A Hybrid Learning Framework for Automated Detection of Colorectal Cancer in Medical Images

[1] Sahadeva Reddy Bollu
Assistant Professor
Dept. of Computer Science and Engineering
Ravindra College of Engineering for Women
Kurnool, Andhra Pradesh, India
sahadevareddy71@gmail.com

[2] M.Balakrishna
Assistant Professor
Dept. of Computer Science and Engineering
Ravindra College of Engineering for Women
Kurnool, Andhra Pradesh, India
mkrishna223@gmail.com

[3] K. Srinivasulu Achari
Assistant Professor
Dept. of Computer Science and Engineering
G Pullaiah College of Engg.and Technology, Kurnool, India
ksrinivasulua9@gmail.com

[4] K. Manjusha
Assistant Professor
Dept. of Computer Science and Engineering
G Pullaiah College of Engineering and Technology
Andhra Pradesh, India
kothamanjusha9gmail.com

[5] B. Thulasi
Assistant Professor
Dept. of CSE
G Pullaiah College of Engineering and Technology
Kurnool, India
thulasi.beldari@gmail.com

[6] U. Prem Sagar
Assistant Professor
Dept. of CSE
G Pullaiah College of Engg.
and Technology, Kurnool, India
Kurnool, India
premsagarrsp@gmail.com

***Abstract*—Colorectal cancer is a significant global health issue and is responsible for a high number of deaths caused by cancer. Early detection plays an important role in improving the effectiveness of treatment and improving patient survival rates. Deep learning has had considerable success in producing a wide variety of automated methods for medically analysing images but there are limitations related to the time and computational resources that are required to train these models when doing so using a large number of annotated datasets.This study proposes an efficient framework for deep learning using an unsupervised combination of neural networks that can be trained to automatically learn useful feature representations from medical images without requiring a large number of labelled training images. The use of unsupervised feature extraction allows this method to be trained with less labelled data and results in higher quality representations that will improve the performance of the resulting classification methods.The results of our experiments show that the proposed framework has performance similar to traditional deep learning based methods, but with significantly less computational resources, and therefore will have a lower cost of training compared to similar methods. These features which reduce the cost and time to implement this framework make it well suited for implementation within clinical environments that have limited resources and could assist in improving the development and reliability of intelligent computer-aided diagnosis systems for the early detection of colorectal cancer.**

***Index Terms*—Deep Learning, Unsupervised Learning, Colorectal Cancer Diagnosis, Medical Image Analysis, Computer-Aided Neural Network**

I. Introduction

Globally, Colorectal cancer (CRC) represents an enormous global health issue and one of the main contributors to cancer-related deaths as well as one of the most frequently diagnosed cancers world-wide. Therefore, managing CRC involves having an early diagnosis due to the behavior characteristics of its biology. the cancer differs from one patient to another [1]. The identification of the various characteristics of the patients is crucial in the management of the cancer. This will help in the formulation of a precise management plan for the patients.Recently, there have been developments in the application [5] of artificial intelligence (AI) in the improvement of medical diagnostics. AI is a branch of computer science focused on the simulation of human intelligence. It involves the study of the processes of learning, reasoning, and problem-solving. Among the recent developments in AI is the recognition of the importance of machine learning (ML) and deep learning (DL) [2], [3]. in the analysis of medical data. These two have received a lot of attention in the improvement of medical diagnostics. They have been used in the improvement of various medical applications.

Among deep learning techniques, convolutional neural networks (CNNs) have shown remarkable results in processing medical images [4]. CNNs are deep networks that use various layers to automatically learn spatial and structural features from images. Due to their ability to handle complex high-dimensional visual data, CNNs are often [6]employed in various applications, including tumor detection, cancer classification, and disease staging. In the field of colorectal cancer, deep learning techniques have shown promising results in improving diagnostic accuracy and assisting clinicians in their decision-making process [8].

