially in environments with small tumor burden and weak heterogeneity signals.

3. Methodology

3.1 Research Design

The research design used in this study is retrospective and quantitative imaging-analysis design, which utilized secondary medical imaging data. This approach involves a methodological framework, which incorporates image preprocessing, region-of-interest (ROI) analysis, radiomics, model, spatial heterogeneity and interpretable machine-learning-based classification, within one pipeline, functional, and reproducible pipeline completion system.

All the analyses were performed with Python in a Jupyter Notebook, which allows performing all processing stages transparently, starting with the ingestion of raw DICOM and ending with the final performance analysis. In order to achieve methodological rigor, patient-wise validation was used to ensure that training and testing data were strictly separated to avoid information leakage across cross-validation folds. The general design aims at interpretability, reproducibility, and robustness, which are critical factors to consider in studies on medical imaging with limited sample sizes.

3.2 Data Collection

3.2.1 Imaging Data

The data consisted of paired magnetic resonance (MR) and computed tomography (CT) brain tumor images that had corresponding expert-validated segmentation masks, and they were in Digital Imaging and Communications in Medicine (DICOM) format. The information was obtained using a publicly available MR–CT brain imaging data set that is intended to aid in quantitative neuroimaging and cross-modality analysis (Abu-Srhan et al., 2021; Al-Kadi et al., 2022). Images and masks were sorted by patient and modality and filenames coded patient identifiers and slice indices to be able to match the image and mask correctly.

The data is represented by two-dimensional image slices (axial) obtained with the help of the RadiAnt DICOM Viewer and converted into the standardized DICOM format with a spatial resolution of 256 x 256 pixels. A total of 179 axial slices of 20 patient volumes were used, 90 MR slices and 89 CT slices, each with a tumor segmentation mask.

A Siemens Verio 3T scanner with a slice thickness of 5.0 mm was used to obtain MR images. Each scan utilized T2-weighted sequences with no contrast enhancement, three fat-saturation pulses, repetition time 2500-4000 ms, echo time 20-30 ms and flip angle 90o/180o. The spatial resolution in-plane was about 0.7 x 0.6 mm 2, which gave a voxel resolution of 0.7 x 0.6 x 5 mm 3. A CT scanner was used to obtain CT images with a slice thickness of 7.0 mm, tube voltage of 130 kV, tube current of 113 to 327 mAs and dose length product of 2.46 mGy cm. An acquisition protocol of a topogram with 64 dual-source was employed and smooth and sharp reconstruction filters were employed. The resultant CT voxel resolution was in the order of 0.6 x 0.6 x 7 mm 3.

A custom Python-based data loader ingested all DICOM files and pixel arrays were extracted directly out of the DICOM objects. Where necessary, images were converted to two-dimensional grayscale representations to make sure there is consistency in processing stages. Preliminary data quality tests confirmed proper image-mask matching, equal spatial resolution and the lack of damaged or inaccessible files and all slices were stored to be analyzed further.

3.2.2 Standardization and Quality Control

All images and masks were brought to the same resolution of 256x 256 pixels to have a uniform spatial representation. Binarization of segmentation masks was done to depict tumor and background classes only. Automated quality control steps ensured that all standardized images had met dimensional constraints and all masks had valid binary values. Fractions of tumor area were calculated to ensure that tumor areas took a physiologically reasonable fraction of each slice of the image.

3.3 Population and Sampling

3.3.1 Study Population

The population of the study was comprised of patients where there was a valid MR or CT image slice with a tumor mask. Even though the extraction of features was done at the slice level, all the modelling and assessment activities were done at the patient level to maintain clinical relevance.

3.3.2 Sampling Strategy

Aggregation of patient level features provided a modeling dataset of 20 modality-specific records of patients representing 14 distinct patients. Supervised learning experiments were only conducted on patients who passed predetermined requirements in downstream labeling. The explicit grouping variables were patient identifiers in cross-validation to make sure that no patient provided data to both training and testing sets.

