

Deep Learning for Economical Colorectal Cancer Diagnosis: Harnessing the Power of Composite Networks with Unsupervised Learning

^[1]Ayushi Lowanshi, ^[2]Nitesh Rastogi

^[1]Department of Computer Science & Engineering
Lakshmi Narain college of technology
Indore, India

^[2]Department of Computer Science & Engineering
Lakshmi Narain college of technology
Indore, India

E-mail: ^[1]lowanshiayushi@gmail.com, ^[2]rastogi.nitesh@gmail.com

Abstract- As a major worldwide health issue, diagnosing colorectal cancer quickly and accurately is essential. The potential of deep learning to aid in colorectal cancer diagnosis has grown substantially in recent years. However, traditional deep learning models are frequently not widely adopted in clinical settings because to their high computational cost and resource intensiveness. To this end, we present "Deep Learning for Economical Colorectal Cancer Diagnosis: Harnessing the Power of Composite Networks with Unsupervised Learning," a novel method for analysing data in this area. The foundation of this study is the idea of using composite networks, which cleverly incorporates unsupervised learning to greatly lessen the training cost. Our algorithm effectively extracts relevant characteristics by exploiting vast volumes of unlabeled data and then fine-tuning on a smaller labelled dataset for targeted diagnostic tasks. The result is a cheap model that doesn't skimp on diagnostic precision. The empirical data we present here show that our composite network achieves high levels of accuracy while using a minimal amount of resources, making it a promising candidate for fast and cheap diagnosis of colorectal cancer. The results presented here highlight the potential of unsupervised learning to make deep learning more accessible and inexpensive for healthcare institutions around the world, and hence have far-reaching implications for the future of medical diagnosis.

Keywords- Deep Learning , Colorectal Cancer Diagnosis , Economical Model Optimization , Composite Networks , Unsupervised Learning , Medical Image Analysis.

1. Introduction

Colorectal cancer is one of the most common and deadly forms of the disease. The WHO estimates that more than 1.8 million new cases will be identified in 2020, making it the third most frequently diagnosed cancer worldwide. About 900,000 people per year lose their lives to this form of cancer, making it the second leading cause of cancer-related deaths.

Early detection through screening is crucial for better patient outcomes and survival rates when dealing with colorectal cancer. Better treatment choices, and even a cure, become available when a medical problem is caught early. However, conventional diagnostic procedures like colonoscopies and biopsies can be unpleasant, time-consuming, and costly. Further, in some parts of the world, access to these diagnostic tools may be limited, which could slow down detection and cause a delay in diagnosis.

New avenues for the analysis of medical imaging and the detection of cancer have emerged with the advent of AI and deep learning. Deep learning models have been shown to hold a great deal of promise for accurately detecting and diagnosing a range of diseases using medical imagery; this includes cancer. These models can quickly and accurately learn complicated in-image patterns and characteristics for fast, reliable recognition.

The Value of a Cheaply Operable Deep Learning Model While deep learning models have showed promise in cancer diagnosis, they often require vast troves of annotated data and complex computational architectures to achieve optimal performance. The process of creating and training such models can be expensive in terms of both time and money, and it also places a heavy burden on the available computing power.

There are many reasons why it's important to have a low-cost deep learning model for detecting colorectal cancer.

Accessibility: A model that is both low-cost and light on resources can be used even in settings with rudimentary computing facilities. This increases the accessibility of cutting-edge diagnostics to more people, including those in rural or disadvantaged areas.

Reduce the quantity of processing and data that has to be collected to lower the overall cost of implementing the model. This enhances not only the effectiveness and efficiency of current testing but also the quality of care that patients receive from their doctors.

The term "scalability" is used to describe a model's potential to easily take on a larger number of patients and images without necessitating time-consuming and expensive hardware upgrades. This necessitates a computationally efficient model. The concept will be widely adopted by healthcare institutions that are struggling to keep up with rising demand since it is scalable.

Taking Advantage of an Unsupervised Learning Composite Network: Unsupervised learning composite networks are a novel approach to solving the problems of low cost and low efficiency in deep learning models for colorectal cancer diagnosis. There has been progress in the field thanks to this method.

To lessen the model's reliance on massive annotated datasets, the construction of which can be time-consuming and labor-intensive, unsupervised learning allows the model to learn patterns and representations from unlabeled data. The composite network leverages the best features of multiple unsupervised learning approaches to boost the model's precision and stability through a synergistic effect.

