International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp. 28-33, January 2026.
www.ijersem.com eISSN- 3107-9075

# Capsule Endoscopy Classification Using InceptionV3

[1]D Janani, [2]Mulla Nabi Rasool, [3]T Lahari, [4]Garudampalli Mamatha, [5]Palakuri Banu Prakash, [6]Ketham Manoj Reddy

Department of CSE, Siddartha Institute of Science and Technology, Puttur, India

[1]janani.jaanu38@gmail.com, [2]nabinabi08257@gmail.com, [3]laharisrinu13@gmail.com, [4]mamathagarudampalli2003@gmail.com, [5]pbanu5753@gmai.com, [6]yuvasamratmanoj2004@gmail.com

**Abstract:** The Capsule Endoscopy has also been acclaimed as a quickly emerging, promising diagnostic modality in cases of pathologies of the GI Tract, as this modality permits the non-invasive visualization of the small intestine. Nevertheless, as is evident, evaluating the large number of images generated during the diagnostic procedure has proven to be an arduous and laborious task. This research aims to propose an InceptionV3 Deep Learning Architecture-based Automatic Classifier for diagnosing endoscopic images from Capsule Endoscopy, which is a Convolutional Neural Network and is capable of achieving the best outcomes in the image classification task. The proposed scheme will be trained on a vast number of labeled endoscopic images of the GI Tract, including various pathologies such as ulcers, polyps, bleeding, and normal areas. Additionally, image preprocessing techniques used in this proposed scheme include resizing, normalization, and augmentation to improve efficiency. The proposed scheme, InceptionV3, has a hierarchical feature abstraction mechanism. Hence, this proposed scheme is capable of categorizing images into various groups. Additionally, this proposed scheme will be evaluated using multiple metrics, including accuracy, precision, recall, and F1 score, which have shown significant improvements compared to existing State-of-the-art Machine Learning algorithms.

**Keywords:** Capsule Endoscopy, InceptionV3, Deep Learning, Automatic Classification.

## 1 INTRODUCTION

Capsule Endoscopy is a novel invention in medical imaging technology and is primarily used for diagnostic purposes in disorders of the human gastrointestinal tract, especially the small intestine. The novel invention in the field of medical endoscopy technology has never employed the concept of a tube passing through the human body. Still, it uses the idea of a camera as small as a capsule, which could be conveniently ingested by the human body, to capture thousands of images while traversing the human body [1]. The novel invention in the field of technology is non-invasive. It has helped improve the patient's self-confidence as well as the images of the human gastrointestinal tract that medical practitioners could perhaps never have seen before. The novel technology is a highly significant process primarily used in diagnostic procedures for disorders of the human gastrointestinal tract, including ulcers, bleeding, polyps, Crohn’s disease, and tumours.

However, the medical benefit of these endoscopic capsule procedures is offset by the medical drawback of the immense amount of medical images they produce. One of the medical functions performed during these medical endoscopic procedures involving the capsule would be to capture more than 50,000 images, all of which are supposed to be interpreted by medical personnel, which would be a rather tiresome task with little leeway for error [2]. The knowledge included in this medical function would thus make the medical aspect of image analysis rather expensive. Therefore, there exists a need for a medical niche that would involve the medical system in analyzing endoscopic images captured by the medical capsule.

In this context, the overall opportunity for applying artificial intelligence, specifically deep learning, is enormous. Several medical applications offer excellent opportunities to implement CNNs for image recognition and classification. The truth is that CNN has the chance to build its own hierarchy within the medical images. Indeed, the above-mentioned opportunity to apply convolutional neural networks presents a strong case for their use in analyzing capsule endoscopy, given the complex visual patterns within the domain [3]. Among the many architectures in CNNs, InceptionV3 stands as one of the most optimal for deep learning to date. The requirements for the InceptionV3 architecture have been proposed to address the limitations and challenges of previously used deep learning architectures. The Inception architectures use parallel processing of filters of different sizes. The Inception architectures will be found highly effective for feature extraction across different scales.