3.4 Data Analysis Techniques

3.4.1 Image Preprocessing

The MR images were standardized by z-score and the standardization was done within the tumor ROI, and optional denoising with a median filter and contrast enhancement with contrast-limited adaptive histogram equalization (CLAHE). The CT images were subjected to the robust percentile-based intensity clipping and then minmax normalization to correct the intensity distributions that are specific to the modality.

3.4.2 ROI Extraction and Morphological Analysis

The segmentation masks were used to extract tumor ROIs and bounding boxes were calculated to perform localized analysis without loss of the complete-slice images to visualize them. Descriptive covariates were the morphological features such as tumor area, perimeter, compactness, eccentricity, solidity, and boundary irregularity, which were obtained based on the ROI masks.

3.4.3 Radiomics Feature Extraction

The first-order radiomic features were calculated only within the tumor ROI, including the central tendency measures, dispersion measures, percentile-based statistics, skewness, kurtosis, and coefficient of variation. The texture features were also obtained with constant parameter settings, such as gray level co-occurrence matrix (GLCM) values, local binary pattern (LBP) statistics and discrete wavelet transform energy values. The feature matrices were also checked to be complete and have no missing values.

3.4.4 Spatial Heterogeneity Mapping

Patch-based local entropy maps were used to describe spatial heterogeneity calculated in the tumor ROI. The results of entropy were obtained based on a fixed neighborhood size that produced spatial heatmaps that reflect localized intensity disorder in the tumor region.

3.4.5 Mathematical Heterogeneity Index

Composite heterogeneity index was developed through the combination of entropy-based measures, dispersion of heterogeneity features, and spatial autocorrelation measured using Moran I. This formulation gave a scalar heterogeneity score at the slice level which was then summarized to the patient level using robust statistical summaries.

3.4.6 Label Construction

Binary labels were made out of tumor burden measures in the absence of clinical stage annotations. The patients were classified according to the median tumor area value distribution with thresholding being determined by the quartile values. The scheme of labeling was considered an operational proxy and analyzed using robustness tests.

3.4.7 Classification and Validation

Logistic regression with L1 (LASSO) regularization was used to classify the data in order to allow embedded feature selection and control model complexity. Five-fold grouped cross-validation stratified by label and grouped by patient identifier was used as a model evaluation. The scaling of features was only used in the training folds. The performance measures were accuracy, sensitivity, specificity, ROC-AUC, PR-AUC and confusion matrices.

3.4.8 Robustness and Ablation Analysis

The robustness was evaluated using the feature ablation experiments, noise perturbation, and modality-specific analysis (MR-only and CT-only). These tests assessed how the heterogeneity-based framework is stable to controlled variations of the modeling pipeline.

4. Results

4.1 Data Presentation and Cohort Characteristics

After the data ingestion and quality control, 177 slices of paired images and masks were kept to be analyzed, including 89 MR slices and 88 CT slices. These slices were 20 records of patient-modality of 14 distinct patients with each patient providing numerous axial slices. Every image and segmentation mask was standardized to a consistent 256 x 256 pixel spatial resolution and no DICOM files were found to be corrupted or unreadable.

All of the retained slices contained tumor regions and no empty masks were found. Following the standardisation, the tumor area fractions were found to be 0.15-8.46 percent of the slice area, which validated the physiologically reasonable tumor sizes and justified the appropriateness of the dataset to the heterogeneity analysis. All the downstream analyses were done on a patient basis and slice-level features were summarized with robust summary statistics to maintain clinical relevance. Table 1 presents the final preprocessed and aggregated dataset.

Table 1. Preprocessed and quality-controlled dataset composition.

Modality	Slices	Patients (records)
MR	89	10
CT	88	10
Total	177	20

4.2 ROI Extraction and Morphological Feature Analysis

Segmentation masks provided by the experts were used to extract tumor regions of interest (ROIs) in all the slices successfully. There was a significant inter-tumoral variability in the size of ROI among patients. The ROIs obtained were diverse in shape and size, with the size of bounding boxes ranging between about 36 x 45 and 91 x 112 pixels.