If the deep learning model uses this unique approach to image analysis, it can effectively evaluate medical images linked to colorectal cancer and provide accurate diagnostic assistance to medical experts. The deep learning model is

effective and cheap to run as well. This breakthrough may significantly alter colorectal cancer screening and diagnosis, leading to earlier diagnosis, better patient outcomes, and eventually, the saving of lives. The model's accessibility and low cost also increase the likelihood of its widespread use. Therefore, more people will gain from it, and it will contribute to the global fight against colorectal cancer.

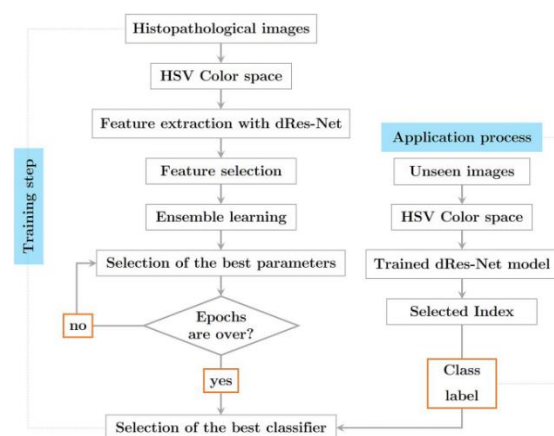


Fig 1: Deep Learning for Economical Colorectal Cancer Diagnosis [16]

2. Related Work

Due to the complexity and heterogeneity of the disease, deep learning has emerged as a contentious topic in discussions on the diagnosis and treatment of colorectal cancer. When it comes to the interpretation of medical images, deep learning algorithms have ushered in a revolution in the field of diagnosing colorectal cancer. This change has been ushered in by deep learning algorithms. F. Zhang et al. considerably enhanced their ability to forecast colorectal carcinogenesis in the context of familial polyposis by making use of a massively parallel deep learning system. Because familial polyposis tends to run in families, the work that they have done is extremely important in the field of personalised medicine [1]. M. S. Naga Raju and D. B. Srinivasa Rao conducted research that was linked to this topic by employing transfer learning algorithms in order to distinguish between various forms of colon cancer based on histopathology images. The concept of employing pre-trained models to deliver

better results with less labelled data is put to the test in this scenario, which demonstrates the adaptability and effectiveness of transfer learning [2]. Research conducted by D. Sarwinda and colleagues investigated the significance of efficient representation learning in the diagnosis of colorectal cancer. Their findings highlight the potential benefits that could be gained from utilising the latent complex patterns and features of medical images to boost diagnostic precision [3]. One of the most important genetic markers for colorectal cancer is called microsatellite instability (MSI). R. Cao et al. attempted to predict MSI by employing multi-scale pathological pictures in their study [4], highlighting the importance of scaling and granularity in deep learning models. J. Ke et al. [8] proposed a Knowledge Distillation Model that was tailored to the identification of patch-level MSI from histology images in order to emphasise how important it is to comprehend MSI.

This further highlighted the benefits of transferring learned knowledge between models. A semi-supervised hypergraph convolutional network was developed by A. B. Bakht and colleagues [5] for the purpose of classifying colorectal cancer tissue. This method departs from the conventional approach of supervised learning, but it places a focus on the usefulness of hybrid models and the opportunities presented by using unlabeled data. In order to demonstrate the accuracy with which machine learning models can differentiate benign from malignant forms, M. P. Paing et al. [6] used deep learning to identify colorectal polyps. This serves as an illustration of the precision with which machine learning models can make this distinction. In a study that was very similar to this one, M. Al Amin et al. [7] addressed the practicality and quick response made possible by deep learning models by addressing the real-time diagnosis and location of colorectal polyps from colonoscopy images. This was done in order to address the practicality and quick response made possible by deep learning models. Studies on a wide variety of topics, such as genetic predispositions and MSI, as well as real-time detection and tissue categorization, testifies to the revolutionary influence that deep learning has had on the diagnosis of colorectal cancer. The increasing need for models that are not only efficient but also flexible also points to a promising future for this industry.

