



This is a repository copy of *From digitized whole-slide histology images to biomarker discovery: A protocol for handcrafted feature analysis in brain cancer pathology.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/227326/>

Version: Published Version

Article:

Lu, X. orcid.org/0009-0005-0537-0355, Ying, Y., Chen, J. et al. (7 more authors) (2025) From digitized whole-slide histology images to biomarker discovery: A protocol for handcrafted feature analysis in brain cancer pathology. *Brain-X*, 3 (2). e70030. ISSN 2835-3153

<https://doi.org/10.1002/brx2.70030>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

From digitized whole-slide histology images to biomarker discovery: A protocol for handcrafted feature analysis in brain cancer pathology

Xuanjun Lu¹  | Yawen Ying² | Jing Chen² | Zhiyang Chen^{3,4,5} | Yuxin Wu^{6,7} | Prateek Prasanna⁸ | Xin Chen⁹ | Mingli Jing¹ | Zaiyi Liu^{3,4} | Cheng Lu^{3,4} 

¹School of Electronic Engineering, Xi'an Shiyou University, Xi'an, Shaanxi, China

²Department of Medical Research, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong, China

³Department of Radiology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong, China

⁴Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangzhou, Guangdong, China

⁵School of Clinical Dentistry, University of Sheffield, Sheffield, UK

⁶Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, Georgia, USA

⁷Department of Computer Science & Informatics, Emory University, Atlanta, Georgia, USA

⁸Department of Biomedical Informatics, Stony Brook University, Stony Brook, New York, USA

⁹Department of Radiology, School of Medicine, Guangzhou First People's Hospital, South China University of Technology, Guangzhou, Guangdong, China

Correspondence

Mingli Jing, Zaiyi Liu and Cheng Lu.
Email: mljingsy@xsyu.edu.cn;
liuzaiyi@gdph.org.cn and
lucheng@gdph.org.cn.

Funding information

National Natural Science Foundation of China, Grant/Award Number: 82272084; National Key R&D Program of China, Grant/Award Number: 2023YFC3402800; Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Grant/Award Number: 2022B121010011; Postgraduate Innovation and Practical Ability Training Program of Xi'an Shiyou University, Grant/Award Number: YCS23114144

Abstract

Hematoxylin and eosin (H&E)-stained histopathological slides contain abundant information about cellular and tissue morphology and have been the cornerstone of tumor diagnosis for decades. In recent years, advancements in digital pathology have made whole-slide images (WSIs) widely applicable for diagnosis, prognosis, and prediction in brain cancer. However, there remains a lack of systematic tools and standardized protocols for using handcrafted features in brain cancer histological analysis. In this study, we present a protocol for handcrafted feature analysis in brain cancer pathology (PHBCP) to systematically extract, analyze, model, and visualize handcrafted features from WSIs. The protocol enabled the discovery of biomarkers from WSIs through a series of well-defined steps. The PHBCP comprises seven main steps: (1) problem definition, (2) data quality control, (3) image preprocessing, (4) feature extraction, (5) feature filtering, (6) modeling, and (7) performance analysis. As an exemplary application, we collected pathological data of 589 patients from two cohorts and applied the PHBCP to predict the 2-year survival of glioblastoma multiforme (GBM) patients. Among the 72 models combining nine feature selection methods and eight machine learning classifiers, the optimal model combination achieved discriminative performance with an average area under the curve (AUC) of 0.615 over 100 iterations under five-fold cross-validation. In the external validation cohort, the optimal model combination achieved a generalization performance with an AUC of 0.594. We provide an open-source code repository (GitHub website: <https://github.com/XuanjunLu/>

Xuanjun Lu, Yawen Ying and Jing Chen contributed equally to this work and shared the first authorship.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Brain-X* published by John Wiley & Sons Australia, Ltd on behalf of Ainoohui Medical Technology.

PHBCP) to facilitate effective collaboration between medical and technical experts, thereby advancing the field of computational pathology in brain cancer.

KEYWORDS

brain cancer pathology, handcrafted features, protocol, whole-slide images

1 | INTRODUCTION

Histopathological slides, recognized as the “gold standard” for tumor diagnosis,^[1] hold significant value not only in the morphological assessment of diseases but also in critical biomedical information such as tumor heterogeneity, micro-environment characteristics, and molecular phenotypes.^[2] In the diagnosis and treatment of brain cancer, histopathological analysis using hematoxylin and eosin (H&E)-stained slides provides indispensable diagnostic evidence for clinical decision-making. However, the traditional diagnostic workflow relies on pathologists' visual inspection of slides under a microscope from low to high magnification. This qualitative analytical approach has inherent limitations. First, subjective interpretation is prone to variability because of differences in experience, leading to diagnostic inconsistency.^[3] Second, conventional examination has difficulty in quantitatively extracting subvisual tissue features, which may include crucial prognostic information.^[4] Third, the efficiency bottleneck of manual analysis becomes apparent when dealing with large numbers of slides.^[5] Thus, an accurate, objective, and interpretable protocol is an important goal in brain cancer pathology.

In recent years, the development of digitized whole-slide image (WSI) technology has revolutionized the field of pathology by enabling permanent digital storage of histopathological slides.^[6] Leveraging WSI, handcrafted features—i.e., features extracted by manually designed algorithms guided by domain-specific prior knowledge and empirical expertise—have been employed to extract attributes for the discovery of biomarkers. These features can derive quantitative prognostic, predictive, and other pathological information from H&E-stained WSIs, potentially transforming precision oncology and improving patient outcomes. Over the past few years, biomarkers based on handcrafted features have been extensively applied in numerous cancers, including head and neck squamous cell carcinoma,^[7] urothelial cancer,^[8] papillary thyroid carcinoma,^[9] hepatocellular carcinoma,^[10,11] lung cancer,^[12,13] oropharyngeal squamous cell carcinoma,^[14] and colorectal cancer.^[15] However, their application in brain cancer remains insufficient and is scarcely reported in the literature.

Current brain cancer research mainly focuses on radiology images,^[16–19] with few studies dedicated to histopathological analysis. Even within the limited pathological investigations, deep learning approaches dominate,^[20] yet their “black-box” nature results in a lack of interpretability, significantly hindering broader clinical translation. The substantial computational resource requirements and intricate preprocessing pipelines associated with deep-learning models pose additional barriers, further limiting their accessibility and practical adoption in clinical settings.

Key points

What is already known about this topic?

- Hematoxylin and eosin (H&E)-stained whole-slide images (WSIs) contain abundant information about cellular and tissue morphology. However, there remains a lack of systematic tools and standardized protocols for using handcrafted features in brain cancer histological analysis.

What does this study add?

- This study presents a protocol for handcrafted feature analysis in brain cancer pathology to systematically extract, analyze, model, and visualize handcrafted features from WSIs, thereby promoting efficient collaboration between medical and technical experts.

In this study, we present a protocol for handcrafted feature analysis in brain cancer pathology (PHBCP) based on H&E-stained WSIs. The protocol represents a simple, flexible, and modular open-source pipeline. We demonstrate the use of the protocol using two cohorts of glioblastoma multiforme (GBM). By following this protocol, medical and technical experts will be able to promote communication and collaboration, develop novel biomarkers, and collectively tackle clinical challenges in brain cancer, ultimately improving patient outcomes.

2 | METHODS

2.1 | Overview of the protocol

The protocol comprises seven main steps (Figure 1): (1) problem definition, (2) data quality control, (3) image preprocessing, (4) feature extraction, (5) feature filtering, (6) modeling, and (7) performance analysis. The problem-definition step specifies the precise clinical objectives to be analyzed. The data quality-control step aims to eliminate slides that contain contamination, artifacts, and other issues. The image-preprocessing step provides a WSI-based standardized preprocessing process, encompassing the region of interest (ROI) acquisition, WSI slicing, and color normalization. The feature-extraction step details the types, roles, extraction, and aggregation approaches for handcrafted features. The feature-filtering step refines a large set of redundant features to identify those most relevant to the label. The modeling step involves

constructing models based on the filtered features to achieve optimal analytical performance. Lastly, the performance-analysis step visualizes important features and conducts downstream analyses. In this study, we use H&E-stained WSIs from two independent cohorts [The Cancer Genome Atlas (TCGA) and The Cancer Imaging Archive (TCIA)]^[21] to demonstrate how to use the PHBCP. The uniqueness of the protocol lies in its use of interpretable handcrafted features, rather than deep learning, to establish the complex relationship between WSI and the clinical question. Subsequently, Sections 2.2–2.8 correspond to steps 1–7, respectively.