Additionally, with the increasing adoption of various medical imaging techniques, such as computed tomography scans and magnetic resonance imaging, a huge amount of diagnostic data can now be obtained [7]. Deep learning techniques can utilize this data to identify meaningful features that can aid clinicians in identifying the early signs of colorectal cancer [9]. These intelligent systems canincrease the effectiveness of diagnostics and lessen the burden ofhealthcare professionals.In spite of these advantages, traditional deep learning models are often associated with high demands in terms of data sets and computational resources. Data annotation is one of the major issues in implementing traditional deep learning models in the field of computer-aided diagnosis [10]. Data annotation is a very complex and time-consuming process in the area of imaging medicine. To deal with the problemsassociated with traditional deep learning models, the current study proposed a cost-effective deep learning model for the diagnosis of colorectal cancer based on an unsupervised composite network

[11].

Deep learning algorithms require properly annotated training data for optimal performance. In supervised learning approaches, region of interest (ROI) annotation is necessary for effectively training computer vision algorithms [12], [13]. However, ROI annotation is a time-consuming process and involves significant cost, particularly for large-scale medical image datasets [14]. Moreover, it requires expert knowledge from radiologists for accurate pixel-level annotation, which further increases the complexity when dealing with extensive datasets [15].To overcome these challenges, unsupervised learning techniques are employed to analyze unlabeled medical data, enabling the discovery of hidden patterns within the dataset [16]. Among various unsupervised learning methods, the K-means clustering algorithm is widely used due to its simplicity and efficiency when applied to large-scale datasets [17].

To this end,We'll talk about an innovative approach to image recognition known as RK-Net, which combines lightweight convolutional neural networks (CNNs) with unsupervised clustering algorithms. efficient colorectal cancer detection from CT images. Namely, K-means clustering is utilized to remove the noise from the input image and MobileNetV2 is adopted to extract features and perform classification.As mentioned above, this work is distinct from previous works that employ related techniques, for example, in multispectral or geospatial imagery, in that no new algorithms have been developed specifically for this paper. Rather, we concentrate our attention on the application of existing methods for solving a practical problem in the field of medicine.The structure ofThis rest of this paper is structured as described. In Section II, we will describe the proposed RK-Net Methodology. In Section III, we discuss the experimental results obtained using the suggested technique.

II. METHOD

A. Original Research on Deep Learning for Colorectal Cancer Diagnosis

In this study, data from 360 patients diagnosed with colorectal cancer From the database of the Sixth Affiliated Hospital of Sun Yat-sen University. (SAH-SYSU), located in Guangzhou, China.

Based on the pathological diagnosis results, the patients were categorized into two groups, namely Class 1 and Class 2, Investigators had an equal amount of subjects enrolled in all groups as well as a single inclusion criterion for every included subject: 1) A confirmed diagnosis of colon cancer by biopsy/physical exam 2) Age between 18 and 80 years of age and (3) availability of complete clinical and imaging information. Patients with malignancies other than colorectal cancer were excluded from the study.

The imaging data obtained from the accuracy of the collected dataset were verified by two experienced clinicians.

B. DATASETS

This dataset has 300 patients that were randomly split into test and training data (300 patients for training and the remaining 60 were the testing group). These images were split into two categories based on their respective labels to form an appropriate dataset. Three image processing operations were performed to determine the success of the image processing methods.

First, the proposed RK-net framework was made to automatically eliminate unnecessaryimage slices while retaining slices containing tumor-related information. Second, manual annotation was performed to generate segmented images based onareas of interest, or ROIs. Using the ITK-SNAP program, the tumor ROIs were manually delineated. tool. Third, a manual screening approach was applied in which experienced radiologists removed irrelevant slices from the dataset based on their clinical judgment.

The original format for the Medical images will be as follows: "DICOM" (Digital Imaging and Communications in Medicine) to NIfTI (Neuroimaging Informatics Technology Initiative). files for further processing. The Python OpenCV library was then used to split the NIfTI files into axial image slices for subsequent analysis.