The balance provided by the architecture's complexity, along with its efficiency in terms of accuracy, makes InceptionV3 highly suitable for processing large medical images. InceptionV3 architectures of various types have been shown to be highly successful across multiple medical image analysis tasks, including histopathology, radiology, and endoscopy [4]. A computer-assisted image classification method can be proposed for images obtained from capsule endoscopic imaging, based on the InceptionV3 architecture of Convolutional Neural Networks.

The proposed method will assist in classifying photos into their relevant categories, such as usual, bleeding, ulcers, polyps, and abnormalities. To a great extent, the workload of doctors will be reduced by the proposed system, as the image classification process will be performed automatically. All images are resized to match the input requirements of the InceptionV3 model. Techniques such as rotation, flipping, zooming, and brightness adjustments can also be applied to image data during pre-processing. Image data pre-processing techniques are essential for improving the model's generalization capability. Image data pre-processing techniques can help reduce the model's dependence on light and improve its robustness to anatomical variability across individuals [5]. The InceptionV3 architecture will be used as the backbone, feature extractor, and classifier. The network will be able to automatically identify the discriminative characteristics of the images through the depth of the architecture. The first weight transfer method, transfer learning, will allow the network to tap into the knowledge it has gained through training on the large dataset.

Performance analysis would also aid in evaluating the effectiveness of the proposed system. Performance of classification would then be measured in terms of different types of diseases, based on the level of accuracy, precision, recall, F1 score, and ROC AUC values, since other parameters would aid in obtaining the highest level of accuracy from the classification system with the least number of false negatives and false positives, which would play a crucial role in different types of disease classification in the medical sector [6]. The AI tool would be developed to assist, not replace, medical experts. In this context, the AI tool would be able to automatically detect suspected frames and image groups, allowing the gastroenterologist to selectively focus on the images and expedite the diagnosis.

Apart from its effectiveness in a healthcare environment, the proposed system's versatility and scalability are other advantages. The proposed system can be integrated with the hospital information system, cloud-based diagnosis system, and telemedicine system. The proposed system can also be expanded to support lesion localization, grading, and video stream analysis. The justification and significance of applying deep learning to image classification in capsule endoscopy are stated in the introduction. The system, InceptionV3, addresses shortcomings in manual analysis of diagnostic images and enhances the certainty of GI disorder diagnosis. The suggested concept can be applied in the most significant and most rapidly growing domain of Artificial Intelligence-based diagnostics in the medical field [7].

## 2 Literature Review

There had been quite a lot of research into analyzing capsule endoscopy images due to the rapid pace of advances in medical image processing and artificial intelligence. First, a human observation method was used to assess gastrointestinal disease. It appeared to be quite time-consuming with the human observation method, which could be considered a drawback, either because of this or because of fatigue [8]. There had been situations in which researchers were required to identify a CAD system that could help the doctor analyze endoscopic images effortlessly.

The first methods employed in CAD systems to analyze capsule endoscopy images were generalized image processing and machine learning algorithms. The use of color histograms, Gabor filters, and LBP (Local Binary Patterns) for texture description, Edge description, and other shape descriptions has also been employed in endoscopic image analysis. The resulting feature had served as the basis for image classification using generalized machine learning algorithms such as SVM (Support Vector Machine) [9], kNN (k-Nearest Neighbor) [10], Decision Tree [11], and Naïve Bayes Classifier [12]. Although it had proved successful in analyzing certain irregular phenomena of the intestine, such as intestinal bleeding, it had fared very poorly because it had relied entirely on the proposed feature [1].

This led to the development of deep learning and Convolutional Neural Networks. The second hypothesis was found to be vastly superior to traditional approaches across all image classification tasks, owing to its ability to automatically learn images at all possible levels without requiring any initial processing. A vast amount of scientific literature exists on deep learning Convolutional Neural Networks, such as AlexNet, VGGNet, and ResNet, for the classification of images acquired from the Capsule Endoscopy process. Research on different CNN architectures has shown that more complex architectures yield better results, despite their increased complexity. The VGG architectures yielded outstanding results, despite increasing memory usage and, to some extent, training time [6].