2.2 | Problem definition

First, the clinical problem is determined, followed by the collection of tissue samples and corresponding pathology reports from patients. Slicing and staining operations were performed on the tissue samples. Stained tissue slides are converted into digital whole-slide histology images using digital scanners for subsequent computational pathology analysis. For example, one may want to interrogate the relationship between the nuclear shape features and the grade of tumor of the central nervous system.

2.3 | Data quality control

The collected WSIs may be scanned by diverse clinical personnel utilizing different scanners across multiple institutions, which inevitably results in heterogeneous image quality. To mitigate the potential influence of these external

variables on the experimental results, it is necessary to exclude substandard slides. However, the manual inspection of image quality in a high-throughput experimental context is often unfeasible. Consequently, a tool referred to as HistoQC^[22] is usually employed as an objective and rapid quality control process, identifying and flagging issues such as reagent contamination, artifacts, tissue folding, and staining irregularities, thereby enabling automated assessment of WSI quality. Combined with a pathological image viewer-QuPath,^[23] substandard data are excluded. Multicenter batch effects, such as stain variations, which may affect the robustness of the model. It is recommended to use Batch Effect Explorer^[24] to unveil the batch effect between cohorts.

2.4 | Image preprocessing

In Section 2.3, we create a tissue mask, which excludes artifacts, tissue folding, and other external influences. In this step, the ROI of the current task is extracted from the tissue mask and split into image tiles of the desired size and magnification, for example, 224×224 pixels at 20x magnification, using the OpenSlide library,^[25] which is a Python library used for processing WSIs. To ensure a relatively dense tissue distribution, only those image tiles containing more than a certain proportion of tissue area are selected, for example, with 80% tissue area. To reduce computational load and avoid subjective selection bias, K-means clustering is performed on the tiles of each WSI to group tiles with similar phenotypes together. To ensure no critical regions are missed, E tiles are selected from each cluster, resulting in $L \times E$ tiles being used to characterize

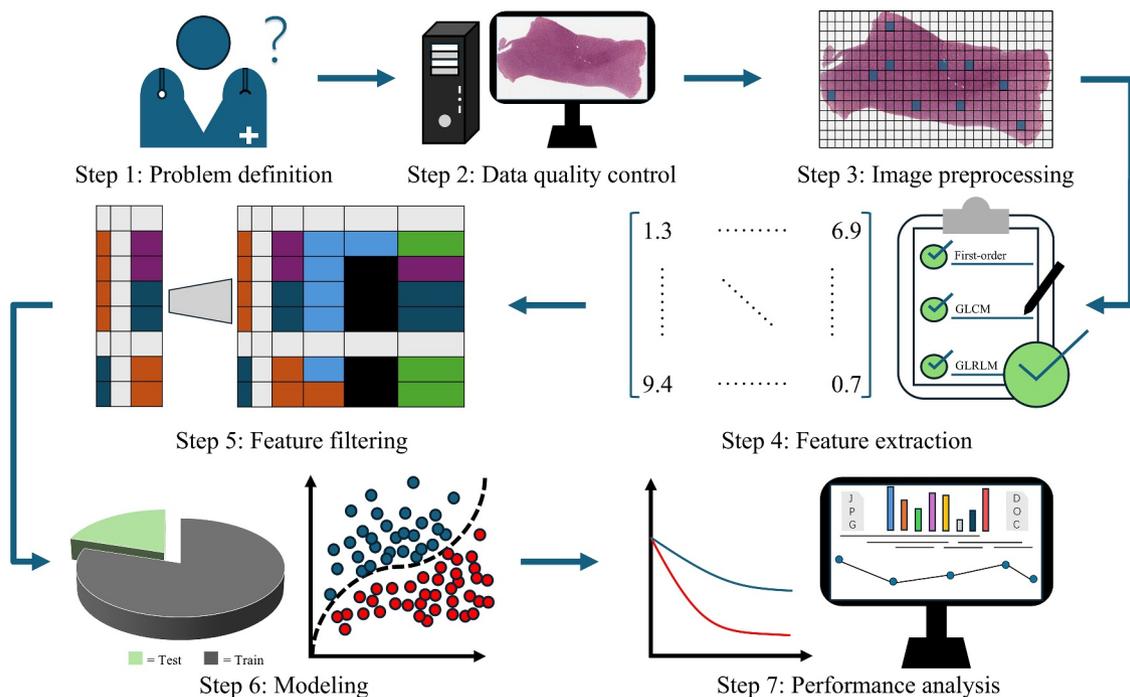


FIGURE 1 Conceptual overview of the protocol. Seven main steps transform images into quantitative feature information, thereby supporting experimental conclusions (Sections 2.2–2.8 correspond to steps 1–7 shown in Figure 1, respectively).

each patient, in which L is the number of clusters. Due to staining variations across different centers, deconvolution-based color normalization^[26–28] is applied to the selected tiles to eliminate color discrepancies between WSIs.

2.5 | Feature extraction

Feature extraction refers to the process of transforming images into quantitatively described feature values. Five types of handcrafted features are provided in this protocol: first-order statistics ($n = 17$), gray level co-occurrence matrix (GLCM) features ($n = 24$), gray level run length matrix (GLRLM) features ($n = 16$), nuclear shape features ($n = 25$), and nuclear texture features ($n = 13$). First-order statistics describe the distribution of pixel intensity values within the tissue region. The GLCM features characterize the frequency of pairs of identical pixel intensity values in the tissue region. The GLRLM features describe the continuity of pixel intensity values over a specified distance in the tissue region. The nuclear shape features are employed to quantify the geometric properties of nuclear contours, thereby reflecting the characteristic patterns of nuclear deformation and cellular morphological changes during tumor progression. The nuclear texture features, by quantifying the heterogeneity and spatial configuration of chromatin distribution within the nucleus, enable an in-depth analysis of the structural distortions in the intranuclear microenvironment during tumor evolution. In total, 95 features can be extracted for each image tile. The details of the five types of features are as follows:

First-order statistics^[29]:

$$\text{Energy} = \sum_{i=1}^{N_p} (X(i) + c)^2 \quad (1)$$

$$\text{Entropy} = - \sum_{i=1}^{N_g} p(i) \log_2(p(i) + \epsilon) \quad (2)$$

$$\text{Minimum} = \min(X) \quad (3)$$

$$\text{The 10}^{\text{th}} \text{percentile of } X \quad (4)$$

$$\text{The 90}^{\text{th}} \text{percentile of } X \quad (5)$$

$$\text{Maximum} = \max(X) \quad (6)$$

$$\text{Mean} = \frac{1}{N_p} \sum_{i=1}^{N_p} X(i) \quad (7)$$

$$\text{Median} = \text{med}(X) \quad (8)$$

$$\text{Interquartile range} = 75^{\text{th}} \text{ percentile} - 25^{\text{th}} \text{ percentile} \quad (9)$$

$$\text{Range} = \max(X) - \min(X) \quad (10)$$

$$\text{Mean absolute deviation} = \frac{1}{N_p} \sum_{i=1}^{N_p} |X(i) - \bar{X}| \quad (11)$$