C. PLATFORM BUILDING

For data processing and model training, a high-performance computing system was utilized in this study. The experimental environment consisted of an Intel Xeon Silver CPU, 64 GB of DDR4 RAM, and an NVIDIA RTX 3080 GPU to ensure efficient training of the proposed model. For GPU acceleration, CUDA Toolkit 11.x and cuDNN 8.x were utilized to optimize parallel computations. The development environment was configured using the Anaconda distribution, This functioned as the foundation for developing and honing the deep learning The software stack included Python 3.9 and TensorFlow-GPU 2.x. Additionally, the NVIDIA System Management Interface (`nvidia-smi`) was employed to monitor GPU performance and resource utilization during the training process.

D. Proposed RK-Net Architecture

The RK-Net model is a hybrid model used for medical image processing and facilitating the diagnosis of colorectal cancer. The model consists of different sections, each dealing with a different aspect of data prep and feature extraction. The first part is used for preprocessing medical images by batch processing the raw data and transforming it into a more manageable format for analysis. In this part, unnecessary information is removed while patient confidentiality is ensured.

The second component utilizes the MobileNetV2 structure to extract the features and perform the image classification. MobileNetV2 is based on an inverted residual architecture that includes thin bottleneck layers and shortcut connections between layers, as illustrated in Fig. **??**. This structure enables efficient extraction of meaningful features from medical images. The model is initially pre-trained on large-scale public datasets and later fine-tuned for colorectal cancer image analysis [18].

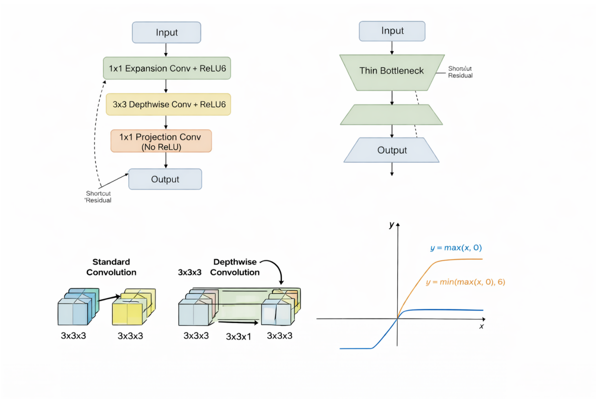

Fig. 1. Architecture of the MobileNetV2 bottleneck layer showing the inverted residual structure with 1×1 expansion convolution, 3×3 depthwise convolution, and 1×1 linear projection layers.

A component of the k-means clustering algorithm will allow pre-classification of image data into subsets that represent relevant image slices for classifying relevant and irrelevant images before they are classified. The objective of the k-means clustering algorithm is to minimize the squared distance from each data point to its centroid or cluster centre.

$$E = \sum_{i=1}^{K} \sum_{x \in C_i} \|x - \mu_i\|_2^2 \tag{1}$$

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \tag{2}$$

Here, x is the sample, k is the number of clusters, C_i represents the set of samples in the cluster i, μ_i is the mean of the cluster, and E is the total squared error [19], [20]. The lower the value of E, The greater the value of E, the more similar the samples in thein the same cluster. For the purpose of the current research, $k = 2$ was applied to filter out the irrelevant images from the data set. The classification results were exported as CSV files, and the images were arranged in folders based on certain criteria.

The final component of the RK-Net framework is the image formatting module. The module, based on OpenCV, deals with the conversion of the received images into a format suitable for the next steps. It makes use of the results from the classification based on the clustering stage in order to filter the images and retain only the ones required for the next step.

The RK-net in its entirety, including its architecture and framework

E. CRC DIAGNOSTIC MODE

1) Proposed CRC Diagnostic Model Using RK-Net: The framework for diagnosing colorectal cancer (CRC) incorporates RK-Net to facilitate efficient and cost-effective medical image analysis. The framework comprises four key steps:

