

Radiomics 特徴量の類似性構造に基づく腫瘍進展のデータ駆動的可視化 —低悪性度グリオーマへの応用—

石井 美桜^{†,††}・崇風 まあぜん^{†††}・内山 良一^{†††,*}

[†]宮崎大学大学院農学工学総合研究科 〒 889-2192 宮崎市学園木花台西 1-1

^{††}宮崎大学医学部附属病院放射線部 〒 889-1692 宮崎市清武町木原 5200

^{†††}宮崎大学工学部情報通信プログラム 〒 889-2192 宮崎市学園木花台西 1-1

*責任著者：内山 良一

(受理日：2026 年 1 月 6 日，採択日：2026 年 4 月 28 日)

Data-Driven Visualization of Tumor Progression Based on Radiomics Feature Similarity: An Application to Low-Grade Glioma

Mio ISHII^{†,††}, Mazen SOUFI^{†††}, and Yoshikazu UCHIYAMA^{†††,*}

[†]Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki,
1-1 Gakuen Kibanadai-nishi, Miyazaki-shi, Miyazaki 889-2192, Japan

^{††}Radiology Section, University of Miyazaki Hospital, 5200 Kihara, Kiyotake-cho, Miyazaki-shi, Miyazaki, 889-1692, Japan

^{†††}Information and Communication Technology Program, Faculty of Engineering, University of Miyazaki,
1-1 Gakuen Kibanadai-nishi, Miyazaki-shi, Miyazaki 889-2192, Japan

*Corresponding author: Yoshikazu UCHIYAMA

(Received on January 6, 2026. In final form on April 28, 2026.)

Abstract: Radiomics aims to quantitatively analyze information embedded in medical images to characterize disease states and clinical outcomes. In practice, however, longitudinal imaging data are sparsely acquired in clinical settings, making it difficult to directly model the temporal continuity of disease progression. In this study, we propose a data-driven analytical framework that represents disease progression based on similarity structures in radiomics feature space, without assuming explicit temporal information or predefined progression models. By modeling lesions from multiple patients as nodes in an unrooted phylogenetic tree, the proposed method enables visualization of disease progression as a pseudo-temporal structure, capturing both continuity and branching patterns of disease formation processes. T2-weighted magnetic resonance images from 35 patients with astrocytoma in the low-grade glioma (TCGA-LGG) dataset were analyzed. Radiomics features were extracted from tumor regions, and a subset of prognostically relevant features was selected using Lasso. Dimensionality reduction and minimum spanning tree algorithms were then applied to construct the phylogenetic representation. The resulting phylogenetic trees showed a tendency for stable and severe cases to be separated when feature selection was applied. Moreover, severe cases were further divided into two subgroups characterized by distinct morphological patterns, and subsequent gene expression analysis revealed differential expression profiles between these subgroups. These results demonstrate that the proposed framework provides a data-driven approach for visualizing disease progression from cross-sectional imaging data and offers a basis for exploratory analysis of phenotypic and genotypic heterogeneity in low-grade glioma.

Keywords : Radiomics, Data-driven analysis, Similarity-based visualization, Disease progression, Low-grade glioma

1. 緒 言

医用画像を定量的に解析し、病態や予後と関連付ける Radiomics 研究は、近年の医用画像分野における重要な研究テーマの一つである[1-6]。従来のコンピュータ支援診断 (computer-aided diagnosis: CAD) が、現在得られる画像から病変の検出や鑑別を主目的としてきたのに対し、Radiomics は画像に内在する情報を用いて治療効果や予後など、将来の状態を推定する点に特徴がある[7-9]。このように、Radiomics は画像診断を「現在の評価」から「将来予測」へと拡張する枠組みとして位置付けられる。

このような将来予測を志向する解析を支える方法論として、近年、医用画像解析分野ではデータ駆動科学のアプローチが注目されている。データ駆動科学は、事前に明示的な

仮定やモデルを置くことなく、観測データに内在する特徴量間の関係性や分布構造を解析することで現象の理解や予測を行う枠組みである。医用画像を高次元の定量データとして扱い、統計解析や機械学習を通じて病態との関連性を探索する Radiomics は、このデータ駆動的アプローチの代表的な応用領域といえる。

しかし、Radiomics 研究において将来の状態を予測するために理想的とされる、過去・現在・未来にわたる十分な時系列画像データを臨床的に取得することは容易ではない。被ばくや医療資源の制約から、利用可能な画像データは多くの場合、時間軸方向の情報が乏しい。さらに、実臨床で蓄積される画像は来院時点の病態を記録したスナップショットであり、発症初期から連続的に取得されたものではない。このため、病態進展の開始時点と画像記録の開始

時点が患者ごとに異なり、単純な時間軸に基づく比較が困難となる。

このような制約下で病態進展を表現する代表的な方法として、ステージングが用いられてきた。類似した画像所見を示す症例を同一のステージに分類することで、Radiomics 特徴量に基づく病態層別化が可能であることが報告されている[10, 11]。しかし、ステージングは現在の病態を離散的に表現する手法であり、ステージ間の連続性や、過去および将来の状態との関係性を直接的に表現することはできない。そのため、病態進展を連続的な過程として捉え、時間情報が不完全なデータからでも進展構造を表現できる新たな解析手法が求められている。