Robust mean absolute deviation

$$= \frac{1}{N_{10-90}} \sum_{i=1}^{N_{10-90}} |X_{10-90}(i) - \bar{X}_{10-90}| \quad (12)$$

$$\text{Root mean squared} = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) + c)^2} \quad (13)$$

$$\text{Skewness} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2} \right)^3} \quad (14)$$

$$\text{Kurtosis} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^4}{\left(\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2 \right)^2} \quad (15)$$

$$\text{Variance} = \frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2 \quad (16)$$

$$\text{Uniformity} = \sum_{i=1}^{N_g} p(i)^2 \quad (17)$$

where X is a set of N_p pixels included in the ROI, \bar{X} is the average of X , c is an optional drift value, ϵ is an arbitrarily small positive number. $P(i)$ is the first-order histogram with N_g discrete intensity levels, and $p(i)$:

$$p(i) = \frac{P(i)}{N_p} \quad (18)$$

GLCM features^[29]:

$$\text{Autocorrelation} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)ij \quad (19)$$

$$\text{Joint average} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)i \quad (20)$$

$$\text{Cluster prominence} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \left(i + j - \mu_x - \mu_y \right)^4 p(i,j) \quad (21)$$

$$\text{Cluster shade} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \left(i + j - \mu_x - \mu_y \right)^3 p(i,j) \quad (22)$$

$$\text{Cluster tendency} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^2 p(i, j) \quad (23)$$

$$\text{Contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i, j) \quad (24)$$

$$\text{Correlation} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) ij - \mu_x \mu_y}{\sigma_x(i) \sigma_y(j)} \quad (25)$$

$$\text{Difference average (DA)} = \sum_{k=0}^{N_g-1} k p_{x-y}(k) \quad (26)$$

$$\text{Difference entropy} = \sum_{k=0}^{N_g-1} p_{x-y}(k) \log_2(p_{x-y}(k) + \epsilon) \quad (27)$$

$$\text{Difference variance} = \sum_{k=0}^{N_g-1} (k - DA)^2 p_{x-y}(k) \quad (28)$$

$$\text{Joint energy} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i, j))^2 \quad (29)$$

$$\text{Joint entropy} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2(p(i, j) + \epsilon) \quad (30)$$

$$\begin{aligned} &\text{Informational measure of correlation 1} \\ &= \frac{HXY - HXY1}{\max(HX, HY)} \end{aligned} \quad (31)$$

$$\begin{aligned} &\text{Informational measure of correlation 2} \\ &= \sqrt{1 - e^{-2(HXY2-HXY)}} \end{aligned} \quad (32)$$

$$\text{Inverse difference moment} = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + k^2} \quad (33)$$

Maximal correlation coefficient

$$= \sqrt{\text{second largest eigenvalue of } Q} \quad (34)$$

$$Q(i, j) = \sum_{k=0}^{N_g} \frac{p(i, k) p(j, k)}{p_x(i) p_y(k)}$$

Inverse difference moment normalized

$$= \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + \left(\frac{k^2}{N_g^2}\right)} \quad (35)$$

$$\text{Inverse difference} = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + k} \quad (36)$$

$$\text{Inverse difference normalized} = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + \left(\frac{k}{N_g}\right)} \quad (37)$$

$$\text{Inverse variance} = \sum_{k=1}^{N_g-1} \frac{p_{x-y}(k)}{k^2} \quad (38)$$

$$\text{Maximum probability} = \max(p(i, j)) \quad (39)$$

$$\text{Sum average} = \sum_{k=2}^{2N_g} p_{x+y}(k) k \quad (40)$$

$$\text{Sum entropy} = \sum_{k=2}^{2N_g} p_{x+y}(k) \log_2(p_{x+y}(k) + \epsilon) \quad (41)$$

$$\text{Sum squares} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i, j) \quad (42)$$

where ϵ is an arbitrarily small positive number, $P(i, j)$ is the co-occurrence matrix, N_g is the number of discrete intensity levels, σ_x is the standard deviation of p_x , σ_y is the standard deviation of p_y , and the other parameters are as follows:

$$p(i, j) = \frac{P(i, j)}{\sum P(i, j)} \quad (43)$$

$$p_x(i) = \sum_{j=1}^{N_g} p(i, j) \quad (44)$$

$$p_y(j) = \sum_{i=1}^{N_g} p(i, j) \quad (45)$$

$$\mu_x = \sum_{i=1}^{N_g} p_x(i) i \quad (46)$$

$$\mu_y = \sum_{j=1}^{N_g} p_y(j) j \quad (47)$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j), i + j = k, k = 2, 3, \dots, 2N_g \quad (48)$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j), |i - j| = k, k = 0, 1, \dots, N_g - 1 \quad (49)$$

$$HX = - \sum_{i=1}^{N_g} p_x(i) \log_2(p_x(i) + \epsilon) \quad (50)$$

$$HY = - \sum_{j=1}^{N_g} p_y(j) \log_2(p_y(j) + \epsilon) \quad (51)$$

$$\text{HXY} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log_2(p(i,j) + \epsilon) \quad (52)$$

$$\text{HXY1} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log_2(p_x(i)p_y(j) + \epsilon) \quad (53)$$

$$\text{HXY2} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j) \log_2(p_x(i)p_y(j) + \epsilon) \quad (54)$$

GLRLM features^[29]:

$$\text{Short run emphasis} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{j^2}}{N_r(\theta)} \quad (55)$$

$$\text{Long run emphasis} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)j^2}{N_r(\theta)} \quad (56)$$

$$\text{Gray level non - uniformity} = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} P(i,j|\theta) \right)^2}{N_r(\theta)} \quad (57)$$

Gray level non - uniformity normalized

$$= \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} P(i,j|\theta) \right)^2}{N_r(\theta)^2} \quad (58)$$

$$\text{Run length non - uniformity} = \frac{\sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_g} P(i,j|\theta) \right)^2}{N_r(\theta)} \quad (59)$$

Run length non - uniformity normalized

$$= \frac{\sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_g} P(i,j|\theta) \right)^2}{N_r(\theta)^2} \quad (60)$$

$$\text{Run percentage} = \frac{N_r(\theta)}{N_p} \quad (61)$$

$$\text{Gray level variance} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)(i - \mu)^2 \quad (62)$$

$$\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)i$$

$$\text{Run variance} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)(j - \mu)^2 \quad (63)$$

$$\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)j$$

$$\text{Run entropy} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta) \log_2(p(i,j|\theta) + \epsilon) \quad (64)$$

$$\text{Low gray level run emphasis} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{i^2}}{N_r(\theta)} \quad (65)$$

$$\text{High gray level run emphasis} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)i^2}{N_r(\theta)} \quad (66)$$

$$\text{Short run low gray level emphasis} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{i^2j^2}}{N_r(\theta)} \quad (67)$$

$$\text{Short run high gray level emphasis} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)i^2}{j^2}}{N_r(\theta)} \quad (68)$$

$$\text{Long run low gray level emphasis} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)j^2}{i^2}}{N_r(\theta)} \quad (69)$$

Long run high gray level emphasis

$$= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)i^2j^2}{N_r(\theta)} \quad (70)$$

where ϵ is an arbitrarily small positive number, N_g is the number of discrete intensity levels, N_r is the number of discrete run lengths, N_p is the number of pixels, $P(i,j|\theta)$ is the run length matrix for a direction, and the other parameters are as follows:

$$N_r(\theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta), 1 \leq N_r(\theta) \leq N_p \quad (71)$$

$$p(i,j|\theta) = \frac{P(i,j|\theta)}{N_r(\theta)} \quad (72)$$

Nuclear shape features are as follows:

$$\text{Area ratio} = \frac{\text{area}}{\text{area}_{\max}} \quad (73)$$

where area is the area of a nucleus, and area_{\max} represents the area of a circle with a radius equal to the maximum Euclidean distance from the centroid of the nucleus to its contour points.

$$\text{Distance ratio} = \frac{\text{distance}_{\text{mean}}}{\text{distance}_{\text{max}}} \quad (74)$$

where $\text{distance}_{\text{mean}}$ is the average Euclidean distance from the centroid to the contour points, and $\text{distance}_{\text{max}}$ is the maximum Euclidean distance from the centroid to the contour points in a nucleus.

