Journal of Advanced Research 75 (2025) 389–404

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Original Manuscript

Generalized deep learning for histopathology image classification using supervised contrastive learning

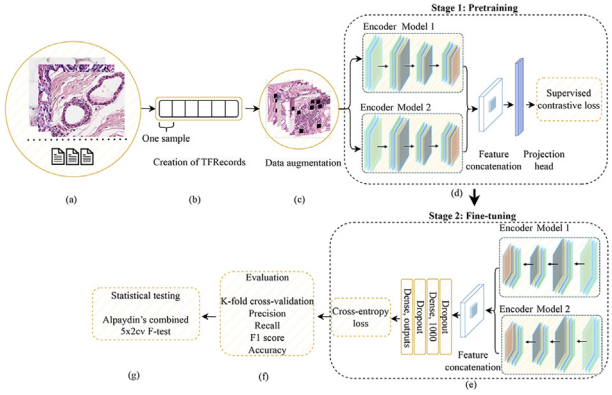
Md Mamunur Rahaman [a], Ewan K.A. Millar [b,c,d], Erik Meijering [a,*]

[a] School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia
[b] Department of Anatomical Pathology, NSW Health Pathology, St. George Hospital, Sydney NSW 2217, Australia
[c] St. George and Sutherland Clinical School, University of New South Wales, Sydney NSW 2052, Australia
[d] Faculty of Medicine & Health Sciences, Western Sydney University, Sydney NSW 2560, Australia

HIGHLIGHTS

- Introduction of HistopathAI, a novel hybrid network structure that merges SCL strategies, specifically tailored for imbalanced data, with a CE loss function to bolster histopathology image classification.
- Introduction of the HDFF technique that amalgamates feature vectors from both EfficientNetB3 and ResNet50, yielding a more comprehensive representation of histopathology images.
- Implementation of a stepwise methodology that transitions from feature learning to classifier learning, resonating with the foundational concept of contrastive learning to refine the normalized embedding space.
- Comprehensive training and testing of our model for diverse histopathology domains, utilizing seven publicly accessible datasets along with one private dataset, resulting in state-of-the-art performance and demonstrating the generalizability of our framework.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 19 June 2024
Revised 7 October 2024
Accepted 10 November 2024
Available online 16 November 2024

ABSTRACT

Introduction: Cancer is a leading cause of death worldwide, necessitating effective diagnostic tools for early detection and treatment. Histopathological image analysis is crucial for cancer diagnosis but is often hindered by human error and variability. This study introduces HistopathAI, a hybrid network designed for histopathology image classification, aimed at enhancing diagnostic precision and efficiency in clinical pathology.

☆ Peer review under responsibility of Cairo University.

* Corresponding author.
E-mail addresses: md_mamunur.rahaman@unsw.edu.au (M.M. Rahaman), ewan.millar@health.nsw.gov.au (E.K.A. Millar), erik.meijering@unsw.edu.au (E. Meijering).

https://doi.org/10.1016/j.jare.2024.11.013

Keywords:
Histopathological image analysis
Contrastive learning
Imbalanced data
Hybrid deep feature fusion
Feature representation
Cancer diagnosis

Objectives: The primary goal of this study is to demonstrate that HistopathAI, leveraging supervised contrastive learning (SCL) and hybrid deep feature fusion (HDFF), can significantly improve the accuracy of histopathological image classification, including scenarios involving imbalanced datasets.
Methods: HistopathAI integrates features from EfficientNetB3 and ResNet50, using HDFF to provide a rich representation of histopathology images. The framework employs a sequential methodology, transitioning from feature learning to classifier learning, mirroring the essence of contrastive learning with the aim of producing superior feature representations. The model combines SCL for feature representation with cross-entropy (CE) loss for classification. We evaluated HistopathAI across seven publicly available datasets and one private dataset, covering various histopathology domains.
Results: HistopathAI achieved state-of-the-art classification accuracy across all datasets, demonstrating superior performance in both binary and multiclass classification tasks. Statistical testing confirmed that HistopathAI's performance is significantly better than baseline models, ensuring robust and reliable improvements.
Conclusion: HistopathAI offers a robust tool for histopathology image classification, enhancing diagnostic accuracy and supporting the transition to digital pathology. This framework has the potential to improve cancer diagnosis and patient outcomes, paving the way for broader clinical application. The code is available on https://github.com/Mamunur-20/HistopathAI.

Introduction

Cancer is one of the leading causes of death globally, with approximately 10 million lives lost in 2020. According to the International Agency for Research on Cancer (IARC), one in five individuals will be diagnosed with cancer during their lifetime [1]. While the most diagnosed cancers in 2020 were breast, lung, and colorectal cancers, the deadliest were lung, colorectal, and liver cancers. Pediatric cancer is also a notable concern, with 400,000 annual diagnoses, and the types of cancers can vary significantly by region [2]. The gravity of these statistics highlights the urgent need for enhanced diagnostic and therapeutic strategies. Early detection is vital, requiring early diagnosis and screening. Early diagnosis focuses on prompt identification for better treatment outcomes, while screening proactively detects cancers or precancerous conditions in those without symptoms. Following any detected abnormalities, thorough investigations are essential to ascertain a diagnosis and facilitate timely treatment.

Histopathological image analysis is the gold standard for cancer diagnosis, traditionally relying on microscopic slide studies [3]. However, despite its long-standing use, evaluations by pathologists can be subjective and prone to errors, with inconsistencies observed both between and within observers. For example, breast cancer subtyping and prostate cancer grading by expert pathologists have shown accuracy rates as low as 57.08% [4] and agreement rates of 57.90% [5], respectively. To enhance accuracy and consistency, there is a growing trend towards digitizing this process, using artificial intelligence (AI) and computer-aided diagnosis (CAD) systems. Notably, convolutional neural networks (CNNs) are making waves in this domain, autonomously extracting vital information from raw images without the need for preprocessing, offering promise for improved histopathological image diagnosis.

Incorporating technological advancements, various image classification methodologies have emerged. The detection and diagnosis of cancers, particularly in their early stages, have been significantly improved with advances in image classification methodologies. Machine learning (ML) techniques, specifically deep learning (DL) using CNNs, have made a transformative impact on histopathology image classification for various types of cancers and pre-invasive lesions [6]. Among them, breast, colorectal, and gastric cancers have seen a vast improvement in diagnostic accuracy due to innovative approaches like transfer learning [7,8], ensemble learning [9], and self-supervised learning [10]. Furthermore, novel algorithms and architecture models have been proposed for optimizing image classification tasks [11–17]. In the context of colorectal cancer, traditional ML models with manually extracted features have demonstrated promising accuracy [18–22], while DL models have further expanded the landscape of diagnostic tools [23–32]. As for gastric cancer, the introduction of DL models [33–38] in histopathology image classification has led to significant improvements over traditional imaging techniques [39–41]. Despite these advancements, challenges remain, particularly in managing imbalanced data, emphasizing varying levels of semantics, and improving the transferability of features across different domains.

While existing methods show promising accuracy, they often face challenges from employing cross-entropy (CE) loss functions and focusing on binary classification, potentially restricting their real-world applicability considering the myriad of cancer variants. Notably, many methods lack generalizability. Traditional methods utilizing CE loss can produce biased classifiers. In contrast, supervised contrastive learning (SCL) loss encourages tighter intraclass features and clear interclass separation (Fig. 1), highlighting its potential for superior accuracy in histopathology image classification. However, most approaches prioritize high-level semantics, neglecting mid- and low-level details. This, combined with the nature of typical loss functions, could unintentionally hinder the transferability of features across domains.

In response to these challenges, we developed HistopathAI, a novel framework that explores SCL strategies tailored to learning from imbalanced data, with the specific aim of enhancing histopathology image classification. HistopathAI employs a hybrid network structure that combines the benefits of a SCL function, which helps in learning image representations, with a CE loss function, which aids in learning classifiers. To make our framework even more robust, we incorporated hybrid deep feature fusion (HDFF). This technique combines feature vectors from EfficientNetB3 and ResNet50, making our representation of histopathology images richer and more comprehensive. This unique structure is grounded in the principle that superior classifiers emerge from superior features. Our methodology progresses gradually from feature learning to classifier learning, reflecting the core idea of SCL: bringing samples of the same class closer together in the embedding space while distancing samples of different classes. HistopathAI is designed to yield less skewed features and, consequently, less biased classifiers.

Studies show that contrastive learning methods effectively capture and adapt low-to-mid level information across various domains [42]. Building on the proposition that minimizing intraclass variation might limit supervised CE models [43], we advocate

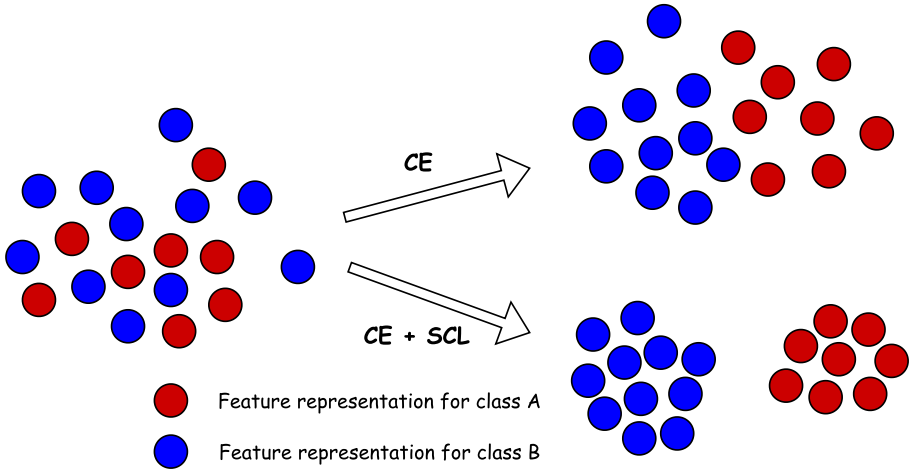


Fig. 1. Comparison of feature learning: CE loss (top right) results in skewed features and biases, while SCL loss (bottom right) emphasizes intraclass cohesion and interclass separation.

an optimal level of intraclass variation for the effective transfer of learned representations to varying domains. Unlike standard CE models, contrastive methods achieve more substantial interclass separation, which likely contributes to their enhanced transferability. Moreover, our comprehensive evaluations, utilizing seven publicly accessible datasets and one private dataset from diverse histopathology domains, show that models using contrastive losses display greater robustness against various image corruptions. They also offer better-calibrated class probabilities, which more precisely mirror the true likelihood of correctness.

The remainder of this paper is organized as follows. After a brief review of relevant literature in the Background section, we present the proposed methodology in the Methods section. The used datasets and experimental approach to evaluate the proposed methodology are described in the Experiments section, while the outcomes are presented in the Results section, which we further reflect on in the Discussion section. Finally, the Conclusion section summarizes the main findings of our study.

Background

Following on from the introduction in the previous section, we briefly review recent advances in histopathology image classification for breast, colon, and gastric cancer diagnosis, reflecting the focus on these cancer types in the public and private datasets used in our experiments (see Table 1 for a summary of these advancements along with their strengths and limitations).

Recent advances in breast histopathology image classification (BHIC)

CNNs have greatly advanced BHIC since 2014, enhancing feature learning and accuracy [6]. Transfer learning, using networks like VGG16 and ResNet50, has been fundamental [7,8]. Contributions also include dark knowledge transfer and multiclass classification with Xception [44,45].