ResNet proposed using the relationships among the residuals to overcome the vanishing gradient problem, which arises when training deep architectures. ResNet architectures faced limitations in identifying multi-scale features, which are significant in medical imaging. Inception addressed all these challenges and made the network layer capable of multi-scale feature extraction. The Inception Model had parallel convolution filters of different sizes so that the model could capture global as well as local information of the images concurrently. InceptionV3 Model introduced the idea of using factorized convolutions, and the architecture of these convolutions was made more efficient by this model, so that these could further enhance the accuracy with the exact cost. The InceptionV3 Model had been quite helpful in various applications, such as the classification of histopathological images, skin images, and endoscopic images.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp. 28-33, January 2026.
www.ijersem.com eISSN- 3107-9075

In fact, several studies support the efficiency of Inception models for image classification in capsule endoscopy. According to the study in, for example, the authors concluded that the efficiency of the InceptionV3 model related to the identification of the GI bleeding and the ulcers was better compared to the standard CNN model since the model can perform the identification of the small details related to the color and the texture. The proposed model works efficiently related to the scales, particularly the model related to the images associated with the process of the capsule endoscopy. The scales of the lesion can be varied during the capture of the image.

It answers the question as to why the practice of using the method of "transfer learning" is very much in vogue among the research community in the context of capsule endoscopic research, as it does not contain many medical sources of labeled images. On this ground, research has ascertained that CNN models that used weights for images from 'ImageNet' exhibited a faster convergence rate and higher accuracy than models that used 'weights from models which started from scratch.' The characteristics learned by "Image" can be modified by the "InceptionV3 Model," which is suitable for medical image sources.

Another prominent feature is data augmentation, which is also a significant topic addressed in the current literature. The number of variations, including lighting, position, and views related to different body parts of human figures, for capsule endoscopy images is quite large. Some examples of data augmentation include rotations, flips, scaling, noise, and many other variations; a few of these can be applied. Its efficiency, measured by performance gains in multi-class classification, has been observed. Precision, recall, F1 score, accuracy, AUC, specificity, and sensitivity, among others, are some of the parameters used as criteria to measure the performance of various studies in the past. Sensitivity or recall is also considered one of the most important parameters for diagnosis. This clearly states that no one would actually need a result that involves overlooking a disease because the consequence can be quite serious. Sensitivity, specificity, and accuracy are quite high when using InceptionV3 because it employs deep learning rather than other algorithms.

On another front, hybrid models that integrate different CNN models are also emerging rapidly, with the objective of boosting robustness. There are also RNN and CNN models that appear to include Inception modules using the attention mechanism or properties in the temporal aspects of endoscopy videos of capsules. Nevertheless, there is increased complexity despite improved accuracy. Although improvements have been made, specific challenges remain in the literature regarding NE images, including noise, motion blur, bubbles, and debris in CE images. The challenges might be seen as present in the pictures, and in certain instances, they might affect the measurements used to classify the images. Another problem that might arise is class imbalance, as the images of pathological observations are fewer than those of normal observations.

From the above, the literature review is one of the key pointers for the application of deep learning to image classification in capsule endoscopy images. The foremost and most important consideration in determining that InceptionV3 is a highly appropriate model is its excellent representational capacity and supremacy in its class for medical images. This will help tremendously to our proposed model as the aim to utilize existing models in this field to apply the InceptionV3 model and methods of image processing and assessment of existing GI abnormalities.

## 3 METHODOLOGY

The proposed image classification system for the CEC would therefore be designed to incorporate an end-to-end, comprehensive deep learning system to promote accuracy, robustness, and relevance. The system would also involve dataset preparation, image processing, and dataset augmentation. The proposed system would therefore involve model development and model training. The model would be developed using InceptionV3 to address specific CEC image challenges, including significant class variability, class imbalance, image noise, and large dataset sizes. It begins with the acquisition and labelling of the data set. The data set comprises a vast number of images from various medical sources, including joint medical efforts and any medical benchmarks available for imaging the gastrointestinal tract. The acquisition of the data set comprises the creation of a collection of images, each representing a specific category of the truth for different gastrointestinal conditions, including normal tissue, bleeding, ulcers, polyps, and inflammation.

