ATTENTION-BASED DEEP LEARNING FOR COLORECTAL CANCER CLASSIFICATION

Fareeda Naaz[1], Dr. I. Samuel Peter James[2], Ruqiya Fatima[3]
[1]PG Scholar, Department of AI&DS, Shadan Women's College of Engineering and Technology, Hyderabad, naazfareeda15@gmail.com
[2]Assoc Professor, Department of AI&DS, Shadan Women's College of Engineering and Technology drisamuelpeterjames@gmail.com
[3]Asst Professor, Department of CSE, Shadan Women's College of Engineering and Technology Ruqiyafatima2716@gmail.com

ABSTRACT
Colorectal cancer (CRC) remains one of the leading causes of cancer-related mortality worldwide, underscoring the critical need for early and accurate diagnosis. While histopathological examination of colorectal tissue is the clinical gold standard, it is labor-intensive and prone to human error. To overcome these challenges, this study proposes an attention-driven, lightweight deep learning model based on MobileNetV2 for the automated classification of CRC histopathology images. The architecture is designed to capture both global and local discriminative features efficiently, leveraging attention refinement modules to focus on diagnostically relevant regions. In addition to the proposed model, a comparative analysis is performed with a baseline Xception architecture, known for its high performance in image classification tasks. Experimental results demonstrate that the MobileNetV2-based attention model achieves comparable or superior accuracy while significantly reducing computational complexity, making it suitable for real-time clinical applications. This approach shows promise in assisting pathologists, reducing diagnostic workload, and enhancing the overall reliability of colorectal cancer detection.
Keywords: Colorectal Cancer (CRC), Histopathological Images, MobileNetV2, Attention-Driven Network, Deep Learning, Image Classification.

1. INTRODUCTION

Colorectal cancer (CRC) is one of the most prevalent and life-threatening malignancies globally, ranking among the top three causes of cancer-related deaths. Early and precise diagnosis is crucial for improving survival rates, as treatment outcomes are significantly better when cancer is detected at an early stage. The current gold standard for CRC diagnosis involves histopathological examination of tissue biopsy slides, which requires expert pathologists to manually analyze the morphology of tissue structures under a microscope. However, this manual process is time-consuming, subjective, and susceptible to intra-observer and inter-observer variability.

With recent advancements in artificial intelligence (AI) and deep learning, there has been a growing interest in automating medical image analysis, particularly in histopathology. Convolutional Neural Networks (CNNs) have shown tremendous success in capturing spatial patterns in images and enabling robust classification. Despite their success, many high-performing models like Xception and ResNet are computationally expensive and not feasible for real-time or resource-constrained environments such as small clinics or mobile diagnostic setups.

To address these challenges, this study proposes an attention-driven lightweight deep learning architecture based on MobileNetV2. MobileNetV2 is known for its efficiency and compact structure, making it suitable for deployment in low-resource settings without compromising performance. To further enhance its ability to detect subtle histological patterns, attention mechanisms are integrated to refine feature maps and focus the model's learning on critical regions of interest. This selective focus helps mimic the pathologist's attention during diagnosis.

The proposed model is evaluated and compared with the existing Xception architecture, which serves as a high-performing baseline. While Xception offers superior feature extraction capabilities due to its depthwise separable convolutions, it requires more memory and computational power. In contrast, the MobileNetV2-based model provides a balanced trade-off between accuracy and efficiency, demonstrating competitive classification performance with significantly reduced complexity.

2. EXISTING SYSTEM

Existing deep learning approaches for colorectal cancer (CRC) classification commonly employ the **Xception** architecture for analyzing histopathological images. Xception is an advanced convolutional neural network based on depthwise separable convolutions, enabling effective extraction of discriminative features while reducing the number of trainable parameters compared to conventional CNNs. Owing to its strong feature extraction capability, Xception has been widely adopted for automated medical image classification and computer-aided diagnosis of colorectal cancer.

Despite its effectiveness, Xception has several limitations when applied to CRC histopathological image classification. The model requires substantial computational resources and memory during training and inference, making deployment challenging in real-time and resource-constrained healthcare environments. Furthermore, it primarily emphasizes global feature extraction and may not effectively

capture fine-grained local tissue characteristics that are essential for accurate diagnosis. These limitations motivate the development of lightweight, attention-guided architectures that can achieve high classification accuracy with reduced computational complexity and faster inference.

