POLITECNICO DI TORINO
Repository ISTITUZIONALE

ShortNeXt: A novel method for accurate classification of colorectal cancer histopathology images

Original
ShortNeXt: A novel method for accurate classification of colorectal cancer histopathology images / Barua, P.D., Tasci, B., Baygin, M., Dogan, S., Tuncer, T., Molinari, F., Salvi, M., Acharya, U.R.. - In: COMPUTER VISION AND IMAGE UNDERSTANDING. - ISSN 1077-3142. - 265:(2026). [10.1016/j.cviu.2026.104672]

Availability:
This version is available at: 11583/3007417 since: 2026-02-06T13:42:14Z

Publisher:
Elsevier

Published
DOI:10.1016/j.cviu.2026.104672

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

19 November 2026

Computer Vision and Image Understanding 265 (2026) 104672

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

ShortNeXt: A novel method for accurate classification of colorectal cancer histopathology images

Prabal Datta Barua [a], Burak Tasci [b], Mehmet Baygin [c], Sengul Dogan [d], Turker Tuncer [d], Filippo Molinari [e], Salvi Massimo [e,*], U. Rajendra Acharya [f,g]

[a] School of Business (Information System), University of Southern Queensland, Australia
[b] Vocational School of Technical Sciences, Firat University, Elazig 23119, Turkey
[c] Department of Computer Engineering, Faculty of Engineering, Erzurum Technical University, Erzurum, Turkey
[d] Department of Digital Forensics Engineering, College of Technology, Firat University, 23119, Elazig, Turkey
[e] Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
[f] School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, Australia
[g] Centre for Health Research, University of Southern Queensland, Australia

ARTICLE INFO

Communicated by Pong Chi Yuen

Keywords:
ShortNeXt
Histopathological image classification
Cancer detection
Convolutional neural network
Digital pathology

ABSTRACT

Cancer is a chaotic disease known as the plague of our age and there are many subtypes of the cancer. Cancer is commonly seen disorder and its mortality rate is very high. Therefore, many researchers have worked/studied on the cancer detection and treatment. To contribute cancer studies according to machine learning, we have presented a new generation convolutional neural network (CNN) termed ShortNeXt in this research. The presented ShortNeXt has inspired by ResNet, ConvNeXt and MobileNet architectures to use the advantages these CNNs together. This model, which aims to extract robust feature map using convolution-based residual blocks, is named ShortNeXt because it incorporates more than one shortcut. The ShortNeXt architecture has four main stages and these stages are: (i) an input/stem, (ii) ShortNeXt, (iii) downsampling, and (iv) output. In this CNN architecture, convolution, batch normalization and the Gaussian Error Linear Unit (GELU) activation functions have been utilized. In this aspect, the implementation of the recommended ShortNeXt is simple. The stem stage uses a 4×4 sized convolution with stride 4 like ConvNeXt and Swin Transformer and this operation is named patchify operation. Additionally, a 2×2 patchify block has been used in the downsampling block. In the ShortNeXt block, an inverted bottleneck has been used, and both 1×1 and 3×3 convolution blocks are employed in the expansion phase. The output layer has increased the number of filters from 768 to 1280 by using pixel-wise convolution, drawing inspiration from MobileNetV2 and a final feature map with a length of 1280 has been obtained by deploying global average pooling (GAP). In the classification phase, fully connected and softmax operators have been used.

To get comparative results about to the recommended ShortNeXt, a publicly available histopathological image dataset has been used and this dataset contains nine classes, and the proposed ShortNeXt has achieved 97.82% and 97.86% validation and test accuracy, respectively. The obtained results and findings openly showcases that ShortNeXt is an effective deep learning method for histopathological image classification for cancer detection/classification.

1. Introduction

Histopathology is an important medical field that looks at tissue samples under a microscope to figure out what diseases a person might have (Jing et al., 2023; Mulrane et al., 2008). Examining tissues closely through a microscope plays a key role in spotting different health problems, especially when it comes to finding and understanding cancer (Gurcan et al., 2009). New digital tools have enabled more detailed examination of tissues, allowing for the analysis of hidden patterns that are not distinguishable or noticeable to the eye (Mungenast et al., 2021). It is known that these advanced capabilities of digital

* Corresponding author.
E-mail addresses: Prabal.Barua@usq.edu.au (P.D. Barua), btasci@firat.edu.tr (B. Tasci), mehmet.baygin@erzurum.edu.tr (M. Baygin), sdogan@firat.edu.tr (S. Dogan), turkertuncer@firat.edu.tr (T. Tuncer), filippo.molinari@polito.it (F. Molinari), massimo.salvi@polito.it (S. Massimo), Rajendra.Acharya@usq.edu.au (U.R. Acharya).

https://doi.org/10.1016/j.cviu.2026.104672
Received 16 September 2024; Received in revised form 16 December 2025; Accepted 25 January 2026
Available online 4 February 2026
1077-3142/© 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

tools have made it easier for doctors in the diagnosis process. In this process, microscopic images are often used. The microscopic images are digitized and input into the system, where they are labeled. The system can detect abnormal cell patterns identified in the images and generate information related to these cells (Linkon et al., 2021). Sorting and categorizing abnormal and normal tissue images is an important process, especially for cancer diagnosis. Standing apart from this, the findings obtained have served as essential guide for medical doctors in application of personalized medication to benefit each patient (Fu et al., 2023). It is a true fact that in the current day, accompanied by the advancement in area of artificial intelligence, computer systems possess the ability to learn autonomously from digital data records and they can create roadmaps for the cancer diagnosis and also for the treatment based on all information they have learned and processed (Yao et al., 2023; Salvi et al., 2021). Computer systems are capable of the examination and interpretation of complex details contained inside tissue images by use of a variety of mathematical formulas and certain learning methods (Elhadary et al., 2023; Naik et al., 2023). These advanced systems work to enhance the precision with which diseases are identified and they are successful in streamlining the whole process of identification. Techniques which take inspiration from the workings of the human brain now maintain the ability for detection of subtle details in the tissue samples that can potentially escape attention of doctor or medical expert (Badawy et al., 2023). The merging of histopathology discipline with machine learning capability marks a very notable shift toward creation of smarter and more effective healthcare solutions about the matter of patient well-being. This is clearly a path for future students and professionals in the field (Piga et al., 2023; Salvi et al., 2020).

The colorectal cancer, abbreviated as CRC, is maintaining a focus of the research because of its high incidence rate and serious consequences for health and the difficulties found in effective diagnosis and treatment of this disease (Wang et al., 2019; Fan et al., 2021). The availability of a large amount of histopathological data concerning CRC stands as a valuable resource for deep learning algorithms. Such algorithms can conduct analysis of digital images and so assist the pathologists in the detection of cancer cells and the relevant lesions. Standing apart from this, Qi et al. (2021) performed a utilization of deep learning to find prognostic spatial features in the CRC microenvironment, specifically employing histopathology images for this purpose. The study of these researchers is highlighting the importance of spatial features in the prognosis of CRC and is considered a further step in the comprehensive understanding of the tumor microenvironment about the matter. The model they utilized, which was based on the VGG19 architecture, achieved a notable success rate of 99.0% during internal validation and a success rate of 95.0% when applied to independent test sets. Another point is concerning the work by Jiao et al. (2021). This team employed the deep learning methods to investigate tumor microenvironment in colon adenocarcinoma. The goal of this kind of approach is focused on the ability to predict the progression of the disease and its subsequent status for patients. The investigation of spatial features is very necessary for the proper assessment of patient status. The results obtained reveal that lymphocyte components in the tumor microenvironment are associated with poor PFI. Li et al. (2023) developed a new deep learning model to predict CRC survival. In this research, two different testing procedures were performed by including only histopathologic features and tissue area features. Consequent to the test procedures, a significant observation was made concerning performance improvement when utilizing the combination of histopathologic and tissue area features. Park et al. (2022) compared two different deep learning methodologies for purpose of predicting microsatellite instability in CRC. This research study maintained the main aim of identifying correlation between genetic abnormalities and CRC. Prezja et al. (2023) dedicated their work toward the improving of existing deep learning architectures that are available in the literature, which enables better classification accuracy in CRC detection. For this specific purpose, new model training strategy was developed. The result showed 95.6% of classification success on the dataset which was used in the study. Kallipolitis et al. (2021) investigated the usefulness of aggregating EfficientNets for the classification of histopathology images. Their developed approach demonstrates the effectiveness of ensemble methods when improving classification accuracy. Accuracy and area under curve metrics for the colon cancer dataset showed high performance across various EfficientNet models, with accuracies ranging from 95.64% and 99.46% and AUC values ranging from 99.13% to 99.93%.