Hybrid models, combining CNNs with graph convolutional networks (GCNs) and generative adversarial networks (GANs), are emerging [11,12]. BHCNet, for instance, achieved a 98.87% accuracy on the BreaKHis dataset [13]. Additionally, self-supervised learning, like the DARC framework, and novel feature extraction methods are advancing histopathological image analysis [46,47].

The field also witnessed remarkable results with deep transfer learning, especially with algorithms like the sparrow search algorithm [14]. Ensemble learning is another area of rapid growth, with architectures such as CSAResNet, DAM-CNN, and DEEP_Pachi enhancing feature learning capabilities [48–61]. Additionally, capsule networks have seen improvements with the incorporation of Res2Net blocks and convolutional layers. Solutions addressing challenges like overfitting, such as DRDA-Net, are also emerging [15,16]. The Dependency-Based Lightweight CNN (DBLCNN) is another remarkable contribution, recognized for its efficiency and performance in multi-classification tasks [17,62,7,63–67].

While BHIC has seen promising advances, challenges remain. Models struggle with generalizability due to variations in image preparation and acquisition. Many studies use private datasets, making performance comparisons difficult. The emphasis is mainly on binary classification, overlooking the importance of multiclass tasks. Additionally, while prioritizing high-level semantics, the value of low-level and mid-level details is often underestimated.

Recent advances in colorectal histopathology image classification (CHIC)

Earlier CHIC studies leveraged ML with handcrafted features, achieving accuracies from 91.3% with texture features [20] to 99.5% using local binary patterns (LBP) and support vector machines (SVM) [18]. An evaluation of the performance of a wide range of texture descriptors and classifiers led to high accuracies of up to 98.6% [22].

The advent of DL brought forward innovative methodologies with notable results. Restricted Boltzmann Machines (RBMs) achieved 96.11% accuracy in a semi-supervised setting [23]. Adaptive CNN models, combined with architectures like DenseNet121, demonstrated high accuracies, with some reaching up to 97.2% on specific datasets [31]. ResNet50 also showcased efficacy, particularly on private datasets [32].

Recent studies proposed methods such as TransPath [68], ConvNeXt-L [69], DenseNet121 [70], and MoCo v2 [71]. While these methodologies advanced the CHIC domain, they show limitations. For instance, some rely heavily on specific architectures, potentially compromising adaptability across varied datasets. Others may not adequately address data imbalances, potentially affecting diagnosis reliability.

Despite the advances, challenges remain. Emphasis on binary classification neglects crucial multiclass categorization due to the multifaceted nature of cancer. Questions on model versatility

Table 1
Summary of related work in histopathology image classification.

Methods	Area	Strengths	Limitations
CNNs (since 2014) [6–8]	Breast	Enhanced feature learning and accuracy. Foundational for further developments. Transfer learning has been effective.	Limited generalizability across diverse datasets. Struggles with imbalanced data. High computational cost. Overfitting on small datasets.
Hybrid Models (e.g., GCNs, GANs) [11,12]	Breast	High accuracy (e.g., 98.87% with BHCNet). Advanced feature extraction methods.	Complexity in model design and training. Limited scalability. Generalization to new data is challenging.
Ensemble Learning (e.g., CSAResNet, DAM-CNN) [48–61]	Breast	Improved robustness and performance. Combines strengths of multiple models. Enhances feature learning capabilities.	High computational complexity. Susceptibility to overfitting. Limited generalizability in diverse environments.
Capsule Networks [15,16]	Breast	Improved performance on multiclass tasks. Better handling of spatial hierarchies.	Computationally intensive. Limited success on imbalanced datasets. Complex to interpret.
ML with Handcrafted Features [20,18,22]	Colorectal	High accuracy with specific feature sets. Foundational techniques for further research.	Poor performance on diverse, large-scale datasets. Limited adaptability and scalability.
Deep Learning (e.g., RBMs, Adaptive CNNs) [23,31,32]	Colorectal	High accuracy, effective for various datasets. Robust performance in semi-supervised settings.	Struggles with data imbalance. Limited generalization to new data. High computational demands.
Recent Methods (e.g., TransPath, ConvNeXt-L, DenseNet121, MoCo v2) [68–71]	Colorectal	Advanced architectures providing state-of-the-art performance.	Generalizability issues across varied datasets. Inadequate handling of imbalanced datasets.
DL Models (e.g., CNNs on WSIs) [33,72,34,36–38]	Gastric	High performance on large-scale data. Robust feature extraction methods.	High dependency on dataset characteristics. Limited success on generalizing to new domains.
VGG16 for GHIC [73,74]	Gastric	Proven architecture, good baseline for comparisons.	May not capture all relevant feature nuances. Struggles with generalization across diverse datasets.
Ensemble Methods (e.g., EfficientNet + DenseNet) [78–80]	Gastric	Combines strengths of multiple architectures. High accuracy.	Resource-intensive. Limited generalization across different environments. Complex to integrate and maintain.

across diverse datasets and imbalanced data issues underscore the need for further research to develop more holistic classification models.

Recent advances in gastric histopathology image classification (GHIC)

Gastric cancer detection has evolved with imaging techniques like gastroscopes [39], X-rays [40], and CT scans [41], but histopathological analysis of H&E tissue slides remains the gold standard, spurring GHIC research.

DL models have been paramount. CNNs on whole slide images (WSIs) [33], random forests for graph feature extraction [72], residual networks [34], and segmentations using DeepLabV3 and ResNet50 on H&E slides [36] have shown promising results, as have multiscale patterns [37] and optimized InceptionV3 networks [38].

Recent methodologies encompass diverse strategies. The potential of VGG16 was examined in GHIC [73,74], while techniques like CCF-GNN and MCLNet offer new perspectives [75,76]. Approaches like MLP IN MLP (MIM) [77], ensemble methods integrating EfficientNet and DenseNet [78], and models like GasHis-Transformer and CAM-VT have provided new insights [79,80].

Notwithstanding great progress, challenges persist, from data dependency to skewed feature extraction, diverse cancer categorization, and cross-domain feature transferability, emphasizing the need for continual research.

Methods

The proposed HistopathAI framework (Fig. 2) enhances histopathology image classification by integrating advanced DL models with a specialized two-phase training strategy for improved diagnostic accuracy. The framework involves several key components: data preprocessing and storage, probabilistic data augmentation, feature extraction using two models, and a dual-stage training process utilizing SCL and CE loss functions. Here we provide a formal description of the computational steps of HistopathAI, the used model architecture, contrastive learning

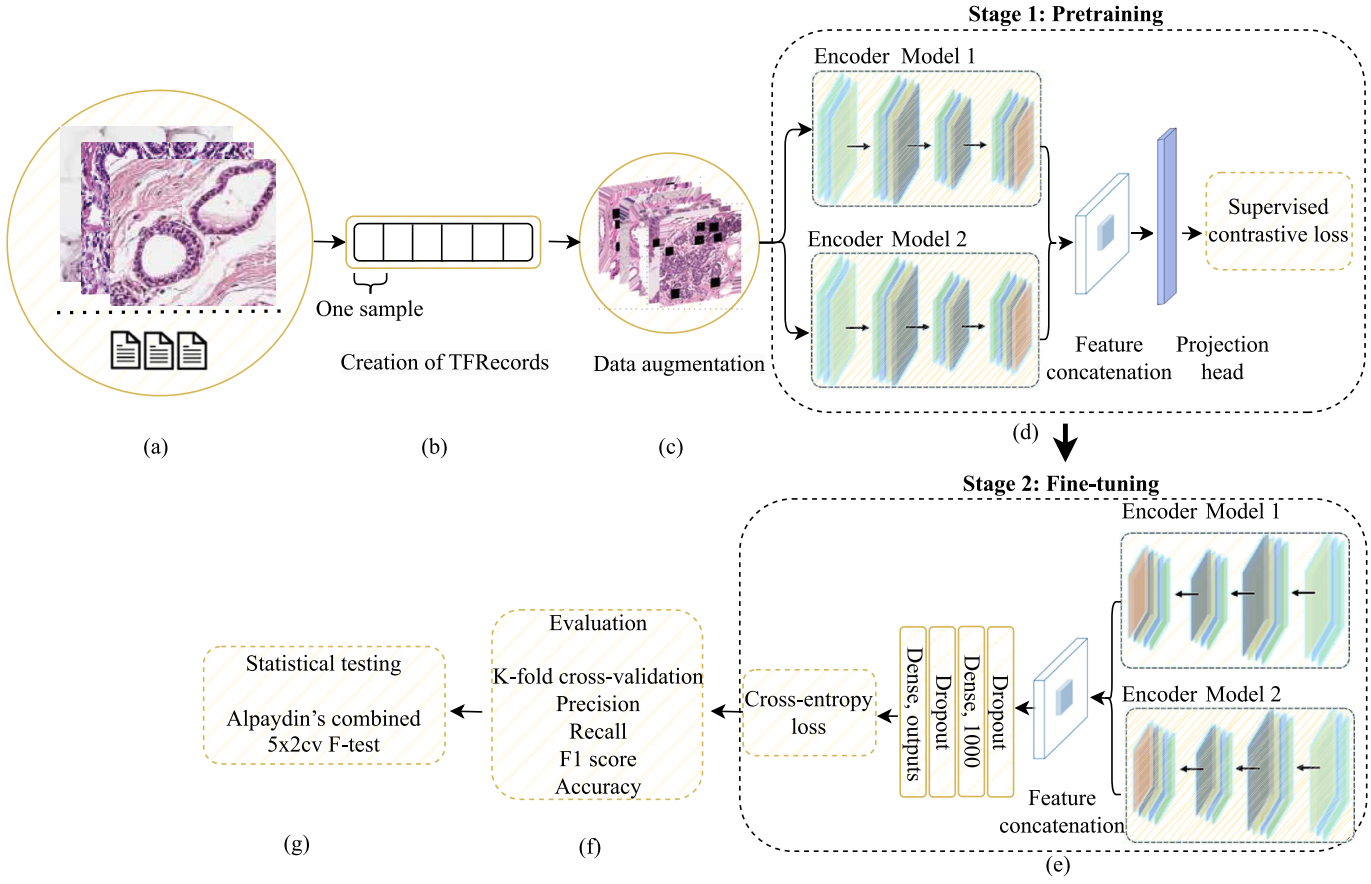


Fig. 2. The HistopathAI framework and evaluation workflow. (a) Image and label acquisition. (b) TFRecord file creation. (c) Probabilistic data augmentation. (d) Stage 1: encoder and projection head pretraining with supervised contrastive loss. (e) Stage 2: classifier training with a frozen encoder. (f) Classification performance evaluation. (g) Robustness and performance statistical tests against other algorithms.

approach, loss function, training strategy, and implementation details.