LIMITATIONS OF EXISTING SYSTEMS

The existing Xception-based approach has several limitations that affect its practical deployment for colorectal cancer classification:

- **High Computational Cost:** The model requires significant computational power and memory during training and inference.
- **Risk of Overfitting:** Performance may degrade when trained on limited histopathological datasets due to overfitting.
- **Limited Local Feature Learning:** The architecture primarily focuses on global feature extraction and may not effectively capture subtle local tissue patterns.
- **Longer Inference Time:** The deep network increases prediction time, making real-time diagnosis more challenging.
- **Limited Suitability for Resource-Constrained Environments:** The large model size and high hardware requirements restrict deployment on mobile, edge, and low-resource healthcare devices.

3. RELATED WORK

This section provides a thorough overview and analysis of several research on CRC classification, with an emphasis on the use of DL approaches. The classification of CRC from Whole Slides Images (WSIs) has been greatly aided by DL approaches, particularly CNNs. While transfer learning and ensemble learning further improve diagnostic accuracy, Generative Adversarial Networks (GANs) provide a way to build deep representations without requiring large volumes of labelled training data. By utilising information from other relevant domains, transfer learning improves the performance of target models in particular areas.

Bychkov et al. developed a method for predicting colorectal cancer (CRC) using tumour tissue microarray samples stained with haematoxylin and eosin (H&E). The VGG-16 CNN model and a recurrent neural network were used in a transfer learning-based method. In order to assess several cutting-edge architectures for CRC classification, such as ResNet, Xception, NASNet, MobileNet, DenseNet, Inception, and VGG, Ohata et al. developed a transfer learning model. The study's best-performing model was determined to be the DenseNet-169 CNN.

In order to improve predictive performance, ensemble learning gathers features using a variety of transformations, applies several algorithms to produce weak predictions, and then combines these findings adaptively using voting systems.

Kallipolitis et al. investigated ensemble classifiers for the categorisation of breast and colon cancer histopathology images using cutting-edge CNN networks. EfficientNetB0–B7, XceptionNet, Inception-V3, VGG-16, and ResNet152-V2 were employed, while EfficientNetB1, B2, and B3 were further used for explainability by generating a guided Grad-CAM image. Tasdemir et al. devised a clinical decision support system for the classification of colorectal polyps. Using a variety of staining normalisation techniques, the ConvNeXt architecture was utilised in an ensemble framework. An ensemble model that combined ResNet-50, Inception-ResNet-V2, NasNet, EfficientNet-B0, and DenseNet-169 was tested by Karthik et al. In addition to image denoising approaches including Bilateral Filtering, Median Filtering, Gaussian Filtering, and Wavelet-based methods, they assessed other stain normalisation techniques like Vahadane, ITK, Reinhard, and Macenko. The best preprocessing method was found to be the combination of bilateral denoising and Macenko staining.

Abnormal regions in WSIs of CRC were categorised and localised by Gupta et al. Several cutting-edge networks, such as VGG-16, ResNet50, GoogleNet, Inception-V3, Inception-V4, and Inception-ResNet-V2, were compared in the study. Among these architectures, the VGG-16 model produced the best outcomes. Histopathology WSIs of the colon and stomach were categorised by Iizuka et al. Using a typical Inception-V3 CNN and a recurrent neural network (RNN), the work categorised the photos into three classes: adenocarcinoma, adenoma, and non-neoplastic. Sarwinda et al. employed ResNet-18 and ResNet-50 architectures to classify colon gland pictures into benign and malignant categories. Hu et al. experimented with binary classification of colon histopathology pictures using VGG16, ResNet-50, Inception-V3, and Vision Transformer (ViT). According to the study's findings, VGG16 was the most successful DL model.

A framework for identifying colorectal H&E slide images was presented by Tavolara et al. The suggested conditional GAN (cGAN), which trains two cGANs—one for the tumour regions and the other for the non-tumor regions—was compared to Inception-V3. Using a GAN-based framework, Sasmal et al. presented a semi-supervised classification model for learning from images of colorectal polyps.

To distinguish between normal and malignant colon samples, Bilal et al. used WSIs stained with H&E normalisation. The investigation of CRC mutations was done using a weakly supervised learning model. Yildirim et al. presented a CNN architecture to classify images of colon cancer into two binary classes: benign and adenocarcinoma. Eleven convolution layers and two fully connected layers make up the developed model. For the categorisation

of CRC H&E-stained histopathology pictures, Mohammed et al. presented a modified ConvMixer block that combines a ViT network and CNN architecture. The work outperforms ViT transformers and CNNs like Inception-V3, ResNet-50, and VGG16 by using channel attention to concentrate on the most informative aspects of the images. The ColonNet model was developed by Iqbal et al. to identify tumours in lung and colon histopathology pictures. The slide images were feature-engineered using depth-wise separable convolutions using a global–local pyramid pattern technique. Therefore, in order to improve the accuracy of CRC classification, a specialised architecture is necessary for extracting complicated information from histological colon pictures.

