HOKKAIDO UNIVERSITY

Title	機械学習モデルを用いた脳腫瘍病理画像解析に関する研究
Author(s)	遠田, 建
Description	配架番号：2977
Degree Grantor	北海道大学
Degree Name	博士(医学)
Dissertation Number	甲第16884号
Issue Date	2026-03-25
DOI	https://doi.org/10.14943/doctoral.k16884
Doc URL	https://hdl.handle.net/2115/99729
Type	doctoral thesis
File Information	ENDA_Ken_abstract.pdf, 論文内容の要旨 https://creativecommons.org/licenses/by/4.0/

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学位論文内容の要旨

博士の専攻分野の名称　博士（医　学）　氏 名　遠田　建

学 位 論 文 題 名

機械学習モデルを用いた脳腫瘍病理画像解析に関する研究
（Studies on Brain Tumor Histopathological Image Analysis Using Machine Learning Models）

【背景と目的】
悪性脳腫瘍は発生頻度は比較的低いものの、適切な治療選択のために正確な診断が極めて重要な疾患群である。主要な悪性脳腫瘍には神経膠腫、原発性中枢神経系リンパ腫（Primary Central Nervous System Lymphoma: PCNSL）、転移性脳腫瘍が含まれ、それぞれ治療方針が大きく異なるため、正確な鑑別診断が患者の予後に直接影響を与える。2021 年 WHO（World Health Organization）中枢神経系腫瘍分類第 5 版では、成人型びまん性神経膠腫を膠芽腫（IDH（Isocitrate Dehydrogenase）野生型）、星細胞腫（IDH 変異型）、乏突起膠腫（IDH 変異型かつ 1p/19q 共欠失）の 3 タイプに分類しているが、分子診断には時間を要するため、初期治療方針決定には形態学的診断が重要な役割を果たしている。

病理診断の標準化と客観化を目指し、近年では機械学習技術、特に病理基盤モデルの応用が期待されている。2023 年以降、DINOv2 アルゴリズムを基盤とした次世代の基盤モデル（UNI、Prov-GigaPath 等）が相次いで登場し、従来モデルを大幅に上回る性能を示している。しかし、これらの基盤モデルを実際の分類タスクに適用する際の最適な転移学習戦略（Fine-tuning（FT）と Linear Probing（LP））については、特に脳腫瘍分類に特化した体系的な比較研究が限定的である。

本研究では、施設内データセットと外部データセット（EBRAINS）を用い、次世代基盤モデル（UNI、Prov-GigaPath）と従来型モデル（CTransPath、ImageNet 事前学習 ViT-L（Vision Transformer-Large）、ResNet-RS 50）について、FT と LP の適用戦略を比較検討し、実臨床環境における基盤モデルの最適な訓練戦略を明らかにすることを目的とした。

【方法】
北海道大学医学部腫瘍病理学教室で 2016 年から 2023 年に収集した 254 症例（膠芽腫 66 例、星細胞腫 29 例、乏突起膠腫 24 例、PCNSL 52 例、転移性腫瘍 81 例、背景組織 5 例）から施設内データセットを構築した。H&E（Hematoxylin and Eosin）染色された永久標本をデジタル化し、腫瘍領域から 512×512 ピクセル以上のパッチに分割、計 368,287 枚のパッチを作成した。外部検証用に EBRAINS データセットから 698 症例を収集した。

5 種類のエンコーダー（UNI、Prov-GigaPath、CTransPath、ViT-L、ResNet-RS 50）を比較し、3 つの訓練条件（FT：エンコーダーと全結合層を同時に更新、LP：事前学習済みエンコーダーを固定し全結合層のみ更新、Random Init：事前学習なしでランダム初期化）を設定した。5 分割交差検証を用い、症例あたりの訓練パッチ数を 500、100、25、10 枚と変化させて訓練と評価を行った。評価指標として、マクロ平均 Recall、パッチ精度、Overall Accuracy を算出し、GradCAM（Gradient-weighted Class Activation Mapping）と UMAP（Uniform Manifold Approximation and Projection）による可視化解析も実施した。

【結果】
施設内データセット（500 パッチ/症例）では、基盤モデル UNI(LP)、UNI(FT)、Prov-GigaPath(LP)が全指標で堅牢な性能を示し、細粒度分類でマクロ平均 Recall 0.87-0.88、パッチ精度 0.93-0.94 を達成した。粗粒度分類では Prov-GigaPath(LP)が最高性能（254 例中 251 例正解）を示した。

外部データセット（EBRAINS）では注目すべき結果が得られた。UNI(LP)が全指標で優れた性能を示し（細粒度：マクロ平均 Recall 0.75±0.01、粗粒度：0.94±0.01）、UNI(FT)を有意に上回った（$p<0.01$）。この逆転現象は次世代基盤モデルに特有で、従来型モデル（ViT-L_{IN}、CTransPath）では依然として FT ≥ LP の関係が維持された。

訓練パッチ数の影響分析では、基盤モデルは 10 パッチ/症例という少量データでも堅牢な性能を示し、UNI(LP)の 10 パッチ条件は、従来モデルの 500 パッチ条件と同等以上の性能を達成した。一方、パッチ数増加に伴う性能向上は基盤モデルで限定的であり、外部データセットでは性能低下も観察された。

UMAP 解析では、事前学習済み UNI は学習前から腫瘍タイプごとの明確な分離を示したが、施設内データセットでの追加学習後は過剰に分離されたクラスタリングを示し、過適合の視覚的証拠となった。GradCAM 解析では、モデルが組織学的に関連する特徴（膠芽腫の微小血管増殖、PCNSL の血管周囲浸潤等）を適切に学習していることが確認された。

【考察】
本研究の重要な発見は、外部データセットにおいて次世代病理基盤モデル UNI で LP が FT を有意に上回った点である。この性能逆転は、大規模データで獲得された汎用的認識能力が、単一施設データでの FT により過適合を起こし、汎化能力を失う「破滅的忘却」を示唆する。UMAP 解析でも、FT 後の過剰なクラスタ分離が視覚的に確認された。さらに、UNI(LP)は 10 パッチ/症例という少量データでも、500 パッチの従来モデルと同等の性能を達成した。LP は計算資源の要求も桁違いに少なく、実臨床での実装可能性が高い。

これらの知見は、病理 AI 開発における「独自データでエンコーダーを訓練する」から「事前学習済みエンコーダーに少量データで問い合わせる」というパラダイムシフトを示している。すなわち、大規模な病理データ収集から、基盤モデルに埋め込まれた汎用的な視覚認識フレームワークを効率的に活用する方向への転換である。

【結論】
本研究は、病理画像分類における次世代病理基盤モデルの適応戦略の包括評価を通じて、外部データセットへの汎化において LP が FT を上回ることを実証し、データセット設計における新たなパラダイムシフトを示した。これらの知見は、より効率的で信頼性の高い医療 AI システムの開発に向けた理論的基盤として貢献するものである。