Formal description

HistopathAI performs the following computational steps (Fig. 3). To efficiently store and access data during training, the input images $X_i \in \mathbb{R}^{W\times H\times C}$, each having C channels with a width of W pixels and a height of H pixels, together with their corresponding labels L_i, are first converted to TensorFlow Record (TFRecord) files:

$$\text{TFRecord}(X_i, L_i) \rightarrow \text{TFRecordfile}.$$

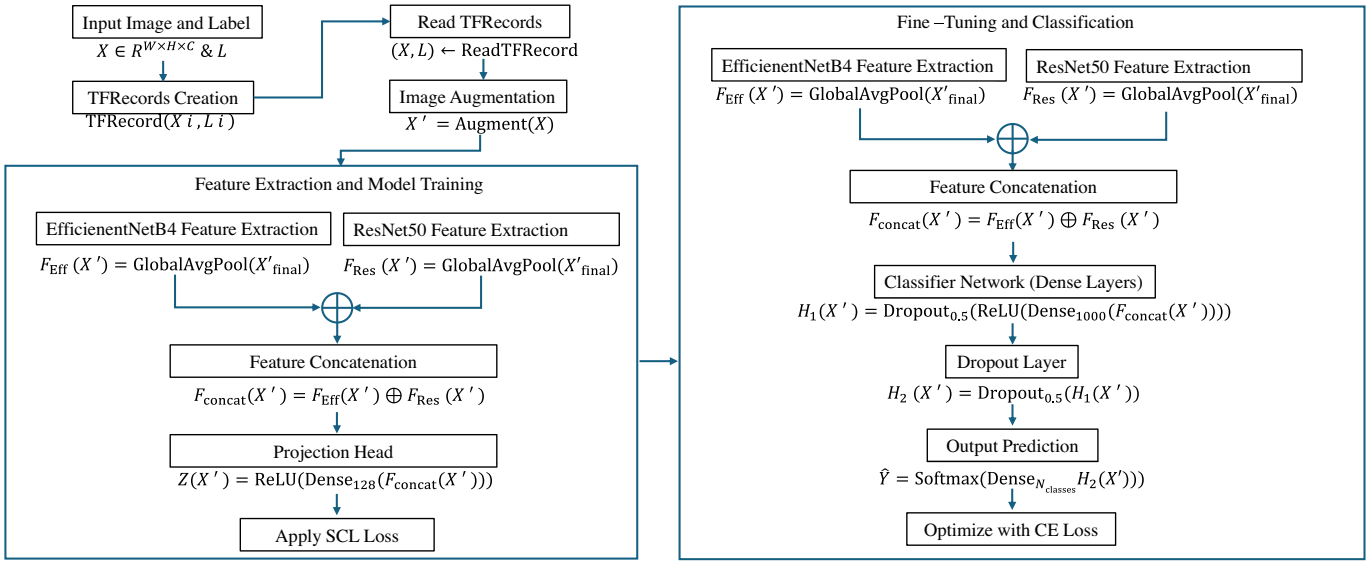


Fig. 3. Formal description of the computational steps of the HistopathAI framework.

When training begins, the images and labels are read from their TFRecord files:

$$(X, L) \leftarrow \text{ReadTFRecord}(\text{TFRecord files}).$$

Next, various augmentations are applied to the images to enhance the model's robustness:

$$X\prime = \text{Augment}(X).$$

The specific augmentations used, including the types of transformations and their probabilities, are detailed in the Probabilistic Augmentation subsection of the Experiments section.

The augmented images $X\prime$ are then passed through two pretrained DL models: EfficientNetB3 and ResNet50. For each layer l of EfficientNetB3, the input $X\prime_l$ is processed successively by convolution (Conv), batch normalization (BatchNorm), and rectified linear unit (ReLU) activation to produce the output $X\prime_{l+1}$, which is the input to the next layer $l+1$:

$$X\prime_{l+1} = \text{ReLU}(\text{BatchNorm}(\text{Conv}(X\prime_l))).$$

The final feature vector after passing through all layers of EfficientNetB3 is obtained by a global average pooling (GlobalAvgPool) operation:

$$F_{\text{Eff}}(X\prime) = \text{GlobalAvgPool}(X\prime_{\text{final}}).$$

The second model, ResNet50, operates using residual blocks, defined as:

$$X\prime_{l+1} = X\prime_l + \text{ReLU}(\text{BatchNorm}(\text{Conv}(\text{ReLU}(\text{BatchNorm}(\text{Conv}(X\prime_l)))))).$$

The final feature vector after passing through all the residual blocks and a global average pooling layer in ResNet50 is:

$$F_{\text{Res}}(X\prime) = \text{GlobalAvgPool}(X\prime_{\text{final}}).$$

The feature vectors from both models are then concatenated (concat):

$$F_{\text{concat}}(X\prime) = F_{\text{Eff}}(X\prime) \oplus F_{\text{Res}}(X\prime).$$

Next, the concatenated feature vector is passed through a projection head, which consists of a multilayer perceptron (MLP) with one hidden layer. The MLP maps the representations from the base encoder to a 128-dimensional latent space, where the contrastive loss is applied. Here, Dense_{128} represents a dense (fully connected) layer with 128 output units. The hidden layer utilizes a ReLU activation function to introduce nonlinearity, enabling the model to learn more complex and discriminative feature representations:

$$Z(X\prime) = \text{ReLU}(\text{Dense}_{128}(F_{\text{concat}}(X\prime))).$$

The SCL loss is then applied to learn robust feature representations. In the fine-tuning stage, the encoder weights are frozen, and a classifier is trained on the learned features.

The concatenated feature vector is passed through the classifier network, which consists of several layers designed to transform the feature vector into class predictions. First, the vector is passed through a dense (fully connected) layer with 1000 output units, denoted as Dense_{1000}, followed by ReLU activation to introduce nonlinearity. To prevent overfitting, a dropout layer with a dropout rate of 0.5, denoted as $\text{Dropout}_{0.5}$, is also applied:

$$H_1(X\prime) = \text{Dropout}_{0.5}(\text{ReLU}(\text{Dense}_{1000}(F_{\text{concat}}(X\prime)))).$$

The output is then passed through another dropout layer with a dropout rate of 0.5 to further refine the feature vector:

$$H_2(X\prime) = \text{Dropout}_{0.5}(H_1(X\prime)).$$

Finally, the transformed vector is passed through another dense layer, $\text{Dense}_{N_{\text{classes}}}$, where N_{classes} is the number of output units corresponding to the number of classes in the classification task. A Softmax activation function is then applied to produce the final class probabilities:

$$\widehat{Y} = \text{Softmax}(\text{Dense}_{N_{\text{classes}}}(H_2(X\prime))).$$

The model's predictions are optimized using the CE loss, described below.

Model architecture

We adopt an integrated approach, leveraging the strengths of EfficientNetB3 and ResNet50 models to enhance histopathology image classification. EfficientNet's advanced scaling capabilities excel in processing intricate histology images, while ResNet's deep residual learning framework captures complex patterns within the data. Their combined feature extraction spans a broad spectrum of histopathological characteristics, enhancing the model's resilience and interpretability.

EfficientNetB3 model

Traditionally, enhancing CNN models involves arbitrary adjustments to depth, width, and resolution, often requiring extensive manual tuning and possibly leading to efficiency losses. EfficientNet [81] addresses this by scaling the network architecture more coherently, balancing depth (d), width (w), and resolution (r) for optimal performance. This is achieved through a scaling technique using a compound coefficient (φ) characterized by

$$\begin{aligned} &d = \alpha^{\varphi}, \quad w = \beta^{\varphi}, \quad r = \gamma^{\varphi} \\ &\text{s.t.} \quad \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2, \quad \alpha \geqslant 1, \quad \beta \geqslant 1, \quad \gamma \geqslant 1. \end{aligned} \tag{1}$$

This systematic scaling in EfficientNet enhances performance in histopathology image classification, ensuring the intricate patterns in the data are effectively processed without added computational burden. Our application of EfficientNetB3 aligns with the aim to robustly represent histological structures while maintaining computational efficiency.

ResNet50 model

Deep neural networks like ResNet50 are beneficial for complex tasks such as histopathology image classification. ResNet models address performance saturation and degradation issues occurring in CNNs beyond a certain depth, by the use of skip connections. This innovation allows ResNet models, including ResNet50, to efficiently manage deeper architectures without performance loss [82].

Mathematically, given a feature map X and a residual function $F(x)$, the skip connection enforces:

$$H(x) = F(x) + X, \quad \text{or} \quad F(x) = H(x) - X. \tag{2}$$

This capability is crucial in our work for capturing intricate patterns in histopathological images, leading to a robust representation of biological structures.

HDFF network

Effective feature representation is crucial for histopathology image analysis. While traditional feature fusion methods have their merits, they also present challenges in utilizing original input features and integrating multiple features effectively [83].

In our work, we employ HDFF, which merges feature vectors from the global average pooling layers of EfficientNetB3 and ResNet50. This results in a 3,584-dimensional composite feature vector combining the strengths of both architectures and offering a rich image representation.

The HDFF network comprises a concatenation layer, dropout layers for regularization, a dense layer with 1,000 units and ReLU

activation for nonlinearity, and a final dense layer for classification. The network structure and operations are mathematically represented as:

$$F^{(i)} = \bigcup_{n=1}^{2} F_n^{(i)}, \tag{3}$$

where $F_n^{(i)}$ denotes the nth feature vector for the ith sample and $F^{(i)}$ the concatenated feature vector. This approach not only enhances the descriptive power of the model but also aligns with our goal of efficiently leveraging information from diverse sources to improve accuracy and robustness.

Contrastive learning

Contrastive learning for histopathology image classification differentiates between "anchor" (pathological tissue of interest), "positive" (similar pathological tissue), and "negative" (different pathological tissue) images in the embedding space. However, in self-supervised learning, the selection of negatives can introduce false negatives, especially detrimental in the critical field of histopathology.

Fully-supervised contrastive learning methods offer a solution [84]. They utilize labeled data to generate positives from the same class and reduce the risk of false negatives. Uniquely, they allow multiple positives per anchor and integrate label information into representation learning, outperforming standard CE loss in large-scale histopathology image classification.

In our framework, we combine EfficientNetB3 and ResNet50 models to enhance the robustness and diversity of histopathological image representations. We concatenate the learned embeddings from the last layers of these models, creating an extensive feature vector for a comprehensive representation. This vector is processed through a projection head to map the features into a lower-dimensional space, conducive for contrastive learning, which is followed by a second training stage focused on classification.

Our dual-model approach, feature concatenation, and two-stage training strategy collectively enhance representation quality and improve feature transferability to the final classification task, facilitating the application of SCL to histopathology image classification.

Loss function

Our framework uses a two-stage training process, leveraging both SCL loss and categorical CE loss.

Supervised contrastive learning loss

The SCL loss aims to enhance the similarity of same-class representations while reducing it for different classes using temperature-scaled cosine similarity [84]:

$$L_{\text{SCL}} = \sum_{i \in I} \frac{-1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(z_i \cdot z_p / \tau)}{\sum_{a \in A(i)} \exp(z_i \cdot z_a / \tau)}, \tag{4}$$

where z_i, z_p, and z_a are the embedding vectors of anchor, positive, and aggregate samples, respectively, $P(i)$ denotes the set of positive sample indices, $A(i)$ includes both positive and negative samples, and τ (set to 0.1) is the temperature parameter controlling embedding similarity sharpness, crucial for differentiating subtle histopathology image details.

The flexibility of SCL in handling multiple same-class samples per batch proves beneficial for histopathology classification, encouraging compact embeddings for each class. Its similarities with N-pairs loss enhance robustness, making our model well-suited for the complex nature of histopathology images. In essence, SCL ensures tailored and effective representation learning for our classification task.

Categorical cross-entropy loss

After pretraining the model using SCL loss, we train a classifier using the standard categorical CE loss. This loss function is designed to handle multiclass classification problems and is expressed as:

$$L_{\text{CE}} = -\sum_{i=1}^{N} \sum_{c=1}^{C} y_{ic} \log(p_{ic}), \tag{5}$$

where N is the number of instances, C the number of classes, y_{ic} the binary indicator for class c of instance i, and p_{ic} the predicted probability of instance i being in class c. This loss ensures robust handling of complex histopathology images, aiding in the generation of powerful classification features.

Training strategy

Our training process consists of two main stages, aimed at building a robust classifier that combines EfficientNetB3 and ResNet50 to analyze histopathology images.