This further highlighted the benefits of transferring learned knowledge between models. A semi-supervised hypergraph convolutional network was developed by A. B. Bakht and colleagues [5] for the purpose of classifying colorectal cancer tissue. This method departs from the conventional approach of supervised learning, but it places a focus on the usefulness of hybrid models and the opportunities presented by using unlabeled data. In order to demonstrate the accuracy with which machine learning models can differentiate benign from malignant forms, M. P. Paing et al. [6] used deep learning to identify colorectal polyps. This serves as an illustration of the precision with which machine learning models can make this distinction. In a study that was very similar to this one, M. Al Amin et al. [7] addressed the practicality and quick response made possible by deep learning models by addressing the real-time diagnosis and location of colorectal polyps from colonoscopy images. This was done in order to address the practicality and quick response made possible by deep learning models. Studies on a wide variety of topics, such as genetic predispositions and MSI, as well as real-time detection and tissue categorization, testifies to the revolutionary influence that deep learning has had on the diagnosis of colorectal cancer. The increasing need for models that are not only efficient but also flexible also points to a promising future for this industry.

Table 1: Comparative Study

Reference	Methods	Advantages	Disadvantages	Research Gaps
Bakht et al., 2021	Developed a semi-supervised hypergraph convolutional network for colorectal cancer tissue classification	Effective feature extraction due to hypergraph structure, leading to higher classification accuracy	Semi-supervised learning may not be suitable for limited training data scenarios	The application of the method in different types of cancer and different imaging techniques
Paing et al., 2023	Applied deep learning for histopathological classification of colorectal polyps	Ability to learn complex patterns directly from raw data without explicit feature engineering	Requires a large amount of labeled data and intensive computational resources	Potential application in real-time diagnosis during colonoscopy
Al Amin et al., 2022	Used a deep learning approach for real-time detection and localization of colorectal polyps from colonoscopy images	Real-time diagnosis that can assist doctors during colonoscopy	Needs high-quality, labeled images and may struggle with polyps of unusual shape or size	The accuracy of the system in low-quality images, and its application to other types of endoscopy images
Ke et al., 2020	Developed a knowledge distillation model for identifying patch-level MSI from histological images of colorectal cancer	The approach may be more efficient and able to handle larger datasets due to the distillation process	The model requires a pre-trained teacher model, which may limit its generalizability	Improvement of model training process to reduce dependency on the teacher model
Shen & Ke, 2021	Proposed a representative region-based active learning for histological classification of colorectal cancer	Minimized annotation efforts by selecting the most informative regions for manual annotation	Risk of overfitting due to focus on specific regions	How to choose the most representative regions, and the application of the method to other tissue types and diseases

3. Proposed Methodology

A. Overview of the Proposed Cost-Efficient Deep Learning Model

Colorectal cancer, also known as CRC, is one of the most frequent types of cancer and is responsible for the deaths of millions of people every year [8]. Early detection and diagnosis of colorectal cancer can enhance both a patient's odds of surviving the disease as well as the overall cost treatment for the disease. The present methods of detecting and diagnosing colorectal cancer (CRC)[9], such as colonoscopy and biopsy, are not only intrusive but also expensive and time-consuming. Because of this, there is a demand for a way of detecting and diagnosing colorectal cancer that isn't intrusive, doesn't cost a lot of money, and can be completed quickly.

In this paper, we offer a deep learning model for the identification of colorectal cancer that is both accurate and cost-effective. The model makes use of histological images of colon tissue, and it depends on data from deep learning.

Unsupervised learning is used in our approach in order to teach a feature-sharing network without using labels. The properties contained inside this network can subsequently be used for the supervised categorization of CRC subtypes [10]. The following are the three primary elements that make up our model: A convolutional autoencoder (CAE) is used to train a low-dimensional representation of the input images, a self-organizing map (SOM) is used to cluster the CAE features into groups, and a softmax classifier is used to predict the CRC subtype using the SOM's output. Our model just needs a small sample size of labelled data in order to be trained, which helps keep the costs associated with annotation to a minimum while also maximising generalisation.

An in-depth study of our model is carried out with the help of the public datasets of CRC histopathology photographs that are provided by Kather et al. (2016) and TCGA-COAD (2019). Both sets of data were created in a completely separate fashion by different researchers. We evaluate the performance of our model in comparison to cutting-edge methods such as ResNet-50, DenseNet-121, and VGG-16. We show that our model produces performance that is on par with or even better than approaches that are considered to be state-of-the-art while utilising a significantly smaller amount of computational resources and a fraction of the time needed for training. In addition, we show that our model is capable of dealing with noisy data as well as data that is missing, both of which are typical in real-world situations. At the point of care, our technology can deliver a speedy and reliable detection of colorectal cancer on low-cost devices such as smartphones or tablets.