Velastegui and Pedersen (2021) investigated about the impact to different color spaces when classifying histological images by employment of CNNs. Their research gives contribution to the understanding of how the processing of color affects models of deep learning inside medical imaging. In their study, it is a true fact that highest overall classification accuracy, which was 93.0%, was achieved using the CIELAB color space structure. The RGB and the XYZ color spaces provided an accuracy of 92.50%, which is slightly less. Another point is concerning Khazaee Fadafen and Rezaee (2023), who developed approach for multi-tissue classification based on an ensemble method using novel hybrid deep learning framework. Their research offers important insights into the effectiveness and capability of hybrid models when performing complex tissue classification tasks. The experiments showed that the correct classification of CRC cases could be done in the range of 98.82% to 99.76% using their research methodology. Standing apart from this, Lv et al. (2022) presented Transurv, which is a transformer-based model that integrates histopathological images and genomic data for survival analysis about CRC. This approach of interdisciplinary nature stands as an example of potential for combining various data types for comprehensive analysis of cancer. Their results showed that the combination of multi-scale features from histopathology and the cross-modal transformer fusion offered the best performance level, achieving a concordance index with value of 0.822. Tsai and Tao (2021) explored various DL techniques for the classifying of CRC tissues. Their study adds to growing body of research about applying DL in histopathological image analysis. The overall nine-class accuracy that was achieved with this approach was close to 99.0% for internal testing set and 94.30% for external testing set. The application of ResNet50 model achieved a significant 99.69% accuracy rate in the same internal testing set and 99.32% in the external testing set. Standing apart from this, when applied to independent dataset with eight classes, the ResNet50 model demonstrated a 94.86% accuracy rate. This rate is significantly higher than the 87.4% best accuracy rate reported by a previous study that used the same specific dataset. Tsai et al. (2023) demonstrated about the matter of how histopathology images could enable prediction of multi-omics aberrations and prognoses in CRC patients. This study gives focus to the potential of histopathological images as a kind of proxy for more comprehensive molecular analyses. Kumar et al. (2023) introduced CRCCN-Net, an automated framework for classification of CRC tissue using the histopathological images. Their study contributes to the automation and efficiency in the diagnosis of CRC. They obtained classification accuracy of 96.26% for the NCT-CRC-HE-100K dataset, which is high performance. Yan et al. (2022) developed a method based on deep contrastive learning for tissue clustering in analysis of histopathology images without annotation requirement. Their research addresses the challenges of manual annotations in the medical imaging field, achieving an accuracy of 94.94%.

1.1. Literature gaps and motivations of our model

Histopathological image classification is one of the methods that is frequently used today, especially for the automatic diagnosis and the detection of cancer. Steps are being taken to develop diagnostic systems that are fast and reliable using the power of computer vision in this field (Elazab et al., 2020; J. et al., 2021). Although models based on deep learning have made significant progress in detection

Table 1
Correspondence between identified challenges and proposed solutions in ShortNeXt.

Challenge	Proposed solution	ShortNeXt module
(1) Computational-accuracy trade-off: High-performing models require excessive parameters (25-138M)	Lightweight architecture with only 8.7M parameters achieving 97.86% accuracy	Overall ShortNeXt architecture with efficient filter progression (96→192→384→768→1280)
(2) Performance-efficiency imbalance in lightweight models	Multiple shortcut connections enabling effective feature learning without parameter overhead	ShortNeXt block with three shortcut connections (main residual, intermediate, expansion shortcuts)
(3) Class-specific performance gaps in critical tissue categories	Multi-scale feature extraction capturing both fine cellular details and broader tissue patterns	Parallel 1 × 1 and 3 × 3 convolution blocks in the expansion phase
(4) Limited external validation and generalizability concerns	Comprehensive validation on independent patient populations	Validated on both NCT-CRC-HE-100K (97.86%) and external CRC-VAL-HE-7K dataset (98.21%)
(5) Real-time clinical applicability requirements	High throughput processing capability with balanced accuracy-efficiency trade-off	1.4 ms inference time, 714 images/second throughput

of cancer (Diao et al., 2023), this area still remains open for development further (Padhi et al., 2023) because challenges exist. Despite these advances, existing methods face several concrete challenges: (1) high-performing models like VGG16 (138.4M parameters) and ResNet-50 (25.6M parameters) achieve accuracies of 94%–96% but require substantial computational resources, limiting deployment in resource-constrained clinical environments; (2) lightweight alternatives such as MobileNetV2 sacrifice classification performance, achieving only 93.23% accuracy; (3) class-specific performance gaps persist, with clinically critical categories like cancer-associated stroma consistently showing lower classification rates; (4) many studies report high internal accuracies but lack validation on independent patient populations, raising generalizability concerns. To provide address to all of these needs, the developments in relevant literature have been followed and a new model is proposed, which draws inspiration from ConvNeXt and ResNet and MobileNetV2 architecture. We extracted features which are considered more robust and more meaningful by the utilization of additional shortcuts functionality. Furthermore, this model has been developed for being lightweight, ensuring the possibility of its easy integration to devices such as mobile telephones and tablets.

Our examination of existing Convolutional Neural Network architectures reveals distinct trade-offs in their design approaches. ResNet successfully enables the training of deep networks through its residual connections, but remains computationally heavy with a large parameter count. ConvNeXt effectively adapts transformer-based design principles to CNNs, but lacks specific optimization for biomedical image classification tasks. MobileNet, while achieving excellent efficiency through its lightweight structure, shows limited capability in capturing the complex patterns characteristic of histopathological tissues.

This balance between computational efficiency and accuracy becomes particularly crucial in clinical settings where analysis extends to Whole Slide Images (WSIs), which typically have dimensions of 100,000 × 100,000 pixels or larger. When processing such large-scale images, even a small increase in per-patch processing time can result in significant delays in whole-slide analysis, potentially impacting clinical workflows where rapid analysis of multiple WSIs is often required. Our primary motivation in conducting this study is the strategic combining of strengths from the aforementioned architectures to develop a model specifically designed for histopathological image classification.