Stage 1: pretraining the encoders and projection head using SCL loss

In the first stage, we aim to derive meaningful features from the images using two encoders: EfficientNetB3 and ResNet50. We concatenate their extracted features and pass them through a projection head—a dense layer with ReLU activation, tailored for the SCL objective. This not only enhances pretraining effectiveness but also ensures a discriminative data representation. The whole system is optimized using SCL loss to minimize intraclass and maximize interclass distances in the embedding space. Complemented by data augmentation and a cosine learning rate schedule with a warm-up, we train the model for 110 epochs, preparing robust feature representations for the next classification stage.

Stage 2: training the classifier with frozen encoder using CE loss

Using the feature representations from Stage 1, we train a classifier using the concatenated features from both encoders. The encoders' weights are frozen and we build a classifier consisting of several dense layers with dropout for regularization. The model is trained to minimize the sparse categorical CE loss, aiming to accurately classify the images into predefined classes. The training process is conducted using stratified 5-fold cross-validation, allowing us to assess the model's performance across different data splits.

Implementation details

Our HistopathAI framework was developed using Python 3.9.16 and TensorFlow 2.10.0. For training, we used a batch size of 16 to balance processing speed and accuracy. The model was trained for 100 epochs to learn the data patterns thoroughly. We chose the Adam optimizer because of its flexible learning rate, which helps in faster learning [85]. We adjusted the learning rate during training using a combination of linear warm-up and cosine decay. It started at a very low value of 1×10^{-8}. During the warm-up phase, the learning rate increased linearly to its maximum value of 3×10^{-5}. After the warm-up, the learning rate followed a cosine decay, but never went below 1×10^{-8}. This approach allowed the model to quickly adapt in the initial stages and refine its learning as training progressed. All our training was done on the Gadi cluster in Australia with a dgxa100 GPU, bolstered by 16 CPU cores and 190 GB RAM allowance. To obtain more reliable performance esti-

mates, we evaluated our model for different data splits, using 5-fold cross-validation.

Experiments

Image datasets

Our HistopathAI framework was rigorously evaluated using a comprehensive collection of seven publicly available datasets and one in–house dataset (Table 2, Fig. 4). These datasets encompass a variety of cancer types, including breast, colorectal, and gastric cancer, as well as a variety of binary and multiclass classification tasks. Together, they aid in the development of a robust classification model.

For the BRACS dataset, in addition to the 7-class task, we also considered a 4-class problem, grouping the original classes into normal (NT), noncancerous (PB + UDH), precancerous (ADH + FEA), and cancerous (DCIS + IC), as well as a binary problem, grouping the original classes into two (NT + PB + UDH versus ADH + FEA + DCIS + IC).

Similarly, for the EBHI dataset, in addition to the 5-class task, we also studied the binary problem, grouping the original classes into benign (normal, polyp, low-grade) versus malignant (high-grade and adenocarcinoma).

Our exclusive in–house dataset contains 13 malignant and 4 benign histopathology slides.

Image preprocessing

Dataset organization

We gathered image data, including filenames and labels, and organized it for analysis. To ensure data integrity, we identified and removed any duplicate images based on their content, striving for a unique dataset.

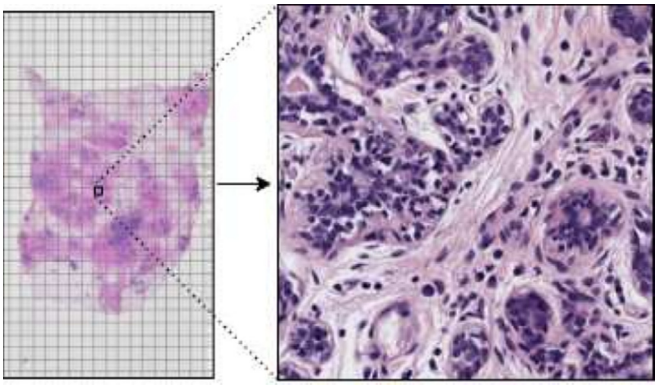


Fig. 4. Sample image from our in–house dataset. Left: Original WSI of 81,920×102,400 pixels. Right: Sample patch of size 1,024×1,024 pixels extracted from the WSI.

The images were preprocessed to enhance consistency and optimize performance for subsequent analysis. This involved decoding, normalization of pixel values to the $[0, 1]$ range, and resizing the images to a standard resolution of 512×512 pixels. The processed data was then efficiently stored in a binary file format, suitable for handling large-scale datasets and accelerating model training times.

To facilitate cross-validation and ensure balanced class representation, the dataset was partitioned using proper stratification, maintaining an equal distribution of classes across different subsets. This approach is advantageous for extensive image analyses across various domains.

Table 2
Details of the datasets used in the study.

Dataset	Reference	Type	#Classes	Classes	#Images	Magnification	Size
BACH	[88]	Breast	4	Normal Benign In Situ Carcinoma Invasive Carcinoma	100 100 100 100	200×	2,040×1,536
BRACS	[46]	Breast	7	Normal Tissue (NT) Pathological Benign (PB) Usual Ductal Hyperplasia (UDH) Flat Epithelial Atypia (FEA) Atypical Ductal Hyperplasia (ADH) Ductal Carcinoma In Situ (DCIS) Invasive Carcinoma (IC)	484 836 517 756 507 790 649	40×	127–17,611×137–9,842
BreaKHis	[89]	Breast	2	Benign Malignant	598 1,370	40×	700×460
EBHI	[90]	Colorectal	5	Normal Polyp Low-grade Intraepithelial Neoplasia High-grade Intraepithelial Neoplasia Adenocarcinoma	61 254 603 130 790	200×	2,048×1,536
GasHisSDB	[73]	Gastric	2	Normal Abnormal	20,160 13,124	20×	160×160
HE-GHI-DS	[79]	Gastric	2	Normal Abnormal	140 560	20×	2,048×2,048
MHIST	[91]	Colorectal	2	Hyperplastic Sessile Adenoma	2,162 990	40×	224×224
In–House		Breast	2	Benign Malignant	3,138 9,018	40×	1.024×1,024

Probabilistic augmentation

To bolster the robustness and performance of our model, we implemented a probabilistic data augmentation strategy, stochastically applying transformations to the input images to diversify our training data and enhance generalization. The following transformations were included (percentages in parentheses indicate the probabilities of applying each augmentation to each image):

- **Shear transformation (20%):** Shears images by 20°.
- **Rotation operations (20%):** Rotates images by ± 45°, with 75% probability for 90°, 180°, and 270°.
- **Flips and transposition (75%):** Randomly flips images and transposes them.
- **Pixel-level transformations (40%):** Adjusts saturation, contrast, and brightness.
- **Central cropping (60%):** Applies random central cropping, resizing images to 50–80% of their original size. A random size crop operation is used for random numbers between 0.3 and 0.6.
- **Cutout (50%):** Randomly removes square regions, accounting for 10–12.5% of the image size.

This strategy ensures diverse training data and that the model never sees the exact same image twice, significantly improving its ability to generalize to unseen data and handle real-world variability in histopathology images.

Evaluation metrics

To validate our histopathology image classification method, we employed a suite of standard evaluation metrics, presented as macro averages from the 5-fold cross-validation in the format: mean ± standard deviation.

- **Precision** (P): The correctly identified positive samples out of all predicted positives, $P = TP/(TP + FP)$.
- **Recall** (R): The correctly identified positive samples out of all actual positives, $R = TP/(TP + FN)$.
- **F1 Score** (F1): The harmonic mean of precision and recall, $F1 = 2 \times P \times R/(P + R)$.
- **Accuracy** (A): The proportion of correctly predicted samples, $A = (TP + TN)/(TP + TN + FP + FN)$.

Here, true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are counts of correctly and incorrectly classified positive and negative samples, respectively.

Statistical testing

To validate the significance of our model's performance against other algorithms, we employed Alpaydm's Combined 5×2cv F-test [86], particularly recommended for small datasets [87]. This ensures that the observed differences in performance are statistically robust and not attributable to random variations. The test is a robust procedure for comparing model performances, applicable to both classifiers and regressors. When comparing two classifiers, labelled 1 and 2, the F-statistic is calculated as:

$$f = \frac{\sum_{i=1}^{5}\sum_{j=1}^{2}(A_{ij})^2}{2\sum_{i=1}^{5}s_i^2} \tag{6}$$

Approximating an F-distribution with 10 and 5 degrees of freedom, this test enhances the credibility of our framework's effectiveness compared to other models. The results not only confirm the differences in performance highlighted by other evaluation metrics but also provide a deeper, statistically grounded understanding of these discrepancies.

Results

Quantitative model performance

The quantitative results of an ablation experiment with different configurations of HistoPathAI show that the proposed HDFF model, combining EfficientNetB3 and ResNet50 and using SCL, performs superiorly (Table 3). The results also show the robustness of the model, with high performance across all datasets and different classification tasks, and comparatively small standard deviations across the different folds of the experiment. As expected, the performance in most multiclass classification tasks is lower than in the binary classification tasks, with metric scores decreasing as the number of classes increase, but still the proposed model shows favorable performance. HistoPathAI performed best on our in–house dataset, with scores close to 100%, suggesting that our dataset was less challenging than most of the public datasets.

Qualitative analysis using t-SNE

Dimensionality reduction techniques such as t-SNE offer insights into a model's ability to generalize. Example visualizations of the embeddings produced by HistopathAI on the BRACS 4-class classification (Fig. 5) distinctly show how the use of SCL loss creates better separated clusters compared to the regular CE loss. Similar patterns were observed in the other datasets. The well-clustered SCL embeddings, particularly in a challenging dataset such as BRACS, validate our model's adeptness in learning intricate patterns. The t-SNE plots reinforce our model's potential for delivering accurate results in practical histopathological settings.

Statistical test results

From the results of the 5x2cv F-test we observe that our HistopathAI framework exhibits dominance over other models (Table 4). For each comparison, an F-statistic is generated from the test, which is compared against a critical value. Here, the critical value is 3.2974, derived from the F-distribution table at a significance level of α=0.1. An F-statistic surpassing the critical value indicates a statistically significant difference. In most cases, HistopathAI showed significant improvement over alternative models, as indicated. The few cases where the improvements were not statistically significant suggest there is still some room for further optimization.

Comparison with previous studies

Finally, we compared the performance of HistopathAI to other methods reported in literature for the considered datasets. Rather than reimplementing or retraining these methods ourselves, we adopted the accuracy from the cited papers, ensuring a direct and fair comparison.

From the results (Table 5) it is evident that many methods face challenges, such as limited model generalizability across datasets, imbalanced data handling, and reliance on private datasets, which hampers reproducibility. There is also a prevalent issue where

Table 3
Performance comparison of different models on various histopathology datasets for different classification tasks.