$$\text{Distance std} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (dc_i - \overline{dc})^2} \quad (75)$$

where dc is the set of the centroid to contour points standardized Euclidean distances in a nucleus, \overline{dc} is the average of dc , and N is the number of the contour points.

$$\text{Distance var} = \text{Distance std}^2 \quad (76)$$

$$\text{Long or short distance ratio} = \frac{\text{disum}_L}{\text{disum}_S} \quad (77)$$

where disum_L is the long-distance sum of contour points in a nucleus. Specifically, it is calculated by uniformly sampling a certain number of points among the contour points with a longer sampling interval and then summing up the Euclidean distances between adjacent sampled points. Similarly, disum_S is the short-distance sum of contour points.

$$\text{Perimeter ratio} = \frac{\text{perimeter}^2}{\text{area}} \quad (78)$$

where perimeter is the perimeter of a nucleus.

$$\text{Smoothness} = \sum_{i=1}^N \left| d_i - \frac{d_{i-1} + d_{i+1}}{2} \right| \quad (79)$$

where d_i is the Euclidean distance from the contour point of a nucleus to the centroid.

$$\text{Invariant moment}_1 = \eta_{20} + \eta_{02} \quad (80)$$

$$\text{Invariant moment}_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (81)$$

$$\text{Invariant moment}_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (82)$$

$$\text{Invariant moment}_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (83)$$

$$\begin{aligned} \text{Invariant moment}_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\ &\quad \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &\quad \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \end{aligned} \quad (84)$$

$$\begin{aligned} \text{Invariant moment}_6 &= (\eta_{20} - \eta_{02}) \\ &\quad \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned} \quad (85)$$

$$\begin{aligned} \text{Invariant moment}_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \\ &\quad \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \\ &\quad \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \end{aligned} \quad (86)$$

where η_{pq} is the normalized central moment^[30] calculated based on the contour point set of a nucleus.

$$\text{Fractal dimension} = \text{slope} \left(\left\{ \lg \left(\frac{1}{r_k} \right), \lg L(r_k) \right\}_{k=1}^m \right) \quad (87)$$

where *slope* is the slope of linear regression, r_k is the k^{th} sampling interval of the contour points in a nucleus, $L(r_k)$ is the fractal length at the k^{th} sampling interval,^[31] and $m = N/2$

$$\text{Fourier descriptor} = z[\text{Re}(Z_0), \text{Re}(Z_1), \dots, \text{Re}(Z_9)] \quad (88)$$

where z is the discrete Fourier transform of a set of contour points in a nucleus, and Z is the Fourier descriptor.

Nuclear texture features are as follows:

$$\text{Contrast energy} = \sum_k (c_k)^2 p_c(c_k) \quad (89)$$

where c_k is the absolute difference of intensity-level pairs, and $p_c(c_k)$ is the normalized co-occurrence probability of the corresponding intensity-level pairs.

$$\text{Contrast inverse moment} = \sum_k \frac{1}{1 + c_k^2} p_c(c_k) \quad (90)$$

$$\text{Contrast average(CA)} = \sum_k c_k p_c(c_k) \quad (91)$$

$$\text{Contrast variance} = \sum_k (c_k - CA)^2 p_c(c_k) \quad (92)$$

$$\text{Contrast entropy} = - \sum_k p_c(c_k) \ln p_c(c_k) \quad (93)$$

$$\text{Intensity average (IA)} = \sum_l m_l p_m(m_l) \quad (94)$$

where m_l is the average of intensity-level pairs, and $p_m(m_l)$ is the normalized co-occurrence probability of the corresponding intensity-level pairs.

$$\text{Intensity variance} = \sum_l (m_l - IA)^2 p_m(m_l) \quad (95)$$

$$\text{Intensity entropy} = -\sum_l p_m(m_l) \ln p_m(m_l) \quad (96)$$

The other features, including entropy, energy, correlation, informational measure of correlation 1, and informational measure of correlation 2, are derived from GLCM features. Note that there are other types of handcrafted features, such as the spatial interaction between histological primitives,^[32,33] that can be integrated into the PHBCP.

In summary, for each WSI, a feature matrix of size $(L \times E) \times N$ can be obtained, in which N is the total number of extracted features. Based on this, users can aggregate and concatenate features using various methods according to their needs, for example, mean, standard deviation, and skewness.

2.6 | Feature filtering

Feature filtering is a critical step in machine learning that identifies the most relevant and informative features from the feature matrix while eliminating redundant and irrelevant features. This process not only reduces the dimensionality of features but also mitigates overfitting, enhances model interpretability, and improves computational efficiency. In this study, the protocol employs comprehensive feature-selection methods to ensure robust and reliable features.

Firstly, to address multicollinearity and reduce feature redundancy, the PHBCP calculates the pairwise Spearman's rank correlation coefficient matrix among all features. Features with a correlation coefficient greater than the threshold, for example, 0.9 is generally used for removing features having more than 90% synchronicity, are removed. Subsequently, the feature matrix is standardized using the Z-score method.^[34] The above steps ensure that only non-redundant features are retained for further analysis.

Secondly, the protocol uses comprehensive feature selection methods to capture diverse aspects of feature importance and interactions. These methods include Lasso regression (LR), random forest (RF), elastic-net (EN), recursive feature elimination (RFE), univariate analysis (UA), minimum redundancy maximum relevance (MRMR), t -test, Wilcoxon rank-sum test (WRST), and mutual information (MI) methods, which are implemented in Python

using scikit-learn, mrmr_selection, and scipy libraries. Here, users can select an appropriate number of features based on the sample size to achieve a suitable predictive performance and avoid overfitting and the curse of dimensionality.

Finally, each feature selection method is integrated with a classifier, and multi-fold cross-validation with user-defined iterations is performed to assess the consistency and reliability of the selected features across multiple data splits.

By combining correlation-based filtering with comprehensive feature selection methods and cross-validation, the protocol provides a robust framework to identify the most discriminative features while minimizing redundancy and overfitting.

2.7 | Modeling

In this step, the PHBCP combines the feature selection methods and classifiers one by one to construct potential models. The protocol employs eight machine learning classifiers, including quadratic discriminant analysis (QDA), linear discriminant analysis (LDA), RF, K-nearest neighbors (KNN), linear support vector machine (LSVM), Gaussian naive bayes (GNB), stochastic gradient descent (SGD), and adaptive boosting (AdaBoost), which are implemented in Python using scikit-learn library. The eight classifiers are implemented in conjunction with the top features selected using the nine feature selection methods. The classifiers are evaluated with multi-fold cross-validation with user-defined iterations within a training cohort. Ultimately, the PHBCP identifies the optimal model combination from the 72 different combinations based on the highest average area under the curve (AUC) across user-defined iterations.

2.8 | Performance analysis

Based on the optimal model combination determined during the modeling phase, one can conduct performance analysis focusing on the visualization of the top features' feature distribution, feature importance, and survival analysis in the external validation cohort.

Firstly, the PHBCP calculates the mean, median, and skewness of the top feature values, and then divides the feature values into equal bins to obtain 10 intervals. The distribution of the top features is visualized using histograms overlaid with kernel density estimation (Gaussian kernel) curves. Secondly, a horizontal bar is used to visualize the selection frequency percentage of the top features across the multi-fold cross-validation with user-defined iterations, highlighting the most important features and understanding their contributions to the optimal model combination. A higher selection frequency indicates a greater predictive contribution to the model and a stronger clinical relevance to

the research question. Finally, the PHBCP locks down the optimal model combination and corresponding top features in the training cohort and conducts survival analysis in the external validation cohort. A Kaplan–Meier curve is used to evaluate, for example, the survival probability between

predicted long- and short-term survival patients. The log-rank test was employed to examine survival differences, indicating the prognostic significance of the categorical variable on the survival endpoint. All tests are two-sided, with the significance level set at 0.05.

TABLE 1 Baseline and clinical characteristics in the training cohort and external validation cohort.