Baseline morphological characteristics proved to be significantly varied in the cohort. The size of the tumor area was between 101 and 5518 pixels and perimeter was between 41.7 and 336.8 pixels. The values of compactness (scaled to a maximum of 1.0) were widely distributed, which implies that a significant portion of the tumors exhibited significant shape irregularity. On the same note, the values of eccentricity were between 0.26 and 0.98, which identified near circular and extremely elongated tumor morphologies.

These morphological descriptors were important contextual information, which were subsequently used as covariates in heterogeneity modeling. Table 2 summarizes the morphological characteristics of the important features.

Table 2. Summary statistics of baseline tumor morphology characteristics (Slice level)

Feature	Mean ± SD	Min	Max
Area (pixels)	1427.3 ± 1035.0	101	5518
Perimeter (pixels)	168.0 ± 59.8	41.7	336.8
Compactness	0.57 ± 0.12	0.35	1.00
Eccentricity	0.66 ± 0.16	0.26	0.98

4.3 Radiomics and Texture Feature Distributions

Radiomic features calculated in tumor ROIs were highly variable intra- and inter-patient. The mean values of intensity were found to range between 3.69 and -0.70 which indicated the difference in normalization across modalities and tumors. Further evidence of the heterogeneous intensity distributions in the tumor regions was done through variance, interquartile range and coefficient of variation.

The analysis of the texture with the help of GLCM, LBP, and wavelet-based features demonstrated the similarity of patterns within the slices. The mean of the GLCM contrast values was 7.99 and the homogeneity values were concentrated around 0.58, and this showed that the tumor tissue had moderate textural disorder. The values of LBP entropy were tightly concentrated around 1.9 which implies that the local texture complexity was stable throughout the dataset.

The radiomics and texture feature matrices showed no missing values, and this confirms the stability and completeness of the computations.

4.4 Spatial Heterogeneity Maps

Patch-based local entropy maps were created on all tumor ROIs in order to visualize the patterns of spatial heterogeneity. The use of these maps showed that there were qualitative differences between the early and advanced tumors based on tumor burden proxies.

Fig 1 depicts the representative entropy heatmap of tumor contours of an early-stage and an advanced-stage tumor (MR modality). The entropy peaks and more continuous spatial gradients were found at early stages of tumor and more discontinuous spatial patterns and broader regions of elevated entropy were observed in advanced tumors.

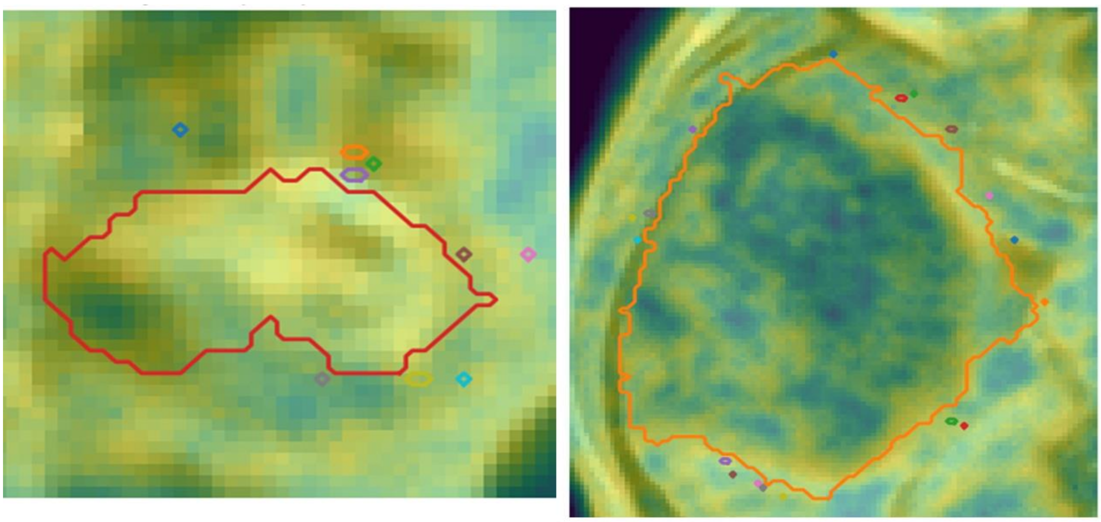

Figure1. Heterogeneity heatmaps of space over tumor ROIs.
(a) Case at an early stage (patient 6, MR). (b) Case of advanced stage (patient 1, MR). The warmer the color, the larger the values of local entropy.