Dataset	Classification	Models	P	R	F1	A
BACH	4-Class	EfficientNetB3	0.87594 ± 0.02862	0.87000 ± 0.02915	0.86793 ± 0.03077	0.87000 ± 0.02915
		ResNet50	0.91633 ± 0.02139	0.91250 ± 0.02092	0.91214 ± 0.02061	0.91250 ± 0.02092
		EfficientNetB3 + SCL	0.88619 ± 0.02286	0.88250 ± 0.02179	0.88198 ± 0.02247	0.88250 ± 0.02179
		ResNet50 + SCL	0.92016 ± 0.01229	0.91750 ± 0.01275	0.91748 ± 0.01268	0.91750 ± 0.01275
		EfficientNetB3 + ResNet50 + SCL	**0.92601 ± 0.02473**	**0.92250 ± 0.02669**	**0.92265 ± 0.02647**	**0.92250 ± 0.02669**
	7-Class	EfficientNetB3	0.76811 ± 0.00693	0.76649 ± 0.00963	0.76662 ± 0.00808	0.77991 ± 0.00886
		ResNet50	0.76909 ± 0.00764	0.76983 ± 0.00813	0.76830 ± 0.00789	0.78145 ± 0.00666
		EfficientNetB3 + SCL	0.77015 ± 0.00752	0.77630 ± 0.00612	0.77054 ± 0.00685	0.78498 ± 0.00664
		ResNet50 + SCL	0.77217 ± 0.00370	0.77551 ± 0.00139	0.77269 ± 0.00258	0.78450 ± 0.00266
		EfficientNetB3 + ResNet50 + SCL	**0.77417 ± 0.01300**	**0.77736 ± 0.00894**	**0.77389 ± 0.01178**	**0.78938 ± 0.01113**
BRACS	4-Class	EfficientNetB3	0.79392 ± 0.01573	0.82081 ± 0.01036	0.80173 ± 0.01479	0.81252 ± 0.01350
		ResNet50	0.81536 ± 0.01186	0.81371 ± 0.00867	0.81400 ± 0.01038	0.82573 ± 0.01224
		EfficientNetB3 + SCL	0.81610 ± 0.00742	0.81347 ± 0.00876	0.81453 ± 0.00804	0.82463 ± 0.00970
		ResNet50 + SCL	0.81660 ± 0.00318	0.82090 ± 0.00862	0.81828 ± 0.00567	0.82970 ± 0.00469
		EfficientNetB3 + ResNet50 + SCL	**0.81956 ± 0.01362**	**0.83419 ± 0.01161**	**0.82501 ± 0.01293**	**0.83345 ± 0.01321**
	2-Class	EfficientNetB3	0.92516 ± 0.02582	0.92534 ± 0.02573	0.92500 ± 0.02593	0.92519 ± 0.02587
		ResNet50	0.93350 ± 0.03520	0.93250 ± 0.03585	0.93285 ± 0.03557	0.93316 ± 0.03538
		EfficientNetB3 + SCL	0.93075 ± 0.02817	0.93062 ± 0.02852	0.93062 ± 0.02840	0.93085 ± 0.02828
		ResNet50 + SCL	0.94099 ± 0.03939	0.93966 ± 0.04048	0.94005 ± 0.04011	0.94036 ± 0.03990
		EfficientNetB3 + ResNet50 + SCL	**0.94514 ± 0.04504**	**0.94410 ± 0.04597**	**0.94445 ± 0.04566**	**0.94473 ± 0.04542**
BreakHis	2-Class	EfficientNetB3	0.99254 ± 0.00287	0.99337 ± 0.00320	0.99295 ± 0.00299	0.99388 ± 0.00260
		ResNet50	0.99388 ± 0.00417	0.99086 ± 0.00585	0.99233 ± 0.00479	0.99337 ± 0.00414
		EfficientNetB3 + SCL	0.99541 ± 0.00203	0.99284 ± 0.00201	0.99411 ± 0.00186	0.99490 ± 0.00161
		ResNet50 + SCL	0.99462 ± 0.00147	0.99246 ± 0.00348	0.99351 ± 0.00222	0.99439 ± 0.00191
		EfficientNetB3 + ResNet50 + SCL	**0.99661 ± 0.00183**	**0.99401 ± 0.00441**	**0.99528 ± 0.00302**	**0.99592 ± 0.00260**
EBHI	5-Class	EfficientNetB3	0.84709 ± 0.04626	0.80080 ± 0.02800	0.81769 ± 0.03409	0.91028 ± 0.01511
		ResNet50	0.83873 ± 0.03667	0.81641 ± 0.02057	0.82508 ± 0.02606	0.90423 ± 0.01045
		EfficientNetB3 + SCL	0.84732 ± 0.05454	0.77606 ± 0.04651	0.80145 ± 0.05210	0.91198 ± 0.01697
		ResNet50 + SCL	0.84518 ± 0.03115	0.82319 ± 0.01998	0.83167 ± 0.02352	0.91592 ± 0.00842
		EfficientNetB3 + ResNet50 + SCL	**0.85220 ± 0.03084**	**0.83414 ± 0.02164**	**0.84083 ± 0.02600**	**0.91930 ± 0.00828**
	2-Class	EfficientNetB3	0.97200 ± 0.00544	0.97176 ± 0.00562	0.97178 ± 0.00562	0.97179 ± 0.00561
		ResNet50	0.97581 ± 0.00333	0.97574 ± 0.00337	0.97573 ± 0.00338	0.97573 ± 0.00338
		EfficientNetB3 + SCL	0.97690 ± 0.00329	0.97687 ± 0.00328	0.97686 ± 0.00327	0.97686 ± 0.00327
		ResNet50 + SCL	0.97971 ± 0.00332	0.97969 ± 0.00331	0.97968 ± 0.00331	0.97968 ± 0.00331
		EfficientNetB3 + ResNet50 + SCL	**0.97971 ± 0.00572**	**0.97970 ± 0.00573**	**0.97969 ± 0.00573**	**0.97969 ± 0.00573**
GasHisSDB	2-Class	EfficientNetB3	0.96460 ± 0.00334	0.94244 ± 0.00620	0.95120 ± 0.00532	0.95441 ± 0.00484
		ResNet50	0.96146 ± 0.00301	0.93622 ± 0.00523	0.94593 ± 0.00458	0.94963 ± 0.00416
		EfficientNetB3 + SCL	0.99145 ± 0.00131	0.99011 ± 0.00133	0.99077 ± 0.00130	0.99119 ± 0.00124
		ResNet50 + SCL	0.97503 ± 0.00582	0.97567 ± 0.00417	0.97527 ± 0.00487	0.97635 ± 0.00471
		EfficientNetB3 + ResNet50 + SCL	**0.99101 ± 0.00079**	**0.99171 ± 0.00089**	**0.99135 ± 0.00078**	**0.99174 ± 0.00074**
HE-GHI-DS	2-Class	EfficientNetB3	0.94718 ± 0.02563	0.97229 ± 0.01945	0.95896 ± 0.02283	0.97281 ± 0.01528
		ResNet50	0.97311 ± 0.01548	0.96964 ± 0.01656	0.97087 ± 0.01149	0.98142 ± 0.00727
		EfficientNetB3 + SCL	0.97525 ± 0.01978	0.98660 ± 0.00746	0.98038 ± 0.01220	0.98713 ± 0.00832
		ResNet50 + SCL	0.98351 ± 0.01844	0.98571 ± 0.01211	0.98452 ± 0.01492	0.98999 ± 0.00969
		EfficientNetB3 + ResNet50 + SCL	**0.98900 ± 0.01283**	**0.99375 ± 0.00668**	**0.99119 ± 0.00817**	**0.99429 ± 0.00535**
MHIST	2-Class	EfficientNetB3	0.87532 ± 0.01340	0.88888 ± 0.01024	0.88090 ± 0.01074	0.89527 ± 0.01031
		ResNet50	0.88006 ± 0.01496	0.88476 ± 0.01743	0.88183 ± 0.01489	0.89749 ± 0.01276
		EfficientNetB3 + SCL	0.88359 ± 0.01540	0.87874 ± 0.01015	0.88052 ± 0.01068	0.89749 ± 0.00997
		ResNet50 + SCL	0.88237 ± 0.01355	0.88852 ± 0.01579	0.88496 ± 0.01367	0.90003 ± 0.01166
		EfficientNetB3 + ResNet50 + SCL	**0.88816 ± 0.01262**	**0.88627 ± 0.01088**	**0.88679 ± 0.01011**	**0.90257 ± 0.00919**
In–House	2-Class	EfficientNetB3	0.99694 ± 0.00152	0.99875 ± 0.00076	0.99784 ± 0.00113	0.99831 ± 0.00088
		ResNet50	0.99801 ± 0.00094	0.99831 ± 0.00061	0.99816 ± 0.00055	0.99857 ± 0.00043
		EfficientNetB3 + SCL	0.99782 ± 0.00132	0.99894 ± 0.00066	0.99838 ± 0.00097	0.99874 ± 0.00075
		ResNet50 + SCL	0.99745 ± 0.00107	0.99866 ± 0.00114	0.99805 ± 0.00106	0.99848 ± 0.00083
		EfficientNetB3 + ResNet50 + SCL	**0.99882 ± 0.00106**	**0.99902 ± 0.00083**	**0.99892 ± 0.00076**	**0.99916 ± 0.00060**

models are predominantly optimized for binary classification, overlooking the nuanced demands of multiclass scenarios. The limitations of CE loss functions, which can induce skewed features and biased classifiers, and concerns about feature transferability across domains, further compound the challenges in the field.

HistopathAI has been strategically designed to address these concerns and indeed outperforms other methods consistently across all considered datasets (Table 5). Unlike many models, it captures both high- and mid-level semantics, is optimized for multiclass tasks, and adopts a 5-fold cross-validation strategy for managing imbalanced data. Additionally, HistopathAI prioritizes reproducibility, comprehensive semantics capture, and enhanced feature transferability.

Computational complexity analysis

The computational complexity analysis (Table 6) reveals key insights into the trade-offs between different model architectures and the impact of incorporating SCL. The computation times shown are based on the MHIST dataset using a 5-fold cross-validation

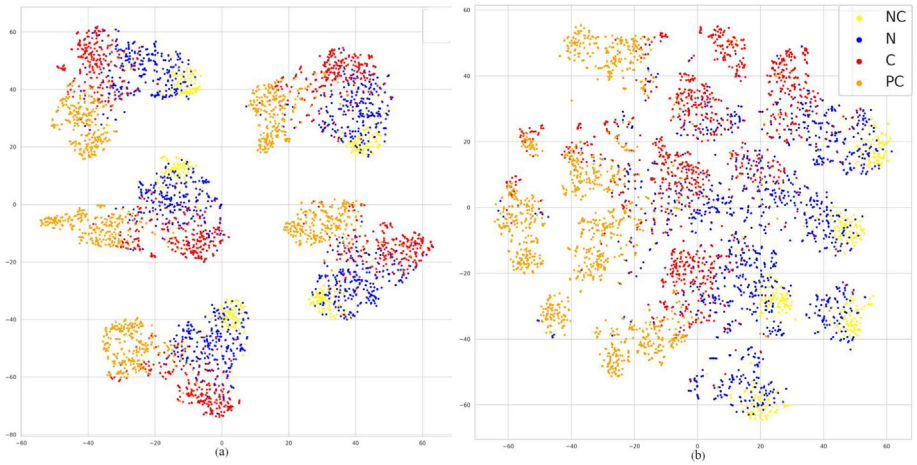


Fig. 5. Example t-SNE embeddings from BRACS 4-class classification. (a) SCL results in distinct, separated clusters, while (b) conventional CE results in dispersed, intermingled representations. This pattern holds consistently across datasets and tasks.

approach, but a similar pattern is observed across all datasets used in this study. EfficientNetB3, due to its high number of trainable parameters, exhibits a significantly longer training time compared to other models. However, when SCL is applied (EfficientNetB3 + SCL), both the training time and the number of trainable parameters in the second stage are notably reduced, indicating the efficiency of SCL in optimizing the training process. A similar pattern is observed with ResNet50 and its SCL-enhanced version (ResNet50 + SCL). The HistopathAI framework, which integrates both networks and SCL (EfficientNetB3 + ResNet50 + SCL), shows a balanced approach, combining a moderate training time with a higher total number of trainable parameters. This suggests that while HistopathAI is more complex, it benefits from the combined strengths of both architectures and the efficiency of SCL, making it a robust choice for achieving efficient and accurate histopathology image classification. Even though the training time is measured in hours, the test time is very low, taking only a few milliseconds per image, making the model highly efficient for real-time applications.