	Training cohort (<i>N</i> = 207)	External validation cohort (<i>N</i> = 57)	<i>p</i>
Age			0.8546
≤ 65	150 (72.5%)	42 (73.7%)	
> 65	57 (27.5%)	15 (26.3%)	
Sex			0.9723
Male	124 (59.9%)	34 (59.6%)	
Female	83 (40.1%)	23 (40.4%)	
Race			<0.0001
White	189 (91.3%)	24 (42.1%)	
Asian	3 (1.4%)	19 (33.3%)	
Other	11 (5.3%)	13 (22.8%)	
Unknown	4 (1.9%)	1 (1.8%)	
History of LGG			
Yes	3 (1.4%)	NA	
No	204 (98.6%)	NA	
Event status			0.4987
Occurred	191 (92.3%)	51 (89.5%)	
Censored	16 (7.7%)	6 (10.5%)	
Survival status			0.5731
Long term (>2 years)	54 (26.1%)	17 (29.8%)	
Short term (≤2 years)	153 (73.9%)	40 (70.2%)	

The *p*-values were calculated by Pearson's Chi-square test.

3 | RESULTS

Given that GBM is the most common and aggressive type of malignant primary brain tumor,^[35,36] we used a GBM survival prediction problem as an exemplary task to demonstrate how to use the PHBCP. First, WSI data and corresponding basic clinical information were obtained from two independent cohorts through The Cancer Genome Atlas (TCGA) (389 cases) and The Cancer Imaging Archive (TCIA, 200 cases).^[21] Subsequently, the entire tissue region was defined as the ROI, with overall survival (OS) established as the endpoint. For patients whose death occurred during the follow-up period, OS of less than or equal to 2 years was classified as short-term survival, while OS greater than 2 years was classified as long-term survival. For censored patients, the final follow-up time was used as the OS, with cases whose OS exceeded 2 years classified as long-term survival, while cases with an OS of less than or equal to 2 years were considered missing information and excluded from the analysis. OS was defined as the time from surgery to death.

In Section 2.3, HistoQC was used to exclude WSIs with fewer than 250,000 usable pixels as well as those exhibiting significant issues such as extensive blurring, tissue folding, reagent contamination, and abnormal staining. The detailed settings are provided in the supplementary parameter settings for HistoQC. For patients with multiple slides, one slide was selected for subsequent analysis based on its image quality using the pathological image viewer-QuPath. Inclusion and exclusion criteria were applied to both cohorts. Inclusion criteria: (1) patients who underwent resection and were confirmed to have GBM through surgical pathological specimens; (2) patients whose OS information was complete; (3) patients who contained

TABLE 2 AUC performance of eight different classifiers with nine different feature selection methods in the training cohort.

	QDA	LDA	RF	KNN	LSVM	GNB	SGD	AdaBoost
LR	0.595 ± 0.022	0.550 ± 0.029	0.597 ± 0.037	0.591 ± 0.028	0.511 ± 0.034	0.601 ± 0.028	0.533 ± 0.038	0.610 ± 0.033
RF	0.593 ± 0.028	0.564 ± 0.025	0.586 ± 0.033	0.580 ± 0.029	0.521 ± 0.039	0.582 ± 0.030	0.536 ± 0.046	0.578 ± 0.032
EN	0.572 ± 0.029	0.547 ± 0.021	0.598 ± 0.038	0.579 ± 0.032	0.527 ± 0.034	0.567 ± 0.027	0.532 ± 0.038	0.584 ± 0.035
RFE	0.585 ± 0.032	0.509 ± 0.032	0.577 ± 0.036	0.583 ± 0.039	0.454 ± 0.034	0.575 ± 0.022	0.513 ± 0.040	0.572 ± 0.034
UA	0.576 ± 0.028	0.565 ± 0.023	0.577 ± 0.034	0.557 ± 0.033	0.530 ± 0.032	0.553 ± 0.028	0.541 ± 0.040	0.590 ± 0.033
MRMR	0.573 ± 0.027	0.566 ± 0.022	0.586 ± 0.032	0.558 ± 0.033	0.540 ± 0.031	0.554 ± 0.026	0.545 ± 0.034	0.595 ± 0.040
<i>t</i> -test	0.572 ± 0.032	0.568 ± 0.020	0.579 ± 0.035	0.562 ± 0.030	0.537 ± 0.029	0.555 ± 0.024	0.535 ± 0.038	0.594 ± 0.035
WRST	0.596 ± 0.031	0.577 ± 0.025	0.588 ± 0.041	0.574 ± 0.036	0.532 ± 0.038	0.587 ± 0.028	0.548 ± 0.036	0.605 ± 0.034
MI	0.594 ± 0.025	0.528 ± 0.032	0.567 ± 0.039	0.615 ± 0.027	0.467 ± 0.036	0.591 ± 0.024	0.531 ± 0.041	0.586 ± 0.037

Note: The bold values represent the AUC and standard deviation of the optimal model combination.

follow-up information. Exclusion criteria: (1) missing H&E-stained WSIs of 20x magnification and (2) histopathological slides that did not meet the standard requirements for analysis. Ultimately, 207 patients from TCGA were included as the training cohort, while 57 patients from TCIA were incorporated as the external validation cohort. Table 1 presents a summary of the basic clinical information and distribution differences between the training cohort and the external validation cohort.

In Section 2.4, the tissue mask generated through HistQC was aligned with the corresponding WSI, and image tiles of 224×224 pixels were extracted at 20x magnification without overlap. The tiles from each WSI were clustered into 10 classes, and 50 tiles were randomly selected from each class to ensure a comprehensive analysis of all regions. Stain normalization was performed on 500 selected tiles for subsequent feature extraction.

In Section 2.5, three types of features were extracted: First-order statistics, GLCM features, and GLRLM features, totaling 57 features. For each WSI, a feature matrix of size 500×57 was obtained, and the feature matrix was averaged to aggregate a 1×57 feature vector.

In Sections 2.6 and 2.7, in order to avoid the curse of dimensionality and overfitting, we set the number of top features to six, based on the experimental experience that the number of selected features should be approximately one-tenth of the number of minority class samples. In this study, the training cohort contained 54 minority class samples and thus the top six features were chosen. The Spearman correlation threshold was set to 0.9. The predictive performance of the model in the training cohort was evaluated by performing 100 iterations of five-fold cross-validation in 72 model combinations to avoid incidental results. The detailed results are presented in Table 2. Table 2