連続的な変化過程を類似性構造に基づいて表現する手法として、生物進化の分野では系統樹が広く用いられている。系統樹は、形態的特徴の類似性に基づいて生物種間の関係性を木構造として表現するものであり、このような構造推定はデータ駆動科学の原型のアプローチと捉えることができる。この考え方を Radiomics 研究に応用すれば、症例を高次元特徴量空間上の点として扱い、症例間の距離構造のみに基づいて病態進展の連続的な関係性を可視化することが可能となる。進化系統樹には、有根系統樹と無根系統樹が存在する (Fig.1) が、病態進展が必ずしも単一の起点から始まるとは限らないことを考慮すると、無根系統樹は病態形成過程の多様性を表現する上で適していると考えられる。

本研究では、明示的な時間情報や進展モデルを仮定せず、Radiomics 特徴量間の類似性構造のみに基づいて病態進展を表現するデータ駆動型の解析の枠組みを提案する。無根系統樹を用いた可視化により、限られた臨床データから病態形成過程の多様性を抽出することを目的とする。

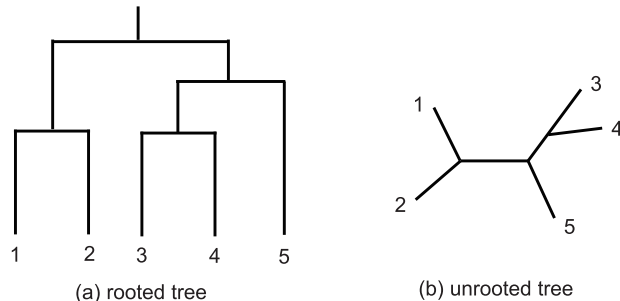


Fig.1 Examples of (a) rooted tree and (b) unrooted tree.

2. 方法

2.1 実験試料

本研究では、The Cancer Imaging Archive (TCIA) に公開されている Low Grade Glioma (TCGA-LGG) データセットを用いた[12]。本データセットには、組織学的悪性度分類においてグレード II に相当する腫瘍として、星細胞腫 41 症例、乏突起星細胞腫 35 症例、乏突起膠細胞腫 66 症例が含まれている。このうち、本研究では、全グリオーマの約 30% を占め、成人男性の大半半球に好発することが知られている星細胞腫に着目して解析に使用した。ただし、経過観察期間が 100 日以内に生存と記録されている 6 症例については、予後評価において経過観察期間が短い打ち切りデータと判断し、解析対象から除外した。その結果、最終的に 35 症例を解析対象とし、治療後に生存していた 25 症例を安定例、治療後に死亡が確認された 10 症例を重篤化例として定義した。なお、打ち切りの閾値を 100 日に設定した理由は、低悪性度グリオーマは一般に進行が緩徐であり、生存期間も長期に及ぶことが知られているためである。このため、経

過観察期間が極めて短い症例では、生存と記録されていても予後を適切に反映しているとは言い難く、生存解析において情報量の乏しい打ち切りデータとなる可能性がある。本研究では、このような極端に短期間の症例を除外する目的で、実務的な閾値として 100 日を設定した。

本データセットには複数の撮像シーケンスが含まれているが、本研究では施設間における撮像プロトコルの不均一性およびデータの欠損を考慮し、すべての症例において一貫して取得されていた T₂ 強調画像を解析対象とした。臨床的には FLAIR 画像やマルチパラメトリック解析の有用性が報告されているが、多施設データにおいてはシーケンス間で信号特性のばらつきが生じやすい。そのため、本研究では単一シーケンスに基づく解析とすることで、データの均一性を確保しつつ、Radiomics 特徴量に基づくデータ駆動型解析を行った。なお、本研究で使用した MR 画像は、いずれも初診時に撮像された画像を対象とした。また、すべての MR 画像については、線形補間を用いて画像サイズを 512 × 512 に統一したものを使用した。これにより、画像のピクセル数 (マトリクスサイズ) を 512 × 512 に統一し、特徴量抽出における画素ベースの計算条件を共通化した。この際に、リサイズのみでは物理的な空間分解能 (mm/pixel) や FOV の差は補正されない点に留意が必要である。

さらに、本研究では、画像データに基づいて分類された重篤化症例の遺伝的特性についても解析を行った。重篤化 10 症例のうち、RNA-Seq データが利用可能であった 8 症例を対象とし、遺伝子発現データとして用いた。遺伝子発現量には、転写産物長およびシーケンス深度による補正が施された FPKM (fragments per kilobase of exon per million reads mapped) の 75 パーセンタイル値である FPKM-UQ を採用した。各症例における遺伝子数は 60,483 個であった。また、Ensembl 遺伝子 ID を HGNC (HUGO Gene Nomenclature Committee) により定義された遺伝子シンボルへ変換し、遺伝子名として用いた。