Discussion

The rising demand for accurate cancer diagnostics, motivated by the global prevalence of the disease, emphasizes the need for intelligent computer-aided procedures. The current global transition to digital workflows offers the opportunity to deliver decision-assisting AI to improve existing traditional diagnostic mechanisms which have often been inconsistent. Our proposed HistopathAI model offers an innovative, systematic, and generalized tool for histopathological image classification. Its distinction rests on two key integrations. First, the employment of SCL with CE ensures that our model benefits from both strong feature representation and accurate classification. Unlike typical self-supervised methodologies, SCL optimizes the use of labeled data to yield high-quality feature representations. Second, the merger of EfficientNetB3 with ResNet50 through the HDFF enables capturing both high- and low-level image features, offering a more comprehensive representation of histopathological images, which enhances the model's ability to discern intricate patterns in the data.

Our experimental results reveal the consistent superiority of HistopathAI across a wide range of datasets, including our in–house collection. The consistent excellence combined with minimal variance solidifies HistopathAI's candidacy as a comprehensive solution for histopathological image classification. It enables optimized laboratory workflows, prioritization of critical cases, and relief from visual and cognitive fatigue for pathologists, thus enhancing diagnostic precision and consistency, subject of course to rigorous local in–house testing and validation with appropriate regulatory and institutional approval for clinical applications. While numerous existing methodologies show dataset-specific achievements, HistopathAI's performance breadth stands out and provides a holistic solution. Contributing to the worldwide challenge to develop AI-based CAD systems for pathology, HistopathAI raises the performance bar and opens the door to full automation of diagnostic procedures.

Nonetheless, the quest for perfection in digital pathology is an ongoing pursuit. Our study is not without limitations. First, while our model demonstrates robust performance across several datasets, its applicability to broader clinical settings may still be limited. Additionally, while HistopathAI is designed to cope with imbalanced data, it may face challenges with more severely imbalanced datasets, which could require further validation. The model's complexity, combining EfficientNetB3 and ResNet50 architectures with SCL and CE loss, may also limit its interpretability and scalability, particularly in resource-constrained environments. Additionally, the 'black box' nature of these models poses a significant challenge in terms of explainability, making it difficult for clinicians to understand and trust the decision-making process, which is crucial for clinical adoption. Furthermore, the variability in staining techniques and slide preparation across different laboratories could affect the model's performance in real-world clinical settings, necessitating additional standardization or domain adaptation efforts.

Future research should explore more complex architectures and investigate the integration of various diagnostic modalities, such as combining histopathological imaging with genomic data, to enhance diagnostic precision and patient outcomes. Another promising direction is the application of transfer learning, where HistopathAI could be fine-tuned on new datasets with limited labeled data, thus improving adaptability to different clinical settings. However, this approach comes with challenges, such as the risk of overfitting and the need for effective domain adaptation strategies to handle variations between source and target datasets. Additionally, HistopathAI could be integrated into clinical workflows by acting as a prescreening tool for pathologists. This integra-

Table 4
Comparison of different model combinations using the 5x2cv F-test. The critical value for all tests is 3.2974. An asterisk (*) indicates statistical significance.

Dataset	Classification	Comparison	F-Statistic
BACH	4-Class	HistopathAI versus ResNet50 + SCL	2.4835
		HistopathAI versus ResNet50	6.9231*
		HistopathAI versus EfficientNetB3	86.2212*
		HistopathAI versus EfficientNetB3 + SCL	5.8209*
	7-Class	HistopathAI versus ResNet50 + SCL	5.5955*
		HistopathAI versus ResNet50	23.3282*
		HistopathAI versus EfficientNetB3	7.9731*
		HistopathAI versus EfficientNetB3 + SCL	9.3220*
BRACS	4-Class	HistopathAI versus ResNet50 + SCL	2.0635
		HistopathAI versus ResNet50	8.7004*
		HistopathAI versus EfficientNetB3	17.6915*
		HistopathAI versus EfficientNetB3 + SCL	2.6426
	2-Class	HistopathAI versus ResNet50 + SCL	2.9088
		HistopathAI versus ResNet50	4.3011*
		HistopathAI versus EfficientNetB3	3.7789*
		HistopathAI versus EfficientNetB3 + SCL	3.2835
BreaKHis	2-Class	HistopathAI versus ResNet50 + SCL	3.1250
		HistopathAI versus ResNet50	3.6667*
		HistopathAI versus EfficientNetB3	2.8312
		HistopathAI versus EfficientNetB3 + SCL	2.1429
EBHI	5-Class	HistopathAI versus ResNet50 + SCL	6.5318*
		HistopathAI versus ResNet50	7.2802*
		HistopathAI versus EfficientNetB3	3.5951*
		HistopathAI versus EfficientNetB3 + SCL	3.9581*
	2-Class	HistopathAI versus ResNet50 + SCL	2.000
		HistopathAI versus ResNet50	4.7222*
		HistopathAI versus EfficientNetB3	8.1207*
		HistopathAI versus EfficientNetB3 + SCL	3.6698*
GasHisSDB	2-Class	HistopathAI versus ResNet50 + SCL	20.2181*
		HistopathAI versus ResNet50	242.4144*
		HistopathAI versus EfficientNetB3	114.9488*
		HistopathAI versus EfficientNetB3 + SCL	3.2974
HE-GHI-DS	2-Class	HistopathAI versus ResNet50 + SCL	2.4692
		HistopathAI versus ResNet50	18.4876*
		HistopathAI versus EfficientNetB3	7.2944*
		HistopathAI versus EfficientNetB3 + SCL	9.9997*
MHIST	2-Class	HistopathAI versus ResNet50 + SCL	2.0102
		HistopathAI versus ResNet50	2.0463
		HistopathAI versus EfficientNetB3	7.1914*
		HistopathAI versus EfficientNetB3 + SCL	2.0013
In–House	2-Class	HistopathAI versus ResNet50 + SCL	3.3334*
		HistopathAI versus ResNet50	3.7678*
		HistopathAI versus EfficientNetB3	4.2224*
		HistopathAI versus EfficientNetB3 + SCL	2.9998

tion would involve prioritizing the diagnosis of malignant cases over normal ones, allowing for faster reporting in high-volume settings and facilitating the early ordering of additional biomarkers or molecular testing as appropriate. Such an approach would not only optimize the workflow efficiency but also potentially improve patient outcomes by ensuring that critical cases are addressed promptly.

Real-world clinical implementation is crucial, and subsequent studies should therefore focus on scalability and address challenges such as model interpretability and robustness to variability in staining techniques across laboratories. Practical deployments should be accompanied by rigorous validation studies to ascertain the generalizability of the models across different demographic and pathological spectra. Feedback from these deployments will be instrumental in bridging the gap between research advancements and practical applicability, ensuring the transition from lab to clinic not only enhances diagnostic accuracy but also aligns with clinical workflows and ultimately improves patient care.

Conclusion

In this study, we introduced HistopathAI, a novel framework for histopathology image classification that combines SCL with HDFF. Our approach addresses the variability and subjectivity inherent in histopathological analysis by leveraging the strengths of SCL and merging feature vectors from EfficientNetB3 and ResNet50 for a richer image representation. Extensive experiments on various datasets demonstrated that HistopathAI consistently outperforms existing methods, particularly in imbalanced data scenarios. The framework's sequential methodology effectively captures high-

Table 5
Performance comparison with existing studies.

Dataset	Classification	Reference	Method	Accuracy (%)
BACH	4-Class	[62]	ResNet	83
		[7]	TL based ResNet18, ResNeXt	87.10
		[63]	Inception-V3	87.50
		[64]	TL based Inception-V3	91.30
		[65]	CNN	90
		[4]	HACT-Net	91
		[66]	Lightweight DNN	86.15
		[67]	BiLSTMs	90
		Ours	HistopathAI	**92.250 ± 2.669**
BRACS	7-Class	[4]	HACT-Net	61.53 ± 0.87
		[46]	DL-based method	55.9
		[47]	Stain-Invariant SSL	59 ± 1.6
		Ours	HistopathAI	**78.938 ± 1.113**
	4-Class	[4]	HACT-Net	69.04 ± 0.46
		Ours	HistopathAI	**83.345 ± 1.321**
	2-Class	[4]	HACT-Net	83.63 ± 0.73
		Ours	HistopathAI	**94.473 ± 4.542**
BreakHis	2-Class	[48]	SA-Net	96
		[49]	TL-Based ResNet50, InceptionV2, Inception ResNetV2	96.24
		[8]	DCNN	97.90
		[13]	BHCNet	98.87
		[12]	TL-Based VGG16 and VGG19	98.10
		[50]	CNN	89.46
		[51]	ResHist	92.52
		[52]	DSoPN	96.54
		[53]	myResNet34	94.03
		[54]	TL-Based CNN-LSTM	99.03
		[55]	CSML, Hashing DNN	99.30
		[56]	SVM, RBF	94.32
		[57]	Ensemble TL	98.90
		[58]	MobileNet-SVM	91.30
		[59]	PLWT	96.9 ± 0.3
		[60]	Cd-dtffNet	98.27
		[61]	CBAM-VGGNet	98.96
		Ours	HistopathAI	**99.592 ± 0.260**
EBHI	5-Class	-	None	-
		Ours	HistopathAI	**91.930 ± 0.828**
	2-Class	[90]	VGG16	95.37
		Ours	HistopathAI	**97.969 ± 0.573**
GasHisSDB	2-Class	[73]	VGG16	96.47
		[75]	CCF-GNN	97.8 ± 1.1
		[76]	MCLNet	97.85
		[74]	VGG16	96.47
		[77]	MLP IN MLP	96.5
		[78]	EfficientNetB0 + B1 + DenseNet121 + 169 + MobileNetV2	99.16
		Ours	HistopathAI	**99.174 ± 0.074**
HE-GHI-DS	2-Class	[79]	GasHis-Transformer	97.97 ± 0.74
		[80]	CAM-VT	90.48
		Ours	HistopathAI	**99.429 ± 0.535**
MHIST	2-Class	[68]	CNN + Trans + TAE + SSL	89.68
		[69]	ConvNeXt-L	88.56 ± 0.67
		[70]	DenseNet121	79.73
		[71]	MoCo v2	85.88
		Ours	HistopathAI	**90.257 ± 0.919**

Table 6
Computational complexity of different models on the MHIST dataset.

Model	Training Time (HH:MM:SS)	Trainable Parameters (Stage 1)	Trainable Parameters (Stage 2)
EfficientNetB3	21:01:59	12,235,234	N/A
EfficientNetB3 + SCL	11:16:49	10,892,968	1,539,002
ResNet50	05:20:27	25,585,594	N/A
ResNet50 + SCL	04:48:28	23,796,864	2,051,002
HistopathAI (EfficientNetB3 + ResNet50 + SCL)	10:43:28	34,689,704	3,587,002

and mid-level semantics, leading to superior intraclass cohesion and interclass separation, and resulting in more accurate and generalizable classifiers.

In conclusion, HistopathAI offers a robust solution for enhancing diagnostic accuracy in digital pathology, with the potential to significantly improve clinical workflows and patient outcomes through more reliable cancer diagnosis.

Credit author statement

Md Mamunur Rahaman: Conceptualization, Methodology, Software, Writing-original draft, Visualization, Validation. **Ewan K. A. Millar**: Data curation, Resources, Visualization, Writing-review & editing, Supervision. **Erik Meijering**: Resources, Formal analysis, Writing-review & editing, Visualiztion, Project administration, Funding acquisition, Supervision.