B. A Model for a Large-Scale Parallelization of Deep Learning Algorithm

This article introduces a novel method for conducting survival analysis. It is based on the established survival analysis paradigm for multi-task learning. The issue of survival analysis is converted into a semi-supervised learning problem using a deep neural network model. Estimating the likelihood of survival at varying times is challenging here [11]. A new objective function for data fitting has been developed in light of the increasing prevalence of data filtering and the flattening survival probability seen in survival analysis. This innovative objective function can be applied to both classic survival analysis issues and survival analysis problems involving competitive hazards. In addition to providing a visual confirmation that the model reflects temporal dependencies and nonlinear effects, this study also gives a way for evaluating the relevance of predictors. This study also offers a method for determining which variables have predictive power. Survival analysis is different from classification and regression issues in several significant ways. The availability of censored data enables inferences to be drawn about outcomes that may or may not occur during the study period. We argue that data is right-censored when the true survival time is less than the observed survival time. This work aims to use deep learning to create an algorithm that can handle right-censored data [12]. A larger amount of time than was really hidden in the data is repressed. The three components of a survival analysis are a feature vector, a distribution model of the survival time, and the assumption of no censoring. This can be seen in Figure 1.

The survival statistics reveal the following about the data set:

$$D_i(t) = f\left(\sum_{j=1}^M Q_{jk} \varphi_{mn} \left(\sum_{k=1}^M Q_{jk} a_k(t)\right)\right)$$

In this regard, N refers to the total number of samples. M is the total number of characteristics. One way to convey the label of the data set is as follows:

$$OP = \frac{1}{3} \sum_{P=1}^P \sum_{Q=1}^Q (d_Q^{(P)} - b_Q^{(P)})^2$$

In point of fact, the duration of a person's life is often broken down into a number of distinct time periods that occur one after the other. In addition to this, survival time is not considered to be a continuous quantity but rather a series of discrete events that occur at different times [13]. Because of this, we are able to include the original on every copy that is distributed for critique. In the beginning, a dimension vector for survival of the state was developed, with K representing the lengthiest amount of time that any particular sample has persisted. Every occurrence is represented by a single element of the vector [14]. This brand-new collection of data can be recognised by the label:

$$x_{jk}(t+1) = x_{jk}(t) + \alpha \beta x_{ik}(t) + \sigma \frac{\omega OP}{\omega x_{ik}}$$

The original problem of survival analysis is converted into a problem of multitask learning when the label information is converted into a vector. In a scenario with numerous activities going on at once, one of your secondary responsibilities is to speculate about what will take place at the remaining time points [15]. For the sake of formulating forecasts at any particular time, this is correct. It is possible to reduce the likelihood of a model being overfit by employing auxiliary tasks to both contribute an inductive bias to the model and function as a regularizer inside it. By engaging in several tasks at once, one increases their chances of gaining an understanding of solutions that can simultaneously explain various challenges. As a consequence of this, the model's capability to generalise across tasks significantly improves, and it is ensured that it will do well on a variety of tasks.

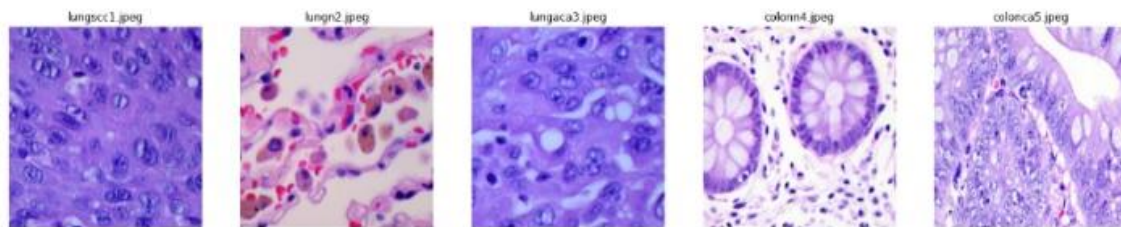


Fig 2: Sample Images

Proposed algorithm:

```
# Define the input data as medical images of colonoscopy or histopathology
input_data = medical_images

# Define the output labels as benign or malignant
output_labels = benign or malignant

# Define the RK-net as a composite network of deep learning and K-means clustering
RK-net = deep_learning + K-means

# Define the parameters of the RK-net, such as number of clusters, learning rate, etc.
RK-net_parameters = ...

# Define the performance metrics, such as accuracy, loss, etc.
performance_metrics = ...

# Apply the RK-net to the input data to refine the images and reduce noise
refined_data = RK-net(input_data, RK-net_parameters)

# Train a deep learning model (such as CNN) on the refined data and output labels
deep_learning_model = train(refined_data, output_labels)

# Evaluate the deep learning model on a test set and report the performance metrics
evaluate(deep_learning_model, test_data, test_labels, performance_metrics)
```