2.2 Radiomics 特徴量の計測

本データセットは多施設・多機種種のデータが混在しているため、Radiomics 特徴量の計測に先立ち、35 症例の MR 画像に対して濃度階調変換を行い、画素値の正規化を実施した。MR 画像には極端に高い画素値を持つノイズが含まれる症例があり、線形濃度階調変換をそのまま適用すると、変換後の最大画素値がノイズの影響を受け、腫瘍領域内の画素値の幅が小さくなる問題が生じる。そこで、各画像の濃度ヒストグラムを作成し、上位 0.05% の画素値を閾値として設定し、それ以上の画素値は 1023 に置換したうえで、画素値の最小値を 0、最大値を 1023 とするように線形濃度階調変換を行い、画素値を補正した。なお、上位 0.05% の画素がノイズであると仮定した数値は、事前の実験に基づき決定した。

濃度階調変換後の 35 症例すべてについて、複数スライスある MR 画像の中から腫瘍面積が最大となるスライスを 1 枚選択し、腫瘍領域を手動で抽出した。本データセットは多施設から収集されており、スライス間隔が不均一な症例が含まれている。不適切な体軸方向の画像補間による特徴量の変質を避けるため、本研究では 3D 解析ではなく 2D 解析を採用した。次に、抽出した腫瘍領域から、合計 298 項目の Radiomics 特徴量を計測した。特徴量の計測には、一般公開されているソフトウェア MaZda[13-15]を使用した。298 項目の内訳は、腫瘍の大きさに関する特徴量 1 項目、ヒストグラム特徴量 9 項目、テクスチャ特徴量 272 項目、解像度 (ウェーブレット変換) に関する特徴量 16 項目である。

特徴量計測時のパラメータは MaZda のデフォルト値を用いた。例えば、テクスチャ特徴量を計算する際に使用する濃度共起行列のパラメータは、濃度階調 16 階調、画素間距離 1~5、方向 0°, 45°, 90°, 135°であった。16 階調への量子化により、機種間の信号強度の絶対値の差を抑制し、腫瘍内部の相対的なテクスチャ構造を抽出するように配慮した。

なお、MaZda は国際的な標準化指針である IBSI (Image Biomarker Standardisation Initiative) に完全準拠したソフトウェアではない。本研究では解析の再現性を担保するため、濃度正規化および階調数指定を IBSI の推奨に準じた形式で行った。本研究の目的は、「個々の特徴量の絶対値の比較」ではなく、「多次元特徴量空間における症例間の相対的な距離構造(類似性)の可視化」に主眼を置いているため、一貫した条件下での特徴量計測により、解析の妥当性は維持されるものとする。

2.3 Radiomics 特徴量の選択

取得した 298 項目の Radiomics 特徴量のすべてが、安定症例と重篤化症例の判別に有用とは限らない。そこで、本研究では least absolute shrinkage and selection (Lasso) [16] を用いて、判別に有用な特徴量を選択した。Lasso は次式で定義される：

$$\hat{\beta}_{\text{lasso}} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2N} \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (1)$$

ここで、 N は症例数、 y_i は i 番目の症例の安定／重篤化情報、 x_j は j 番目の Radiomics 特徴量、 β_0 は定数項、 $\lambda \geq 0$ は縮小度合いを制御する正則化パラメータ、 p は特徴量の総数を表す。パラメータ β_j は、(1) 式の最適化問題を解くことで求められる。なお、本研究では R の glmnet パッケージを用いて解析を行ったため、式(1)の第 1 項には同パッケージの実装に準じて $1/2N$ による正規化を適用している。本研究では、10-fold クロスバリデーションを用いて、平均二乗誤差が最小となるを決定した。その結果、 β_j がゼロでない Radiomics 特徴量が 3 項目選択され、これらを安定症例と重篤化症例を判別するための有用な Radiomics 特徴量として採用した。

2.4 擬時系列解析

多次元の Radiomics 特徴量を 2 次元空間に縮約することで、系統樹の可視化が可能となる。本研究では次元縮約の手法として、古典的多次元尺度構成法 (Multidimensional Scaling, MDS) [16] と t-SNE [17] を採用し、その効果を比較した。MDS は主座標分析 (Principal coordinate analysis) と呼ばれ、データ間のユークリッド距離を保持するように低次元空間へ線形変換する手法である。具体的には、まず入力 i と j のユークリッド距離から距離行列 d_{ij} を作成し、重心が原点となるように座標変換行列 z_{ij} を求める：

$$z_{ij} = -\frac{1}{2} \left(d_{ij}^2 - \sum_{i=1}^n \frac{d_{ij}^2}{n} - \sum_{j=1}^n \frac{d_{ij}^2}{n} + \sum_{i=1}^n \sum_{j=1}^n \frac{d_{ij}^2}{n^2} \right) \quad (2)$$

その後、行列 z_{ij} の固有ベクトルを用いて低次元空間上の座標を決定する。MDS は線形変換であるため、多次元空間でデータが非線形に分布している場合、類似するデータを低次元空間でも近くに保つことは難しい。一方、t-SNE は非線形次元縮約法であり、高次元空間のデータ間距離が低次元空間でもなるべく保持されるよう変換する。局所的な構造が維持されるため、高次元空間で非線形に分布するデータでも、低次元空間上で類似データを近くに配置できる。解析には R の Rtsne パッケージ (version 0.17) を使い、ハイパー

パラメータは症例数 ($n=35$) を考慮して Perplexity を 10 に設定した。学習率および最大反復回数は、同パッケージのデフォルト値である 200 および 1000 を採用し、コスト関数の収束を確認した。

