On the Classification Accuracy of Wireless Capsule Endoscopy Images for Small Intestine Bleeding Detection Using Convolutional Neural Network

Ziyad k. Farej, Amer F. Sheet, and Noora M. Sheet *

Department of Medical Instrumentation Techniques Engineering, Technical Engineering College of Mosul / Northern Technical University, Mosul, Iraq
Email: drziyad.farej@ntu.edu.iq (Z.K.F.); dr.amerfarhan@ntu.edu.iq (A.F.S.); noora.mazin@ntu.edu.iq (N.M.S.)
*Corresponding author

***Abstract*—The detection of bleeding in the gastrointestinal tract is a critical use of Wireless Capsule Endoscopy (WCE), a valuable diagnostic tool for examining the whole gastrointestinal tract, particularly the hard to reach small intestine. However, manually analyzing a large volume of WCE images is time-consuming, burdensome, and susceptible to human errors. The purpose of this research paper is to develop and assess a Convolutional Neural Network (CNN) model that can automatically analyze and classify WCE images to detect bleeding. This model is projected on a hardware (Raspberry Pi 4 model B) platform to evaluate its classification process practically. A database consisting of 892 WCE images is employed to train and evaluate the model in terms of the metrics such as accuracy, precision, recall, and F1-Score. Two dataset cases are considered (one with 70% & 30% and the other with 80% & 20%) for training and validation, respectively. The findings exhibit promising results, with a macro-average precision, recall, and F1-Score of (0.98611, 0.98438, and 0.98502) for Case 1, (0.99474, 0.99412, and 0.99440) for Case 2, respectively. Furthermore, the suggested model obtains accuracies of 0.98507 and 0.99441 for Case 1 and Case 2 respectively. Finally, the hardware of the developed model is tested with 12 s duration video (which has 56 images) and classified all the images (11 normal 45 abnormal) within this video correctly.**

***Keywords*—wireless capsule endoscopy, bleeding, convolutional neural network, deep learning, small intestine**

I. Introduction

Wireless capsule endoscope invented by Iddan et al. has enabled a pain-free approach to diagnosing the digestive tract [1]. Given Imaging created the first commercial wireless capsule endoscopy system, PillCam, in the year 2000. The Food and Drug Administration (FDA) authorized it in 2001 [2]. Wireless capsule endoscopy has emerged as a valuable detection device for examining the digestive tract as a whole, especially the small bowel, due to the difficulty of accessing it with traditional endoscopes. Where WCE is non-invasive, easy for the patient, and can reach regions with restricted access without requiring surgery [3, 4]. A wireless capsule endoscope is a small, camera-equipped, swallowable gadget that captures images as it travels through the digestive system and then transmits them wirelessly to the sensors, from which they are transmitted to a patient-carried recorder that archives the images. This procedure continues for around 8 hours, or until the wireless capsule endoscope batteries die. This results in a very large amount of data about 60,000 images being transferred to a workstation [5, 6]. These images provide clinicians with vital information regarding the presence of abnormalities such as bleeding, ulcers, and lesions. Among these, bleeding is a critical condition that requires prompt identification and intervention [7].

Small-Bowel Bleeding (SBB), which can be overt or covert, accounts for 5–10% of all digestive tract bleeding and is a clinical manifestation of various gastrointestinal disorders, including ulcers, polyps, tumors, and vascular malformations [8, 9]. The exact diagnosis of bleeding in WCE images plays a critical role in the early detection and treatment of gastrointestinal disorders [10]. Manually examining a huge number of images obtained from a WCE test is time-consuming, cumbersome, and prone to human error. Consequently, there is a greater need for hemorrhaging automated detection systems in order to enhance the diagnostic reliability and effectiveness of Wireless Capsule Endoscopy (WCE).

This research aims to build and evaluate the proposed Convolutional Neural Network (CNN) model capable of detecting small bowel bleeding using automated analysis of WCE images, thus aiding clinicians in the early detection and treatment of small intestine hemorrhage. This research aims to solve the limitations of the literature, such as insufficient accuracy, high computational costs, reliance on specific features such as color, and the lack of generalization of their models across different types of wireless capsule endoscopies.

Manuscript received May 28, 2024; revised July 15, 2024; accepted August 6, 2024; published December 25, 2024.