1.2. Novelties and contributions

The fundamental innovation of ShortNeXt lies in its triple-shortcut residual learning mechanism—a design philosophy not explored in existing CNN architectures. Unlike ResNet's single residual connection, ShortNeXt introduces three complementary shortcuts operating at different abstraction levels within a single block, systematically enhancing gradient flow and feature learning while maintaining computational efficiency. The technical innovations of ShortNeXt can be summarized as follows:

- Triple-shortcut block design: Three strategically positioned shortcut connections (main residual, intermediate, and expansion shortcuts) within each block improve gradient flow and enable more effective feature learning compared to conventional single-shortcut designs.
- Parallel multi-scale feature extraction: The simultaneous use of 1 × 1 (point-wise) and 3 × 3 (spatial) convolutions during the expansion phase enables the capture of both fine cellular details and broader tissue patterns in histopathological images.
- Optimized hybrid architecture: ConvNeXt's patchify strategy (4 × 4 stem convolution with stride 4) and MobileNet's efficient output block design (1 × 1 expansion to 1280 channels with GAP) are strategically integrated for histopathological image analysis.

As a result, ShortNeXt is not simply a combination of existing architectures, but rather an integrated design solution where each component serves a specific purpose, optimized for the unique challenges of large-scale histopathological image analysis.

The main contributions of our paper are:

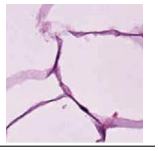
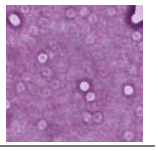
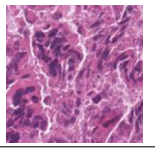
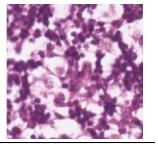
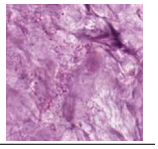
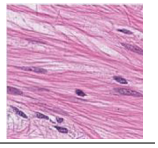
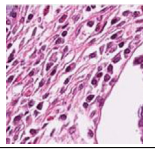
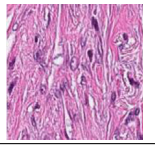
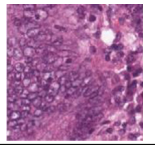
- A lightweight CNN model achieving state-of-the-art accuracy (>97%) with only 8.7 million parameters—approximately one-third of ResNet-50 (25.6M)—enabling efficient processing of large-scale WSIs in clinical settings.
- Comprehensive validation demonstrating robust generalizability: 97.86% on internal test set (NCT-CRC-HE-100K) and 98.21% on external validation set (CRC-VAL-HE-7K) from independent patient populations.
- Interpretable diagnostic support through Grad-CAM visualization, enhancing transparency and clinical applicability of the model's decision-making process.

To explicitly demonstrate how each proposed contribution addresses the specific challenges identified in the literature, Table 1 presents the correspondence between the identified problems and ShortNeXt's targeted solutions.

2. Dataset

In this research, we have used a publicly available cancer image dataset and this dataset contains 100,000 images with nine distinct classes (Kather et al., 2018). These classes include (i) debris, (ii) lymphocytes, (iii) colorectal adenocarcinoma epithelium, (iv) mucus, (v) adipose, (vi) smooth muscle, (vii) normal colon mucosa, (viii) background, and (ix) cancer-associated stroma. The dataset was systematically divided into training and test subsets. Each image in the dataset has a resolution of 224 × 224 pixels and is represented in a three-channel color format (RGB). The uniformity in image size ensures consistency in input data for the DL model, improving its ability to learn and generalize from the dataset. The attributes and characteristics of this dataset have been given in Table 2.

Table 2
Overview of the used histopathological image dataset.

No	Class name	Training	Test	Total	Sample image
1	Adipose	7807	2600	10,407	
2	Background	7926	2640	10,566	
3	Debris	8636	2876	11,512	
4	Lymphocytes	8669	2888	11,557	
5	Mucus	6672	2224	8896	
6	Smooth muscle	10,152	3384	13,536	
7	Normal colon mucosa	6575	2188	8763	
8	Cancer-associated stroma	7836	2610	10,446	
9	Colorectal adenocarcinoma epithelium	10,740	3577	14,317	
Total		75,013	24,987	100,000	

3. The proposed ShortNeXt

The main objective of this research is to present the recommended ShortNeXt CNN model and this CNN model is a lightweight and high accurate CNN model. In order to propose this model, we have inspired by the ConvNeXt CNNs. The proposed ShortNeXt contains four main phases and these are: (i) the stem/input block, (ii) the main block, (iii) the downsampling block, and (iv) the output block.

In the stem block, a patchify block like ConvNeXt has been utilized and in this phase, we have used a 4×4 sized convolution with stride 4 to create $56 \times 56 \times 96$ sized tensor from $224 \times 224 \times 3$ sized image.

In the main ShortNeXt block, we have inspired by the ConvNeXt block and we have used four convolution operators. The essential objective of this block is to create a robust feature maps. In this stage, a depth-wise convolution with a kernel of 3×3 has been used and we have used convolution based residuals. After that 3×3 and 1×1 convolutions have been used to increasing number of filters four times. In the scaling phase, a 1×1 convolution block has been used. The multiple shortcut connections in this block inspired the name "ShortNeXt". The ShortNeXt block architecture implements three distinct shortcut connections, each serving a specific function:

Shortcut 1 (Main Residual Connection): This is the main shortcut connection extending directly from the input of the block to the output of the block. This connection operates similarly to the classical residual connection in ResNet. The input tensor is added directly to the output

Table 3
Shortcut connection details in the ShortNeXt block.

Shortcut no	Starting point	Ending point	Connection type	Weights	BN/activation
1	Block input	Block output	Identity	None	None
2	DWConv output	Before contraction	Identity	None	None
3	1 × 1 Conv output + 3 × 3 Conv output	Merge	Element-wise sum	Yes (Conv weights)	Yes (BN+GELU per branch)

without any transformation (identity mapping). No weights, batch normalization, or activation functions are used in this connection.

Shortcut 2 (Intermediate Shortcut Connection): This connection extends from the output of the 7 × 7 depthwise convolution layer to the output of the expansion phase. This connection is also implemented as identity mapping and contains no additional parameters.

Shortcut 3 (Expansion Shortcut): The outputs of the 1 × 1 and 3 × 3 convolution layers applied in parallel during the expansion phase are summed element-wise. Both branches have their own batch normalization and GELU activation functions.

The mathematical formulation of the ShortNeXt block can be defined as follows:

Let the input tensor be x. First, the 7 × 7 depthwise convolution is applied:

$$x_1 = \text{GELU}\left(BN(DWConv_{7x7}(x))\right) \tag{1}$$

Then, parallel convolutions are applied in the expansion phase:

$$x_2 = \text{GELU}\left(BN(Conv_{1x1}(x_1))\right) + \text{GELU}\left(BN(Conv_{3x3}(x_1))\right) \tag{2}$$

In the contraction phase, the number of channels is restored to the original value using 1 × 1 convolution:

$$x_3 = BN\left(Conv_{1x1}\left(x_2\right)\right) \tag{3}$$

Finally, the output is obtained by adding the shortcut connections:

$$Output = x_3 + x_1 + x \tag{4}$$

Here, "$+x_1$" represents the second shortcut, and "$+x$" represents the first (main) shortcut. Table 3 summarizes the characteristics of the shortcut connections in the ShortNeXt block.

Having explained the core ShortNeXt block in detail, we now describe the remaining components of the architecture. In the Downsampling block, a patchify block has been used to decrease the width and height of the tensor and increase depth of this tensor.

The concluding (output) section of ShortNeXt, the output block, inspired by MobileNet (Sandler et al., 2018). We have used a pixel-wise convolution operator to increase number of features from 768 to 1280. After that, the features have been obtained deploying global average pooling (GAP) operator. By deploying fully connected (FC) layer, number of the outputs have been determined and softmax generates the classification outputs.

Fig. 1 demonstrates the general block diagram of the recommended ShortNeXt for better clarification with details of the phases.