Even though DL for CRC classification has made great strides, there are still a number of unanswered research questions. The majority of current research relies on pre-trained models that aren't tailored for CRC detection, which limits their capacity to capture the distinctive histopathological characteristics required for accurate classification. Furthermore, local and semi-local patterns—which are crucial for identifying minute variations in histopathological images—are frequently overlooked by many of these methods, which primarily concentrate on global feature extraction. Moreover, existing models' capacity to highlight the most informative areas of the image is limited by the absence of efficient attention mechanisms that concurrently integrate spatial and channel attention. The necessity for an optimised architecture specifically designed for CRC classification is highlighted by these restrictions, which reduce the total discriminative capability of current models.

4. PROPOSED SYSTEM

MobileNetV2 is used as the primary feature extraction framework in the suggested approach for classifying colorectal cancer (CRC) from histopathology images. MobileNetV2 is a lightweight and effective convolutional neural network created especially for resource-constrained applications without sacrificing performance, in contrast to the heavier Xception model utilised in the current system. Its architecture makes use of inverted residual blocks and depthwise separable convolutions, which make it ideal for real-time medical picture analysis.

MobileNetV2 is integrated into the dual-track deep learning architecture to capture both local fine-grained features and global context of tissue samples. To improve diagnostic focus, the system integrates attention mechanisms that direct the network toward clinically relevant regions, improving both precision and interpretability. This combination of attention refinement and lightweight design reduces computational overhead while maintaining high accuracy, making the proposed system ideal for deployable and scalable solutions in medical imaging diagnostics, particularly for early and automated colorectal cancer detection.

Proposed System Advantages:

1. **Lightweight Architecture**
   The proposed system utilizes the MobileNetV2 backbone, which significantly reduces computational complexity, model size, and memory requirements while maintaining effective feature extraction capability.
2. **Improved Classification Accuracy**
   The integration of an attention mechanism enables the network to focus on diagnostically important tissue regions, resulting in improved feature representation and an overall classification accuracy of approximately **98%**.
3. **Efficient Feature Extraction**
   The proposed network effectively extracts both local tissue characteristics and global contextual information from histopathological images, improving the discrimination between different colorectal tissue classes.
4. **Reduced Computational Cost**
   Compared with conventional deep learning architectures such as Xception, the proposed system requires fewer computational resources, making it suitable for systems with limited hardware capabilities.
5. **Faster Inference**
   The lightweight architecture enables rapid processing and prediction of histopathological images, supporting timely clinical decision-making.
6. **Better Generalization**
   Image preprocessing and augmentation techniques improve the robustness of the model and reduce the possibility of overfitting when trained on medical image datasets.
7. **Scalability**
   The proposed system can be deployed in hospitals, diagnostic laboratories, cloud-based healthcare platforms, and portable medical devices due to its lightweight and efficient design.
8. **Computer-Aided Diagnosis**
   The proposed framework assists pathologists by providing reliable and automated colorectal cancer classification, thereby reducing diagnostic workload and supporting accurate clinical diagnosis.

5. MODULES

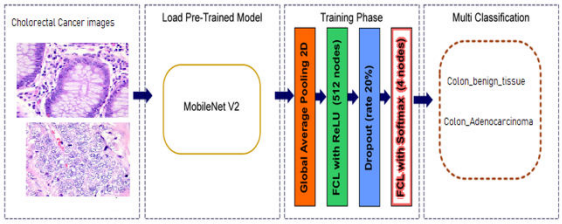
1. **Data Collection**
   Colorectal cancer histopathology image datasets from reputable sources, such NCT-CRC-HE-100K or private clinical datasets, are the main emphasis of this module. For reliable training, it guarantees that the data gathered includes a variety of cancer kinds and stages.
2. **Data Preprocessing**
   Image resizing, normalisation, stain normalisation, augmentation (rotation, flipping,

contrast adjustment), and patch creation are all included in this package. The dataset is prepared such that deep learning models can use it.