Research questions include:

1. How can the proposed model be fine-tuned to balance accuracy and computational efficiency?
2. What is the effect of differences in the partitioning of training and validation data on the performance of the CNN system?
3. Does the model achieve excellent generalization to images of another type of wireless capsule endoscope?

Following is how the rest of the sections are arranged: Section II covers some research contributions on detecting gastrointestinal bleeding. The suggested approach is explained in Section III, Section IV displays the suggested model's hardware and Section V contains an analysis of the study's results with a discussion. Conclusions are included in Section VI.

II. Literature Review

The urgent need to discover gastrointestinal bleeding required researchers' to develop techniques for its detection. Pan *et al.*'s [11] study aimed to diagnose bleeding in capsule endoscopy images by measuring color similarity coefficients with two-color vector similarity. As a result, it was applied in RGB color space, yielding specificity and sensitivity values of 90% and 97%, respectively. However, it should be noted that most instances of image misclassification occur when old hemorrhages are present in WCE images due to the darkness of the bleed regions. Mathew et al. aimed to recognize bleeding from healthy regions in 332 wireless endoscopic capsule images by proposing an efficient method for correct bleeding detection that integrated the contourlet transform and Local Binary Pattern (LBP). When applying the k-Nearest Neighbour (k-NN) classifier in the CIE XYZ color space, the results showed an accuracy rate of 96.38% [12].

In 2017, Jia and Meng presented an approach to the detection of WCE bleeding that blends Handcrafted (HC) and CNN features. In their framework, which consists of three phases: the extraction of features, feature integration, and classification. Either CNN derived, handcrafted features were collected from the input frame. Afterwards, features from both categories were combined using an identified strategy. The aggregated feature vector served as an input for classification. The last decision was achieved using a Softmax classifier. The result recall was 0.9100, the precision was 0.9479, and the F1-Score was 0.9285. The number of images used in their study was unequal. The dataset was unbalanced (1500 WCE images, involving 300 bleeding images and 1200 normal images) [13].

In their study in 2018, Tuba *et al.* [14] aimed to propose an approach for detecting hemorrhage in WCE images by utilizing features like texture and color. Their technique is region-based, describing each area using HSI and CIE Lab color spaces with a consistent local binary pattern. The support vector machine classifies regions into 3 groups: background, hemorrhage, or non-bleeding, based on these features. The average Dice Similarity Coefficient (DSC) achieved was 0.85, with a misclassification error of 0.092.

Mamun *et al.* [15] developed a computerized system for recognizing specific images of hemorrhage generated by Wireless Capsule Endoscopy (WCE). For image analysis, the system utilized color thresholding and morphological operations. On the basis of statistical feature vectors in the hue, saturation, and value (HSV) color space, an additional classifier called Quadratic Support Vector Machine (QSVM) was applied to divide the images into those that were bleeding and those that were not. The obtained accuracy was 95.8%, and precision was 80%.

In the study by Coelho *et al.* [16], which looked for ways to recognize and segment red lesions within the small intestine, the U-Net architecture for deep learning was evaluated. The investigation gave an accuracy rate of 95.88%.

In 2021, Saraiva *et al.* [17] developed CNN based on the Xception model with weights trained on ImageNet for automatically identifying blood in large bowel capsule endoscopy frames. Blood was identified by the CNN with 93.2% specificity, 93.8% positive predictive value, 99.8% negative predictive value, and 99.8% sensitivity. The overall accuracy of CNN was 96.6%.

In 2023, Sreejesh [18] proposed a color feature extraction technique to distinguish between hemorrhage frames and normal frames, with localization of the hemorrhage regions. The technique was based on a two-fold method. The first technique made complete use of the information regarding colors in WCE pictures and applied the K-means clustering technique to the pixel-representated frames to produce the cluster centers, which defined WCE pictures as color histograms based on words. The classification performance results were achieved using the YCbCr color space, cluster number 80, with SVM. The accuracy of the classification was 95.75% with an AUC of 0.9771. The second approach introduced a 2-phase saliency map extraction technique to detect areas of hemorrhage. The first step saliency map was generated using a different color channel mixer, while the next stage saliency map was acquired using visual contrast. After that, a suitable fusion approach and threshold, localized the bleeding spots. The ability of the method to generalize to data from another type of WCE has not been tested.