To understand the sequential operations of the proposed ShortNeXt, the transition table of this CNN model is also provided in Table 4.

ShortNeXt has approximately 8.7 million learnable parameters and this number of parameters clearly demonstrated that the presented CNN is a lightweight CNN model. The overall mathematical definition of the recommended ShortNeXt is also demonstrated below.

$$ShortNeXt: F = (96, 192, 384, 768, 1280), R = (2, 2, 4, 2, 1) \tag{5}$$

where F: number of filters and R: number of repetitions.

4. Experimental results

In order to implement the recommended model, we have used a personal workstation (PW) with 128 GB of main memory, a 3 GHz processor, and a GPU with 3072 CUDA cores optimized for deep learning tasks. In this PW, the used operating system is Windows 11.

We developed the proposed ShortNeXt by deploying MATLAB (version 2023a) programming environment and we used deep network designer in this model. To create the presented ShortNeXt, 170 operators and 199 connections have been utilized.

To train the presented SohrtNeXt, the utilized key hyperparameters are: initial learning rate of 0.01 and the use of the Stochastic Gradient Descent with Momentum (sgdm) solver. The training was conducted over 100 epochs with a mini-batch size of 128. Furthermore, an L2 Regularization factor of 10^{-4} was applied to prevent overfitting and ensure the model's generalizability. The dataset was split into training and validation sets in a 70:30 ratio, allowing for thorough training while reserving a substantial portion of data for validation purposes. The effectiveness of the training process and the model's learning progression were systematically tracked and illustrated in Fig. 2.

After the training phase, the proposed model achieved a training accuracy of 100%, a validation accuracy of 97.82%, and a final validation loss of 0.0892. These results and Fig. 2 demonstrate that the proposed model did not overfit. Data augmentation techniques such as flipping, rotation, color jitter, and elastic deformation were intentionally omitted during the training process for several reasons. First, the dataset contains 100,000 images with sufficient natural variation for histopathological image classification. Second, achieving high classification performance without data augmentation directly demonstrates the intrinsic learning capacity and generalization ability of ShortNeXt. Third, the training and validation curves shown in Fig. 2 indicate no overfitting, confirming robust model generalization even without augmentation. Finally, data augmentation techniques might introduce artificial variations that could cause the model to deviate from real clinical data distributions. Our objective was to ensure the model learns authentic histopathological image characteristics. Additionally, we calculated the test results, and the computed test confusion matrix is depicted in Fig. 3.

To comprehensively evaluate the classification performance of the proposed ShortNeXt model, additional evaluation metrics commonly used in medical image analysis were calculated alongside the standard metrics. Table 5 presents the detailed performance results of the ShortNeXt model on the test set.

As shown in Table 5, our proposed ShortNeXt model achieved over 97% in overall classification performance for the test dataset. Cohen's Kappa coefficient measures how well the classifier performs compared to random chance. The Kappa value calculated for the ShortNeXt model is 0.9759, which corresponds to the "almost perfect agreement" category. The Matthews Correlation Coefficient (MCC), which is considered a reliable performance indicator even for imbalanced datasets, was calculated as 0.9759 for ShortNeXt. Additionally, the macro-averaged specificity was calculated as 99.73%, indicating that the model exhibits excellent performance in correctly identifying negative samples across all classes. These comprehensive evaluation metrics demonstrate the robust and reliable classification capability of the proposed ShortNeXt model. Table 6 lists the class-wise recall, precision, F1-scores and specificity achieved on the test set, where the highest metrics are highlighted in bold.

The highest sensitivity score was also observed in Adipose class, which reached 99.77% class-wise sensitivity. It is a true fact that model also showed an exemplary classification performance in the Adipose class exceeding 99%. Standing apart from this, the Background class achieved 99.85% recall, which is a class-wise classification accuracy, and 99.74% F1 score. This shows stability of the ShortNeXt approach

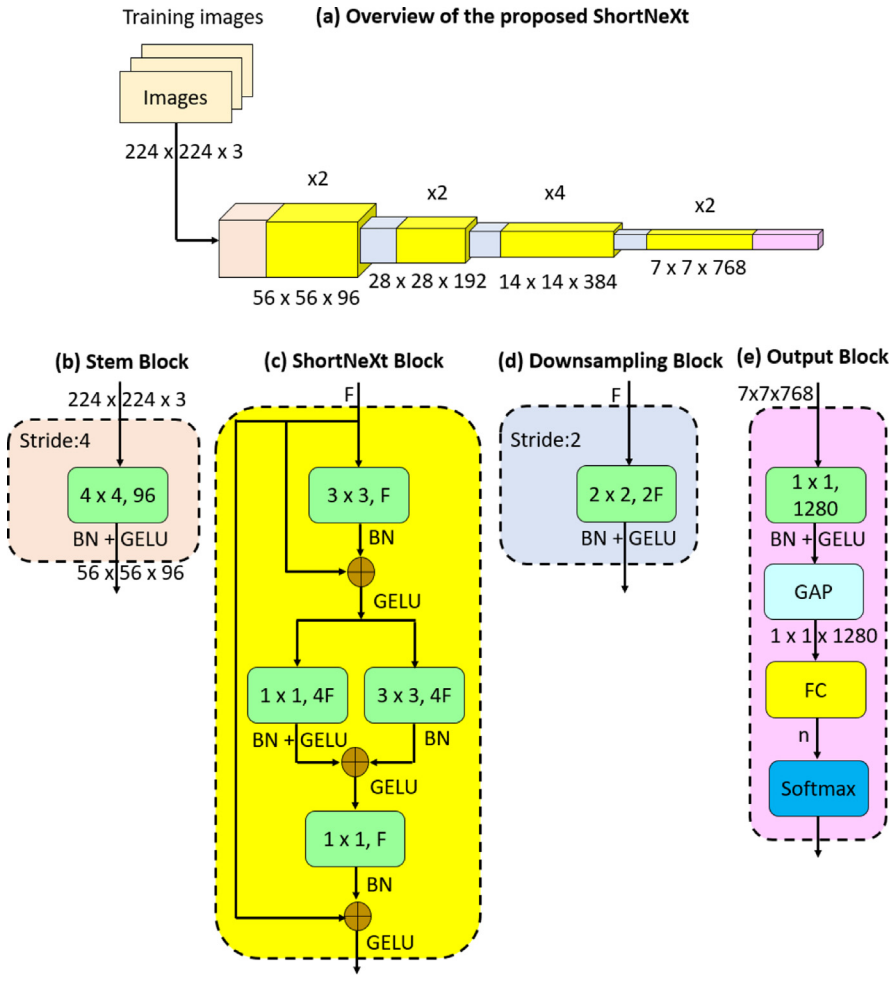
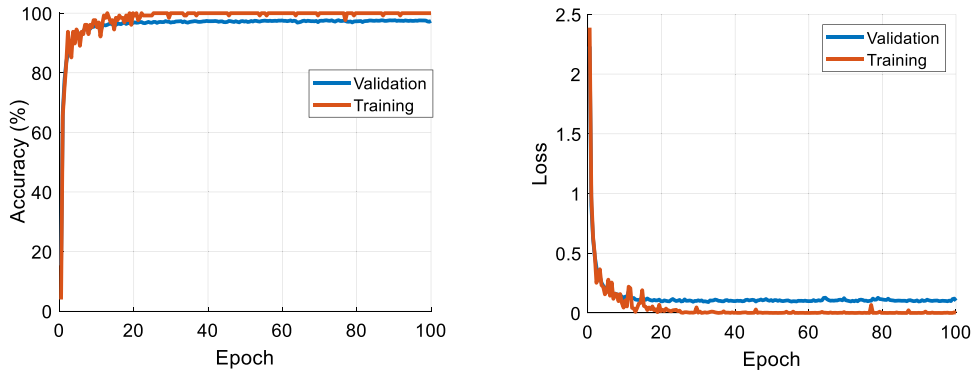


Fig. 1. Overview of the presented ShortNeXt. Herein, n: number of classes, BN: batch normalization, GELU: Gaussian Error Linear Unit, F: number of filters, FC: fully connected.

