DEEPINSIGHT: ENDOSCOPIC IMAGE-BASED COLORECTAL CANCER CLASSIFICATION USING SEQUENTIAL CNN ARCHITECTURE

[1]Muthusamy P, [2]Arthi R, [3]Mowleshwaran K, [4]Arunkumar S, [5]Yasar S

[1]Professor, Department of CSE(Cyber Security), Muthayammal Engineering College.

[2]Assistant Professor, Department of CSE(Cyber Security), Muthayammal Engineering College.

[3,4,5]Student, Department of CSE(Cyber Security), Muthayammal Engineering College.

[1]mpmmuthu6@gmail.com, [2]arthiramcse@gmail.com, [3]mouly842@gmail.com,
[4]arunsaravanan0120@gmail.com, [5]iamyasar110@gmail.com

ABSTRACT - Colorectal cancer (CRC) is one of the most common and life-threatening cancers worldwide, where early detection significantly improves survival rates. Conventional diagnostic approaches such as colonoscopy and histopathological examination rely heavily on manual interpretation by clinicians, making the process time-consuming. This paper presents an artificial intelligence–based computer-aided diagnosis (CAD) system for automated colorectal cancer classification using endoscopic images. The proposed system employs a Sequential Convolutional Neural Network (CNN) architecture to classify colorectal images into normal, polyp, and cancerous categories. Image preprocessing techniques such as resizing, noise removal, normalization, and data augmentation are applied to enhance feature learning and improve model performance. Experimental results demonstrate improved classification accuracy, reduced false detection rates, and real-time diagnostic support.

KEYWORDS - Colorectal Cancer, Deep Learning, Convolutional Neural Network, Endoscopic Imaging, Medical Image Classification, Computer-Aided Diagnosis, Artificial Intelligence

1. INTRODUCTION

The Colorectal cancer is a major global health concern and ranks among the leading causes of cancer-related mortality. According to global cancer statistics, the incidence of colorectal cancer continues to rise due to lifestyle changes, aging populations, and dietary habits. Early detection plays a critical role in improving patient survival rates, as colorectal cancer is highly treatable when identified in its initial stages. Colonoscopy is the most widely used screening technique for colorectal cancer. However, manual inspection is highly dependent on the expertise of the clinician and may result in missed detections, particularly for small, flat, or early-stage lesions. Histopathological examination, though accurate, is labor-intensive and time-consuming, often delaying diagnosis and the

treatment. Recent advancements in artificial intelligence and deep learning have shown significant potential in automating medical image analysis. This study proposes a deep learning–based framework that leverages CNNs to automate colorectal cancer classification from endoscopic images, reducing diagnostic subjectivity and improving clinical efficiency.

1.1 SCOPE OF THIS PROJECT

This project aims to:

- Develop an AI-based colorectal cancer classification system using endoscopic image data.
- Implement a Sequential Convolutional Neural Network (CNN) architecture for automated diagnosis.
- Apply image preprocessing and data augmentation techniques to improve model generalization.
- Accurately classify colorectal conditions into normal, polyp, and cancerous categories.
- Reduce diagnostic errors caused by human subjectivity.
- Provide real-time diagnostic assistance during endoscopic procedures.
- Offer a scalable framework adaptable to other gastrointestinal disease classification tasks.

The proposed system is intended to support clinicians, researchers, and healthcare institutions by enhancing diagnostic accuracy and screening efficiency.

2. PROPOSED WORK

Several studies have explored artificial intelligence techniques for colorectal cancer detection and classification.

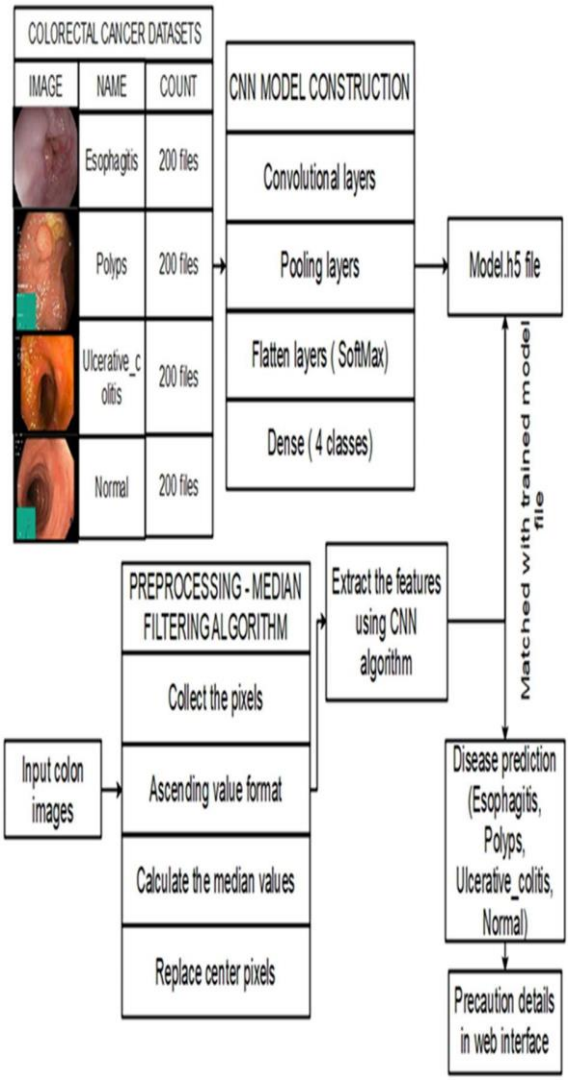
Recent research has focused on deep learning-based approaches, particularly CNNs, for automated colorectal cancer diagnosis. Studies have demonstrated that CNNs outperform traditional methods by learning discriminative features directly from image data.