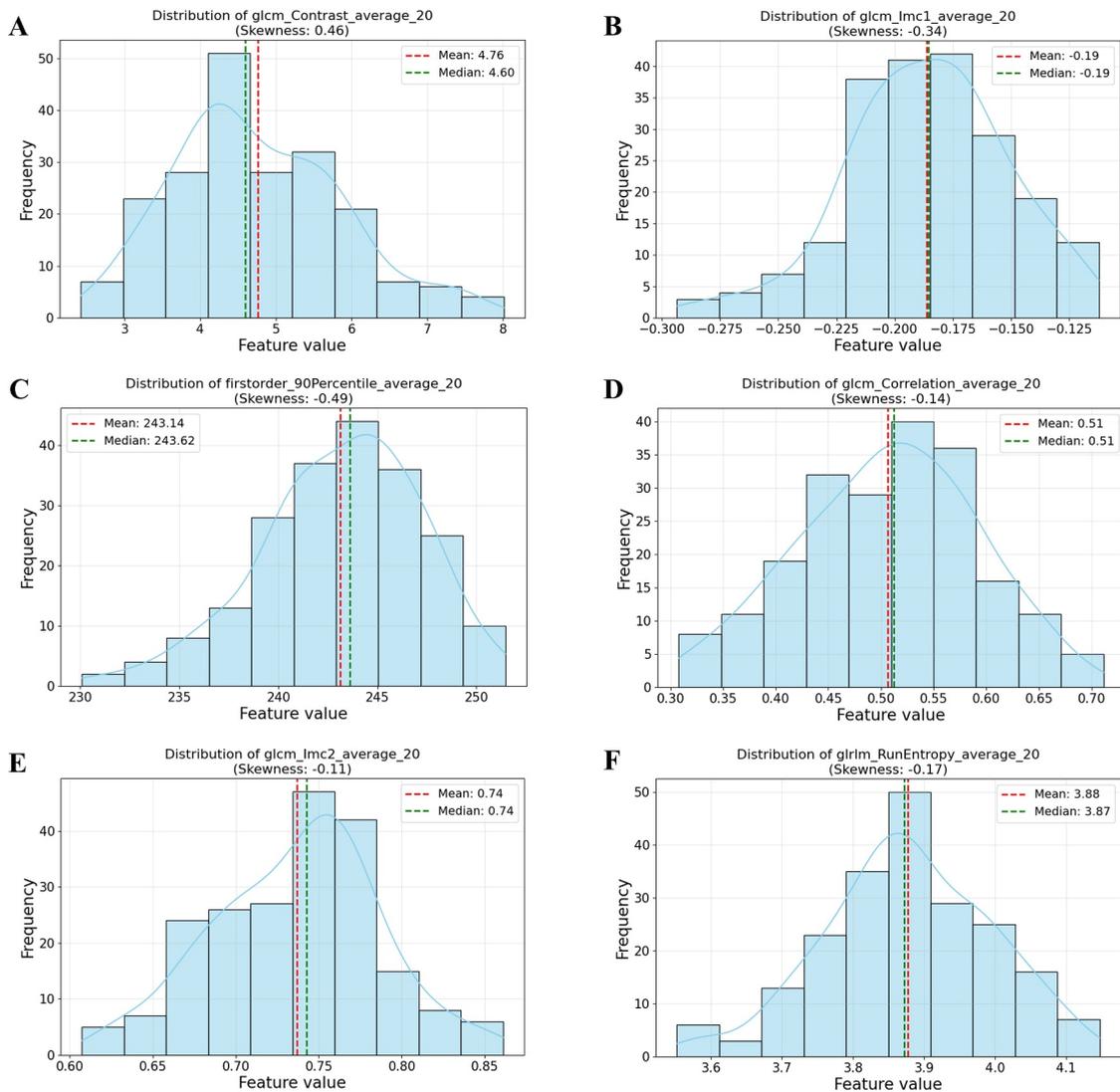


FIGURE 2 The distribution of the top six feature values in the training cohort (A–F) approximate a normal distribution. The x-axis represents the binned feature intervals (10 bins), and the y-axis indicates the frequency of samples.

shows that the optimal model combination was MI-KNN (AUC = 0.615 ± 0.027). The results for accuracy and F1 score are presented in Tables S1 and S2, respectively.

The top six features and their distribution from the MI-KNN model combination in Section 2.8 are illustrated in Figure 2. Six features closely approximated a normal distribution. The normal distribution indicates that these features were relatively stable within the patients, reflecting limited interindividual variability. The features shown in Figure 3 were used to analyze the contribution of the top 12 selected features to the MI-KNN model combination. The top two features, `glcm_Contrast_average_20` and `glcm_Imc1_average_20`, were considered the most relevant to patient outcomes because of their highest selection frequency, underscoring their potential as biomarkers for GBM survival prediction. Finally, in the survival analysis of the external validation cohort, long-term survival patients showed higher survival probabilities compared with short-term survival patients, with statistically significant differences between the two groups (Figure 4), indicating that the constructed classification model had significant predictive value for survival endpoints in GBM patients. Based on the top six features of the MI-KNN model combination in the training cohort, a KNN classifier was trained. In the external validation cohort, the classification performance was observed with an AUC of 0.594, an ACC of 0.754, and an F1 score of 0.848.

4 | DISCUSSION

In this paper, we develop and present a PHBCP. The presented protocol, termed PHBCP, offers a systematic, modular, and open-source framework and provides WSI processing and analysis guidelines in brain cancer. The results and methodology outlined in this protocol demonstrate its potential to enhance the discriminability and efficiency of brain cancer prediction and prognosis.

Features can be primarily categorized into two types: handcrafted features and deep learning-derived features. Handcrafted features are those extracted through manually designed algorithms, typically based on domain-specific knowledge or experience, such as texture features, statistical features, and geometric features.^[32] Given an input, these features yield a fixed and interpretable output. In contrast, deep learning-derived features are primarily learned automatically from data by deep learning models, without the need for manual design. Examples include features extracted by ResNet,^[37] CONCH^[38] and UNI.^[39] In practical application, models based on handcrafted features can provide interpretable and clinically relevant insights, which are essential for building trust among medical and technical experts. In contrast, although deep learning models have shown excellent performance in many tasks, these models typically rely on large amounts of data and

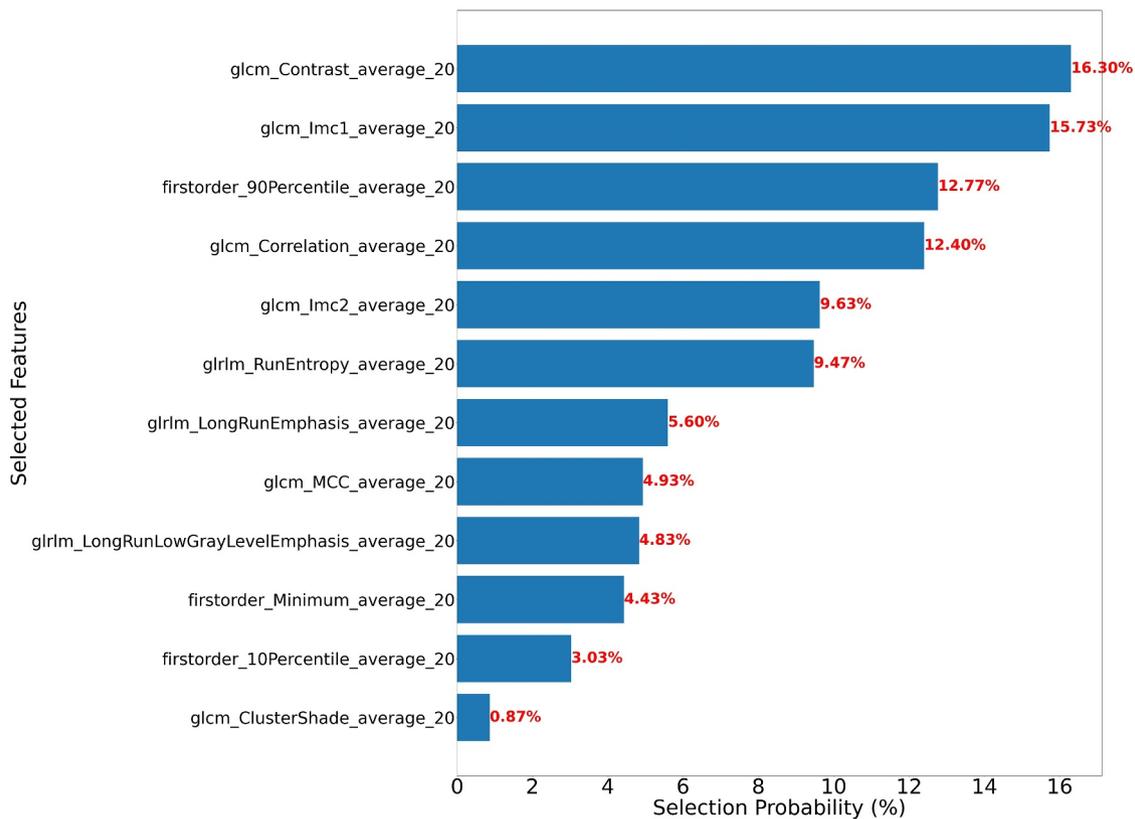


FIGURE 3 The top 12 features at selection frequency and their percentage contributions to the MI-KNN model combination across 100 iterations of five-fold cross-validation.