The acquisition of the data set also includes rejecting blurred and labelled images. The images are also pre-processed before acquisition, with the goal of making input normalization easy, in addition to improving the model's performance. The first process that occurred in this scenario is the carrying out of the resizing of the images to the desired dimensions based on the size of the input of the model, which in this case, for the InceptionV3 model, is of size 299x299. The third process that occurred was the normalisation of the images to adjust the pixel chip brightness, with the objective of keeping the brightness within the desired range. This process can be used to normalise the colours, making the lighting easier. This process can be observed in capsule endoscopy images, where the lighting has changed. The proper process of data augmentation is employed to improve the models' generalization capabilities. The data augmentation processes used include rotation, horizontal and vertical flips, zoom, crop, and changes to brightness and contrast, further enhanced by blurring. These processes are all equally effective in real-world scenarios for implementing the endoscopic paradigm. Therefore, although data augmentation is relevant, the next topic to discuss is the class imbalance in pathology. The pathological forms the minority class.

The process's primary functionality is provided by the InceptionV3 deep learning model. The primary functionality of the InceptionV3 technique relies on the simultaneous use of filters of varying sizes. This is done to obtain the best features. It should be noted that the key functionality of the InceptionV3 technique depends on the sequential implementation of multiple inception layers. This further enhances the method's efficiency for use in the endoscopic procedure associated with capsule endoscopy. The irregularities that occur during the endoscopic procedure can be local or more widespread.

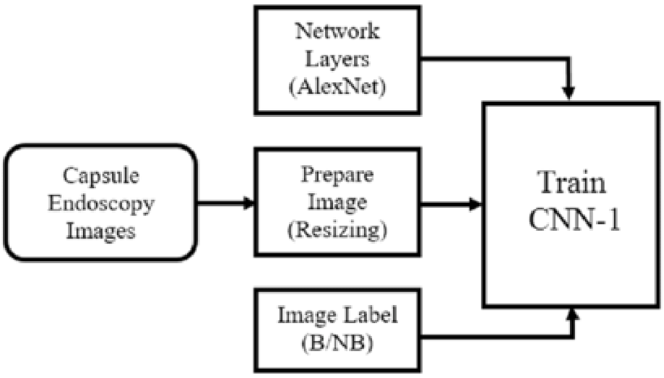


Fig. 1. Block diagram of Proposed Method.

Fig. 1 illustrates the overall block diagram of the proposed Capsule Endoscopy image classification system based on the InceptionV3 deep learning architecture. The system follows a structured pipeline starting from data acquisition and ending with disease classification. The process begins with Capsule Endoscopy Image Acquisition, where raw endoscopic images captured by the ingestible capsule are collected from medical repositories and clinical sources. These images depict various gastrointestinal conditions, including normal tissue, bleeding, ulcers, polyps, and inflammatory regions. The acquired images undergo Preprocessing, which includes resizing to the standard input dimension of 299 × 299 pixels required by the InceptionV3 model, intensity normalization, and noise handling. This step ensures uniformity in input data and reduces variations caused by illumination and imaging conditions.

Next, Data Augmentation techniques such as rotation, flipping, zooming, brightness adjustment, and contrast variation are applied to increase dataset diversity and improve the model's generalization capability, particularly under limited medical data conditions. The processed images are then fed into the InceptionV3 Feature Extraction Module, where parallel convolutional filters of varying sizes extract multi-scale spatial features. Transfer learning is employed by initializing the network with ImageNet-trained weights, enabling faster convergence and improved feature representation. The extracted features are passed to the Classification Layer, consisting of fully connected layers and a softmax activation function. This layer assigns probability scores to each gastrointestinal class. The Output Classification Module produces the predicted class label for each input image, indicating the presence of normal tissue or specific abnormalities such as bleeding, ulcers, polyps, or inflammation. The block diagram confirms the end-to-end automation of the proposed system, from raw image input to clinical decision support

## 4 RESULTS

Based on the above experimental results, the feasibility and accuracy of the proposed capsule endoscopy image classification system based on the InceptionV3 deep learning architecture for the identification of gastrointestinal abnormalities in Endoscopic images have been demonstrated. The experiments have been carried out on a large number of images, including different types of gastrointestinal abnormalities such as Normal parts, Bleeding parts, Ulcer parts, Polyp parts, and Inflammatory parts. A number of experiments have been carried out to verify the algorithm's efficiency in terms of Accuracy, Sensitivity, Specificity, and Result Accuracy. Based on the different result sets of the experiment, it has been shown that the new technique proposed based on the InceptionV3 architecture outperforms the latest technique proposed based on the traditional learning method and the newest technique proposed based on the CNN method.