Meena *et al.* [19] developed a convolutional neural network for automatically diagnosing brain tumors using Magnetic Resonance Imaging (MRI) images, with a classification accuracy of 99.28% on the Br35H dataset. The research significantly reduced overfitting by using dropout regularization and the Adam optimizer, and it thoroughly examined model performance using measures such as accuracy, recall, precision, and F1-Score. The dependence on a single dataset restricts the capacity to extend the results to other kinds of medical imaging and diagnostic challenges.

In recent years, there has been significant progress in the field of computerized vision, particularly in the application of deep learning methods such as Convolutional Neural Networks (CNN). These techniques have demonstrated impressive performance in

various tasks, such as recognizing things, segmenting images, and medical image analysis [20, 21]. CNNs have the ability to automatically learn discriminatory features from raw image data, which makes them highly suitable for detecting bleeding in WCE images [22].

III. Proposed CNN Model

The large number of images captured by wireless capsule endoscopy can present challenges for doctors in terms of time consumption, fatigue, and potentially decreased diagnostic accuracy. This study aims to detect small bowel bleeding using a dataset of WCE images and a deep convolutional neural network. We propose a method that can accurately detect infected images in a dataset and aid physicians in making rapid diagnoses. Fig. 1 illustrates the structure of the proposed approach.

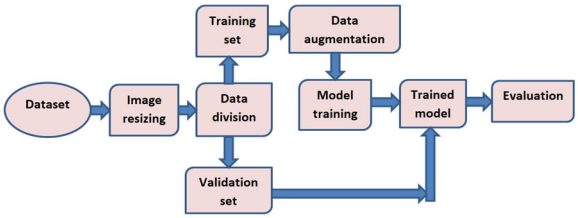


Fig. 1. The proposed CNN method.

For the purpose of conducting the approach, an 11th Generation Intel(R) Core(TM) i5 processor with 8 GB of Random Access Memory (RAM) was used on Microsoft's Windows 10 operating system, with the Python programming language and Anaconda Navigator software, and on a Jupyter notebook.

A. Dataset

In this research paper, the Kvasir-Capsule dataset is used, which can be accessed through the Open Science Framework (OSF). The dataset contains 44,228 labeled images stored using the PNG format, which are divided into 13 different classes. The dataset has a total size of circa 94.7 GB. In this study, images of the bleeding category and the normal mucosa category were used, and the data were balanced. The wireless capsule endoscopy used to obtain this data is the Olympus EC-S10 endocapsule. Fig. 2 illustrates an example of the normal and bleeding WCE images taken from the dataset. Images that are normal do not have any bleeding. There are 446 bleeding images and 446 normal images among the 892 total images.

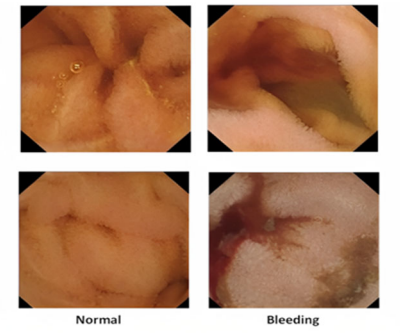


Fig. 2. Data sample for both the normal and bleeding classes.

The dataset images have been resized to (300×300) pixels and then split into 624 training images and 268 validation images with case 1 and divided into 713 training images and 179 validation images with Case 2 for the purpose of analysis, as depicted in Table I.

TABLE I. Shows the Training and Validation Counts in the Dataset

Dataset	Case 1	Case 2
Total images	892	892
Training images	624 (70% of total images)	713 (80% of total images)
validation images	268 (30% of total images)	179 (20% of total images)

In order to augment the training dataset and enhance its variety, a set of augmentation techniques is applied to the training images, such as image rotation, shearing, zooming, shifting, and fill mode. Subsequently, the model is trained using the augmented training images, and its performance is evaluated through validation on an independent set of validation images.

B. Architecture of the Proposed CNN

Convolutional Neural Networks (CNNs) have gained vital importance in medical research due to their ability to analyze complex medical data, particularly medical images. This study aims to build and evaluate the proposed CNN model that is capable of automated WCE image analysis and detect infected images.