Table 4
Overview of the proposed ShortNeXt.

Layer	Input	Operation	Output
Stem/Input	224 × 224 × 3	4 × 4, 96, BN + GELU, stride: 4	56 × 56 × 96
ShortNeXt 1	56 × 56 × 96	$\begin{bmatrix} 7\times7, 96 \\ (1\times1, 384)+(3\times3384) \\ 1\times1, 96 \end{bmatrix}\times2$	56 × 56 × 96
Downsampling 1	56 × 56 × 96	2 × 2, 192, BN + GELU, stride: 2	28 × 28 × 192
ShortNeXt 2	28 × 28 × 192	$\begin{bmatrix} 7\times7, 192 \\ (1\times1, 768)+(3\times3768) \\ 1\times1, 96 \end{bmatrix}\times2$	28 × 28 × 192
Downsampling 2	28 × 28 × 192	2 × 2, 384, BN + GELU, stride: 2	14 × 14 × 384
ShortNeXt 3	14 × 14 × 384	$\begin{bmatrix} 7\times7, 384 \\ (1\times1, 1536)+(3\times3, 1536) \\ 1\times1384 \end{bmatrix}\times4$	14 × 14 × 384
Downsampling 3	14 × 14 × 384	2 × 2, 768, BN + GELU, stride: 2	7 × 7 × 768
ShortNeXt 4	7 × 7 × 768	$\begin{bmatrix} 7\times7, 768 \\ (1\times1, 3072)+(3\times3, 3072) \\ 1\times1768 \end{bmatrix}\times2$	7 × 7 × 768
Output block	7 × 7 × 768	1 × 1, 1280, BN + GELU, global average pooling, fully connected layer, softmax, classification	Outcomes
Total learnable parameters			8.7 Million



(a) Accuracy (b) Loss values

Fig. 2. Training and validation curves of the proposed ShortNeXt on the used 100K histopathological image dataset.

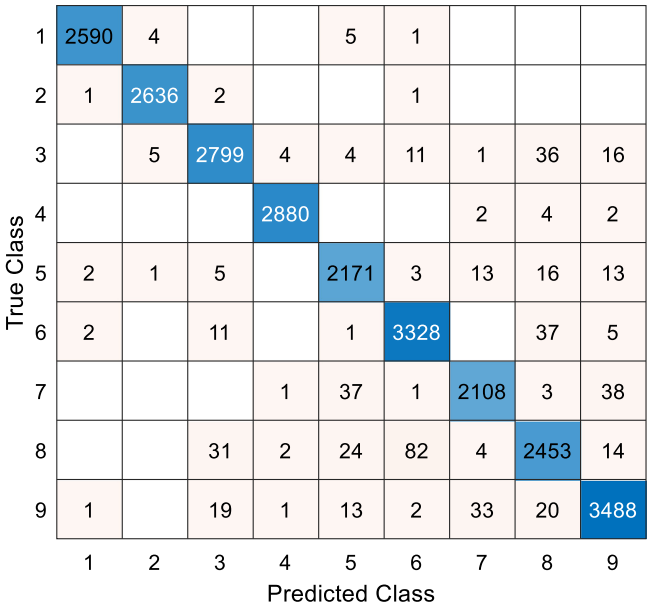


Fig. 3. Test confusion matrix of the proposed ShortNeXt. Herein, 1: Adipose, 2: Background, 3: Debris, 4: Lymphocytes, 5: Mucus, 6: Smooth muscle, 7: Normal colon mucosa, 8: Cancer-associated stroma, 9: Colorectal adenocarcinoma epithelium.

Table 5
Comprehensive evaluation metrics of ShortNeXt on the test set.

Performance metric	Overall result (%)
Accuracy	97.86
Recall (Sensitivity)	97.81
Precision	97.85
F1-score	97.83
Specificity (macro-avg)	99.73
Balanced Accuracy	97.81
Geometric Mean	97.80
Cohen's Kappa	97.59
Matthews Correlation Coefficient (MCC)	97.59

for specific classifications. In this aspect, the best detected classes by deploying ShortNeXt are Adipose and Background.

However, Cancer-associated stroma class achieved 93.98% in class-wise recall, 95.48% in precision, and 94.73% in F1-score. Table 6 clearly demonstrated that all classes attained over 93.5% class-wise performances. These results obviously demonstrated the robust and high classification performances of the recommended ShortNeXt CNN.

To comprehensively evaluate the computational efficiency of the proposed ShortNeXt model, metrics such as the number of parameters, computational complexity (FLOPs), and inference time were calculated. Experiments were conducted on an NVIDIA RTX 4090 GPU (16,384 CUDA cores, 24 GB GDDR6X memory).

The total floating-point operations for the ShortNeXt model were calculated as 2.1 GFLOPs. The average inference time for a single 224×224 image was measured as 1.4 ms. This value indicates that the model can process approximately 714 images per second. The obtained results are given in Table 7.

Table 7 compares the computational efficiency of ShortNeXt with other popular CNN architectures. The results show that ShortNeXt achieves higher accuracy with approximately 3 times fewer parameters and 2 times less computational complexity compared to ResNet-50. When compared to MobileNetV2, ShortNeXt requires more computation but offers significantly higher classification performance. These findings clearly demonstrate that ShortNeXt strikes an optimal balance between high accuracy and computational efficiency, making it suitable for real-time clinical diagnosis support systems.

5. Discussions

In order to demonstrate classification ability of ShortNeXt, the public histopathological image dataset was utilized in our experiments. Analytical observations demonstrate that ShortNeXt model not only obtains generally high performance in task of classification, but also exhibits superior performance in the categorizing specific classes like Adipose and Background and Lymphocytes. For these specific classes, the accuracy rate achieved was over 99%. Another point is concerning the comparison with available literature methods about this matter. Table 8 provides analysis comparison including the results of state-of-the-art methods and also the proposed model.

As it can be seen from Table 8, ShortNeXt model has attained high test accuracy rate of 97.86% compared to SOTA methods. Furthermore, the fact that we did not utilize any fine-tuning procedure to obtain the results must be evaluated as another demonstration of the success of the method.

To evaluate the generalization capability of the proposed ShortNeXt model on independent patient populations, the model was also tested on an external dataset. For this purpose, the CRC-VAL-HE-7K (Kather et al., 2019) external validation set, which originates from the same source as the NCT-CRC-HE-100K (Kather et al., 2018) dataset but consists of entirely different patients, was utilized. This dataset contains 7180 histopathological images obtained from 50 different colorectal adenocarcinoma patients and provides an ideal testing environment for assessing the model's performance on independent patient data. Table 9 presents a comparison of ShortNeXt with state-of-the-art methods on the CRC-VAL-HE-7K dataset.

Table 6
Class-wise evaluation metrics on the test set.