3. **Baseline Model Integration (Xception)**
   As a benchmark, the current Xception architecture is implemented in this module. It aids in highlighting enhancements by contrasting performance data with the suggested lightweight model.
4. **Proposed Model Construction (MobileNetV2 + Attention)**
   The core module integrates attention mechanisms (such as Squeeze-and-Excitation or Self-Attention) with the lightweight MobileNetV2 architecture. Model design, compilation, and optimisation are handled by this module.
5. **Training and Validation**
   Both models are trained on the preprocessed dataset using this module. It makes use of suitable batch sizes, optimisers, loss functions, and evaluation metrics such as accuracy, precision, recall, F1-score, and AUC.
6. **Performance Evaluation and Comparison**
   Here, graphs, confusion matrices, and performance tables are used to compare the outcomes of the suggested model with the baseline (Xception). It demonstrates increases in model accuracy, speed, and resource consumption.
7. **Visualization and Interpretability**
   To visualise model focus zones, this optional yet useful module incorporates tools like Grad-CAM or attention heatmaps. It makes the model more explainable, which is important in the medical field.

6. TECHNIQUES USED IN PROJECT

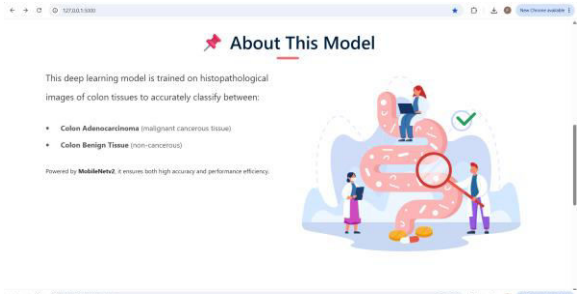
6.1. MOBILENETV2

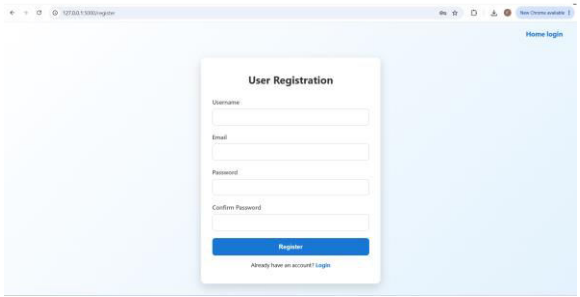


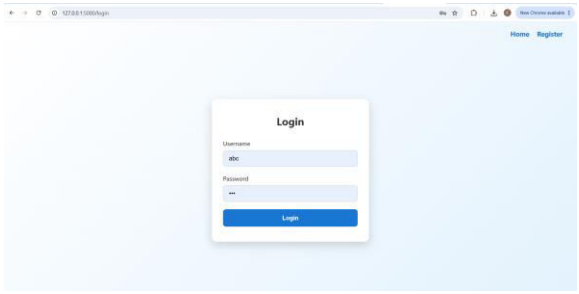
System Architecture

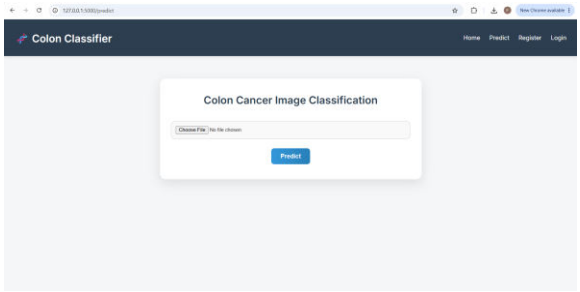
For real-time medical picture processing, the suggested system makes use of the MobileNetV2 architecture, a highly effective convolutional neural network that is optimised for speed and low computational cost. MobileNetV2 maintains good accuracy while drastically lowering the number of parameters and calculations by including inverted residuals and linear bottlenecks. For applications involving the categorisation of histopathology images, where deployment in real-world clinical settings is crucial, this lightweight design is particularly appropriate. An attention mechanism is incorporated into the architecture to further improve the diagnostic accuracy of the system. By automatically focusing on significant areas of the image that hold vital diagnostic information, this method enhances feature relevance and classification performance.