The basic components of the CNN model are convolutional layer, pooling layer, nonlinear activation layer, flattening layer, and dropout layer. The convolutional layer, also known as the principal layer in the CNN model, extracts feature maps from input images. The pooling layer decreases the spatial dimensions of feature maps generated by convolutional layers. This is accomplished by reducing the feature maps while keeping the most critical information. Max pooling is the most popular pooling procedure, which picks the maximum value inside each pooling window. Activation functions such as Rectified Linear Unit (ReLU) are applied to the feature maps after each convolutional or pooling layer. CNN is considerably easier and quicker using ReLU. For all negative values, ReLU is 0, whereas positive numbers stay constant. The flattening layer takes the multi-dimensional feature maps produced by the convolutional and pooling layers and reshapes them into a one-dimensional vector or array to prepare the data for the subsequent fully connected layers. Overfitting is a significant challenge when training deep neural networks, and it occurs when multiple neurons consistently find identical results. Therefore, a dropout layer can be employed to selectively drop neurons in the network, which helps prevent the network from relying too heavily on specific neurons, reducing overfitting. Dropout rates of 0.2 and 0.5 have been applied.

The proposed CNN consists of layers like the input layer, two convolutional layers, two max-pooling layers, one batch normalization, three activation function layers, two dropout layers, one flatten layer, and 3 dense layers.

The model input data is (300×300×3), representing an image with 300 pixel heights and 300 pixel widths and red, green, and blue channels. The first Conv2D layer applies 64 filters and (3×3) kernel dimensions with the ReLU activation function. Max pooling layer with a pooling size of 2×2 and a batch normalization layer are employed with the axis = −1 parameter. The second Conv2D layer employs 32 filters (3×3) kernel size, and the ReLU activation function. Max pooling layer with a pooling size of 2×2. The ReLU activation layer is used after every Conv2D layer and fully connected layer except for the last dense layer, after which the softmax activation function is used. Dropout layers with a 20% and 50% dropout rate are used. Dropout layers and batch normalization are used to improve the training process and to combat overfitting. By randomly dropping neurons during training, dropout layers prevent the model from becoming too reliant on specific features. A flattened layer is applied to convert the three-dimensional output of the second pooling layer into a vector before moving it to the three dense layers, It transforms the output of the second pooling layer (73×73×32) into a flat vector (170,528). This step prepares the data for classification by the dense layers. The design of the suggested model is displayed in Fig. 3.

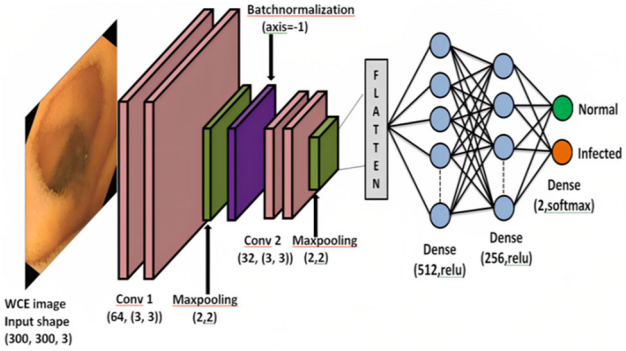


Fig. 3. Design for the proposed CNN.

The details of the proposed network have been depicted in Table II.

TABLE II. THE SUGGESTED C NN ARCHITECTURE PARAMETERS

Layer (Kind)	No. Kernel	Kernel Dimensions	Output Shape	No. Param
Input	-	-	(300, 300, 3)	0
Conv 1 (Conv2D)	64	3×3	(298, 298, 64)	1792
Pooling 1 (MaxPooling2D)	-	2×2	(149, 149, 64)	0
(Batch Normalization)	-	-	(149, 149, 64)	256
Conv 2 (Conv2D)	32	3×3	(147, 147, 32)	18,464
Pooling 2 (MaxPooling2D)	-	2×2	(73, 73, 32)	0
Dropout 1	-	-	(73, 73, 32)	0
(Flatten)	-	-	(170, 528)	0
Dense 1	-	-	(512)	87,310,848
Dropout 2	-	-	(512)	0
Dense 2	-	-	(256)	131,328
Dense 3	-	-	(2)	514
Output (Activation)	-	-	(2)	0