系統樹は、症例間の類似度に基づいて構築した。本研究では、グラフ理論における最小全点木 (Minimum Spanning Tree, MST) [18] を用いた。MST の構築においては、次元縮約後の 2 次元座標ではなく、Radiomics 特徴量からなる高次元特徴量空間におけるユークリッド距離を用いた。すなわち、各症例間の距離は、選択された Radiomics 特徴量ベクトル間のユークリッド距離として定義し、この距離に基づいて類似度行列を作成した後、MST アルゴリズムを適用して系統樹の構造を決定した。その後、得られた接続関係を MDS または t-SNE により得られた 2 次元空間上に写像し、可視化を行った。具体的には、 x 番目の症例と y 番目の症例の類似度を Radiomics 特徴量のユークリッド距離で定量化した：

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_j - y_j)^2} \quad (3)$$

ここで、 j は Radiomics 特徴量の総数である。例えば、すべての Radiomics 特徴量を用いる場合は $j=298$ であり、3 つの選択された Radiomics 特徴量を用いる場合は $j=3$ である。この計算を全症例の組み合わせで行い、類似度行列を作成した後、MST アルゴリズムを適用して系統樹の枝を決定した。最終的に、2 次元空間に投影された症例を直線で結ぶことで、紙上での系統樹として可視化した。

2.5 病変の表現型と遺伝型の関係分析

Radiomics は、医用画像に基づく形態的特徴(表現型)と、遺伝子やタンパク質などの分子レベルの情報(遺伝型)を関連付けて解析することを目的とした研究領域であり、Radio と Omics を組み合わせた造語である。したがって、Radiomics 研究においては、画像から得られる病変の表現型が、腫瘍の分子生物学的特性をどの程度反映しているかを検証することが重要である。

前節で述べた擬時系列解析および形態系統樹の可視化により、重篤化症例が 2 つの群に分かれる傾向が観察された。そこで、本研究では、形態的特徴に基づいて分類されたこれら 2 群の腫瘍が、遺伝子発現レベルにおいても異なる特性を有しているかを検討するため、遺伝子発現データを用いた解析を行った。まず、RNA-Seq により得られた遺伝子発現量を入力とし、Lasso を適用することで、2 群を分類するのに寄与すると考えられる遺伝子を選択した。その結果、4 つの遺伝子が特徴的な遺伝子として選択された。次に、これら 4 遺伝子の発現量を用いて階層的クラスタリングを行い、ヒートマップとして可視化した。これにより、形態系統樹に基づいて分類された重篤化症例の群が、遺伝子発現パターンの観点からも特徴付けられるかを評価し、画像に基づく表現型と分子レベルの遺伝型との関連性を探索的に分析した。

3. 実験結果

Fig.2 に、すべての Radiomics 特徴量を用いて作成した MDS および t-SNE による系統樹を示す。全特徴量を用いた場合、腫瘍の大きさ、ヒストグラム、テクスチャなど多様な観点から類似した病変が近接して配置される傾向が確認された。一方で、安定症例と重篤化症例は明確に分離されず、両群が混在して分布する傾向が見られた。また、MDS による系統樹では、類似した症例を結ぶ枝が交差しており、

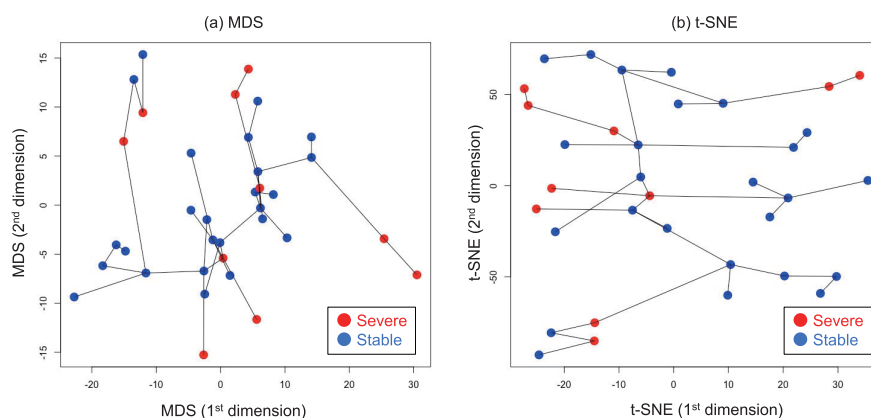


Fig.2 Phylogenetic trees obtained using 298 radiomics features as input data. (a) Output using MDS. (b) Output using t-SNE.

Table 1 Three radiomics features selected by Lasso.

