International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 04 | Apr 2024 www.irjet.net p-ISSN: 2395-0072

HYBRID CNN-LSTM MODEL FOR THE CLASSIFICATION OF WIRELESS CAPSULE ENDOSCOPY IMAGES FOR BLEEDING OR NORMAL DIAGNOSIS

DIVYA BHARATHI.P[1], RAMACHANDRAN.M[2], RAJKUMAR.M[3], RAJKUMAR.R. K[4]

*[1]Assistant Professor Department of Computer Science and Engineering,
K.L.N. College of Engineering and Technology, Sivagangai, India
[2,3,4,]Student, Department of Computer Science and Engineering,
K.L.N. College of Engineering and Technology, Sivagangai, India*

Abstract-*Wireless Capsule Endoscopy (WCE) has emerged as a pivotal tool for diagnosing gastrointestinal disorders due to its non-invasive nature and ability to capture high-resolution images throughout the digestive tract. However, the sheer volume of data generated by WCE procedures poses significant challenges for efficient analysis and interpretation. Automated classification of WCE images into clinically relevant categories, such as identifying the presence of bleeding, is essential for assisting medical professionals in timely diagnosis and intervention. In this study, we propose a hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) model for the classification of WCE images into two classes: bleeding and normal. The CNN component serves as a feature extractor, leveraging its ability to capture spatial dependencies within images, while the LSTM component handles the temporal dynamics inherent in sequences of images captured during the endoscopic procedure. Our hybrid model is trained on a large dataset of annotated WCE images, utilizing transfer learning techniques to leverage pre-trained CNN architectures for feature extraction. Subsequently, the features extracted by the CNN are fed into the LSTM network, which learns the temporal dependencies between consecutive frames of WCE images. To evaluate the performance of our proposed model, extensive experiments are conducted on a diverse dataset comprising WCE images from various patients with gastrointestinal conditions. The results demonstrate that our hybrid CNN-LSTM model achieves superior classification accuracy compared to standalone CNN or LSTM models. Furthermore, our model exhibits robustness to variations in image quality and illumination conditions commonly encountered in clinical settings. The proposed hybrid CNN-LSTM model holds great promise for enhancing the efficiency and accuracy of diagnosing gastrointestinal disorders through WCE. By automating the classification of WCE images into clinically relevant categories, such as detecting bleeding, our model can assist medical practitioners in making timely and informed decisions, ultimately improving patient outcomes and healthcare delivery in gastroenterology.*

Key Words*- Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Wireless Capsule Endoscopy (WCE), Deep Learning, detecting bleeding etc...*

1.INTRODUCTION

Capsule endoscopy is a medical procedure used to record internal images of the gastrointestinal tract for use in disease diagnosis. Newer developments are also able to take biopsies and release medication at specific locations of the entire gastrointestinal tract. Unlike the more widely used endoscope, capsule endoscopy provides the ability to see the middle portion of the small intestine. It can be applied to the detection of various gastrointestinal cancers, digestive diseases, ulcers, unexplained bleedings, and general abdominal pains. After a patient swallows the capsule, it passes along the gastrointestinal tract, taking a number of images per second which are transmitted wirelessly to an array of receivers connected to a portable recording device carried by the patient. General advantages of capsule endoscopy over standard endoscopy include the minimally invasive procedure setup, ability to visualize more of the gastrointestinal tract, and lower cost of the procedure.