The proposed model, whose output is the probability of 0 or 1, uses a categorical cross-entropy loss function. It consists of a softmax activation along with a cross-entropy loss used to determine the difference between the predicted and true classes. The definitions of the categorical crossentropy (CCE) loss function and softmax activation function can be obtained from Eqs. (1) and (2).

$$CCE = -\sum_i^C t_i \log(f(s)_i) \quad (1)$$

$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}} \quad (2)$$

Here t_i represents the real label (1 for hemorrhage and 0 for normal), $f(s)_i$ represents the predicted probability of class i after applying the softmax activation, and i is an index that runs from 1 to C, where C indicates the number of classes or categories. The raw score (s), which is often the output of a neural network's final layer before applying the softmax, and j represents the index used for summations over the classes when calculating the softmax probabilities and the cross-entropy loss. For training, an adaptive moment estimation (Adam) optimizer is being applied, which affects the parameter weight and learning rate to decrease the learning model's loss.

C. Metrics of Evaluation

Accuracy, recall, precision, and the F1-Score are used to assess deep learning's performance model. TP is a number when a model expects that an image is normal, and the image's real label is also normal; When the model guesses an infected image with a real infected label, it indicates TN. FP is defined when the model predicts that an image is normal but the true label of the image is infected. The False Negative (FN) refers to the situation in which the model predicts an image to be infected while the true label for that image is actually normal. The following is the mathematical representation:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

IV. HARDWARE OF THE PROPOSED CNN MODEL

The proposed CNN model is analyzed and designed using the Python language. This model is trained and validated using the Kvasir-Capsule dataset (Olympus EC-S10 endocapsule). The proposed model is projected onto the hardware prototype (Raspberry Pi 4 Model-B) system, as shown in Fig. 4. A video with 12-second duration is used in the testing process of the projected model. This video, which contains 56 images, is part of an OMOM wireless capsule endoscope video. The images were

formatted to 300×300 pixels so that they are consistent with the size of the trained and validated images. The results of the testing process for the hardware of the projected model show that there are 11 normal images and 45 abnormal images.

Fig. 4. The overall hardware suggested system.

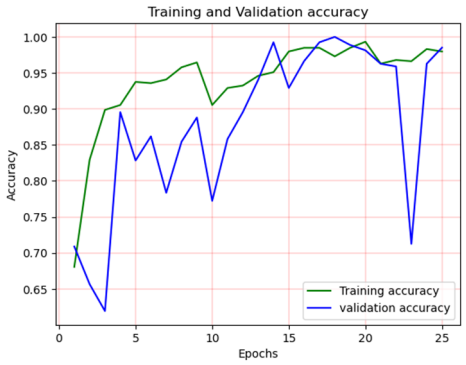
V. Result and Discussion

Our study is mostly about using a proposed Convolutional Neural Network (CNN) to find images that contain bleeding using the wireless capsule endoscopy images dataset. This dataset consists of 446 images that show bleeding and 446 images that show normal. Out of the total 892 images, 624 (in Case 1) and 713 (in Case 2) are employed for training purposes.

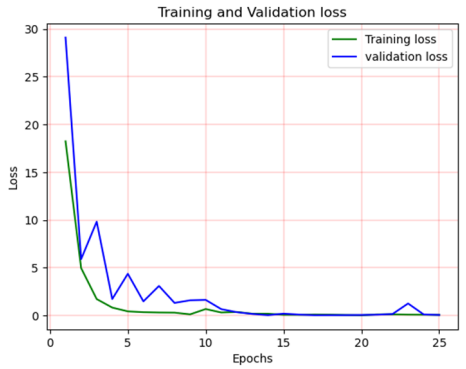
To increase the training dataset, data augmentation techniques were used. The validation data, comprising 268 images in Case 1 and 179 in Case 2, was then applied to evaluate the trained approach. During the training process, the model required 25 epochs, training and validation accuracy and loss were monitored throughout the training process. As shown, the model started with accuracies of 0.6807 in Case 1, 0.7034 in Case 2, and gradually improved over the epochs, reaching a final training accuracies of 0.9797 in Case 1, 0.9853 in Case 2, and a validation accuracies of 0.9851 in Case 1, 0.9944 in Case 2. Similarly, the loss decreased over the epochs, indicating that the model learned to make more accurate predictions.