	Feature	Category	Description
#1	S(0,1)Correlat	Texture	Correlation obtained from co-occurrence matrix S(0,1) is the between-pixels distance
#2	WavEnHL_s-1	Resolution	Wavelet coefficient at scale 1 H: high-pass filter, L: low-pass filter
#3	WavEnHH_s-3	Resolution	Wavelet coefficient at scale 3 H: high-pass filter

低次元空間上での病態進展の解釈が困難であった。

Table 1 に、Lasso によって安定症例と重篤化症例の判別に有用と判断された3つのRadiomics特徴量を示す。選択された特徴量は、濃度共起行列に基づくテクスチャ特徴量が1項目、ウェーブレット係数に基づく解像度関連特徴量が2項目であった。Table 1 で選択された3つの特徴量は、単なる画像統計量としての意味にとどまらず、腫瘍の微細構造や病態生理学的特性を反映している可能性がある。S(0,1)Correlat は、濃度共起行列に基づく相関を表し、腫瘍内での画素間の構造的類似性や均一性の程度を示す指標である。また、WavEnHL_s-1 および WavEnHH_s-3 は、ウェーブレット変換による高周波成分のエネルギーを表し、腫瘍内の細胞密度の微細な不均一性や微小壊死、マクロファージ浸潤と関連することが報告されている[19]。これらの知見は、本研究で選択されたRadiomics特徴量が腫瘍の形態的特徴だけでなく、病態生理学的背景をある程度反映していることを示唆している。

Fig.3 は、これら3つの特徴量のみを用いて作成したMDSおよびt-SNEによる系統樹を示す。Fig.2の結果と比較すると、安定症例と重篤化症例が空間的に分離される傾向がより明瞭に観察された。この実験では、t-SNEの堅牢性の検

証をするために、乱数を変更して複数回の試行を行った。個々の症例の座標値は変動するものの、MSTによって接続される症例間の隣接関係および「安定群から重篤化群へ向かう」という大きなマクロ構造が維持されることを確認した。

Fig.4 に、Fig.3のt-SNEによる配置と対応する重篤化症例のMR画像を示す。また、各症例の腫瘍面積およびIDH mutation statusも併せて表示した。その結果、腫瘍面積およびIDH変異の有無はいずれも系統樹上の分岐構造やクラスターと明確な対応関係を示さなかった。重篤化症例は2つの群に分かれ、形態的に類似した表現型を持つ症例がそれぞれ近接して配置されていた。以下では、これらの2群をType 1 および Type 2 と呼ぶ。

Fig.5 に、遺伝子発現データが利用可能であった8症例の重篤化症例を対象とした遺伝子解析の結果を示す。遺伝子発現量を入力としたLassoにより、Type 1 と Type 2 の群を分類する4つの遺伝子が選択された。さらに、これら4遺伝子の発現量を用いて階層的クラスタリングを行った結果、表現型に基づいて分類されたType 1 および Type 2 の群が遺伝子発現パターンにおいても明確に分離された。

以上の結果から、本手法は、Radiomics特徴量に基づく腫

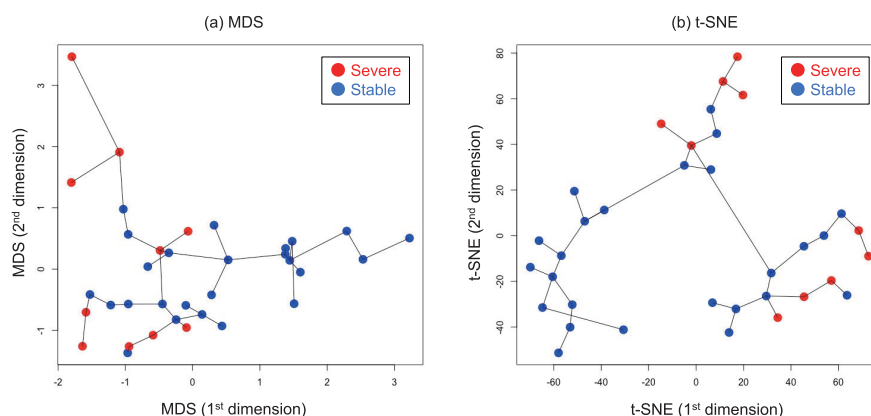


Fig.3 Phylogenetic trees obtained using three radiomics features selected by Lasso. (a) Output using MDS. (b) Output using t-SNE.

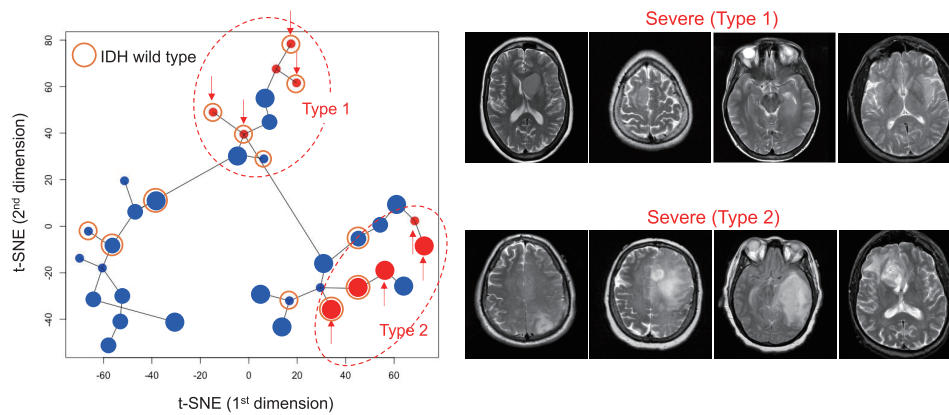


Fig.4 Result of displaying eight severe cases in the phylogenetic tree obtained by t-SNE using three radiomic features. The size of the circles in the phylogenetic tree represents tumor size categorized into three levels, and information on IDH mutation status is also indicated.