Compliance with ethics requirements

Ethical approval for the use of breast cancer WSIs in our in–house dataset was provided by the Human Research Ethics Committee at Prince of Wales Hospital Sydney, with the approval number HREC/17/POWH/389–17/176. This study was conducted in accordance with the guidelines of the committee and with respect for the subjects' data. Due to ethical restrictions, this dataset is not publicly available, but it was used in this study to demonstrate the potential of the proposed framework beyond public datasets.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

EKAM is supported by a Researcher Exchange and Development in Industry (REDI) Fellowship from MTPConnect and MRFF Australia. The authors thank the National Computational Infrastructure (NCI) Australia for providing access to the Gadi high-performance computing system, which contributed to the research results reported in this paper.

References

[1] Sung Hyuna, Ferlay Jacques, Siegel Rebecca L, Laversanne Mathieu, Soerjomataram Isabelle, Jemal Ahmedin, Bray Freddie. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021.
[2] World Health Organization. Cancer, 2023. Accessed on: 2023–06-26.
[3] Rahaman Md Mamunur, Millar Ewan KA, Meijering Erik. Breast cancer histopathology image-based gene expression prediction using spatial transcriptomics data and deep learning. Scient Rep 2023;13(1):13604.
[4] Pati Pushpak, Jaume Guillaume, Foncubierta-Rodríguez Antonio, Feroce Florinda, Maria Anna Anniciello, Scognamiglio Giosue, et al. Hierarchical graph representations in digital pathology. Med Image Anal 2022;75:102264.
[5] Yang Ching-Wei, Lin Tzu-Ping, Huang Yi-Hsiu, Chung Hsiao-Jen, Kuo Junne-Yih, Huang William JS, et al. Does extended prostate needle biopsy improve the concordance of Gleason scores between biopsy and prostatectomy in the Taiwanese population? J Chin Med Assoc 2012;75(3):97–101.
[6] Li Y, Wu J, Wu Q. Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning. IEEE Access 2019;7:21400–8.
[7] Cao H, Bernard S, Heutte L, Sabourin R. Improve the performance of transfer learning without fine-tuning using dissimilarity-based multi-view learning for breast cancer histology images. In: International Conference on Image Analysis and Recognition (ICIAR). p. 779–87.
[8] Xie J, Liu R, Luttrell IV J, Zhang C. Deep learning based analysis of histopathological images of breast cancer. Front Genet 2019;10:80.
[9] Karthik R, Menaka R, Siddharth MV. Classification of breast cancer from histopathology images using an ensemble of deep multiscale networks. Biocybern Biomed Eng 2022;42(3):963–76.
[10] Li Junjian, Liu Jin, Yue Hailin, Cheng Jianhong, Kuang Hulin, Bai Harrison, Wang Yuping, Wang Jianxin. DARC: Deep adaptive regularized clustering for histopathological image classification. Med Image Anal 2022;80:102521.
[11] Gao Zhiyang, Zhiyang Lu, Wang Jun, Ying Shihui, Shi Jun. A convolutional neural network and graph convolutional network based framework for classification of breast histopathological images. IEEE J Biomed Health Inform 2022;26(7):3163–73.
[12] Huynh MBT, Truong VH. Fusing of deep learning, transfer learning and gan for breast cancer histopathological image classification. In: In International Conference on Computer Science, Applied Mathematics and Applications (ICCSAMA). p. 255–66.
[13] Jiang Y, Chen L, Zhang H, Xiao X. Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. PLoS One 2019;14(3):e0214587.
[14] K Shankar, Ashit Kumar Dutta, Sachin Kumar, Gyanendra Prasad Joshi, and Ill Chul Doo. Chaotic sparrow search algorithm with deep transfer learning enabled breast cancer classification on histopathological images. Cancers, 14 (11):2770, 2022.
[15] Hayder A Khikani, Naira Elazab, Ahmed Elgarayhi, Mohammed Elmogy, and Mohammed Sallah. Breast cancer classification based on histopathological images using a deep learning capsule network. arXiv:2208.00594, 2022.
[16] Chattopadhyay Soham, Dey Arijit, Singh Pawan Kumar, Sarkar Ram. DRDA-Net: Dense residual dual-shuffle attention network for breast cancer classification using histopathological images. Comput Biol Med 2022;145:105437.
[17] Wang Chaoqing, Gong Weijun, Cheng Junlong, Qian Yurong. DBLCNN: Dependency-based lightweight convolutional neural network for multi-classification of breast histopathology images. Biomed Signal Process Control 2022;73:103451.
[18] Linder N, Konsti J, Turkki R, Rahtu E, Lundin M, Nordling S, Haglund C, Ahonen T, Pietikäinen M, Lundin J. Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. Diagn Pathol 2012;7(1):1–11.
[19] Jiao L, Chen Q, Li S, Xu Y. Colon cancer detection using whole slide histopathological images. In: World congress on medical physics and biomedical engineering. p. 1283–6.
[20] Peyret R, Bouridane A, Al-Maadeed SA, Kunhoth S, Khelifi F. Texture analysis for colorectal tumour biopsies using multispectral imagery. In: International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). p. 7218–21.
[21] Chaddad A, Desrosiers C, Bouridane A, Toews M, Hassan L, Tanougast C. Multi texture analysis of colorectal cancer continuum using multispectral imagery. PLoS One 2016;11(2):e0149893.
[22] Kather JN, Weis C-A, Bianconi F, Melchers SM, Schad LR, Gaiser T, Marx A, Zöllner FG. Multi-class texture analysis in colorectal cancer histology. Sci Rep 2016;6(1):1–11.
[23] Sari CT, Gunduz-Demir C. Unsupervised feature extraction via deep learning for histopathological classification of colon tissue images. IEEE Trans. Med. Imag. 2019;38(5):1139–49.
[24] J. Malik, S. Kiranyaz, S. Kunhoth, T. Ince, S. Al-Maadeed, R. Hamila, and M. Gabbouj. Colorectal cancer diagnosis from histology images: a comparative study. arXiv:1903.11210, 2019.
[25] Yoon H, Lee J, Oh JE, Kim HR, Lee S, Chang HJ, Sohn DK. Tumor identification in colorectal histology images using a convolutional neural network. J Digit Imag 2019;32(1):131–40.
[26] Dif N, Elberrichi Z. A new deep learning model selection method for colorectal cancer classification. Int. J. Swarm Intell. Res. 2020;11(3):72–88.
[27] Nguyen H-G, Blank A, Lugli A, Zlobec I. An effective deep learning architecture combination for tissue microarray spots classification of H&E stained colorectal images. In: In IEEE International Symposium on Biomedical Imaging (ISBI). p. 1271–4.
[28] Ghosh S, Bandyopadhyay A, Sahay S, Ghosh R, Kundu I, Santosh K. Colorectal histology tumor detection using ensemble deep neural network. Eng. Appl. Artif. Intell. 2021;100:104202.
[29] Hamida AB, Devanne M, Weber J, Truntzer C, Derangère V, Ghiringhelli F, Forestier G, Wemmert C. Deep learning for colon cancer histopathological images analysis. Comput. Biol. Med. 2021;136:104730.
[30] E.F. Ohata, J.V.S. das Chagas, G.M. Bezerra, M.M. Hassan, V.H.C. de Albuquerque, and P.P. Reboucas Filho. A novel transfer learning approach for the classification of histological images of colorectal cancer. J. Supercomput., 77:9494–9519, 2021.
[31] Sarkar T, Hazra A, Das N. Classification of colorectal cancer histology images using image reconstruction and modified DenseNet. In: In International Conference on Computational Intelligence in Communications and Business Analytics (CICBA). p. 259–71.
[32] Sarwinda D, Paradisa RH, Bustamam A, Anggia P. Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. Procedia Comput. Sci. 2021;179:423–31.
[33] Sharma H, Zerbe N, Klempert I, Hellwich O, Hufnagl P. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. Comput. Med. Imaging Graph. 2017;61:2–13.
[34] Liu B, Yao K, Huang M, Zhang J, Li Y, Li R. Gastric pathology image recognition based on deep residual networks. In IEEE Annual Computer Software and Applications Conference (COMPSAC) 2018;2:408–12.
[35] Wang S, Zhu Y, Yu L, Chen H, Lin H, Wan X, Fan X, Heng P. RMDL: Recalibrated multi-instance deep learning for whole slide gastric image classification. Med. Image Anal. 2019;101549:58.