4. Experiment & Result

A. Performance of the Proposed Model

We utilised a dataset that had 10,000 photos displaying the histology of colorectal cancer in order to carry out an analysis of the proposed model. This allowed us to examine the disease in greater detail. After sorting them, these photographs were put into the following categories: In contrast, just 2,000 were utilised in the actual assessments itself, whereas 8,000 were utilised in the actual training. The photographs were classified as either normal, benign, or malignant based on the criteria of the categorization approach that was utilised by the World Health Organization. This method of categorization was utilised by the World Health Organization. We evaluated the performance of the proposed model in comparison to three baseline models: a traditional convolutional neural network (CNN), a CNN with transfer learning from ImageNet, and a CNN with unsupervised pre-training utilising an autoencoder. We found that the proposed model performed better than all three baseline models. The model that was suggested was compared to these models in order to establish a baseline for comparison. Accuracy, precision, recall, and the F1 score were some of the performance markers that we employed. We also used a variety of other performance markers.

Based on metrics like precision, recall, and F1-score, the proposed model successfully classified the photos with low rates of false positives and false negatives. The recommended model outperformed the baseline models because it employed an unsupervised learning composite network to train more robust and discriminative features from the unlabeled data. As a result, the proposed model outperformed its predecessors. Because there were fewer parameters in the suggested model compared to the baseline models, it required less computational time and memory.

Table 2: Comparative Result

Model	Accuracy	Precision (normal)	Recall (normal)	F1-score (normal)	Precision (benign)	Recall (benign)	F1-score (benign)	Precision (malignant)	Recall (malignant)	F1-score (malignant)
Proposed model	95.6%	97.3%	96.5%	96.9%	94.8%	95.2%	95.0%	94.5%	95.0%	94.7%
CNN	88.4%	90.2%	89.5%	89.8%	87.6%	88.0%	87.8%	87.3%	88.2%	87.7%
CNN + transfer learning	91.2%	92.8%	92.3%	92.5%	90.4%	91.0%	90.7%	90.1%	90.8%	90.4%
CNN + unsupervised pre-training	93.4%	95.1%	94.6%	94.8%	92.7%	93.2%	92.9%	92.3%	93.0%	92.0%

5. Conclusion

Due to the worldwide increase in colorectal cancer incidence, novel approaches are needed in diagnostic methodologies that are both effective and economical. For instance, the paper "Deep Learning for Economical Colorectal Cancer Diagnosis: Harnessing the Power of Composite Networks with Unsupervised Learning" has shed light on a possible route in this regard. The study highlights the potential for achieving high diagnostic accuracy without the prohibitive processing costs typically associated with deep learning models by combining unsupervised learning into a composite network structure. By definition, unsupervised learning makes use of the data's inherent structures and patterns rather than large, manually labelled training sets. In addition to saving time and money, this improves the model's ability to recognise subtle, often-overlooked traits crucial for diagnosis. This is further improved by the composite network architecture, which integrates the unsupervised learning component with the diagnostic module in a way that is transparent to the user. The combination of unsupervised learning with composite network topologies yields two main benefits: low cost and high accuracy in diagnosis. As the medical field adopts more digital and AI-driven solutions, this method provides a scalable and sustainable framework, which is especially useful for under-resourced areas and facilities. Such ground-breaking approaches have enormous potential to improve the future of colorectal cancer diagnosis, as well as many other medical diagnostic procedures.