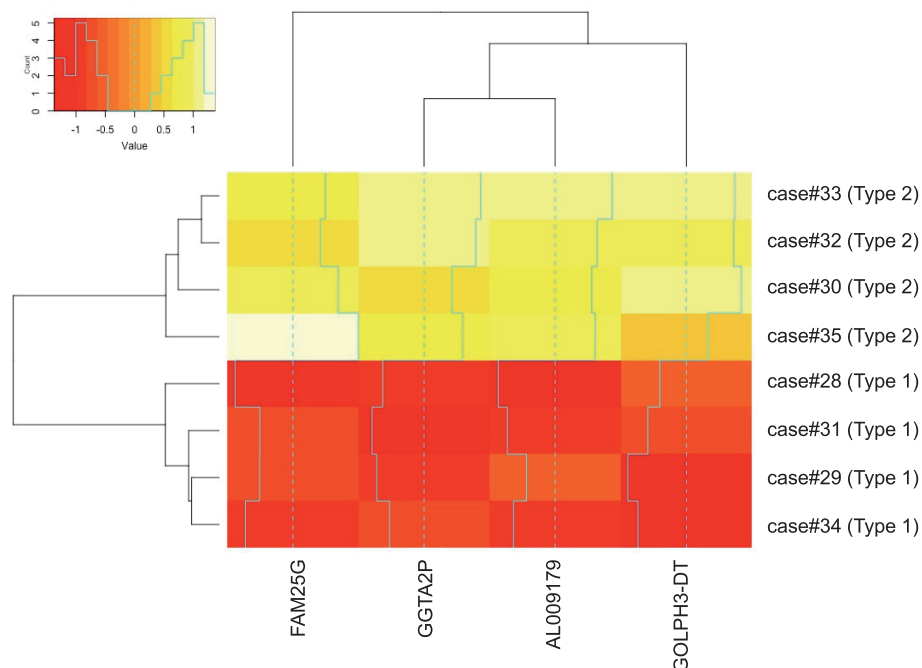


Fig.5 Heatmap obtained from RNA-Seq of eight severe cases. We can explore the genes that lead to disease-relevant variations in the imaging phenotypes between Type 1 and Type 2.

瘍進展の表現型分類のみならず、その表現型と関連する遺伝的特徴を探索的に解析するための枠組みとして有用であることが予備的に示唆された。ただし、遺伝子解析のサンプルサイズが限られているため、得られた結果は仮説生成的な段階に留まる点に留意が必要である。

4. 考 察

Fig.2 および Fig.3 に示したように、MDS を用いて系統樹を構築した場合、類似した症例を結ぶ枝が交差する結果となった。これは、高次元の Radiomics 特徴量空間において、症例間の分布が非線形構造を有していることを反映していると考えられる。線形変換に基づく MDS では、高次元空間における局所的な類似関係を低次元空間上で十分に保持することが難しく、その結果として病態進展の解釈が困難になった可能性がある。一方、t-SNE は高次元空間における局所的な距離関係を低次元空間上で保持する非線形次元縮約法であり、本研究においては症例間を結ぶ枝の交差が抑制された。本研究では、MST の構築は高次元特徴量

空間における距離に基づいて行っており、t-SNE はその構造を可視化するための手段として用いている。したがって、t-SNE の大域的距離構造の非保存性は、MST の構造自体には影響を与えない。一方で、t-SNE による 2 次元配置は高次元空間における距離関係を完全には保持しないため、可視化された配置における症例間距離や枝の長さは、実際の特徴量空間での距離を厳密に反映するものではない点に留意が必要である。これらの結果から、系統樹の可視化を目的とした次元縮約手法としては、局所構造の保持に優れた t-SNE がより適していると考えられる。

類似画像検索や症例検索の分野では、画像特徴量間の距離に基づく類似度定義が広く用いられている[20-22]。しかし、Fig.2 に示したように、298 項目すべての Radiomics 特徴量を用いてユークリッド距離を算出した場合、安定症例と重篤化症例を明確に分離することはできなかった。これは、予後判別に寄与しない特徴量が多数含まれることで、判別に有用な情報が距離計算の過程で平均化され、群間差が不明瞭になったためと考えられる。これに対し、Fig.3 に示すように、Lasso により安定症例と重篤化症例の判別に有

用な特徴量を選択した上で系統樹を構築した場合、両群が分離する傾向が観察された。この結果は、選択された3つのRadiomics特徴量に、予後と関連する情報が含まれている可能性を示唆しており、系統樹構築においては解析目的に応じた特徴量選択が重要であることを示している。Fig.3に示したt-SNEに基づく系統樹では、おおよそ3つの群が形成され、末端に近づくにつれて重篤化症例の割合が増加する傾向が認められた。また、系統樹には短い枝や方向性の異なる幹が複数存在しており、病態形成過程が単一の経路ではなく、複数の異なる進展様式を有している可能性が示唆される。このように、系統樹表現を用いることで、病態進展の連続性や分岐構造を考慮した解析が可能となる。一方で、Fig.4で得られた系統構造は、腫瘍サイズやIDH変異といった既知の臨床・分子指標とは独立した分布を示した。この結果は、本手法が単純な病勢の進行や既知の分子サブタイプ分類を直接的に反映したものではないことを示唆する。したがって、本研究における「擬似時間」は、生物学的時間経過を一義的に表すものではなく、Radiomics特徴量に基づく症例間の類似性から構成される潜在的な構造として解釈する必要がある。このような潜在構造は、未同定の分子機構や腫瘍微小環境の違いを反映している可能性があり、今後の検証が必要である。