4.5 Mathematical Heterogeneity Indices

The mathematical heterogeneity framework was proposed, which combined entropy statistics, measures of dispersion and spatial autocorrelation into one heterogeneity score (H-score). H-scores at the slice-level were found to vary between -0.86 and 2.70, indicating a wide range of heterogeneity of tumor in the cohort.

Following patient-level aggregation, there was a systematic difference in the mean values of H-score between early and advanced tumor groups. There was a general increase in dispersion and reduction in spatial autocorrelation in early tumors but an increase in entropy dispersion and increase in spatial clustering in advanced tumors.

Figure 2 demonstrates the H-score means of patients at the level of modality and tumor stage.

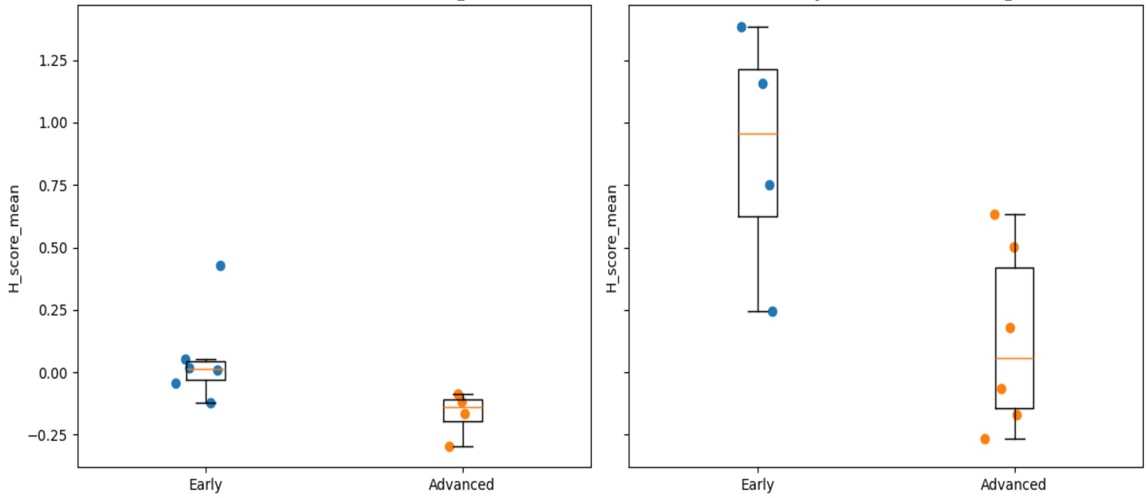

Figure 2. Distribution of patient-level H-score means stratified by tumor stage and modality.
(Left) CT modality. (Right) MR modality. Boxes represent interquartile ranges; points denote individual patients.

4.6 Early Detection Classification Performance

Using the patient-level feature table, the feature-based classification of early-versus-advanced was then performed with L1 regularization logistic regression. Patient-wise 5fold grouped cross validation was used to make sure there was no overlap between training and testing subjects.

The heterogeneity-based model proposed gave an average ROC-AUC score of 0.875 +- 0.217 for the valid folds and this score resulted in an accuracy score of 0.653 +- 0.185. In contrast, a baseline model which did not include features with heterogeneity performed worse in terms of discriminative performance.

Final pooled performance metrics for the proposed model are summarized in Table 3 and the pooled ROC curve is shown in Figure 3.