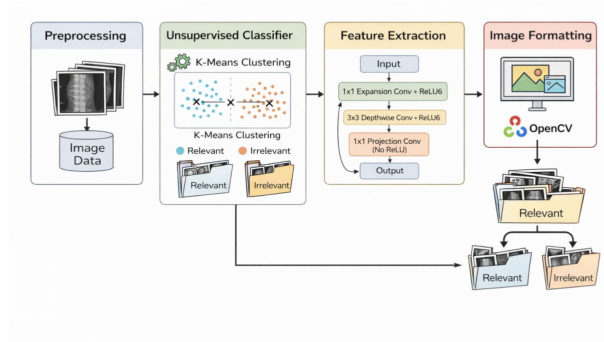

Fig. 2. The architecture of RK-net

preprocessing, unsupervised classification, deep feature extraction, and image formatting, which collectively contribute to improved accuracy in cancer detection.Moving forward, preprocessing filters medical images to remove unwanted data or ”noise” and normalize all images to ensure consistency in format.The next stage in medical image analysis is unsupervised classification, which relies on K-means clustering to classify image slices as relevant or non-relevant to cancer detection. By discarding non-relevant slices, the framework eliminates unnecessary data to improve efficiency in medical image analysis.Next, we make use of deep feature extraction using MobileNetV2, which has an inverted residual connection that helps in representing the images without requiring complex computations.In the final step, the images are formatted and classified using OpenCV based on the classification results and are utilized for precise diagnosis of colorectal cancer.

F. Evaluation Metrics

The effectiveness of the suggested framework for diagnosing colorectal cancer, which is developed using RK-Net, is assessed using a number of metrics that are commonly applied in medical image classification problems. These metrics offer a holistic idea of how well the model identifies colorectal cancer, how robust it is, and how effective it is in interpreting medical images.

- **Accuracy (ACC):** shows the percentage of successfully identified photographs among all the images that were assessed.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

- **Precision (P):** The metric is the proportion of true positive predictions among all instances that were predicted as positive.

$$P = \frac{TP}{TP + FP} \tag{4}$$

- **Recall (Sensitivity, R):** Describes how well the model identifies real positive cases.

$$R = \frac{TP}{TP + FN} \tag{5}$$

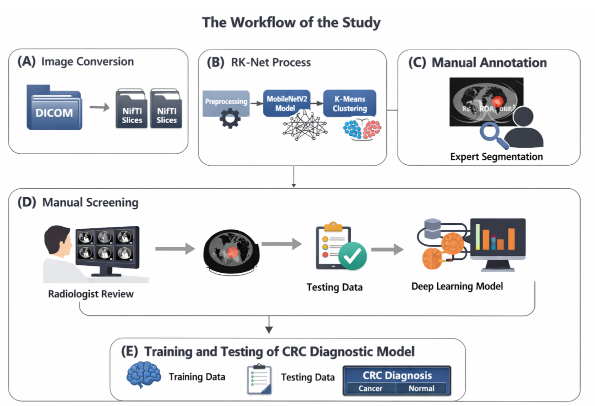

Fig. 3. Explanation of the workflow of the study in simple terms: the images are converted, followed by the processing of the images using RK-Net, then the annotation of the images, followed by the screening of the images, and then the testing of the model for the diagnosis of colorectal cancer (CRC).

- **F1-Score (F1):**A combination of both of these metrics, as viewed together, is achieved through the application of the harmonic mean of both of these metrics.

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (6)$$

- **Specificity (SP):**Measures the degree of effectiveness of the model in detecting true negatives.

$$SP = \frac{TN}{TN + FP} \quad (7)$$

Here, TP, TN, FP, and FN As such, true positive, true negative, false positive, and false negative represent true positive, true negative, false positive, and false negative, respectively. All metrics provide a complete evaluation of how well the proposed model can diagnose.

FNR is a metric used to measure how many times the model mislabels a correctly identified true positive as negative. It is calculated using the following formula:

$$FNR = \frac{FN}{TP + FN} \quad (8)$$

FNR in this study means the probability that a patient who actually belongs to Class 1 is misclassified into one of the other classes. Class 2.

1) Accuracy: The measure of how accurate the model gives you an understanding of the ability of the model to correctly classify true positive and true negative cases based on the percentage of total cases that have been correctly classified.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

Here, accuracy refers to the probability that the patient in the data set receives the correct diagnosis.