7. RESULTS









8. CONCLUSION

This project proposes an effective and attention-enhanced deep learning framework for the histopathological image-based categorisation of colorectal cancer (CRC). The model successfully captures both local and global tissue properties while retaining computational efficiency by combining attention mechanisms with the lightweight MobileNetV2 architecture. By emphasising diagnostically significant areas, this approach not only increases classification accuracy but also makes the data easier to understand. The suggested model exhibits a solid balance between performance and resource consumption when compared to heavier models like Xception, which makes it perfect for deployment in real-time, low-resource medical scenarios. The study demonstrates how AI-assisted diagnostic tools might lessen pathologists' workloads and enhance early cancer diagnosis. Although the findings are encouraging, this study also creates opportunities for further developments in clinical validation, scalability, and multi-modal integration. All things considered, this effort makes a substantial contribution to the development of deep learning-based colorectal cancer diagnostic systems that are easily accessible, comprehensible, and highly effective.

9. FUTURE ENHANCEMENT

The suggested attention-driven lightweight deep learning model has demonstrated encouraging results in the classification of colorectal cancer (CRC) using histopathological images; however, a number of improvements can be made in the future to further improve its performance, scalability, and clinical usability. First, the model is currently trained on static image patches; in the future, it can be extended to handle whole-slide images (WSIs) with region-based attention and tiling techniques to better mimic clinical diagnosis; integration of multi-modal data, such as merging histopathological images with clinical records and genomic information to create a more comprehensive diagnostic system; and finally, a self-supervised or semi-supervised learning approach to leverage large amounts of unlabelled data to reduce reliance on expensive manual annotations. The model's capacity to learn global context could be improved by adding vision transformers (ViT) or hybrid CNN-transformer architectures. Accessibility in remote and resource-constrained places would also be improved by implementing the model as a web or mobile application for real-time diagnosis. Explainability modules, like Grad-CAM or attention heatmaps, can be enhanced to produce visual interpretations that are easier for pathologists to understand. In order to distinguish between different stages or kinds of colorectal anomalies, the model can also be trained for multi-class classification. Transfer learning can be investigated for fine-tuning the model on similar diseases such as esophageal or stomach cancer in order to guarantee adaptability. The method can be made more scalable for embedded systems by using model optimisation techniques like pruning, quantisation, and edge AI deployment. In order to evaluate the model on a variety of real-world datasets, future study may possibly involve cooperation with clinical institutes. Lastly, active learning frameworks can be used, enabling the model to continuously improve its performance by interactively asking the pathologist to categorise challenging cases. Together, these improvements would bring this system closer to practical clinical integration and more influence in the field of AI-powered cancer diagnosis.

REFERENCES

[1] Y. Xi and P. Xu, ‘‘Global colorectal cancer burden in 2020 and projections to 2040,’’ Transl. Oncol., vol. 14, no. 10, Oct. 2021, Art. no. 101174, doi:10.1016/j.tranon.2021.101174.

[2] C. Mattiuzzi, F. Sanchis-Gomar, and G. Lippi, ‘‘Concise update on colorectal cancer epidemiology,’’ Ann. Transl. Med., vol. 7, no. 21, p. 609, Nov. 2019, doi: 10.21037/atm.2019.07.91.

[3] A. G. Russo, A. Andreano, A. Sartore-Bianchi, G. Mauri, A. Decarli,S. Siena, ‘‘Increased incidence of colon cancer among individualsyounger than 50 years: A 17 years analysis from the cancer registry of the municipality of Milan, Italy,’’ Cancer Epidemiology, vol. 60, pp. 134–140, Jun. 2019, doi: 10.1016/j.canep.2019.03.015.

[4] N. Murphy, V. Moreno, D. J. Hughes, L. Vodicka, P. Vodicka, E. K. Aglago, M. J. Gunter, and M. Jenab, ‘‘Lifestyle and dietary environmental factors in colorectal cancer susceptibility,’’ Mol. Aspects Med., vol. 69, pp. 2–9,Oct. 2019, doi: 10.1016/j.mam.2019.06.005.

[5] A. O. Adigun et al., ‘‘Causes and prevention of early-onset colorectal cancer,’’ Cureus, vol. 15, no. 9, 2023, Art. no. e45095, doi:10.7759/cureus.45095.

[6] W. Hamilton and S. E. R. Bailey, ‘‘Colorectal cancer in symptomatic patients: How to improve the diagnostic pathway,’’ Best Pract. Res. Clin. Gastroenterology, vol. 66, Oct. 2023, Art. no. 101842, doi: 10.1016/j.bpg.2023.101842.

[7] C. L. Niedzwiedz, L. Knifton, K. A. Robb, S. V. Katikireddi, and D. J. Smith, ‘‘Depression and anxiety among people living with and beyond cancer: A growing clinical and research priority,’’ BMC Cancer,vol. 19, p. 943, 2019, doi: 10.1186/s12885-019-6181-4.