さらに、Fig.5に示した遺伝子発現解析では、病態進展の様式が異なるType 1およびType 2に関連する4つの遺伝子が選択された。そのうち、GOLPH3-DTはGOLPH3遺伝子由来の転写産物であり、GOLPH3は複数のがん種において悪性化や予後不良と関連するオンコジーンとして報告されている。膠芽腫の一部症例においても高発現が報告されており、腫瘍細胞の増殖促進との関連が指摘されている[23]。また、初期の研究では、GOLPH3の発現と病理学的腫瘍グレードとの関連が示唆されている[24]。これらの知見を踏まえると、本研究で同定された表現型の違いとGOLPH3関連遺伝子との関連性については、今後の詳細な検証により生物学的意義が明らかになる可能性がある。しかしこの結果は、極めて少数例での探索的解析であるため、この遺伝子発現パターンの差異は予備的な知見として捉えるべきであり、多施設コホートによる大規模な分子生物学的検証が待たれる。

本研究で提案した教師なし学習による系統樹解析は、従来の教師あり学習に基づく高精度な予後予測モデルとは異なる独自の臨床的意義を有する。第一に、決定プロセスが不明瞭になりがちなディープラーニング等の手法と比較し、症例間の幾何学的な類似性構造を可視化することで、予測結果に対する直感的な説明可能性を提示できる点である。第二に、データ駆動型アプローチにより、既存の医学的分類体系に縛られない新たな知見の探索が可能となる点である。本研究において、重篤化症例が表現型および遺伝型の両面でType 1およびType 2という異なる特性を持つ群として抽出されたことは、従来の画一的な悪性度分類では見逃されていた病態の不均一性を示唆している。これは、個々の表現型に最適化された治療戦略を選択する精密医療の進展に寄与し得ると考えられる。第三に、病態を離散的なステージに分類するのではなく、連続的な進展プロセス(Trajectory)として表現することで、既存の臨床分類では捉えきれない「境界領域」に位置する症例の特定が可能となる。例えば、系統樹の分岐点付近に位置する症例や、安定例のクラスターから重篤化例の枝へと移行しつつある症例を同定することは、早期介入が必要な高リスク患者のスクリーニングにおいて極めて有用な指標となり得る。第四に、腫瘍の表現型に基づく細分化は、腫瘍の形成・進展機序の理

解や、新規抗腫瘍薬開発の標的選定においても有益な視点を提供し得ると考えられる。

本研究のリミテーションとして、解析対象が35症例と限られている点が挙げられる。しかし、本研究は、大規模データを前提とした予測モデルの構築を目的とするものではなく、明示的な時間情報や病態進展モデルを仮定せずに、Radiomics特徴量間の類似性構造から病態進展の様相を可視化するデータ駆動型の解析の枠組みの有効性を検証することを主眼としている。この観点から、本研究で得られた系統構造は、症例数の制約下においても病態形成過程の多様性を表現し得ることを示す概念実証として位置付けられる。今後、多施設・多症例データを用いることで、本枠組みの安定性および一般化可能性をさらに検証する必要がある。また、T₂強調像の単一スライスのみを解析対象とした点もリミテーションとして挙げられる。先行研究[25]が指摘するように、造影T₁像や拡散強調像を組み合わせた3次元マルチパラメトリック解析は、腫瘍の細胞密度や血管新生を多角的に評価する上で不可欠である。本研究の結果は、あくまでT₂強調像が反映する浸潤や水腫性変化の類似性に基づく進展構造を示したものであり、他のモダリティを統合することで系統樹の分岐構造がさらに詳細化される可能性がある。

本研究で用いたMR画像はすべて初診時に撮像されたものであり、症例間で撮像タイミングのばらつきは存在しない。したがって、本研究で得られた系統樹構造は、治療経過中の時間的変化を反映したものではなく、同一時点における症例間の形態的類似性に基づいて構成されたものである。このため、本手法により得られる擬似時間軸は、実時間に対応する病態進行過程を直接表すものではなく、初診時における腫瘍の進展段階や悪性度の違いを反映した“状態空間上の連続性”として解釈されるべきである。本研究で構築した系統樹は、症例間の距離構造に基づく相対的な関係性を表現するものであり、病態進展の実時間的順序や因果関係を直接的に推定するものではない。一方で、臨床現場において連続的な時系列画像データの取得が困難であるという現実的制約を考慮すると、本手法は、断片的に取得されたスナップショット画像のみから病態進展の連続性や分岐構造を再構成するための実用的かつ汎用的な解析戦略を提示するものといえる。