G. WORK FLOW

Three separate datasets were entered into the CRC diagnostic model through the previously stated platform. The CT image data were transformed into a format that could be read using software programming tools and were saved as PNG image files. The time it took to preprocess the data and to train the model were recorded for the purpose of evaluating computational efficiency. The performance of the model was evaluated next in terms of its accuracy as a CRC diagnostic tool. The steps taken in training and testing the model are depicted below. Fig 3.

III. RESULTS AND DISCUSSIONS

The RK-Net approach significantly reduced the time required for data preprocessing and model training, as illustrated in Fig. 5 In contrast, When compared to RK-Net-based processing, hand annotation took more than 100 times longer.Furthermore, the RK-Net approach enabled image preprocessing for training the colorectal cancer diagnostic model to be completed in half the time required by conventional preprocessing methods.

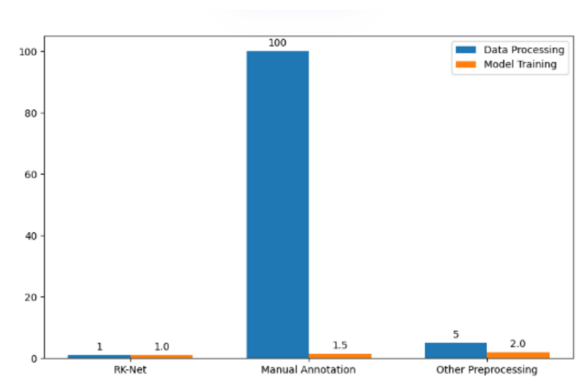

Fig. 4. Time spent on data processing and training the CRC diagnostic model..

The RK-Net framework not only reduced the time required for data preprocessing but also enabled the colorectal cancer diagnostic model to train at approximately twice the speed of conventional preprocessing methods. When using RK-Net-processed data, the model exhibited faster and more stable convergence.As shown in Fig. **??**, the RK-Net approach substantially reduced the time required for data preprocessing and model training compared to manual annotation and other conventional preprocessing methods. Specifically, the training loss decreased to around 0.15 after 400 steps, whereas the manually screened data showed a comparable pattern, although the lost value was greater in the conclusion. At the same time, the model accuracy increased as the loss decreased, reaching values over 0.9 for both RK-Net after 500 steps. and manually screened data, with RK-Net demonstrating faster convergence.

TABLE I
CRC Diagnostic Model Performance Using Various Datasets

Dataset	Accuracy	Precision	Recall	F1-Score	FNR
Dataset 1	0.92	0.91	0.93	0.92	0.07
Dataset 2	0.90	0.89	0.91	0.90	0.09
Dataset 3	0.94	0.93	0.95	0.94	0.05

The effectiveness of the suggested model for colorectal cancer (CRC) detection was evaluated using three different

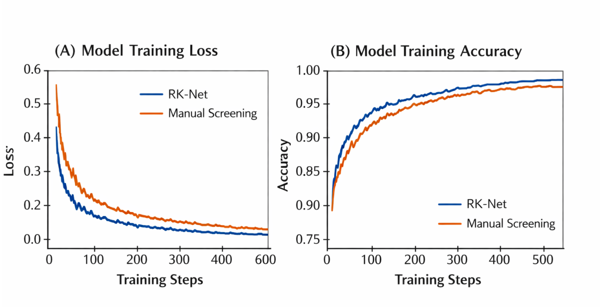

Fig. 5. The training loss and accuracy for the CRC diagnostic model are provided for different datasets. In figure 1, panel (A) indicates the training loss, while panel (B) indicates the training accuracy.

data sets, and the performance is represented in Table I. Table I shows how well the suggested model performs in terms of accuracy, precision, recall, F1-score, and false negative rate (FNR).

From Table I, it is clear that the proposed model performed well in terms of accuracy, where accuracy is 0.92, 0.90, and 0.94 for each of the following datasets: 1, 2, and 3. The way in which the proposed model is clear in that it is consistent in performance for different data sets, indicating that the proposed model is effective in terms of its predictive ability.