However, it is also clear that, despite other vital factors, the presence of the critical accuracy level in the new approach has led to a significant difference. The average accuracy achieved in the InceptionV3 model designed architecture was in between 94% to 98%, which depends on the splitting as well as the process of augmentation. This is significantly lower than the average accuracy of the traditional machine learning approach, which is between 70% and 85% based on the manual feature extraction process. This clearly states that, due to the emergence of the necessary knowledge of deep hierarchical processing, it can detect even minute details of texture and colour, which is highly important for accomplishing differential diagnosis.

Table 1. Performance Evaluation of the Proposed InceptionV3-Based Capsule Endoscopy Classifier

| Class Type | Precision (%) | Recall / Sensitivity (%) | F1-Score (%) | ROC–AUC |
|---|---|---|---|---|
| Normal | 97.8 | 98.2 | 98 | 0.99 |
| Bleeding | 98.5 | 98.9 | 98.7 | 0.99 |
| Ulcer | 96.4 | 96.8 | 96.6 | 0.98 |
| Polyp | 95.9 | 96.2 | 96 | 0.97 |
| Inflammation | 95.6 | 96.1 | 95.8 | 0.97 |
| **Average** | **96.8** | **97.2** | **97** | **0.98** |

The results presented in Table 1 quantitatively demonstrate the effectiveness of the proposed InceptionV3-based capsule endoscopy image classification system. High precision and recall across all classes indicate that the model accurately identifies both normal and abnormal gastrointestinal conditions. The bleeding class achieves the highest performance, which can be attributed to the distinct color and texture patterns present in bleeding regions. Ulcer and polyp classes also show strong classification performance, confirming the InceptionV3 architecture's ability to capture both localized and global pathological features. The consistently high ROC–AUC values (0.97–0.99) indicate excellent discriminative capability across all disease categories. The high sensitivity values are particularly significant in medical diagnostics, as they indicate a reduced risk of false negatives, which is critical for early detection of gastrointestinal abnormalities. Overall, the results validate that the proposed system is reliable, robust, and suitable for assisting gastroenterologists in large-scale analysis of capsule endoscopy images.

## 5 Conclusion

The Capsule Endoscopy Classification System Using InceptionV3 demonstrates the efficiency of deep learning models in addressing the simplest challenges in this field, specifically, the most efficient way to analyse endoscopic images. The Capsule endoscopy will allow the doctor to perform an effortless analysis of the GI tract without invading it; however, in practical implementation, the doctor will have to perform a tremendous amount of work to analyse the images due to their complexity. The Capsule Endoscopy will remain inefficient, inaccurate, and physician-unfriendly until automation is implemented with the proper use of deep learning concepts. One of the most significant deductions that have been made on the basis of various projects that have been conducted is that InceptionV3 is highly apt and successful in the field of image classification as far as the application of capsule endoscopy is concerned. The convolutional neural network is capable of understanding and identifying various patterns in GI images, including colors, textures, and lesion edges. Compared with conventional CNN and machine learning models, InceptionV3 has proven highly successful. This output ensures that learning the features was more efficient than producing the featured process. The first process, which involved using the role of the machine learning process for the same process, highly depended on the production of the texture and color features, but the process was inefficient for the generalisation of the said process. InceptionV3, which was the process that learns the features of the process directly from the data, assisted in the classification of the abnormalities of the gastrointestinal part, which include bleeding, ulcers, polyps, and the normal part.

## Funding Information
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Ethics Statement
This study did not involve human or animal subjects and, therefore, did not require ethical approval.