5. 結 論

本研究では、Radiomics特徴量とグラフ理論を組み合わせて、症例間の類似性構造に基づいて病態進展を表現する系統樹解析手法を提案した。本手法は、明示的な時間情報や進展モデルを仮定することなく、断片的に取得された医用画像データから、病態進展を擬似的な時間軸として再構成・可視化できる点に特徴がある。さらに、系統樹構造を用いることで、病態形成過程を連続的かつ分岐を伴う多様な経路として捉えることが可能となり、病態進展ダイナミクスの探索的・定量的解析への道を開いた。本解析の枠組みは、Radiomicsに基づく予後予測や表現型・遺伝型統合解析に向けた基盤的手法として、医用画像におけるデータ駆動型病態理解を拡張するものである。

謝 辞

本研究の一部は、JSPS科研費基盤研究C(課題番号24K15794)にて行われた。

参考文献

- [1] Sanghani P, Ang BT, King NKK, et al: Overall survival prediction in glioblastoma multiforme patients from volumetric, shape and texture features using machine learning, *Surg Oncol.*, 27(4), 709-714, 2018.
- [2] Wu G, Shi Z, Chen Y, et al.: A sparse representation-based radiomics for outcome prediction of higher grade gliomas, *Med Phys*, 46(1), 250-261, 2019.
- [3] Wang Y, Hu Z, Wang, H: The clinical implications and interpretability of computational medical imaging (radiomics) in brain tumors, *Insights Imaging*, 16, 77, 2025.
- [4] 木下学：脳腫瘍におけるこれまでのRadiomicsと今後について, *Medical Imaging Technology*, 41(2), 73-77, 2023.
- [5] タキンキン, 北川真歩, 工藤興亮：脳腫瘍の診療におけるRadiomics, *Radiogenomicsの活用, 臨床放射線*, 68(9), 853-859, 2023.
- [6] 内山良一：Radiomicsの再整理と新たな可能性, *医用画像情報学会雑誌*, 41(2), 37-40, 2024.
- [7] Doi K: Computer-aided diagnosis in medical imaging: historical review, current status and future potential, *Comput. Med. Imaging Graph*, 31, 198-211, 2007.
- [8] Giger ML: Computerized analysis of images in the detection and diagnosis of breast cancer, *Semin. Ultrasound C.T.M.R.*, 25, 411-418, 2004.
- [9] Li Q, Nishikawa RM: *Computer-Aided Detection and Diagnosis in Medical Imaging*, CRC Press, FL, 2015.
- [10] Cho HH, Lee SH, Kim J, et al.: Classification of the glioma grading using radiomics analysis, *PeerJ*, 6:e5982, 2018.
- [11] Kickingereider P, Neuberger U, BonekampD, et al.: Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma, *Neuro Oncol.*, 20(6), 848-857, 2018.
- [12] TCGA-LGG Database. Available at <https://www.cancerimagingarchive.net/collection/tcga-lgg/>. Accessed 24 Dec 2025.
- [13] Technical University of Lodz. Available at <https://qmazda.p.lodz.pl>. Accessed 24 Dec 2025.
- [14] Szczypiński PM, Strzelecki M, Materka A, et al: MaZda-a software package for image texture analysis, *Comput Methods Programs Biomed*, 94, 66-76, 2009.
- [15] Strzelecki M, Szczypinski P, Materka A, et al: A software tool for automatic classification and segmentation of 2D/3D medical images, *Nucl Instruments Methods Phys Res*, 702, 137-140, 2013.
- [16] Hastie T, Tibshirani R, Friedman J: *The elements of statistical learning, data mining, inference and prediction*, second edition, Springer, New York, 2009.
- [17] Maaten LVD, Hinton G: Visualizing data using t-SNE, *Journal of machine learning research*, 9, 2579-2605, 2008.
- [18] Diestel R: *Graph theory*, 5th edition, Springer-Verlag, Heidelberg, 2016.
- [19] Li G, Li L, Li Y, et al: An MRI radiomics approach to predict survival and tumour-infiltrating macrophages in gliomas. *Brain*, 145 (3), 1151-1161, 2022.
- [20] Swett HA, Mutalik PG, Neklesa VP, et al: Voice-activated retrieval of mammography reference images, *Journal of Digital Imaging*, 11, 65-73, 1998.
- [21] Qi H, Snyder WE: Content-based image retrieval in picture archiving and communications system, *Journal of Digital Imaging*, 12, 81-83, 1999.
- [22] Muramatsu C, Li Q, Suzuki K, et al: Investigation of psychophysical measure for evaluation of similar images for mammographic masses: Preliminary results, *Medical Physics*, 32, 2295-2304, 2005.
- [23] Sechi S, Frappaolo A, Karimpour-Ghahnavieh A, et al.: Oncogenic Roles of GOLPH3 in the Physiopathology of Cancer, *Int J Mol Sci*, 31;21(3):933, 2020.
- [24] Li XY, Liu W, Chen SF, et al.: Expression of the Golgi phosphoprotein-3 gene in human gliomas: a pilot study, *J Neurooncol*. 2011 105(2), 159-63, 2011.
- [25] Hao M, Yan J, Wang X, et al.: Survival prediction in gliomas based on MRI radiomics combined with clinical factors and molecular biomarkers, *PeerJ*, 2025 Aug 20; 13:e19906.