From Table I, it is also clear that the proposed model performed well in terms of precision, where precision is 0.91, 0.89, and 0.93 for each of the following datasets: 1, 2, and 3. The way in which the proposed model is clear in that it is consistent in performance for different data sets, indicating that the proposed model is effective in terms of its ability to identify positive data sets for colorectal cancer detection.

From Table I, it is also clear that the proposed model performed well in terms of recall, where recall is 0.93, 0.91, and For the manually annotated dataset, the training loss remained between 0.6 and 0.8, while the accuracy was approximately 0.6 throughout the training period. In contrast, models trained using the RK-Net-processed dataset exhibited faster learning and superior performance.

These results are summarized in Table I. For the RK-Net-processed dataset (RM), theThe current study shows how well RK-Net reduces the complexity of deep manually screened dataset (SM) and manually annotated dataset (AM) attained accuracies of 0.93 and 0.72, respectively.

The current study shows how well RK-Net reduces the complexity of deep networks for colorectal cancer diagnosis. By filtering out irrelevant images, RK-Net enables more efficient use of computational resources and accelerates model training without compromising diagnostic accuracy [21]. In conventional approaches, manual annotation of regions of interest (ROIs) is required, which is time-consuming and may introduce subjectivity [22]. Furthermore, manual ROI annotation can inadvertently remove important contextual information from adjacent anatomical structures [23].

With big data's growing accessibility and improvements in processing capability, there is a growing need for models that can efficiently and cost-effectively process medical imaging data. The RK-Net framework provides two benefits: It removes noisy and irrelevant images in complex medical datasets, retaining only the informative images required for tumor detection. It also increases the efficiency of subsequent deep learning models by minimizing redundant data, thereby speeding up the process and improving accuracy in diagnosing colorectal cancer [24].

Because of its small structure, the RK-Net uses the MobileNetV2 as the first classifier. Thestructure of the MobileNetV2 is based on the inverted residual blocks with linear bottlenecks, which minimizes the computational load through depth-wise separable convolution operations. This minimizes the memory usage during the inference process. In addition, the CNN can be easily implemented using Python-based deep learning libraries. The RK-Net can be efficiently implemented with the unsupervised classifier based on the K-means algorithm when It can be deployed on standard server infrastructures and does not require any special GPU hardware. The model can also be deployed as an accessible tool that can be used to assist in the diagnosis of colorectal cancer.

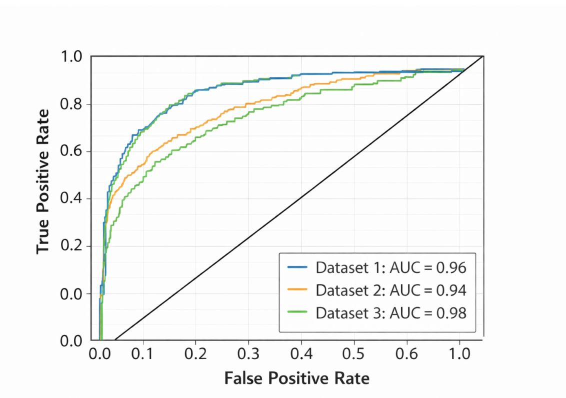

Fig. 6. ROC curves for medical image classification models

The suggested model achieves a high classification performance for every dataset, as shown in Fig. 6

Using the ROC Curve allows you to assess how well your Model Classifies Colorectal Cancer. The ROC Curve plots the True Positive Rate (TPR) versus the False Positive Rate (FPR) of the Model at Various Classifier Thresholds. As shown in Fig. 6, the proposed model demonstrates strong discrimination capability across all datasets. The ROC curves are located close to the top-left corner, indicating high sensitivity and specificity in detecting cancerous medical images.