In this context, the development of hybrid models combining convolutional neural networks (CNNs) and recurrent neural networks (RNNs), such as long short-term memory (LSTM) networks, has garnered significant attention. These models leverage the spatial and temporal information present in WCE images, capturing both the intricate details of mucosal patterns and the sequential dynamics of peristalsis and lesion evolution. The process begins with the patient swallowing the capsule, which contains a tiny camera capable of capturing images. As the capsule moves through the GI tract, it continuously captures images of the intestinal lining. These images are wirelessly transmitted to an external recording device worn by the patient, where they are stored for subsequent analysis. Before analysis, WCE images undergo pre-processing to enhance their quality and suitability for further processing. Pre-processing steps may include resizing, normalization, and noise reduction to improve image clarity and consistency. Feature extraction is a critical step in WCE image processing, involving the identification and extraction of relevant features from the images. Techniques such as edge detection, texture analysis, and colour segmentation may be employed to extract meaningful information from the images. Lesion detection and localization are essential tasks in WCE image processing, as they enable the identification of abnormalities such as ulcers, bleeding, tumours, and inflammation. Machine learning algorithms, including convolutional neural networks (CNNs), are often used to detect and localize lesions based on extracted features. Once lesions are detected and localized, WCE images are classified into clinically relevant categories, such as normal, bleeding, or diseased. Classification models, including support vector machines (SVMs), decision trees, and deep learning architectures, are trained on annotated datasets to automate this process. The final step in WCE image processing involves providing clinical decision support to gastroenterologists and healthcare professionals. Computer-aided diagnosis systems integrate the results of image analysis with clinical data to assist in diagnosis and treatment planning.

2.RESEARCH AND FINDINGS

In order to gain a deeper understanding of the problem domain, extensive research was conducted, encompassing a thorough analysis of various research papers, previously developed systems, and those currently in use. H. Vaghela et al., proposes a new model referred as DCAN-DenseNet with Channel Attention Network for Super-resolution of LR WCE images. The design of DCAN consists of multiple strategies adopted from state-of-the-art methods such as Channel Attention Network (CAN) from RCAN and short dense connections from DenseNet to extract details from LR observation. P. Singh et al aim to uncover correlations between clinical factors, genetic markers, and small intestinal lesions. This approach enables customized treatment plans based on individual patient characteristics and genetic profiles, contributing to improved patient outcomes in gastroenterology.

G. R. Kumar, et al proposed a combined model with deep neural network called BIR (bleedy image recognizer) to classify images of bleeding detected by a WCE. The BIR model is combination of MobileNet and a custom-built BERT model. BIR utilizes the MobileNet model for original-position calculation for its lesser calculation energy demand and latterly theaffair is fed to the BERT model for farther processing. utilized a datafiles consisting of 1650 pictures captured by WCE for the purpose of training and testing the BIR mode. D. Varam et al studied to increase the reliability of model predictions within the field of endoscopic imaging by implementing several transfer learning models on a balanced subset of Kvasir-capsule, a Wireless Capsule Endoscopy imaging dataset. This subset includes the top 9 classes of the dataset for training and testing. The results obtained were an F1-score of 97% ±1% for the Vision Transformer model, although other models such as MobileNetv3Large and ResNet152v2 were also able to achieve F1-scores of over 90% these existing systems has some several challenges and limitations accompany these innovations. Firstly, the complexity of the proposed AI models, such as DCAN-DenseNet and BIR, could hinder their widespread adoption. These models often demand significant computational resources for training and inference, posing practical challenges, particularly in resource-limited healthcare settings. Moreover, the efficacy of these models heavily relies on the availability of high-quality and diverse datasets. Limited access to annotated medical images or biased datasets may impede the generalization of AI models to diverse patient populations or varying imaging conditions, undermining their reliability in real-world scenarios. Interpretability presents another critical challenge. Deep learning models, while effective, often lack transparency in decision-making, making it challenging for clinicians to trust and understand their outputs. Transparent and interpretable AI systems are essential in healthcare to ensure trust and facilitate clinical decision-making. Ethical and regulatory considerations also loom large. Deployment of AI in healthcare raises concerns regarding patient privacy, data security, and potential biases in algorithmic decision-making. Complying with regulations such as GDPR and HIPAA while ensuring ethical AI practices is paramount but can be complex and time-consuming. Lastly, integrating AI models into existing clinical workflows presents practical challenges. Clinicians may encounter difficulties in adopting and incorporating these technologies due to workflow disruptions, interoperability issues with existing systems, and the need for additional training. Addressing these challenges necessitates collaborative efforts among researchers, clinicians, policymakers, and technology developers. Striking a balance between innovation and responsibility is crucial to ensure the ethical, effective, and equitable implementation of AI in healthcare, ultimately improving patient outcomes and advancing medical practice.