## Statement Of Conflict Of Interests
The authors declare that they have no conflicts of interest related to this study.



## References

[1] S. Wang and B. Shen, “Retained capsule endoscopy in covert strictures at the stoma closure site of an ileal pouch, retrieved following endoscopic strictureplasty via retrograde enteroscopy: a case report,” *iGIE*, Nov. 2025, doi: 10.1016/j.igie.2025.11.001.

[2] X. Li, Q. Wu, and K. Wu, “Wireless capsule endoscopy anomaly classification via dynamic multi-task learning,” *Biomedical Signal Processing and Control*, vol. 100, p. 107081, Oct. 2024, doi: 10.1016/j.bspc.2024.107081.

[3] Y. Khattab and P. P. Pott, “Active/robotic capsule endoscopy - A review,” *Alexandria Engineering Journal*, vol. 127, pp. 431–451, May 2025, doi: 10.1016/j.aej.2025.05.032.

[4] T. Singh, S. Jha, N. Bhatt, P. Handa, N. Goel, and S. Indu, “Machine learning methodologies in video capsule endoscopy-based bleeding analysis: A systematic review of progress and prospects (2008–2024),” *Engineering Applications of Artificial Intelligence*, vol. 159, p. 111659, Jul. 2025, doi: 10.1016/j.engappai.2025.111659.

[5] M. A. Manivasagam, S. S. Ram, C. L. Reddy, E. R. Charan, P. V. Charan, and M. Prabhash, "MQTTNET-IDS: Deep-Fuzzy Fusion for Intelligent Threat Detection," *International Journal of Emerging Research in Science Engineering and Management*, vol. 1, no. 6, pp. 55–62, Dec. 2025, doi: 10.58482/ijersem.v1i6.7.

[6] P. Eidler, U. Kopylov, and O. Ukashi, "Capsule endoscopy in inflammatory bowel disease," *Gastrointestinal Endoscopy Clinics of North America*, vol. 35, no. 1, pp. 73–102, Sep. 2024, doi: 10.1016/j.giec.2024.07.002.

[7] V. C. Gogineni, J.-M. Braun, B. Schelde-Olesen, G. Baatrup, and E. S. Nadimi, "Enhancing polyp characterization in colon capsule endoscopy using ResNet9-KAN," *Knowledge-Based Systems*, vol. 329, p. 114415, Sep. 2025, doi: 10.1016/j.knosys.2025.114415.

[8] C. Houdeville *et al.*, "Toward automated small bowel capsule endoscopy reporting using a summarizing machine learning algorithm: The SUM UP study," *Clinics and Research in Hepatology and Gastroenterology*, vol. 49, no. 1, p. 102509, Nov. 2024, doi: 10.1016/j.clinre.2024.102509.

[9] V. Gopi, K. Gnanasree, G. Dharshini, M. C. Teja, V. Murali, and S. N. Kumar, "Detection of Phishing Websites using Novel Machine Learning Fusion Approach," *International Journal of Emerging Research in Science Engineering and Management*, vol. 1, no. 6, pp. 11–18, Dec. 2025, doi: 10.58482/ijersem.v1i6.2.

[10] N. Hotta *et al.*, "Nationwide multicenter prospective study on the usefulness, safety, and acceptability of colon capsule endoscopy in Japan," *Gastrointestinal Endoscopy*, vol. 101, no. 5, pp. 1051-1063.e7, Nov. 2024, doi: 10.1016/j.gie.2024.11.004.

[11] A. Surekha, P. Aswini, M. S. Kumar, V. Manisha, V. N. Reddy, and B. N. K. Reddy, "AI-based food recognition and nutrient prediction," *International Journal of Emerging Research in Science Engineering and Management*, vol. 1, no. 6, pp. 44–54, Dec. 2025, doi: 10.58482/ijersem.v1i6.6.

[12] Y.-J. Liao *et al.*, "Clinical application and feasibility of capsule endoscopy in children at a medical center in central Taiwan," *Journal of the Formosan Medical Association*, vol. 124, no. 6, pp. 569–573, Jun. 2024, doi: 10.1016/j.jfma.2024.06.012.