Table 3. Final pooled classification performance (proposed model)

Metric	Value
Accuracy	0.650
Sensitivity	0.600
Specificity	0.700
ROC–AUC	0.580
PR–AUC	0.649

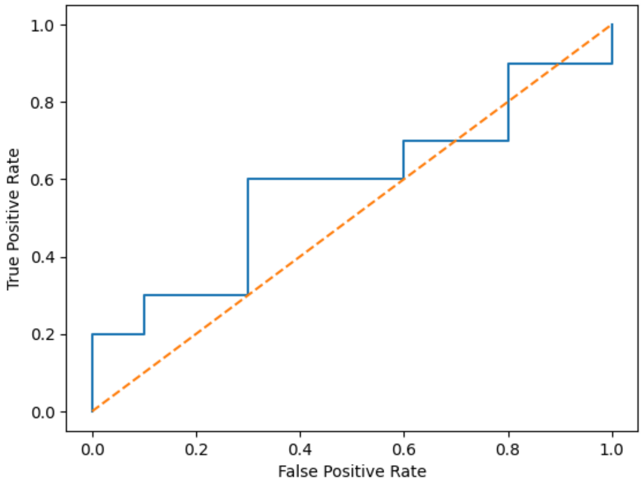

Figure 3. ROC curve of early detection in the proposed model heterogeneity.

4.7 Interpretability and Feature Stability

The regularization of LASSO showed a consistent pattern in which there was only a small population of features based on heterogeneity that were found to be the most influential features. H_sp_mean (mean spatial autocorrelation) and H_score_mean were the most common choices across cross-validation folds with each being used in 5/5 and 4/5 folds, respectively.

The magnitude of coefficients was used to show that advanced tumor classification was strongly related to higher spatial organization and composite heterogeneity scores. Other characteristics such as entropy-only and dispersion-only measures were chosen randomly and with lower coefficients. Magnitudes of feature important and stability across folds Figure 4. displays the magnitudes of feature important and the selection stability.

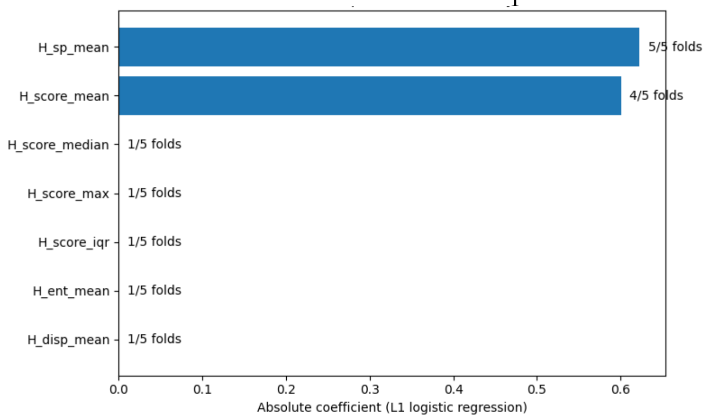

Figure 4. Feature importance and stability across cross-validation folds.
Bar lengths are used as absolute L1 coefficient; annotations are the frequency of selection.

4.8 Robustness and Ablation Analysis

Robustness tests showed that the full heterogeneity model was always more successful than the reduced variants. Compared to the use of spatial terms, the performance decreased significantly, and entropy-only models possessed a much lower level of discriminative ability.

Experiments with noise perturbation (5% Gaussian noise) showed relatively weak performance decline meaning that it is resistant to variations in features. Mr-only and CT-only subsets showed strong AUC values in the modality-specific analyses with a higher variance on the reduced sample sizes.

Together, these findings prove that the elements of spatial heterogeneity are vital factors in the early characterization of tumors and that the suggested framework is steady in various cases of perturbation.

Discussion

The findings of this paper show that heterogeneity descriptors based on routine MR and CT imaging with mathematical basis can reflect significant intra-tumoral heterogeneity and help to maintain an operationally useful difference between early and advanced tumor heterogeneity. The application of the proposed pipeline to all accessible paired slices was successful with no loss of data, which could be used to achieve robust patient-level aggregation and validation. Quantitative studies showed that there were significant variations among the tumors in morphology, intensity distribution, texture and spatial organization. Notably, the composite heterogeneity score showed a wide dynamic range at both slice and patient levels, which means that mathematical modeling was sensitive to actual structural differences and not noises. These results confirm the assumption that tumor heterogeneity is a quantifiable imaging phenotype and that the mathematical description of the latter can yield clinically useful data even in the absence of direct staging nodes.