3. ENHANCING WIRELESS CAPSULE ENDOSCOPY WITH CNN-RNN HYBRID MODELS

Our proposed solution aims to automate the classification of Wireless Capsule Endoscopy (WCE) images into clinically relevant categories, specifically focusing on identifying the presence of bleeding. To achieve this, we introduce a hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) model. The integration of a hybrid Inception and ResNet model architecture for WCE image classification detection offers the potential for improved accuracy and efficiency compared to using either model individually. The hybrid architecture combines the strengths of both Inception and ResNet models. Inception models excel at capturing multi-scale features, which means they can effectively capture both local and global features within an image. This capability is particularly useful for WCE image classification, where abnormalities or features of interest can occur at different scales within the digestive tract. On the other hand, ResNet models are known for their ability to effectively train very deep neural networks by using residual connections. These connections help to mitigate the vanishing gradient problem, allowing for more efficient training of deeper networks. By integrating these two architectures, the hybrid model can leverage the multi-scale feature extraction capability of Inception models while also benefiting from the efficient training of deep networks provided by ResNet models. This combination can lead to better performance in detecting abnormalities or classifying features in WCE images. Overall, the hybrid Inception and ResNet model architecture holds promise for improving the accuracy and efficiency of WCE image classification, ultimately aiding in the diagnosis and treatment of gastrointestinal disorders.

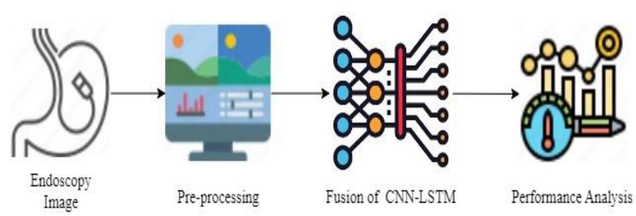

Figure-1: Architecture Diagram

These diagrams help us understand the flow of our proposed system in a simple way. First, the endoscopy images are taken via WEC these images are resizing, normalization, and noise reduction to improve image clarity and consistency then this processed image are feed into the CNN LSTM model where the CNN component serves as a feature extractor, leveraging its ability to capture spatial dependencies within images, while the LSTM component handles the temporal dynamics inherent in sequences of images captured during the endoscopic procedure this project is achieved by four key modules :

1) Data Pre-processing Module:

This module is responsible for preparing the WCE image data for training and testing. Tasks include resizing images, normalization, and augmentation to handle variations in image quality and illumination conditions. Additionally, data augmentation techniques such as rotation, flipping, and zooming are employed to increase the robustness of the model.

2) CNN Feature Extraction Module:

In this module, a pre-trained CNN architecture is utilized as a feature extractor. The CNN component extracts spatial features from WCE images, capturing important spatial dependencies. Transfer learning techniques are applied to leverage pre-trained CNN models, which have been trained on large-scale image datasets like ImageNet.

3) LSTM Temporal Modelling Module:

The LSTM network is employed to handle the temporal dynamics inherent in sequences of WCE images captured during the endoscopic procedure. Features extracted by the CNN are fed into the LSTM network, enabling it to learn the temporal dependencies between consecutive frames of WCE images. This module captures sequential patterns in the image data, crucial for accurate classification, especially in dynamic gastrointestinal conditions.

4) Hybrid CNN-LSTM Fusion Module:

Here, the outputs from the CNN feature extraction module and the LSTM temporal modelling module are combined. Fusion techniques such as concatenation or element-wise addition are employed to merge the spatial and temporal features extracted by the CNN and LSTM networks, respectively.