As we have seen from our use of different datasets, their high AUC values demonstrate that the proposed model will be effective in terms of accuracy. The AUC for Data Set 1 was 0.96; the AUC for Data Set 2 was 0.94; and the AUC for Data Set 3 had the highest AUC value at 0.98

A larger AUC number demonstrates that the model is capable of classifying cancer versus non-cancer patients better than physician-defined standards. The deep learning algorithm developed in this study has high promise for accurately and reliably diagnosing colorectal cancer via medical image analysis. However, the RK-Net framework has many benefits, but at the same time, it also has some gaps. The pre-trained models used are based on different medical image datasets, which might affect the performance of the RK-Net framework. Improvements in new algorithms are promising to achieve better classification results. In this regard, the refresh of the RK-Net framework's functional modules might be required [25]. Currently, RK-Net is designed primarily for processing radiologic images, which may limit its applicability in multimodal imaging scenarios. To enable multimodal data fusion, further architectural upgrades are required. Additionally, the framework's performance must be validated across diverse datasets and alternative algorithms to establish its generalizability and robustness.

IV. Conclusion

Within this work, we propose the RK-net composite network, which combines elements of both deep learning and unsupervised learning for polishing radiologic images. RK-net has the ability to filter out irrelevant images, thereby reducing human factors that may influence the quality of the images. Besides removing the burden of manually screening images, RK-net ensures that the quality of the input images is high, thus improving the performance of deep neural networks. This provides an avenue for refining images within the medical field and building improved deep learning models.

V. Future Research Work

There are several avenues for extending the RK-Net framework, and some of them include the following:

One possible direction for extending the RK-Net framework is to experiment with the latest and best-in-class deep learning architectures, including their hybrids, EfficientNet, and Vision Transformers hybrids, which combine CNN and Transformers, to improve feature extraction and classification accuracy.The second possible direction for extending the RK-Net framework is to extend the framework to handle multimodal medical data, such as CT, MRI, and histopathology images, together to improve the accuracy of medical diagnosis.The third possible direction for extending the RK-Net framework is to make better use of the large amounts of unlabeled medical data through the incorporation of self-supervised and semi-supervised learning.The final possible direction for extending the RK-Net framework is to validate the RK-Net framework with larger and diverse clinical data from various medical centers.

References

[1] H. Sung et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA-Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.

[2] N. H. C. O. China, "Chinese protocol of diagnosis and treatment of colorectal cancer (2020 edition)," *Zhonghua Wai Ke Za Zhi*, vol. 58, no. 8, pp. 561–585, 2020, doi: 10.3760/cma.j.cn112139-20200518-00390.

[3] S. Harada and D. Morlote, "Molecular pathology of colorectal cancer," *Adv. Anatomic Pathol.*, vol. 27, no. 1, pp. 20–26, Jan. 2020, doi: 10.1097/PAP.0000000000000247.

[4] S. Shalev-Shwartz, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge Univ. Press, 2014.

[5] G. Rompianesi, F. Pegoraro, C. D. Ceresa, R. Montalti, and R. I. Troisi, "Artificial intelligence in the diagnosis and management of colorectal cancer liver metastases," *World J. Gastroenterol.*, vol. 28, no. 1, pp. 108–122, Jan. 2022, doi: 10.3748/wjg.v28.i1.108.

[6] K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti, and A. Madabhushi, "Artificial intelligence in digital pathology—New tools for diagnosis and precision oncology," *Nature Rev. Clin. Oncol.*, vol. 16, no. 11, pp. 703–715, Nov. 2019, doi: 10.1038/s41571-019-0252-y.

[7] S. Kuntz et al., "Gastrointestinal cancer classification and prognostication from histology using deep learning: Systematic review," *Eur. J. Cancer*, vol. 155, pp. 200–215, Sep. 2021, doi: 10.1016/j.ejca.2021.07.012.

[8] C. Wong et al., "MRI-based artificial intelligence in rectal cancer," *J. Magn. Reson. Imag.*, 2022, doi: 10.1002/jmri.28381.

[9] M. A. Van De Wiel, M. Neerincx, T. E. Buffart, D. Sie, and H. M. Verheul, "ShrinkBayes: A versatile R-package for analysis of count-based sequencing data in complex study designs," *BMC Bioinf.*, vol. 15, no. 1, p. 116, Dec. 2014, doi: 10.1186/1471-2105-15-116.