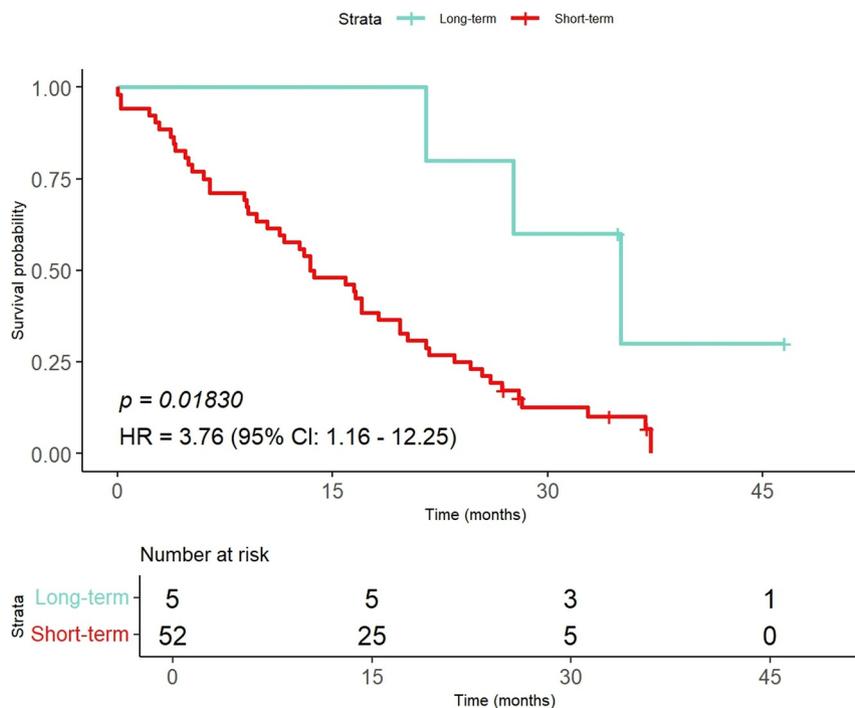


FIGURE 4 The Kaplan–Meier curve of GBM patients from the external validation cohort was stratified by the MI-KNN model combination into long-term and short-term survival groups.

cannot extract discriminative features from small samples. Additionally, the internal feature representations of deep learning models are complex, their decision-making processes are opaque, and it is difficult to explain their reasoning logic, resulting in poor interpretability. It is also challenging to incorporate specific medical prior knowledge into these deep learning models. These issues limit the widespread application of deep learning-derived features in clinical practice and, to some extent, hinder efficient collaboration between medical and technical experts. In this paper, PHBCP indicates the importance of these features in discovering novel biomarkers and improving the understanding of tumor heterogeneity, a key challenge in brain cancer research.

In the exemplary task of predicting 2-year survival in GBM, four of the top six features were related to GLCM, one feature was associated with first-order statistics, and one feature was connected to GLRLM. The `glcm_Contrast_average_20` feature was identified as the most prognostically relevant image feature because of its highest selection frequency. This feature is utilized to quantify the local intensity variation. Through analysis, it was observed that the magnitude of contrast is often closely correlated with the area and distribution of tumors and necrotic regions. This correlation can be used to explain the reasoning logic behind survival prediction using image features. When combined with multiomics data, the biological significance underlying these images can be further elucidated.

By providing a step-by-step guide, the protocol enables seamless collaboration between medical and technical

experts, fostering the development of innovative solutions to clinical problems. In this paper, the 2-year survival prediction in GBM serves merely as an exemplary task. Researchers can also conduct other brain cancer-related analyses based on PHBCP, such as isocitrate dehydrogenase (IDH) mutation analysis. The open-source nature of the protocol ensures its accessibility to a wide range of researchers, promoting reproducibility and scalability across different institutions and datasets.

The protocol has limitations. Although we have established a protocol for handcrafted features in brain cancer pathology, it does not encompass all types of handcrafted features. However, other researchers can add comprehensive handcrafted features to the PHBCP. As new pathological insights emerge, the reliance on handcrafted features may require continuous refinement. We anticipate that future contributions from medical and technical experts will enhance and expand this protocol.

5 | CONCLUSION

The protocol presented in this study is a significant step forward in the analysis of handcrafted features for brain cancer pathology. By providing a structured and collaborative framework, it empowers pathologists and clinicians to harness histopathological data for improved brain cancer care. We anticipate that this protocol will serve as a valuable resource for the scientific community, driving innovation and promoting the diagnosis and treatment of brain cancer.

AUTHOR CONTRIBUTIONS

Xuanjun Lu: Methodology; data curation; validation; writing—original draft; software; investigation; formal analysis; visualization; funding acquisition. **Yawen Ying:** Methodology. **Jing Chen:** Data curation. **Zhiyang Chen:** Data curation. **Yuxin Wu:** Software; methodology. **Prateek Prasanna:** Data curation. **Xin Chen:** Resources; writing—review and editing. **Mingli Jing:** Writing—review and editing; resources. **Zaiyi Liu:** Writing—review and editing; resources. **Cheng Lu:** Conceptualization; methodology; project administration; funding acquisition; writing—review and editing; resources; software; writing—original draft.

ACKNOWLEDGMENTS

This study was supported by National Natural Science Foundation of China (82272084), National Key R&D Program of China (2023YFC3402800), Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application (2022B1212010011), and Post-graduate Innovation and Practical Ability Training Program of Xi'an Shiyong University (YCS23114144). The authors would like to thank the support provided by MediAI Hub, an advanced medical image analysis software developed and maintained by MediaLab. TCIA data used in this publication were generated by the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Image and data of The Cancer Genome Atlas (TCGA) are publicly available in <https://portal.gdc.cancer.gov/>. The Cancer Imaging Archive (TCIA) data used in this publication were generated by the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC).

ETHICS STATEMENT

Image and data of TCGA and TCIA are publicly available,^[21] and do not require ethics approval.