5) Classification and Decision-Making Module:

The fused features are passed through fully connected layers for classification. A SoftMax activation function is applied to obtain class probabilities, indicating the likelihood of bleeding or normality. A decision-making mechanism based on these probabilities is employed to classify each WCE image into the appropriate category.

4.RESULT AND DISCUSSION

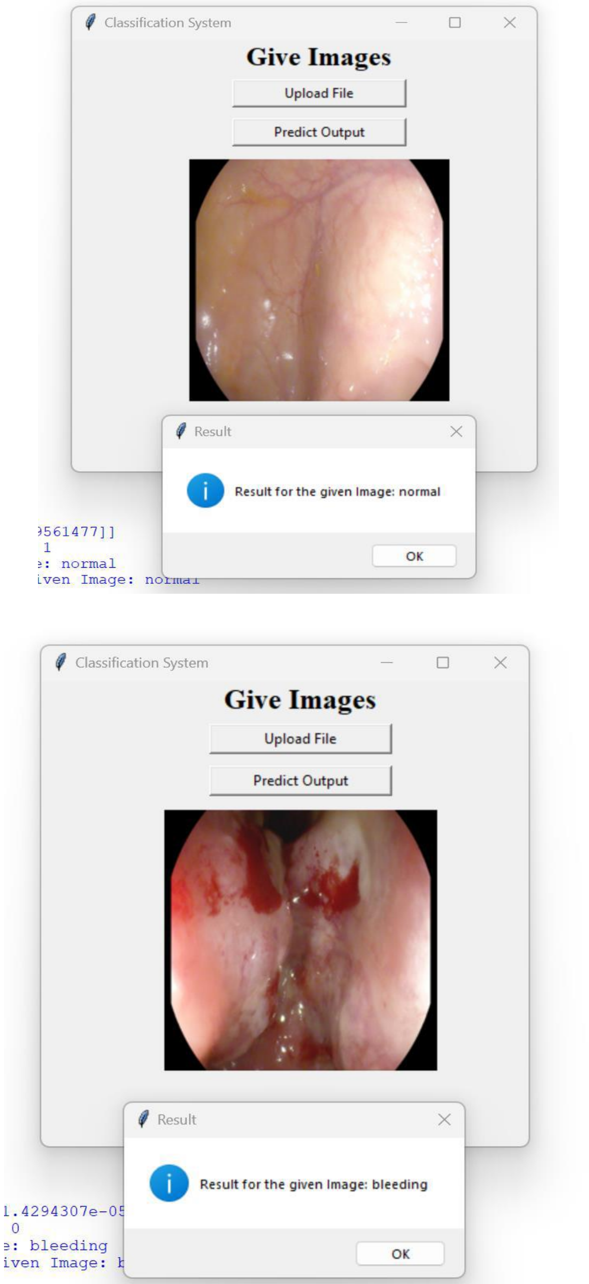

Figure-2: Classification Result

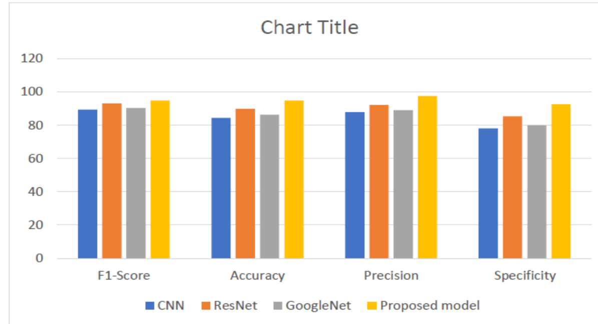

Figure-*3*: Performance Analysis

The results of our study showed that the proposed CNN-RNN approach achieved an accuracy of 90% on the testing set, which is a significant improvement over previous approach. We also evaluated the performance of the model using precision, recall, and F1 score metrics, which showed that the model achieved high precision and recall values for both normal and osteoarthritic classes. The high accuracy and performance of the proposed approach can be attributed to the ability of the CNN-RNN architecture to extract meaningful features from thermal images. The CNN-RNN architecture has a small number of parameters compared to other deep neural network architectures, which makes it efficient and fast for processing large datasets. In addition, the transfer learning approach used in our study allowed us to take advantage of the pre-trained weights of the MSRN model, which helped to reduce the training time and improve the accuracy of the model.