The training process showed that the model was able to generalize well to unseen data, as shown by the high validation accuracy. This suggests that the model is not overfitting the training data and has learned to capture important patterns and features that generalize to new instances. Fig. 5(a) and Fig. 5(b) for Case 1 and Fig. 5(c) and Fig. 5(d) for Case 2 illustrate the graphs for model accuracy and loss curves during the training and validation stages.

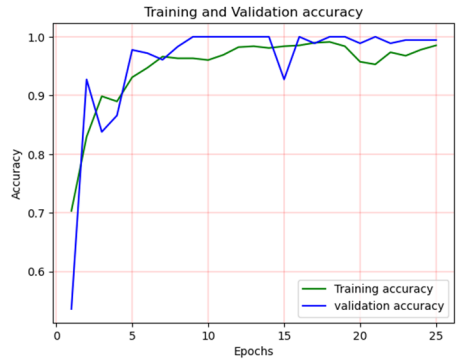
For Case 1, Table III shows the model's performance in both classes, exhibiting precision, recall, and F1-Score values above 0.96 for both classes (normal (0) and infected (1)). These results indicate the proposed model has the capability to accurately identify instances of Class 0. The model displays a high ability to correctly describe instances in Class 1. Considering the overall accuracy of the model, computed as 0.98507, it is seen that the model accurately classified the majority of instances. This remarkable accuracy suggests that the model effectively distinguishes between the two classes present in the dataset. Table IV illustrates the model performance for Case 2.



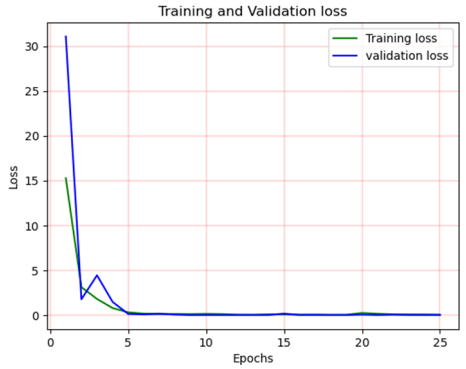
(a)



(b)



(c)



(d)

Fig. 5. Accuracy and loss of proposed model (a) accuracy and (b) loss curve of the suggested model for Case 1; (c) accuracy; (d) loss curve of the suggested model for Case 2.

TABLE III. CNN'S SUGGESTED RESULTS FOR CASE 1

Category	Precision	Recall	F1-Score	Support
Normal	0.97222	1.00000	0.98592	140
Infected	1.00000	0.96875	0.98413	128
Accuracy		0.98507		268

TABLE IV. RESULTS OF THE PROPOSED CNN FOR CASE 2

Category	Precision	Recall	F1-Score	Support
Normal	0.98947	1.00000	0.99471	94
Infected	1.00000	0.98824	0.99408	85
Accuracy		0.99441		179

As seen in Fig. 6, for Case 1 the proposed method correctly predicts 264 of 268 images. The confusion matrix indicates that the model makes a prediction that four images are normal, whereas the true class is infected.

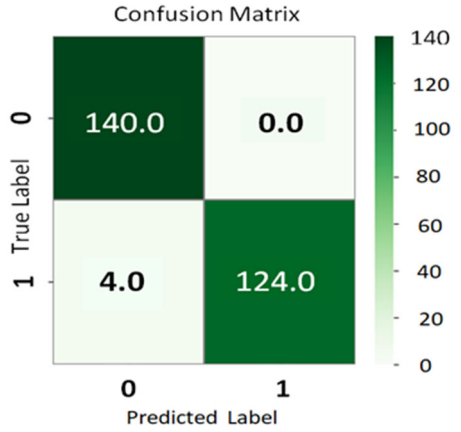


Fig. 6. Confusion matrix of the proposed CNN, with the normal class denoted by 0, while the infected class is represented by 1.

As shown in Fig. 7, the suggested technique successfully predicts 178 from 179 images in Case 2. The confusion matrix shows that the model classifies a single image as normal when it is actually infected.

The two cases efficacy is compared to three other bleeding detection methods described in the scientific literature. Table V displays the comparison findings.

Based on the comparison, with regards to accuracy rate, the suggested model outperforms previous studies, achieving accuracies of 98.5% in Case 1 and 99.4% in case 2 for bleeding detection. This indicates the effectiveness of our model in accurately classifying infected images in wireless endoscopic capsule images. However, it's crucial to take into account additional elements, including the dataset used and potential differences in experimental setups and conditions that could impact the direct comparison of the results.