[10] C. Lovis, "Unlocking the power of artificial intelligence and big data in medicine," *J. Med. Internet Res.*, vol. 21, no. 11, Nov. 2019, Art. no. e16607, doi: 10.2196/16607.

[11] Y. Wang, X. He, H. Nie, J. Zhou, P. Cao, and C. Ou, "Application of artificial intelligence to the diagnosis and therapy of colorectal cancer," *Amer. J. Cancer Res.*, vol. 10, no. 11, pp. 3575–3598, 2020.

[12] H. Qiu, S. Ding, J. Liu, L. Wang, and X. Wang, "Applications of artificial intelligence in screening, diagnosis, treatment, and prognosis of colorectal cancer," *Current Oncol.*, vol. 29, no. 3, pp. 1773–1795, Mar. 2022, doi: 10.3390/curroncol29030146.

[13] Y. Hu et al., "Computed tomography-based deep-learning prediction of neoadjuvant chemoradiotherapy treatment response in esophageal squamous cell carcinoma," *Radiother Oncol.*, vol. 154, pp. 6–13, Jan. 2021, doi: 10.1016/j.radonc.2020.09.014.

[14] D. Jin et al., "DeepTarget: Gross tumor and clinical target, volume segmentation in esophageal cancer radiotherapy," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101909, doi: 10.1016/j.media.2020.101909.

[15] M. A. Berbís, J. Aneiros-Fernández, F. J. M. Olivares, E. Nava, and A. Luna, "Role of artificial intelligence in multidisciplinary imaging diagnosis of gastrointestinal diseases," *World J. Gastroenterol.*, vol. 27, no. 27, pp. 4395–4412, Jul. 2021, doi: 10.3748/wjg.v27.i27.4395.

[16] K. Raza and N. K. Singh, "A tour of unsupervised deep learning for medical image analysis," *Current Med. Imag. Formerly Current Med. Imag. Rev.*, vol. 17, no. 9, pp. 1059–1077, Sep. 2021, doi: 10.2174/1573405617666210127154257.

[17] A. Walker and P. Surda, "Unsupervised learning techniques for the investigation of chronic rhinosinusitis," *Ann. Otology, Rhinology Laryngology*, vol. 128, no. 12, pp. 1170–1176, Dec. 2019, doi: 10.1177/0003489419863822.

[18] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.

[19] A. K. Jain, "Data clustering: 50 years beyond K-means," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*, 2008, pp. 3–4.

[20] C. Chinrungrueng and C. H. Sequin, "Optimal adaptive K-means algorithm with dynamic adjustment of learning rate," *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 157–169, Jan. 1995, doi: 10.1109/72.363440.

[21] H. Xiao et al., "A review of deep learning-based three-dimensional medical image registration methods," *Quant. Imag. Med. Surgery*, vol. 11, no. 12, pp. 4895–4916, 2021, doi: 10.21037/qims-21-175.

[22] A. Kleppe, O.-J. Skrede, S. De Raedt, K. Liestøl, D. J. Kerr, and H. E. Danielsen, "Designing deep learning studies in cancer diagnostics," *Nature Rev. Cancer*, vol. 21, no. 3, pp. 199–211, Mar. 2021, doi: 10.1038/s41568-020-00327-9.

[23] C. Le Berre et al., "Application of artificial intelligence to gastroenterology and hepatology," *Gastroenterology*, vol. 158, no. 1, p. 76, 2020, doi: 10.1053/j.gastro.2019.08.058.

[24] R. Bommasani et al., "On the opportunities and risks of foundation models," 2022, arXiv:2108.07258.

[25] X. Gao, F. Shi, D. Shen, and M. Liu, "Task-induced pyramid and attention GAN for multimodal brain image imputation and classification in Alzheimer's disease," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 1, pp. 36–43, Jan. 2022, doi: 10.1109/JBHI.2021.3097721.