The proposed work addresses these challenges by employing a lightweight Sequential CNN architecture combined with effective preprocessing and augmentation strategies to achieve reliable performance in clinical settings.

This paper proposes an AI-driven framework for automated colorectal cancer classification using endoscopic images.

The Sequential CNN architecture is designed to efficiently learn spatial and texture-based features from endoscopic images. The framework aims to assist clinicians by providing accurate predictions and reducing reliance on manual interpretation. The proposed system is optimized for real-time usage during endoscopic examinations.

2.1 System Architecture



3.1 Data Processing & Preprocessing

To ensure accurate feature learning and reliable colorectal disease classification, the Deepinsight framework incorporates the following data processing and preprocessing steps:

- **Image Data Collection:** Collects endoscopic colon images from publicly available colorectal cancer datasets, including four classes: Esophagitis, Polyps, Ulcerative Colitis, and Normal images.
- **Input Image Acquisition:** Accepts raw colonoscopy images as input for further processing.
- **Feature Extraction Using CNN:** Extracts discriminative spatial features from preprocessed images using sequential CNN layers.
- **Disease Prediction:** Classifies input images into Esophagitis, Polyps, Ulcerative Colitis, or Normal categories.

3.2 CNN-Based Classification and Disease Prediction

3.2.1 CNN-Based Feature Extraction and Classification:

The CNN-based module employs a sequential architecture to extract discriminative features from preprocessed endoscopic images. Convolution and pooling layers capture essential spatial patterns while reducing complexity, and dense layers with SoftMax activation enable accurate multi-class colorectal disease classification.

3.2.2 Colorectal Disease Prediction and Model Deployment:

The disease prediction module uses the trained CNN model to classify endoscopic images into Esophagitis, Polyps, Ulcerative Colitis, or Normal categories. The model is stored in H5 format, and results with precautionary details are displayed through a web-based interface.

3.3 TOOLS AND LIBRARIES

The Deepinsight framework utilizes the following tools and technologies:

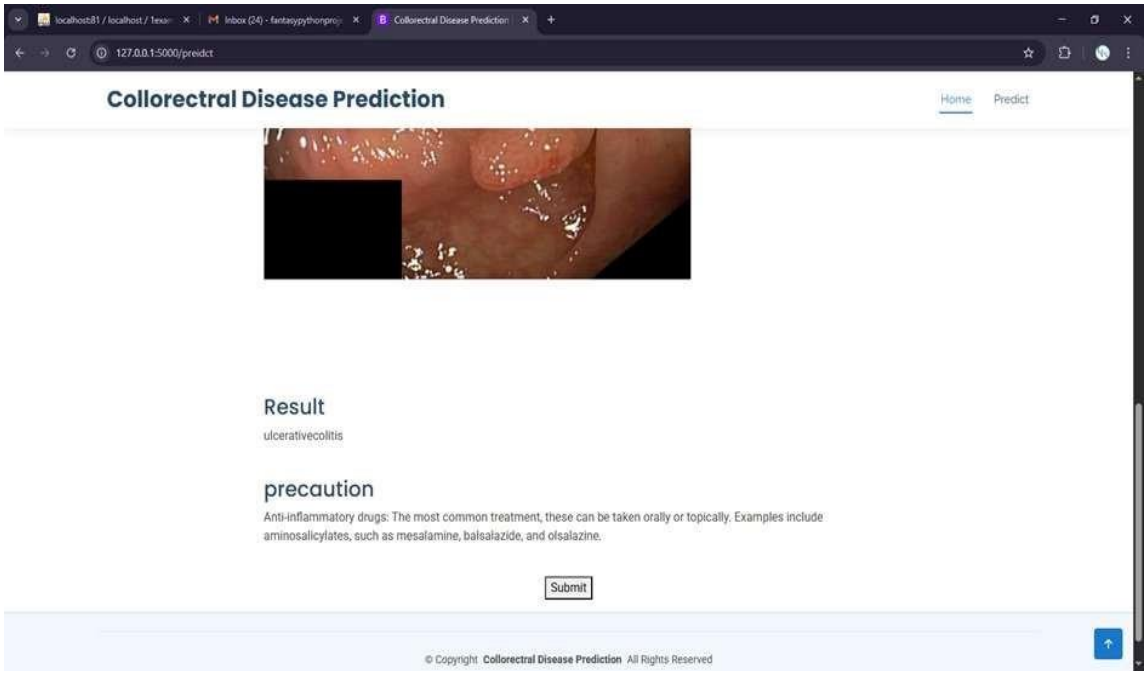
- **Python:** Primary programming language used for image processing, model development, and system integration.
- **TensorFlow & Keras:** Deep learning libraries used to design, train, and deploy the Sequential CNN architecture.
- **OpenCV:** Used for image preprocessing tasks such as resizing, noise removal, and enhancement.
- **NumPy:** Supports numerical computations and matrix operations for image data handling.
- **Matplotlib:** Used for visualizing images, training performance, and evaluation metrics.
- **Scikit-learn:** Provides utilities for dataset splitting and performance evaluation metrics.
- **PyCharm:** Integrated Development Environment (IDE) used for implementing and testing the project.