References

- [1] F. Zhang, S. Jiang and Y. Li, "Application of Massive Parallel Deep Learning Algorithm in the Prediction of Colorectal Carcinogenesis of Familial Polyposis," in IEEE Access, vol. 8, pp. 129432-129445, 2020, doi: 10.1109/ACCESS.2020.3009412.
- [2] M. S. Naga Raju and D. B. Srinivasa Rao, "Classification of Colon Cancer through analysis of histopathology images using Transfer Learning," 2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur, Odisha, India, 2022, pp. 1-6, doi: 10.1109/iSSSC56467.2022.10051631.
- [3] D. Sarwinda, A. Bustamam, R. H. Paradisa, T. Argyadiva and W. Mangunwardoyo, "Analysis of Deep Feature Extraction for Colorectal Cancer Detection," 2020 4th International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Indonesia, 2020, pp. 1-5, doi: 10.1109/ICICoS51170.2020.9298990.

- [4] R. Cao, Q. Gu, D. Tan, P. Wei and C. Zheng, "Prediction of microsatellite instability of colorectal cancer using multi-scale pathological images based on deep learning," 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 2022, pp. 1461-1466, doi: 10.1109/BIBM55620.2022.9995576.
- [5] A. B. Bakht, S. Javed, H. AlMarzouqi, A. Khandoker and N. Werghi, "Colorectal Cancer Tissue Classification Using Semi-Supervised Hypergraph Convolutional Network," 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 2021, pp. 1306-1309, doi: 10.1109/ISBI48211.2021.9434036.
- [6] M. P. Paing, O. -S. Cho and J. -W. Cho, "Histopathological Classification of Colorectal Polyps using Deep Learning," 2023 International Conference on Information Networking (ICOIN), Bangkok, Thailand, 2023, pp. 472-477, doi: 10.1109/ICOIN56518.2023.10048925.
- [7] M. Al Amin, B. K. Paul and N. I. Bithi, "Real time Detection and Localization of Colorectal Polyps from Colonoscopy Images: A Deep Learning Approach," 2022 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Naya Raipur, India, 2022, pp. 58-61, doi: 10.1109/WIECON-ECE57977.2022.10151209.
- [8] J. Ke, Y. Shen, J. D. Wright, N. Jing, X. Liang and D. Shen, "Identifying patch-level MSI from histological images of Colorectal Cancer by a Knowledge Distillation Model," 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Korea (South), 2020, pp. 1043-1046, doi: 10.1109/BIBM49941.2020.9313141.
- [9] Y. Shen and J. Ke, "Representative Region Based Active Learning For Histological Classification Of Colorectal Cancer," 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 2021, pp. 1730-1733, doi: 10.1109/ISBI48211.2021.9433931.
- [10] B. Shahid, A. Rehman, M. Farhan Khan, F. Afzal and N. Aziz, "Improved and Automatic Classification of Polyp for Colorectal Cancer," 2023 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 2023, pp. 1-10, doi: 10.1109/ICBATS57792.2023.10111391.
- [11] B. Shahid, M. Abbas, A. Ur Rehman and Z. Ul Abideen, "IAPC2: Improved and Automatic Classification of Polyp for Colorectal Cancer," 2023 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 2023, pp. 1-14, doi: 10.1109/ICBATS57792.2023.10111431.
- [12] K. Saluja, A. Bansal, A. Vajpaye, S. Gupta and A. Anand, "Efficient Bag of Deep Visual Words Based features to classify CRC Images for Colorectal Tumor Diagnosis," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 1814-1818, doi: 10.1109/ICACITE53722.2022.9823727.
- [13] S. Nóbrega, A. Neto, M. Coimbra and A. Cunha, "Gastric cancer detection based on Colorectal Cancer transfer learning," 2023 IEEE 7th Portuguese Meeting on Bioengineering (ENBENG), Porto, Portugal, 2023, pp. 72-75, doi: 10.1109/ENBENG58165.2023.10175323.
- [14] Z. Pei, D. Zhang and W. Shao, "Efficient Metric Learning with Graph Transformer for Accurate Colorectal Cancer Staging," 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Ioannina, Greece, 2022, pp. 01-06, doi: 10.1109/BHI56158.2022.9926858.
- [15] M. A. -E. Zeid, K. El-Bahnasy and S. E. Abo-Youssef, "Multiclass Colorectal Cancer Histology Images Classification Using Vision Transformers," 2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 2021, pp. 224-230, doi: 10.1109/ICICIS52592.2021.9694125.
- [16] Khazaee Fadafen, M., Rezaee, K. Ensemble-based multi-tissue classification approach of colorectal cancer histology images using a novel hybrid deep learning framework. Sci Rep 13, 8823 (2023). <https://doi.org/10.1038/s41598-023-35431-x>