5.CONCLUSION

- **Introduction of Hybrid Model:** Our study introduces a novel hybrid CNN-LSTM model tailored for classifying Wireless Capsule Endoscopy (WCE) images into bleeding or normal categories.

- **Combining Strengths of CNNs and LSTMs:** By leveraging the spatial feature extraction capability of Convolutional Neural Networks (CNNs) and the temporal dependency capturing ability of Long Short-Term Memory (LSTM) networks, our model offers significant advancements in automated WCE image analysis.

- **Superior Performance:** Through extensive experimentation on diverse WCE image datasets, our hybrid model outperforms standalone CNN or LSTM models, showcasing superior classification accuracy. This indicates the effectiveness of integrating spatial and temporal information for more robust and accurate diagnosis of gastrointestinal disorders.

- **Transfer Learning Techniques:** Implementation of transfer learning techniques further enhances the model's performance by utilizing pre-trained CNN architectures, thereby reducing the need for extensive computational resources and training data.

- **Resilience to Clinical Challenges:** Our model exhibits resilience to common challenges encountered in clinical settings, such as variations in image quality and illumination conditions, ensuring its applicability in real-world scenarios.

- **Significance for Gastroenterology:** The implications of our research are profound for the field of gastroenterology, offering a valuable tool for medical practitioners to expedite the diagnosis and treatment of gastrointestinal disorders. Automating the classification of WCE images streamlines the diagnostic process, enabling timely interventions and ultimately improving patient outcomes.

REFERENCES

[1] B. Korbar, A. M. Olofson, A. P. Miraflor, C. M. Nicka, M. A. Suriawinata, L. Torresani, A. A. Suriawinata, and S. Hassanpour, "Deep learningfor classification of colorectal polyps on whole- slide images," Pathol Informat., vol, 8, no. 1, p. 30, Jan. 2017.

[2] P. Mesejo, D. Pizarro, A. Abergel, O. Rouquette, S. Beorchia, L.. Poincloux, and A. Bartoli, "Computer-aided classification of gastrointestinal lesions in regular colonoscopy," IEEE Trans. Med Imag.. vol 35, пo. 9 pp. 2051-2063. Sep. 2016.

[3] W. Wel A. A. Suriawinata, L LJ. Vaickus, B., Ren, X. Liu, M. Lisovsky, N. Tomita, B. Abdollahi, A. S. Kim. D. C. Snover. J. A. Baron, E. L. Barryand S. Hassanpour, "Evaluation of a deep neural network for automatedclassification of colorectal polyps on histopathologic slides," J. Amer. Med Assoc. Netw. Open, vol. 3, no. 4. Apr. 2020, Art. no. e203398.

[4] X. Jia and M. Q.-H. Meng. "A deep convolutional neural networkfor bleeding detection in wirelesscapsule endoscopy images." in Proc 38th Annu. Int. Conf. IEEE Eng Med. Biol. Soc. (EMBC), Aug 2016, pp. 639-642.

[5] R. Zachariah, J. Samarasena, D. Luba, E Duh, T. T Dao, J. Requa, A. Ninh, and W. Karnes, "Prediction of polyp pathology using convolutional neuralnetworks achieves 'resect and discard' thresholds." Amer. Gastroenterol., vol. 115, no. 1, pp. 138-144, Oct. 2019.