Class	Recall	Precision	F1-score	Specificity
Adipose	99.62%	99.77%	99.69%	**99.97%**
Background	**99.85%**	99.62%	**99.74%**	99.96%
Debris	97.32%	97.63%	97.48%	99.69%
Lymphocytes	99.72%	99.72%	99.72%	99.96%
Mucus	97.62%	96.27%	96.94%	99.63%
Smooth Muscle	98.35%	97.05%	97.70%	99.53%
Normal Colon Mucosa	96.34%	97.55%	96.94%	99.77%
Cancer-associated Stroma	93.98%	95.48%	94.73%	99.48%
Colorectal Adenocarcinoma Epithelium	97.51%	97.54%	97.53%	99.59%
Macro Average	**97.81%**	**97.85%**	**97.83%**	**99.73%**

Table 7
The comparison of computation efficiency.

Model	Parameter (M)	FLOPs (G)	Inference Time (ms)	Throughput (img/s)	Accuracy (%)
VGG16	138.4	15.5	3.8	263	94.30
ResNet-50	25.6	4.1	1.9	526	96.16
MobileNetV2	3.4	0.3	0.8	1250	93.23
ConvNeXt-Tiny	28.6	4.5	2.3	435	96.50
ShortNeXt (Ours)	8.7	2.1	1.4	714	97.86

Table 8
Comparative results.

Authors	Method	Accuracy (%)
Kather et al. (2019)	VGG19 CNN	94.30
Ghosh et al. (2021a)	Ensemble DL model	96.16
Liang et al. (2020)	Multi-scale feature fusion convolutional neural network (MFF-CNN) with shearlet transform	96.00
Kumar et al. (2023)	Cancer classification convolutional neural network (CRCCN-Net)	96.26
Wang et al. (2017)	Bilinear CNN	92.60
Shankar et al. (2023)	1D-CNN	96.80
Our model	ShortNeXt	97.86

Table 9
Comparison of ShortNeXt with state-of-the-art methods on the CRC-VAL-HE-7K external validation set.

Method	Type	Accuracy (%)
DenseNet based solution (Khvostikov et al., 2021)	CNN	92.9
CONCH (ViT-Base) (Ming et al., 2023)	Transformer	93.0
VGG19 based solution (Kather et al., 2019)	CNN	94.3
Inception-v3 based solution (Wang et al., 2021)	CNN	94.8
ResNet-50 based solution (Sun et al., 2023)	CNN	94.8
VGG16 based solution (Anju and Vimala, 2023)	CNN	95.3
iBOT (ViT-Large) (Filiot et al., 2023)	Transformer	95.8
DINO (ViT) (Kang et al., 2023)	Transformer	95.9
Ensemble of 4 models (Ghosh et al., 2021b)	Ensemble	96.16
Ensemble of 5 models (Kumar et al., 2023)	Ensemble	96.26
CTransPath (Swin Transformer) (Wang et al., 2022)	Transformer	96.52
DeepCMorph (Ignatov et al., 2024)	CNN	96.99
EfficientNet-B0 (Ignatov and Malivenko, 2024)	CNN	97.73
ShortNeXt (Ours)	CNN	98.21

The results demonstrate that ShortNeXt significantly outperforms transformer-based models, including ViT-Base, ViT-Large, and Swin Transformer. Specifically, ShortNeXt achieved 98.21% accuracy, surpassing CONCH (ViT-Base) (Ming et al., 2023) at 93.0%, iBOT (ViT-Large) (Filiot et al., 2023) at 95.8%, and CTransPath (Swin Transformer) (Wang et al., 2022) at 96.52%. Moreover, as a single model, ShortNeXt outperformed ensemble methods, exceeding both the Ensemble of 4 models (Ghosh et al., 2021b) (96.16%) and Ensemble of 5 models (Kumar et al., 2023) (96.26%), demonstrating that high performance can be achieved without relying on ensemble approaches.

These external validation results confirm ShortNeXt's consistent and high classification performance across both training data and independent patient populations, supporting its reliable applicability in real clinical scenarios.

Standing apart from the primary classification tasks, we applied the technique known as Gradient-Weighted Class Activation Mapping, often abbreviated as Grad-CAM, to the images for achievement of explainable results. Grad-CAM stands as a technique generating heatmaps. This highlighting is focusing on specific regions which were most effective for the model's predictions about the matter. The interpretable results were obtained from these heat maps and they are presented in Fig. 4 for a set of sample images categorized by their class type.

The ShortNeXt approach for adipose tissue images is focusing onto the white spaces for identification of distinguishing features. In the case of the other classifications, it is observed that the model's focus is toward the granular areas and the nuclear structures for the successful identification. Concerning the analysis related to background textures, the proposed ShortNeXt focuses attention on structures similar to noise and on the textural pixels for the purpose of identifying important features. This method enables the increasing of the classification accuracy and provides clarity in the decision making process concerning ShortNeXt. Another point is regarding the lymphocyte class, where the model maintains an emphasis on the cell clusters. Within the muscle tissue, ShortNeXt places its focus on stroma areas. Nevertheless, for the stroma associated with cancer, the necessary operation is the analysis of the cells and the surrounding stroma simultaneously. The capability to visualize these focused areas aids understanding of why prediction was made. This level of explainability stands as a vital consideration, especially within the medical field. It is a true fact that this process may provide insight for medical professionals about the model decision-making procedure and potentially assist pathologists in understanding and confirming the AI-supported diagnostic suggestions and it is also highlighting the transparency of model.

We have discussed the important points of the proposed ShortNeXt-based cancer detection model below:

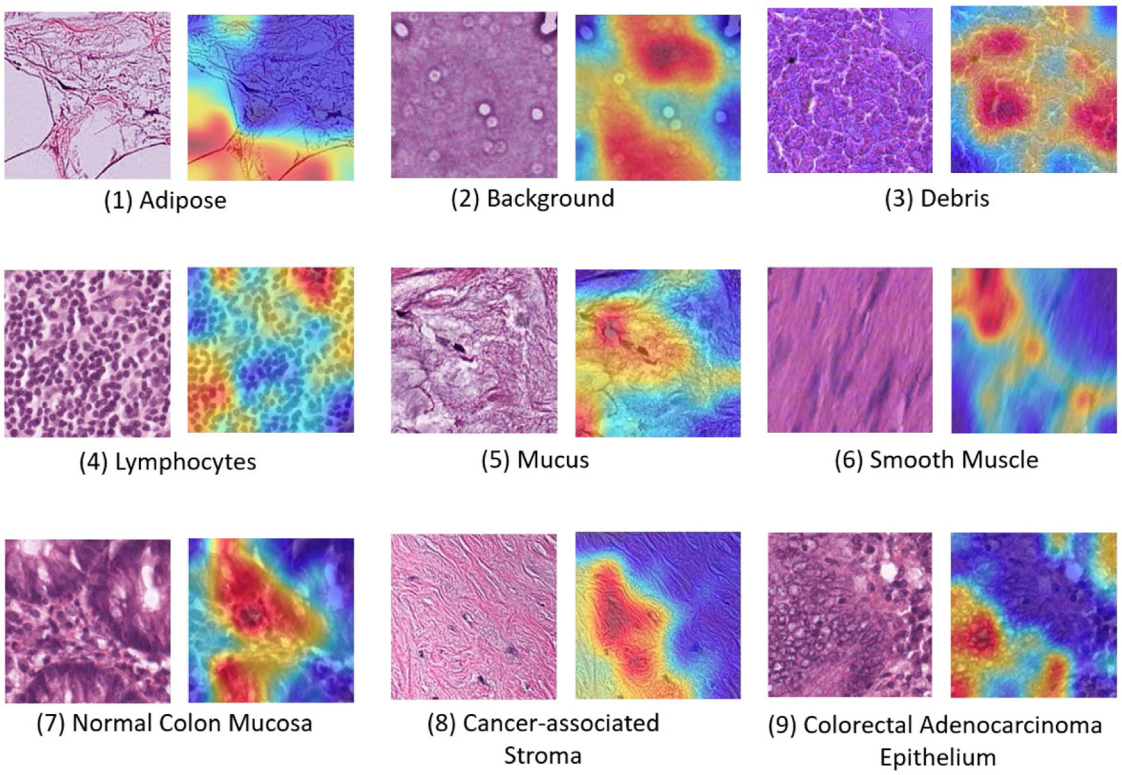


Fig. 4. Grad-CAM results for sample images by class.