ORCID

Xuanjun Lu  <https://orcid.org/0009-0005-0537-0355>

Cheng Lu  <https://orcid.org/0000-0002-7651-3924>

REFERENCES

- El Nahhas OSM, Van Treeck M, Wölflein G, et al. From whole-slide image to biomarker prediction: end-to-end weakly supervised deep learning in computational pathology. *Nat Protoc.* 2025;20(1):293-316. <https://doi.org/10.1038/s41596-024-01047-2>
- Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng.* 2009;2:147-171. <https://doi.org/10.1109/RBME.2009.2034865>
- Elmore JG, Longton GM, Carney PA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA.* 2015;313(11):1122. <https://doi.org/10.1001/jama.2015.1405>
- Bhargava R, Madabhushi A. Emerging themes in image informatics and molecular analysis for digital pathology. *Annu Rev Biomed Eng.* 2016;18(1):387-412. <https://doi.org/10.1146/annurev-bioeng-112415-114722>
- Caicedo JC, Cooper S, Heigwer F, et al. Data-analysis strategies for image-based cell profiling. *Nat Methods.* 2017;14(9):849-863. <https://doi.org/10.1038/nmeth.4397>
- Pantanowitz L, Sinar JH, Henricks WH, et al. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the college of American pathologists pathology and laboratory quality center. *Arch Pathol Lab Med.* 2013;137(12):1710-1722. <https://doi.org/10.5858/arpa.2013-0093-CP>
- Corredor G, Nag R, Toro P, et al. 1205 AI-informed integration of spatial architecture of collagen and tumor-infiltrating lymphocytes from H&E images predicts response to immunotherapy in head and neck squamous cell carcinoma patients. In: *Regular and Young Investigator Award Abstracts.* BMJ Publishing Group Ltd; 2024. <https://doi.org/10.1136/jitc-2024-SITC2024.1205.A1344>
- Hammouda K, Tokuyama N, Corredor G, et al. 546 AI-informed features of the spatial architecture of tumor-infiltrating lymphocytes predict the benefit of immune checkpoint inhibitors in urothelial cancer. In: *Regular and Young Investigator Award Abstracts.* BMJ Publishing Group Ltd; 2024. <https://doi.org/10.1136/jitc-2024-SITC2024.0546.A620>
- Monabbati S, Fu P, Asa SL, et al. Machine vision-detected peritumoral lymphocytic aggregates are associated with disease-free survival in patients with papillary thyroid carcinoma. *Lab Invest.* 2024;104(12):102168. <https://doi.org/10.1016/j.labinv.2024.102168>
- Lin S, Yong J, Zhang L, et al. Applying image features of proximal paracancerous tissues in predicting prognosis of patients with hepatocellular carcinoma. *Comput Biol Med.* 2024;173:108365. <https://doi.org/10.1016/j.combiomed.2024.108365>
- Chen Z, Xie T, Chen S, et al. AI-based tumor-infiltrating lymphocyte scoring system for assessing HCC prognosis in patients undergoing liver resection. *JHEP Rep.* 2025;7(2):101270. <https://doi.org/10.1016/j.jhepr.2024.101270>
- Lin H, Pan X, Feng Z, et al. Automated whole-slide images assessment of immune infiltration in resected non-small-cell lung cancer: towards better risk-stratification. *J Transl Med.* 2022;20(1):261. <https://doi.org/10.1186/s12967-022-03458-9>
- Wang Y, Pan X, Lin H, et al. Multi-scale pathology image texture signature is a prognostic factor for resectable lung adenocarcinoma: a multi-center, retrospective study. *J Transl Med.* 2022;20(1):595. <https://doi.org/10.1186/s12967-022-03777-x>
- Lu C, Koyuncu C, Corredor G, et al. Feature-driven local cell graph (FLoCK): new computational pathology-based descriptors for prognosis of lung cancer and HPV status of oropharyngeal cancers. *Med Image Anal.* 2021;68:101903. <https://doi.org/10.1016/j.media.2020.101903>
- Zhao K, Li Z, Yao S, et al. Artificial intelligence quantified tumour-stroma ratio is an independent predictor for overall survival in resectable colorectal cancer. *EBioMedicine.* 2020;61:103054. <https://doi.org/10.1016/j.ebiom.2020.103054>
- Ismail M, Prasanna P, Bera K, et al. Radiomic deformation and textural heterogeneity (R-DepTH) descriptor to characterize tumor field effect: application to survival prediction in glioblastoma. *IEEE Trans Med Imag.* 2022;41(7):1764-1777. <https://doi.org/10.1109/TMI.2022.3148780>
- Zeng X, Gong J, Li W, Yang Z. Knowledge-driven multi-graph convolutional network for brain network analysis and potential biomarker discovery. *Med Image Anal.* 2025;99:103368. <https://doi.org/10.1016/j.media.2024.103368>
- Song X, Li J, Qian X. Diagnosis of glioblastoma multiforme progression via interpretable structure-constrained graph neural networks. *IEEE Trans Med Imag.* 2023;42(2):380-390. <https://doi.org/10.1109/TMI.2022.3202037>
- Tang Z, Xu Y, Jin L, et al. Deep learning of imaging phenotype and genotype for predicting overall survival time of glioblastoma patients.

- IEEE Trans Med Imag.* 2020;39(6):2100-2109. <https://doi.org/10.1109/TMI.2020.2964310>
20. Guo J, Xu P, Wu Y, et al. CroMAM: a cross-magnification attention feature fusion model for predicting genetic status and survival of gliomas using histological images. *IEEE J Biomed Health Inform.* 2024;28(12):7345-7356. <https://doi.org/10.1109/JBHI.2024.3431471>
 21. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). The clinical proteomic tumor analysis consortium glioblastoma multiforme collection (CPTAC-GBM) (version 16) [dataset]; 2018. The Cancer Imaging Archive. <https://doi.org/10.7937/K9/TCIA.2018.3RJE41Q1>
 22. Janowczyk A, Zuo R, Gilmore H, Feldman M, Madabhushi A. HistQC: an open-source quality control tool for digital pathology slides. *JCO Clin Cancer Inform.* 2019(3):1-7. <https://doi.org/10.1200/CCI.18.00157>
 23. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep.* 2017;7(1):16878. <https://doi.org/10.1038/s41598-017-17204-5>
 24. Wu Y, Xu X, Cheng Y, et al. BEE is an open-source tool that evaluates batch effects in medical images to enable multicenter studies. *Cancer Res.* 2025;85(2):218-230. <https://doi.org/10.1158/0008-5472.CAN-23-3846>
 25. Goode A, Gilbert B, Harkes J, Jukic D, Satyanarayanan M. OpenSlide: a vendor-neutral software foundation for digital pathology. *J Pathol Inf.* 2013;4(1):27. <https://doi.org/10.4103/2153-3539.119005>
 26. Macenko M, Niethammer M, Marron JS, et al. A method for normalizing histology slides for quantitative analysis. In: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro.* IEEE; 2009:1107-1110. <https://doi.org/10.1109/ISBI.2009.5193250>
 27. Van Eycke YR, Allard J, Salmon I, Debeir O, Decaestecker C. Image processing in digital pathology: an opportunity to solve inter-batch variability of immunohistochemical staining. *Sci Rep.* 2017;7(1):42964. <https://doi.org/10.1038/srep42964>
 28. Xu J, Xiang L, Wang G, et al. Sparse Non-negative Matrix Factorization (SNMF) based color unmixing for breast histopathological image analysis. *Comput Med Imag Graph.* 2015;46:20-29. <https://doi.org/10.1016/j.compmedimag.2015.04.002>
 29. Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77(21):e104-e107. <https://doi.org/10.1158/0008-5472.CAN-17-0339>
 30. Hu M-K. Visual pattern recognition by moment invariants. *IEEE Trans Inf Theory.* 1962;8(2):179-187. <https://doi.org/10.1109/TIT.1962.1057692>
 31. Mandelbrot B. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science.* 1967;156(3775):636-638. <https://doi.org/10.1126/science.156.3775.636>
 32. Ji MY, Yuan L, Jiang XD, et al. Nuclear shape, architecture and orientation features from H&E images are able to predict recurrence in node-negative gastric adenocarcinoma. *J Transl Med.* 2019;17(1):92. <https://doi.org/10.1186/s12967-019-1839-x>
 33. Lu C, Bera K, Wang X, et al. A prognostic model for overall survival of patients with early-stage non-small cell lung cancer: a multicentre, retrospective study. *Lancet Digit Health.* 2020;2(11):e594-e606. [https://doi.org/10.1016/S2589-7500\(20\)30225-9](https://doi.org/10.1016/S2589-7500(20)30225-9)
 34. Andrade C. Z scores, standard scores, and composite test scores explained. *Indian J Psychol Med.* 2021;43(6):555-557. <https://doi.org/10.1177/02537176211046525>
 35. Huse JT, Holland EC. Targeting brain cancer: advances in the molecular pathology of malignant glioma and medulloblastoma. *Nat Rev Cancer.* 2010;10(5):319-331. <https://doi.org/10.1038/nrc2818>
 36. Sturm D, Bender S, Jones DTW, et al. Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. *Nat Rev Cancer.* 2014;14(2):92-107. <https://doi.org/10.1038/nrc3655>
 37. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE; 2016:770-778. <https://doi.org/10.1109/CVPR.2016.90>
 38. Lu MY, Chen B, Williamson DFK, et al. A visual-language foundation model for computational pathology. *Nat Med.* 2024;30(3):863-874. <https://doi.org/10.1038/s41591-024-02856-4>
 39. Chen RJ, Ding T, Lu MY, et al. Towards a general-purpose foundation model for computational pathology. *Nat Med.* 2024;30(3):850-862. <https://doi.org/10.1038/s41591-024-02857-3>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lu X, Ying Y, Chen J, et al. From digitized whole-slide histology images to biomarker discovery: a protocol for handcrafted feature analysis in brain cancer pathology. *Brain-X.* 2025;3:e70030. <https://doi.org/10.1002/brx2.70030>