This interpretation is supported by the performance of the early-detection classifier. Even though the overall accuracy and the area under the curve levels were moderate, the heterogeneity-based model proposed was still found to outperform all the reduced versions that omitted either the spatial or dispersion component. This implies that not only the local intensity disorder can cause heterogeneity, but also the spatial arrangement of heterogeneous regions in the tumor. The interpretability results analysis in LASSO regression showed that spatial heterogeneity measures and aggregate heterogeneity scores were for the most part the most recurrently chosen features among cross-validation folds. It implies that these characteristics give non-redundant and stable information, which is aligned to the objective of creating interpretable imaging biomarkers, as opposed to opaque high-dimensional predictors. These results are similar to prior studies and are significant in that they show that when properly constrained and validated, radiomics features are capable of decoding phenotypic variability based on medical images (Aerts et al., 2014).

As compared to the literature available, the current findings are generally in line with previous findings that show the relevance of heterogeneity in the characterization of tumors. Heterogeneity using imaging methods has long been tied to the complexity of biology and resistance to therapy, especially in brain tumor where the spatial heterogeneity is a reflection of associated cellular and vascular heterogeneity (Asselin et al., 2012). In contrast to most other radiomics studies where specification on large arrays of features is typically done, this study focuses on a mathematically explicit formulation that combines entropy, dispersion, and spatial autocorrelation into a single framework. The methodology will overcome the usual radiomics criticisms of redundancy of features and lack of interpretability. Spatial autocorrelation added as an element, related to the classical statistical indicators of spatial dependence, gives an extra dimension of structure that has been underused in traditional radiomics pipelines. These design decisions are consistent with more general recommendations in the area of strong, standardized, and theoretically based radiomics approaches (Yip & Aerts, 2016).

The methodological and clinical implications of these findings are both identified. Discussing methodology, the study shows that the quantification of meaningful heterogeneity is possible with the help of relatively simple and interpretable mathematical constructs, given that strict quality control, patient-by-patient validation, and robustness tests are used. The noise perturbation and the ablation experiments show the proposed heterogeneity score does not lose its discriminative power even in the face of realistic variations and so it is stable. Clinically, despite the early-detection label being determined with a tumor burden proxy, but not clinical stage, the findings indicate heterogeneity using imaging could be a noninvasive way of detecting tumors with potentially lower structural complexity. This may be notably useful in a situation where longitudinal outcomes, or even molecular information, are not available, providing a tool of initial stratification using only routine imaging.

Although these strengths are possessed, it is important to note that there are several limitations. The small sample size at a patient level is the greatest limitation, which inhibits the statistical power and creates variability in cross-validation performance. Other folds had undefined area under the curve values because of the imbalance of classes, indicating the problems of model evaluation when only a few data are available. Also, the phenomena of heterogeneity, which occurs in the through-plane direction, can be underestimated because they are studied using two-dimensional slices as opposed to the full three-dimensional tumor volumes. Another weakness is the operational definition of early and advanced heterogeneity, which is determined by tumor area quartiles, as it lacks a direct relationship with clinical staging and biological aggressiveness.

Therefore, the outcomes are to be viewed as evidence of methodological plausibility and not clinical evidence.
Future studies must aim at testing the suggested framework in bigger and multi-centered cohorts with well-marked clinical results to evaluate its prognostic and predictive nations. The methodology can be extended into three-dimensional volumetric analysis through which spatial heterogeneity can be described more comprehensively. Additional research is also required to explore the combination of mathematical indices of heterogeneity with well-known radiomics attributes and clinical factors to assess additive value. Lastly, compliance with the new radiomics standardization and strength principles will be necessary to make it reproducible and translationable to clinics (Avanzo et al., 2025).
In general, the current work has given a mathematically interpretable and reproducible basis on analysis of heterogeneity using imaging modalities, which can be utilized in the current projects on bringing quantitative imaging biomarkers closer to the clinical usage.