- We have proposed a new CNN model (ShortNeXt) inspired by ConvNeXt and ResNet.
- The proposed ShortNeXt attained over 97% test and validation accuracies.
- This CNN is a lightweight CNN.
- Adipose, Background, and Lymphocytes classes, surpassing 99% accuracy. This indicates ShortNeXt's capability to handle diverse tissue types with high classification performances.
- The recommended ShortNeXt attained over 93.5% class-wise classification performances for all classes.
- By deploying Grad-CAM, we have presented interpretable results. These results clearly demonstrated the reliable results of the recommended ShortNeXt.

6. Limitations and future directions

Our work, while demonstrating promising results, presents several limitations. First, although we focused on biomedical image classification using one of the largest and most diverse public datasets in the literature, the model's performance could be further validated on broader datasets such as ImageNet. Second, ShortNeXt was designed and evaluated specifically for histopathological image classification, without investigating its adaptability to other computer vision tasks like object detection or segmentation. Additionally, our validation was limited to colorectal cancer histopathology images, leaving unexplored the model's performance on different imaging modalities and cancer types.

Future research directions could extend the validation of ShortNeXt to diverse image collections like CIFAR while exploring its adaptation for object detection tasks. This could be achieved by integrating the ShortNeXt backbone into established detection frameworks like YOLO or Faster R-CNN, potentially combining its multi-scale feature extraction capabilities with Feature Pyramid Network structures to detect pathological structures of varying sizes.

The application of ShortNeXt to different imaging modalities such as CT, MR, and X-ray represents another significant opportunity, though this would require careful consideration of modal-specific challenges. These include resolution differences between high-resolution histopathological images and typically lower-resolution CT/MR slices, the transition from RGB to single-channel grayscale information, and fundamental differences in texture characteristics between cellular and anatomical structures. Addressing these challenges would likely require specific input layer modifications and transfer learning strategies.

Furthermore, ShortNeXt shows potential as a diagnostic support tool in clinical workflows, and its principles could be extended to guide therapeutic strategies in precision medicine applications. These future directions highlight the potential broader impact of our work while acknowledging the necessary adaptations and validations required for different clinical applications.

7. Conclusions

In this research undertaking, for purpose of colorectal cancer diagnosis, a new technique named ShortNeXt was developed about the matter of classification of open-access histopathology images collected. The model which is developed maintains structure that fundamentally takes inspiration from both the ResNet and also the ConvNeXt architectures. Furthermore, it is included in this model multi-shortcut connections for the strengthening of feature extraction process even more. The ShortNeXt architecture which was engineered within the scope of this investigation has approximately 8.7 million parameters. When we make comparison of this architecture with methods available in the literature, it stands as lightweight at its core and it provides solution of high performance. In testing procedures that involved use of 100,000 images, the ShortNeXt model achieved an accuracy rate of 97.86%. Another point is concerning that the developed architecture exhibited a detection success over 99% for the Adipose and Background and Lymphocytes classes. The classification accuracies based on classes generally are above the 93.5%. For providing the explainability of

the model in this work that was executed, the Grad-CAM algorithm was employed. With this algorithm, heat maps were supplied for the regions that represent the highest important image areas in respect to classification within the model predictions. These heat maps give various clues toward the decision-making procedure of the ShortNeXt architecture, which is of a lightweight structure. The results that were obtained demonstrate that the ShortNeXt architecture is qualified to be used in real-time applications as a model that is accurate and explainable and also lightweight.

CRediT authorship contribution statement

Prabal Datta Barua: Writing – review & editing, Methodology, Data curation. **Burak Tasci:** Writing – review & editing, Visualization, Formal analysis. **Mehmet Baygin:** Writing – review & editing, Validation. **Sengul Dogan:** Writing – original draft, Visualization. **Turker Tuncer:** Writing – review & editing, Supervision. **Filippo Molinari:** Writing – review & editing, Supervision. **Salvi Massimo:** Writing – review & editing. **U. Rajendra Acharya:** Writing – review & editing, Supervision, Conceptualization.

Funding

This research received no external funding.

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.

Data availability

Data will be made available on request.