4. PROGRAM

```
import os
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense,
Dropout
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.optimizers import Adam
import matplotlib.pyplot as plt
import warnings

warnings.filterwarnings('ignore')

IMAGE_SIZE = (200, 200)
BATCH_SIZE = 64
EPOCHS = 20  # Increased for better learning
LEARNING_RATE = 0.0001  # Lower for stable training

# Define dataset paths
train_dir = os.path.abspath('Dataset/train')

if not os.path.exists(train_dir):
    raise ValueError(f"Dataset directory '{train_dir}' not found! Check your
file path.")

train_datagen = ImageDataGenerator(
    rescale=1/255,
    rotation_range=20,
    width_shift_range=0.2,
```

```
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    fill_mode='nearest',
    validation_split=0.2  # Splitting some data for validation
)

train_generator = train_datagen.flow_from_directory(
    train_dir,
    target_size=IMAGE_SIZE,
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='training'
)
```

5. RESULTS AND OUTPUT



Prediction Result: The system analyzes the uploaded colonoscopy image and predicts the detected condition.

- **Disease Identified:** The model classifies the condition as Ulcerative Colitis.
- **Precautionary Guidance:** Recommends anti-inflammatory treatment options such as aminosalicylates.
- **Clinical Support Output:** Provides decision support to assist clinicians with early diagnosis and care planning.

6. CONCLUSION

This paper presents an AI-based colorectal cancer classification system using a Sequential CNN architecture. The proposed framework enhances diagnostic accuracy, reduces human dependency, and provides real-time clinical support. The system demonstrates strong potential for integration into routine colorectal cancer screening and contributes to advancements in AI-driven medical diagnostics.

REFERENCES

[1] Lei, Z. N., Liu, Y., Zhang, J. Y., and Fu, L., "ABCB1-dependent collateral sensitivity of multidrug-resistant colorectal cancer cells to the survivin inhibitor MX106-4C," Drug Resistance Updates, vol. 73, pp. 101065, 2024.

[2] Maida, M., Macaluso, F. S., and Ianiro, G., "Screening and Surveillance of Colorectal Cancer: A Review of the Literature," Cancers, vol. 16, no. 15, pp. 2746, 2024.

[3] Papavassiliou, A. G., and Delle Cave, D., "Novel Therapeutic Approaches for Colorectal Cancer Treatment," International Journal of Molecular Sciences, vol. 25, no. 4, pp. 2228, 2024.

[4] Corrales, D., García, M., and Gómez, J., "Colorectal cancer risk mapping through Bayesian networks," Computer Methods and Programs in Biomedicine, vol. 237, pp. 108407, 2024.

[5] White, M. T., and Sears, C. L., "The microbial landscape of colorectal cancer," Nature Reviews Microbiology, vol. 22, no. 4, pp. 240–254, 2024.

[6] Dunne, P. D., and Arends, M. J., "Molecular pathological classification of colorectal cancer—An update," Virchows Archiv, vol. 484, no. 2, pp. 273–285, 2024.

[7] Litjens, G., Kooi, T., Bejnordi, B. E., et al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017.

[8] Esteva, A., Kuprel, B., Novoa, R. A., et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[9] Shin, H. C., Roth, H. R., Gao, M., et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1285–1298, 2016.

[10] Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., et al., "Convolutional neural networks for medical image analysis: Full training or fine tuning?" IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1299–1312, 2016.