[36] Song Zhigang, Zou Shuangmei, Zhou Weixun, Huang Yong, Shao Liwei, Yuan Jing, Gou Xiangnan, Jin Wei, Wang Zhanbo, Chen Xin, Ding Xiaohui, Liu Jinhong, Chunkai Yu, Calvin Ku, Liu Cancheng, Sun Zhuo, Gang Xu, Wang Yuefeng, Zhang Xiaoqing, Wang Dandan, Wang Shuhao, Wei Xu, Davis Richard C, Shi Huaiyin. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. Nat. Commun. 2020;11(1):1–9.
[37] Kosaraju S, Hao J, Koh H, Kang M. Deep-Hipo: Multiscale receptive field deep learning for histopathological image analysis. Methods 2020;179:3–13.
[38] Iizuka O, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep learning models for histopathological classification of gastric and colonic epithelial tumours. Sci. Rep. 2020;10(1):1–11.
[39] Zhu R, Zhang R, Xue D. Lesion detection of endoscopy images based on convolutional neural network features. In: In International Congress on Image and Signal Processing (CISP). p. 372–6.
[40] Ishihara K, Ogawa T, Haseyama M. Detection of gastric cancer risk from X-ray images via patch-based convolutional neural network. In: In IEEE International Conference on Image Processing (ICIP). p. 2055–9.
[41] Li R, Li J, Wang X, Liang P, Gao J. Detection of gastric cancer and its histological type based on iodine concentration in spectral CT. Cancer Imaging 2018;18 (1):1–10.
[42] Islam Ashraful, Chen Chun-Fu Richard, Panda Rameswar, Karlinsky Leonid, Radke Richard, Feris Rogerio. A broad study on the transferability of visual representations with contrastive learning. In: In IEEE/CVF International Conference on Computer Vision (ICCV). p. 8845–55.
[43] Nanxuan Zhao, Zhirong Wu, Rynson WH Lau, and Stephen Lin. What makes instance discrimination good for transfer learning? arXiv:2006.06606, 2020.
[44] Sepahvand Majid, Abdali-Mohammadi Fardin. Overcoming limitation of dissociation between MD and MI classifications of breast cancer histopathological images through a novel decomposed feature-based knowledge distillation method. Comput. Biol. Med. 2022;145:105413.
[45] Zabit Hameed, Begonya Garcia-Zapirain, José Javier Aguirre, and Mario Arturo Isaza-Ruget. Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network. Sci. Rep., 12 (1):15600, 2022.
[46] Nadia Brancati, Anna Maria Anniciello, Pushpak Pati, Daniel Riccio, Giosuè Scognamiglio, Guillaume Jaume, Giuseppe De Pietro, Maurizio Di Bonito, Antonio Foncubierta, Gerardo Botti, Maria Gabrani, Florinda Feroce, and Maria Frucci. BRACS: A dataset for breast carcinoma subtyping in H&E histology images. Database, 2022:baac093, 2022.
[47] Alexandre Tiard, Alex Wong, David Joon Ho, Yangchao Wu, Eliram Nof, Stefano Soatto, and Saad Nadeem. Stain-invariant self supervised learning for histopathology image analysis. arXiv:2211.07590, 2022.
[48] Bolei Xu, Jingxin Liu, Xianxu Hou, Bozhi Liu, Jon Garibaldi, Ian O Ellis, Andy Green, Linlin Shen, and Guoping Qiu. Look, investigate, and classify: a deep hybrid attention method for breast cancer classification. In IEEE International Symposium on Biomedical Imaging (ISBI), pages 914–918, 2019.
[49] Bhuiyan Md Nuruddin Qaisar, Shamsujjoha Md, Ripon Shamim H, Proma Farhin Haque, Khan Fuad. Transfer learning and supervised classifier based prediction model for breast cancer. In: Big Data Analytics for Intelligent Healthcare Management. Elsevier; 2019. p. 59–86.
[50] Saxena S, Shukla S, Gyanchandani M. Pre-trained convolutional neural networks as feature extractors for diagnosis of breast cancer using histopathology. Int. J. Imaging Syst. Technol. 2020;30(3):577–91.
[51] Gour Mahesh, Jain Sweta, Sunil Kumar T. Residual learning based CNN for breast cancer histopathological image classification. Int. J. Imaging Syst. Technol. 2020;30(3):621–35.
[52] Jiasen Li, Jianxin Zhang, Qiule Sun, Hengbo Zhang, Jing Dong, Chao Che, and Qiang Zhang. Breast cancer histopathological image classification based on deep second-order pooling network. In International Joint Conference on Neural Networks (IJCNN), pages 1–7, 2020.
[53] Chuhan Hu, Sun Xiaoyan, Yuan Zhenming, Yingfei Wu. Classification of breast cancer histopathological image with deep residual learning. Int. J. Imaging Syst. Technol. 2021;31(3):1583–94.
[54] Srikantamurthy Mahati Munikoti, Rallabandi VP, Dudekula Dawood Babu, Natarajan Sathishkumar, Park Junhyung. Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning. BMC Med. Imaging 2023;23(1):1–15.
[55] Pratiher Sawon, Chattoraj Subhankar. Diving deep onto discriminative ensemble of histological hashing & class-specific manifold learning for multi-class breast carcinoma taxonomy. In: In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). p. 1025–9.
[56] Atban Furkan, Ekinci Ekin, Garip Zeynep. Traditional machine learning algorithms for breast cancer image classification with optimized deep features. Biomed. Signal Process. Control 2023;81:104534.
[57] Zheng Yuchao, Li Chen, Zhou Xiaomin, Chen Haoyuan, Hao Xu, Li Yixin, Zhang Haiqing, Li Xiaoyan, Sun Hongzan, Huang Xinyu, Grzegorzek Marcin. Application of transfer learning and ensemble learning in image-level classification for breast histopathology. Intell. Med. 2023;3(02):115–28.
[58] Ogundokun Roseline Oluwaseun, Misra Sanjay, Akinrotimi Akinyemi Omololu, Ogul Hasan. MobileNet-SVM: A lightweight deep transfer learning model to diagnose BCH scans for IoMT-based imaging sensors. Sensors 2023;23(2):656.
[59] Junjie Wang, Hao Quan, Chengguang Wang, and Genke Yang. Pyramid-based self-supervised learning for histopathological image classification. Comput. Biol. Med., page 107336, 2023.
[60] Wang Pin, Li Pufei, Li Yongming, Wang Jiaxin, Jin Xu. Histopathological image classification based on cross-domain deep transferred feature fusion. Biomed. Signal Process. Control 2021;68:102705.
[61] Ijaz Areesha, Raza Basit, Kiran Iqra, Waheed Abdul, Raza Aadil, Shah Habib, Aftan Sulaiman. Modality specific CBAM-VGGNet model for the classification of breast histopathology images via transfer learning. IEEE Access 2023;11:15750–62.
[62] Awan R, Koohbanani NA, Shaban M, Lisowska A, Rajpoot N. Context-aware learning using transferable features for classification of breast cancer histology images. In: In International Conference on Image Analysis and Recognition (ICIAR). p. 788–95.
[63] Vang YS, Chen Z, Xie X. Deep learning framework for multi-class breast cancer histology image classification. In: In International Conference on Image Analysis and Recognition (ICIAR). p. 914–22.
[64] Yan Rui, Ren Fei, Wang Zihao, Wang Lihua, Zhang Tong, Liu Yudong, Rao Xiaosong, Zheng Chunhou, Zhang Fa. Breast cancer histopathological image classification using a hybrid deep neural network. Methods 2020;173:52–60.
[65] Roy Kaushiki, Banik Debapriya, Bhattacharjee Debotosh, Nasipuri Mita. Patch-based system for classification of breast histology images using deep learning. Comput. Med. Imaging Graph. 2019;71:90–103.
[66] Waqas Muhammad, Maul Tomas, Liao Iman Yi, Ahmed Amr. Lightweight deep network for the classification of breast cancer histopathological images. In: In International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). p. 1–6.
[67] Tripathi Suvidha, Singh Satish Kumar, Lee Hwee Kuan. An end-to-end breast tumour classification model using context-based patch modelling–a BiLSTM approach for image classification. Comput. Med. Imaging Graph. 2021;87:101838.
[68] Xiyue Wang, Sen Yang, Jun Zhang, Minghui Wang, Jing Zhang, Junzhou Huang, Wei Yang, and Xiao Han. TransPath: Transformer-based self-supervised learning for histopathological image classification. In International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), pages 186–195, 2021.
[69] Springenberg Maximilian, Frommholz Annika, Wenzel Markus, Weicken Eva, Ma Jackie, Strodthoff Nils. From modern CNNs to vision transformers: Assessing the performance, robustness, and classification strategies of deep learning models in histopathology. Med. Image Anal. 2023;87:102809.
[70] Tabrizchi Sepehr, Nezhadi Ali, Angizi Shaahin, Roohi Arman. AppCiP: Energy-efficient approximate convolution-in-pixel scheme for neural network acceleration. IEEE J Emerg Sel Top Circ Syst 2023;13(1):225–36.
[71] Kang Mingu, Song Heon, Park Seonwook, Yoo Donggeun, Pereira Sérgio. Benchmarking self-supervised learning on diverse pathology datasets. In: In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). p. 3344–54.
[72] Sharma Harshita, Zerbe Norman, Böger Christine, Wienert Stephan, Hellwich Olaf, Hufnagl Peter. A comparative study of cell nuclei attributed relational graphs for knowledge description and categorization in histopathological gastric cancer whole slide images. In: In IEEE International Symposium on Computer-Based Medical Systems (CBMS). p. 61–6.
[73] Weiming Hu, Chen Li, Xiaoyan Li, Md Mamunur Rahaman, Jiquan Ma, Yong Zhang, Haoyuan Chen, Wanli Liu, Changhao Sun, Yudong Yao, Hongzan Sun, and Marcin Grzegorzek. GasHisSDB: A new gastric histopathology image dataset for computer aided diagnosis of gastric cancer. Comput. Biol. Med., 142:105207, 2022.
[74] Weiming Hu, Chen Haoyuan, Liu Wanli, Li Xiaoyan, Sun Hongzan, Huang Xinyu, Grzegorzek Marcin, Li Chen. A comparative study of gastric histopathology sub-size image classification: From linear regression to visual transformer. Front Med 2022;9:1072109.
[75] Hongxiao Wang, Gang Huang, Zhuo Zhao, Liang Cheng, Anna Juncker-Jensen, Máté Levente Nagy, Xin Lu, Xiangliang Zhang, and Danny Z Chen. CCF-GNN: A unified model aggregating appearance, microenvironment, and topology for pathology image classification. IEEE Trans. Med. Imaging, 42(11), 3179–3193, 2023.
[76] Xuanshuo Fu, Liu Shuyong, Li Chao, Sun Jingbo. MCLNet: An multidimensional convolutional lightweight network for gastric histopathology image classification. Biomed. Signal Process. Control 2023;80:104319.
[77] Liu Mingyang, Li Liyuan, Wang Haoran, Guo Xinyu, Liu Yunpeng, Li Yuguang, Song Kaiwen, Shao Yanbin, Fei Wu, Zhang Junjie, Sun Nao, Zhang Tianyu, Luan Lan. A multilayer perceptron-based model applied to histopathology image classification of lung adenocarcinoma subtypes. Front Oncol 2023;13:1172234.
[78] Ming Ping Yong, Yan Chai Hum, Khin Wee Lai, Ying Loong Lee, Choon-Hian Goh, Wun-She Yap, and Yee Kai Tee. Histopathological gastric cancer detection on GasHisSDB dataset using deep ensemble learning. Diagnostics, 13 (10):1793, 2023.
[79] Haoyuan Chen, Chen Li, Ge Wang, Xiaoyan Li, Md Mamunur Rahaman, Hongzan Sun, Weiming Hu, Yixin Li, Wanli Liu, Changhao Sun, Shiliang Ai, and Marcin Grzegorzek. GasHis-Transformer: A multi-scale visual transformer approach for gastric histopathological image detection. Pattern Recognit., 130:108827, 2022.
[80] Fan Zizhen, Xiangchen Wu, Li Changzhong, Chen Haoyuan, Liu Wanli, Zheng Yuchao, Chen Jing, Li Xiaoyan, Sun Hongzan, Jiang Tao, Grzegorzek Marcin, Li Chen. CAM-VT: A weakly supervised cervical cancer nest image identification approach using conjugated attention mechanism and visual transformer. Comput Biol Med 2023;162:107070.

[81] Tan M, Le QV. EfficientNet: Rethinking model scaling for convolutional neural networks. In: In International Conference on Machine Learning (ICML). p. 6105–14.
[82] He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. Deep residual learning for image recognition. In: In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). p. 770–8.
[83] Jian Yang, Jing-yu Yang, David Zhang, and Jian-feng Lu. Feature fusion: parallel strategy vs. serial strategy. Pattern Recognit., 36(6):1369–1381, 2003.
[84] Khosla Prannay, Teterwak Piotr, Wang Chen, Sarna Aaron, Tian Yonglong, Isola Phillip, Maschinot Aaron, Liu Ce, Krishnan Dilip. Supervised contrastive learning. Adv Neural Inform Process Syst 2020;33:18661–73.
[85] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv:1412.6980, 2014.
[86] Alpaydm Ethem. Combined 5×2cv F-test for comparing supervised classification learning algorithms. Neural Comput 1999;11(8):1885–92.
[87] Sebastian Raschka. Model evaluation, model selection, and algorithm selection in machine learning. arXiv:1811.12808, 2018.
[88] Guilherme Aresta, Teresa Araújo, Scotty Kwok, Sai Saketh Chennamsetty, Mohammed Safwan, Varghese Alex, Bahram Marami, Marcel Prastawa, Monica Chan, Michael Donovan, Gerardo Fernandez, Jack Zeineh, Matthias Kohl, Christoph Walz, Florian Ludwig, Stefan Braunewell, Maximilian Baust, Quoc Dang Vu, Minh Nguyen Nhat To, Eal Kim, and Paulo Aguiar. BACH: Grand challenge on breast cancer histology images. Med. Image Anal., 56:122–139, 2019.
[89] Spanhol Fabio A, Oliveira Luiz S, Petitjean Caroline, Heutte Laurent. A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng 2015;63(7):1455–62.
[90] Hu Weiming, Li Chen, Rahaman Md Mamunur, Chen Haoyuan, Liu Wanli, Yao Yudong, et al. EBHI: A new enteroscope biopsy histopathological H&E image dataset for image classification evaluation. Phys Med 2023;107:102534.
[91] Wei Jerry, Suriawinata Arief, Ren Bing, Liu Xiaoying, Lisovsky Mikhail, Vaickus Louis, et al. A petri dish for histopathology image analysis. In International Conference on Artificial Intelligence in Medicine (AIME) 2021; p. 11–24.