References

Anju, T.E., Vimala, S., 2023. Finetuned-VGG16 CNN model for tissue classification of colorectal cancer. Springer, pp. 73–84.
Badawy, M., Balaha, H.M., Maklad, A.S., Almars, A.M., Elhosseini, M.A., 2023. Revolutionizing oral cancer detection: An approach using aquila and gorilla algorithms optimized transfer learning-based CNNs. Biomimetics 8, 499.
Diao, S., Luo, W., Hou, J., Lambo, R., Al-Kuhali, H.A., Zhao, H., et al., 2023. Deep multi-magnification similarity learning for histopathological image classification. IEEE J. Biomed. Health Inform. 27, 1535–1545.
Elazab, N., Soliman, H., El-Sappagh, S., Islam, S.R., Elmogy, M., 2020. Objective diagnosis for histopathological images based on machine learning techniques: Classical approaches and new trends. Math. 8, 1863.
Elhadary, M., Elshoeibi, A.M., Badr, A., Elsayed, B., Metwally, O., Elshoeibi, A.M., et al., 2023. Revolutionizing chronic lymphocytic leukemia diagnosis: A deep dive into the diverse applications of machine learning. Blood Rev. 101134.
Fan, A., Wang, B., Wang, X., Nie, Y., Fan, D., Zhao, X., et al., 2021. Immunotherapy in colorectal cancer: current achievements and future perspective. Int. J. Biol. Sci. 17, 3837.
Filiot, A., Ghermi, R., Olivier, A., Jacob, P., Fidon, L., Camara, A., et al., 2023. Scaling self-supervised learning for histopathology with masked image modeling. MedRxiv. 2023-07.
Fu, X., Liu, S., Li, C., Sun, J., 2023. MCLNet: An multidimensional convolutional lightweight network for gastric histopathology image classification. Biomed. Signal Process. Control. 80, 104319.
Ghosh, S., Bandyopadhyay, A., Sahay, S., Ghosh, R., Kundu, I., Santosh, K., 2021a. Colorectal histology tumor detection using ensemble deep neural network. Eng. Appl. Artif. Intell. 100, 104202.
Ghosh, S., Bandyopadhyay, A., Sahay, S., Ghosh, R., Kundu, I., Santosh, K.C., 2021b. Colorectal histology tumor detection using ensemble deep neural network. Eng. Appl. Artif. Intell. 100, 104202.
Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B., 2009. Histopathological image analysis: A review. IEEE Rev. Biomed. Eng. 2, 147–171.
Ignatov, A., Malivenko, G., 2024. NCT-CRC-HE: Not All Histopathological Datasets are Equally Useful. Springer, 300-17.
Ignatov, A., Yates, J., Boeva, V., 2024. Histopathological image classification with cell morphology aware deep neural networks. pp. 6913–6925.
J., De Matos, Ataky, S.T.M., de Souza Britto Jr, A., Soares de Oliveira, L.E., Lameiras Koerich, A., 2021. Machine learning methods for histopathological image analysis: A review. Electron. 10, 562.
Jiao, Y., Li, J., Qian, C., Fei, S., 2021. Deep learning-based tumor microenvironment analysis in colon adenocarcinoma histopathological whole-slide images. Comput. Methods Programs Biomed. 204, 106047.
Jing, Y., Li, C., Du, T., Jiang, T., Sun, H., Yang, J., et al., 2023. A comprehensive survey of intestine histopathological image analysis using machine vision approaches. Comput. Biol. Med. 107388.
Kallipolitis, A., Revelos, K., Maglogiannis, I., 2021. Ensembling EfficientNets for the classification and interpretation of histopathology images. Algorithms 14, 278.
Kang, M., Song, H., Park, S., Yoo, D., Pereira, S., 2023. Benchmarking self-supervised learning on diverse pathology datasets. pp. 3344–3354.
Kather, J.N., Halama, N., Marx, A., 2018. 100,000 histological images of human colorectal cancer and healthy tissue. http://dx.doi.org/10.5281/zenodo.1214456.
Kather, J.N., Krisam, J., Charoentong, P., Luedde, T., Herpel, E., Weis, C.-A., et al., 2019. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. PLoS Med. 16, e1002730.
Khazaee Fadafen, M., Rezaee, K., 2023. Ensemble-based multi-tissue classification approach of colorectal cancer histology images using a novel hybrid deep learning framework. Sci. Rep. 13, 8823.
Khvostikov, A., Krylov, A., Mikhailov, I., Malkov, P., Danilova, N., 2021. Tissue type recognition in whole slide histological images. p. 50.
Kumar, A., Vishwakarma, A., Bajaj, V., 2023. Crccn-net: Automated framework for classification of colorectal tissue using histopathological images. Biomed. Signal Process. Control. 79, 104172.
Li, Y.-J., Chou, H.-H., Lin, P.-C., Shen, M.-R., Hsieh, S.-Y., 2023. A novel deep learning-based algorithm combining histopathological features with tissue areas to predict colorectal cancer survival from whole-slide images. J. Transl. Med. 21, 731.
Liang, M., Ren, Z., Yang, J., Feng, W., Li, B., 2020. Identification of colon cancer using multi-scale feature fusion convolutional neural network based on shearlet transform. IEEE Access. 8, 208969-77.
Linkon, A.H.M., Labib, M.M., Hasan, T., Hossain, M., 2021. Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: An extensive study. Inform. Med. Unlocked 24, 100582.
Lv, Z., Lin, Y., Yan, R., Wang, Y., Zhang, F., 2022. Transsurv: transformer-based survival analysis model integrating histopathological images and genomic data for colorectal cancer. IEEE/ACM Trans. Comput. Biol. Bioinform.
Ming, Y.L., Bowen, C., Drew, F.K.W., Richard, J.C., Liang, I., 2023. Towards a visual-language foundation model for computational pathol-ogy. arXiv preprint.
Mulrane, L., Rexhepaj, E., Penney, S., Callanan, J.J., Gallagher, W.M., 2008. Automated image analysis in histopathology: a valuable tool in medical diagnostics. Expert. Rev. Mol. Diagn. 8, 707–725.
Mungenast, F., Fernando, A., Nica, R., Boghiu, B., Lungu, B., Batra, J., et al., 2021. Next-generation digital histopathology of the tumor microenvironment. Genes 12, 538.
Naik, D.A., Mohana, R.M., Ramu, G., Lalitha, Y.S., SureshKumar, M., Raghavender, K., 2023. Analyzing histopathological images by using machine learning techniques. Appl. Nanosci. 13, 2507–2513.
Padhi, A., Agarwal, A., Saxena, S.K., Katoch, C., 2023. Transforming clinical virology with AI, machine learning and deep learning: a comprehensive review and outlook. pp. 1–11, VirusDisease.
Park, J., Chung, Y.R., Nose, A., 2022. Comparative analysis of high-and low-level deep learning approaches in microsatellite instability prediction. Sci. Rep. 12, 12218.
Piga, I., L'Imperio, V., Capitoli, G., Denti, V., Smith, A., Magni, F., et al., 2023. Paving the path toward multi-omics approaches in the diagnostic challenges faced in thyroid pathology. Expert. Rev. Proteom.
Prezja, F., Äyrämö, S., Pölönen, I., Ojala, T., Lahtinen, S., Ruusuvuori, P., et al., 2023. Improved accuracy in colorectal cancer tissue decomposition through refinement of established deep learning solutions. Sci. Rep. 13, 15879.
Qi, L., Ke, J., Yu, Z., Cao, Y., Lai, Y., Chen, Y., et al., 2021. Identification of prognostic spatial organization features in colorectal cancer microenvironment using deep learning on histopathology images. Med. Omics 2, 100008.
Salvi, M., Mogetta, A., Meiburger, K.M., Gambella, A., Molinaro, L., Barreca, A., et al., 2020. Karpinski score under digital investigation: A fully automated segmentation algorithm to identify vascular and stromal injury of donors' kidneys. Electron. 9, 1644.
Salvi, M., Molinari, F., Iussich, S., Muscatello, L.V., Pazzini, L., Benali, S., et al., 2021. Histopathological classification of canine cutaneous round cell tumors using deep learning: A multi-center study. Front. Vet. Sci. 8, 640944.
Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018. pp. 4510–4520.
Shankar, S.A., Shivakumar, M., Prakash, K., Koundinya, P.S., 2023. Occupancy grid map for a multi-robot system using LiDAR. SN Comput. Sci. 4, 196.
Sun, K., Chen, Y., Bai, B., Gao, Y., Xiao, J., Yu, G., 2023. Automatic classification of histopathology images across multiple cancers based on heterogeneous transfer learning. Diagn. 13, 1277.
Tsai, P.-C., Lee, T.-H., Kuo, K.-C., Su, F.-Y., Lee, T.-L.M., Marostica, E., et al., 2023. Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. Nat. Commun. 14, 2102.

Tsai, M.-J., Tao, Y.-H., 2021. Deep learning techniques for the classification of colorectal cancer tissue. Electron. 10, 1662.

Velastegui, R., Pedersen, M., 2021. The impact of using different color spaces in histological image classification using convolutional neural networks. In: 2021 9th European Workshop on Visual Information Processing. EUVIP, IEEE, pp. 1–6.

Wang, W., Kandimalla, R., Huang, H., Zhu, L., Li, Y., Gao, F., et al., 2019. Molecular Subtyping of Colorectal Cancer: Recent Progress, New Challenges and Emerging Opportunities. Elsevier, pp. 37–52, Seminars in cancer biology.

Wang, C., Shi, J., Zhang, Q., Ying, S., 2017. Histopathological image classification with bilinear convolutional neural networks. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBC, IEEE, pp. 4050–4053.

Wang, X., Yang, S., Zhang, J., Wang, M., Zhang, J., Yang, W., et al., 2022. Transformer-based unsupervised contrastive learning for histopathological image classification. Med. Image Anal. 81, 102559.

Wang, K.-S., Yu, G., Xu, C., Meng, X.-H., Zhou, J., Zheng, C., et al., 2021. Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence. BMC Med. 19, 76.

Yan, J., Chen, H., Li, X., Yao, J., 2022. Deep contrastive learning based tissue clustering for annotation-free histopathology image analysis. Comput. Med. Imaging Graph. 97, 102053.

Yao, J., Tran, S.N., Sawyer, S., Garg, S., 2023. Machine learning for leaf disease classification: data, techniques and applications. Artif. Intell. Rev. 1–46.