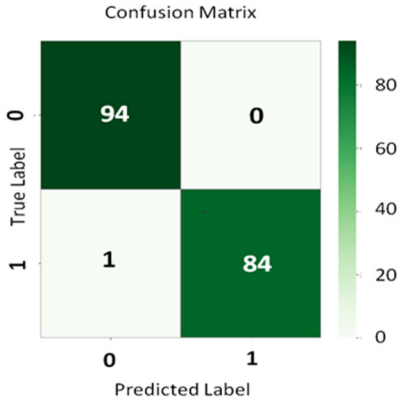


Fig. 7. The suggested CNN's confusion matrix, with the normal class marked by 0 and the infected class denoted by 1.

TABLE V. DISPLAYS THE RESULTS OF PREVIOUS STUDIES COMPARED TO THE PROPOSED MODEL

Ref.	Year	Aim	AI type	Results
Mathew *et al.* [12]	2015	Bleeding detection	k-Nearest Neighbour (k-NN) classifier	Accuracy 96.38%
Coelho *et al.* [16]	2018	Recognize and segment red lesions within the small intestine	U-Net	Accuracy rate of 95.88%
Mamun *et al.* [15]	2021	Detection of small intestinal hemorrhage	Quadratic Support Vector Machine (QSVM)	Accuracy 95.8%
Sreejesh [18]	2023	Bleeding frame and region automatic detection	SVM	Accuracy 95.75% AUC 0.9771
Nakada *et al.* [23]	2023	Utilize RetinaNet to diagnose erosions and ulcers, vascular lesions, and tumours in WCE imaging	Deep neural network	The mean accuracies for Erosions and ulcers: 93%; Vascular lesions 96%; Tumours: 92%
Chu *et al.* [24]	2023	Utilize CNN segmentation method for angiodysplasias detection	Resnet-50	Pixel accuracy of 99%
Sahafi *et al.* [25]	2024	Polypoid lesion identification in WCE images	YOLO-V8 Network	Precision: 98%, Recall: 97.9%.
Al-Ali *et al.* [26]	2024	Detect lesions in colonoscopy videos for Ulcerative colitis severity assessment	Linear models in color spaces	92.29 ± 0.443% specificity; 88.59 ± 2.984% sensitivity for bleeding detection and 58.22 ± 0.393% specificity; 81.68 ± 4.173% sensitivity for ulcer detection.
Proposed (Case 1)	2024	Bleeding detection	CNN	Accuracy rate 98.5%
Proposed (Case 2)	2024	Bleeding detection	CNN	Accuracy rate 99.4%

VI. Conclusion

Our research focused on developing and evaluating a convolutional neural network model to automatically examine Wireless Capsule Endoscope (WCE) images and identify small intestine bleeding. With a dataset of 892 WCE images, the trained CNN model achieved excellent results. It exhibited macro-average precision, recall, and F1-Score values of 0.98611, 0.98438, and 0.98502 for Case 1, while it showed 0.99474, 0.99412, and 0.99440 for Case 2, respectively, indicating its big ability to detect bleeding. Furthermore, compared to previous studies with accuracies of 0.98507 for Case 1 and 0.99441 for Case 2, our model demonstrated superior performance. The video used in the process of testing the proposed hardware CNN model is 12 s and has 56 images. The system has classified all the images correctly (11 normal and 45 abnormal). By implementing this CNN model, diagnostic accuracy in capsule endoscopy can be enhanced, saving time, reducing workload, and minimizing human errors for physicians. This automation enables early detection, appropriate management, and improved patient outcomes in cases of small bowel bleeding. In the future, researchers will benefit from conducting thorough studies with bigger and more different data sets. It is important to investigate how this automated analysis tool may be integrated with medical professionals' expertise and experience.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Noora collected the dataset and wrote this paper with the assistance, guidance, and supervision of Ziyad and Amer. Ziyad and Amer chose the idea of the work and contributed to the scientific assignment and theoretical background of the basic steps in this paper. All authors had approved the final version.

Acknowledgment

The authors would like to express their gratitude to the Northern Technical University for allowing them to work in the laboratories of the College of Technical Engineering in Mosul and for their support, guidance, and use of their equipment throughout the duration of this research project.