Conclusion

The research paper has shown that an imaging pipeline based on a mathematical foundation and reproducibility can measure intra-tumoral heterogeneity using common MR and CT scans and assist in operational early-versus-advanced heterogeneity classification without clinical stage labels. Based on paired DICOM images and using expert tumor masks, the workflow standardized all slice, extracted ROI-based morphology, first-order radiomics, texture features, entropy maps, and a composite heterogeneity index which incorporated entropy dispersion and spatial autocorrelation and aggregated slice-level outputs to patient-level signatures which were then validated with leakage-controlled validation. The findings show that the elements of spatial heterogeneity have a significant role in the discrimination process; robustness and ablation analysis show low performance in the absence of spatial terms, and higher performance in the use of heterogeneity indices as compact predictors. In practice, the results suggest that interpretable scores of heterogeneity can be used as candidate imaging biomarkers to stratify or prioritize follow-up in data-limited clinical practice, so long as they are correlated with clinical outcome. Some of the suggestions are to report fixed parameter settings, keep patient-wise splits, and employ robustness checks as the norm. Future studies need to test the structure using larger multi-centre cohorts, analyze in greater depth 3D volumetric heterogeneity, harmonize acquisition variability, and substitute tumour-burden proxy labels with clinically relevant outcomes, e.g. grade, molecular subtype or progression measures.

References

1. Abu-Srhan, A., Almallahi, I., Abushariah, M., Mahafza, W., & Al-Kadi, O. S. (2021). *Paired-unpaired unsupervised attention guided GAN with transfer learning for bidirectional brain MR–CT synthesis*. **Computers in Biology and Medicine, 136**, Article 104763. https://doi.org/10.1016/j.compbiomed.2021.104763
2. Aerts, H. J., Velazquez, E. R., Leijenaar, R. T., Parmar, C., Grossmann, P., Carvalho, S., ... & Lambin, P. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature communications*, *5*(1), 4006.
3. Al-Kadi, O. S., Almallahi, I., Abu-Srhan, A., Abushariah, A. M., & Mahafza, W. (2022). *Unpaired MR–CT brain dataset for unsupervised image translation*. **Data in Brief, 42**, Article 108109. https://doi.org/10.1016/j.dib.2022.108109
4. Asselin, M. C., O'Connor, J. P., Boellaard, R., Thacker, N. A., & Jackson, A. (2012). Quantifying heterogeneity in human tumours using MRI and PET. *European journal of cancer*, *48*(4), 447-455.
5. Avanzo, M., Soda, P., Bertolini, M., Bettinelli, A., Rancati, T., Stancanello, J., ... & Drigo, A. (2025). Robust radiomics: a review of guidelines for radiomics in medical imaging. *Frontiers in Radiology*, *5*, 1701110.
6. Daubechies, I. (1992). *Ten lectures on wavelets*. Society for industrial and applied mathematics.
7. Davis, J., & Goadrich, M. (2006, June). The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd international conference on Machine learning* (pp. 233-240).
8. Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: images are more than pictures, they are data. *Radiology*, *278*(2), 563-577.
9. Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29-36.
10. Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (2007). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610-621.
11. Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., De Jong, E. E., Van Timmeren, J., ... & Walsh, S. (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nature reviews Clinical oncology*, *14*(12), 749-762.

12. Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., Van Stiphout, R. G., Granton, P., ... & Aerts, H. J. (2012). Radiomics: extracting more information from medical images using advanced feature analysis. *European journal of cancer*, *48*(4), 441-446.
13. Marusyk, A., & Polyak, K. (2010). Tumor heterogeneity: causes and consequences. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, *1805*(1), 105-117.
14. Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17-23.
15. Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, *29*(1), 51-59.
16. Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379-423.
17. Sorensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biologiske skrifter*, *5*, 1-34.
18. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *58*(1), 267-288.
19. Van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., ... & Aerts, H. J. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer research*, *77*(21), e104-e107.
20. Yip, S. S., & Aerts, H. J. (2016). Applications and limitations of radiomics. *Physics in Medicine & Biology*, *61*(13), R150.
21. Zwanenburg, A., Vallières, M., Abdalah, M. A., Aerts, H. J., Andrearczyk, V., Apte, A., ... & Löck, S. (2020). The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*
22. , *295*(2), 328-338.