[8] K. S. Wang et al., ‘‘Accurate diagnosis of colorectal cancer based onhistopathology images using artificial intelligence,’’ BMC Med., vol. 19,p. 76, 2021, doi: 10.1186/s12916-021-01942-5.

[9] L. Adlung, Y. Cohen, U. Mor, and E. Elinav, ‘‘Machine learning in clin
ical decision making,’’ Med, vol. 2, no. 6, pp. 642–665, Jun. 2021, doi:10.1016/j.medj.2021.04.006

[10] M. Rana and M.Bhushan, ‘‘Machine learning and deep learning approachfor medical image analysis: Diagnosis to detection,’’ Multimedia Tools Appl., vol. 82, no. 17, pp. 26731–26769, Jul. 2023, doi: 10.1007/s11042022-14305-w.

[11] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, ‘‘Medical image analysis using deep learning algorithms,’’ Frontiers Public Health, vol. 11, 2023,
Art. no. 1273253, doi: 10.3389/fpubh.2023.1273253.

[12] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta,and A. A. Bharath, ‘‘Generative adversarial networks: An overview,’’10.1109/MSP.2017.2765202.

[13] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He,‘‘A comprehensive survey on transfer learning,’’ Proc. IEEE, vol. 109,no. 1, pp. 43–76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.

[14] D. Bychkov, N. Linder, R. Turkki, S. Nordling, P. E. Kovanen, C. Verrill,M. Walliander, M. Lundin, C. Haglund, and J. Lundin, ‘‘Deep learningbased tissue analysis predicts outcome in colorectal cancer,’’ Sci. Rep.,vol. 8, no. 1, p. 3395, Feb. 2018, doi: 10.1038/s41598-018-21758-3.

[15] E. F. Ohata, J. V. S. D. Chagas, G. M. Bezerra, M. M. Hassan,V. H. C. De Albuquerque, and P. P. R. Filho, ‘‘A novel transfer learning approach for the classification of histological images of colorectal cancer,’’ J. Supercomput., vol. 77, no. 9, pp. 9494–9519, Sep. 2021, doi:
10.1007/s11227-020-03575-6.

[16] X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, ‘‘A survey on ensemble learning,’’ Frontiers Comput. Sci., vol. 14, no. 2, pp. 241–258, Apr. 2020, doi: 10.1007/s11704-019-8208-z.

[17] A. Kallipolitis, K. Revelos, and I. Maglogiannis, ‘‘Ensembling Efficient Nets for the classification and interpretation of histopathology images,’’ Algorithms, vol. 14, no. 10, p. 278, Sep. 2021, doi: 10.3390/a14100278.

[18] S. B. Yengec-Tasdemir, Z. Aydin, E. Akay, S. Dogan, and B. Yilmaz, ‘‘Improved classification of colorectal polyps on histopathological images with ensemble learning and stain normalization,’’ Comput. Methods Programs Biomed., vol. 232, Apr. 2023, Art. no. 107441, doi:
10.1016/j.cmpb.2023.107441.49373VOLUME 13, 2025A. S. Bisht et al.: DeepCRC-Net: An Attention-Driven DL Network for Colorectal Cancer Classification

[19] R. Karthik, V. Thalanki, and P. R. T. S. Yadav, ‘‘Deep learning-based histopathological analysis for colon cancer diagnosis: A comparative study of CNNandtransformer models with image preprocessing techniques,’’ in Proc. Int. Conf. Intell. Syst. Design Appl., in Lecture NotesinNetworksand Systems, Jan. 2024, pp. 90–101, doi: 10.1007/978-3-031-64813-7_11.

[20] P. Gupta, Y. Huang, P. K. Sahoo, J.-F. You, S.-F. Chiang, D. D. Onthoni, Y.-J. Chern, K.-Y. Chao, J.-M. Chiang, C.-Y. Yeh, and W.-S. Tsai,‘‘Colon tissues classification and localization in whole slide images using deep learning,’’ Diagnostics, vol. 11, no. 8, p. 1398, Aug. 2021, doi:
10.3390/diagnostics11081398.

[21] O. Iizuka, F. Kanavati, K. Kato, M. Rambeau, K. Arihiro, and M. Tsuneki, ‘‘Deep learning models for histopathological classification of gastric and colonic epithelial tumours,’’ Sci. Rep., vol. 10, no. 1, p. 1504, Jan. 2020, doi: 10.1038/s41598-020-58467-9.