References

[1] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, no. 6785, pp. 417–418, 2000. https://doi.org/10.1038/35013140

[2] P. R. Slawinski, K. L. Obstein, and P. Valdastri, "Emerging issues and future developments in capsule endoscopy," *Tech. Gastrointest. Endosc.*, vol. 17, no. 1, pp. 40–46, 2015. https://doi.org/10.1016/j.tgie.2015.02.006

[3] M. W. Alam, S. S. Vedaei, and K. A. Wahid, "A fluorescence-based wireless capsule endoscopy system for detecting colorectal cancer," *Cancers (Basel).*, vol. 12, no. 4, 2020. doi: 10.3390/cancers12040890

[4] L. Biancone *et al.*, "Wireless capsule endoscopy and small intestine contrast ultrasonography in recurrence of Crohn's disease," *Inflamm. Bowel Dis*., vol. 13, no. 10, pp. 1256–1265, 2007. https://doi.org/10.1002/ibd.20199

[5] C. M. Pablo Laiz, Jordi Vitri`a, Hagen Wenzek and S. S. Fernando Azpiroz, "WCE Polyp Detection with Triplet based Embeddings Pablo," *Comput. Med. Imaging Graph*. 127120, 2020. https://doi.org/10.1016/j.compmedimag.2020.101794

[6] R. Ezatian, D. Khaledyan, K. Jafari, M. Heidari, A. Z. Khuzani, and N. Mashhadi, "Image quality enhancement in wireless capsule endoscopy with Adaptive Fraction Gamma Transformation and Unsharp Masking filter," in *Proc. 2020 IEEE Glob. Humanit. Technol. Conf. GHTC 2020*, 2020. https://doi.org/10.1109/GHTC46280.2020.9342851

[7] M. Sharif, M. Attique Khan, M. Rashid, M. Yasmin, F. Afza, and U. J. Tanik, "Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images," *J. Exp. Theor. Artif. Intell*., vol. 33, no. 4, pp. 577–599, 2021. https://doi.org/10.1080/0952813X.2019.1572657

[8] G. Sreejesh and School, "Bleeding frame and region detection in the wireless capsule endoscopy video," *Int. J. Hum. Comput. Intell*., vol. 02, 2023. https://doi.org/10.1109/JBHI.2015.2399502

[9] J. R. Kuo, S. F. Pasha, and J. A. Leighton, "The clinician's guide to suspected small bowel bleeding," *Am. J. Gastroenterol.*, vol. 114, no. 4, pp. 591–598, 2019. https://doi.org/10.1038/s41395-018-0424-x

[10] S. Sunitha and S. S. Sujatha, "An improved bleeding detection method for Wireless capsule endoscopy (WCE) images based on AlexNet," in *Proc. 2021 3rd Int. Conf. Signal Process. Commun. ICPSC 2021*, no. May, pp. 11–15, 2021. https://doi.org/10.1109/ICSPC51351.2021.9451699

[11] G. Pan, F. Xu, and J. Chen, "Bleeding detection in wireless capsule endoscopy using color similarity coefficient," *Appl. Mech. Mater.*, vol. 195–196, pp. 307–312, 2012. https://doi.org/10.4028/www.scientific.net/AMM.195-196.307

[12] M. Mathew and V. P. Gopi, "Transform based bleeding detection technique for endoscopic images," in *Proc. 2nd Int. Conf. Electron. Commun. Syst. ICECS 2015*, 2015, pp. 1730–1734. https://doi.org/10.1109/ECS.2015.7124882

[13] X. Jia and M. Q. H. Meng, "Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, 2017, pp. 3154–3157. doi: 10.1109/EMBC.2017.8037526

[14] E. Tuba, S. Tomic, M. Beko, D. Zivkovic, and M. Tuba, "Bleeding detection in wireless capsule endoscopy images using texture and color features," in *Proc. 2018 26th Telecommun. Forum, TELFOR 2018*, 2018, pp. 1–4. doi: 10.1109/TELFOR.2018.8611939

[15] A. Al Mamun, M. S. Hossain, P. P. Em, A. Tahabilder, R. Sultana, and M. A. Islam, "Small intestine bleeding detection using color threshold and morphological operation in WCE images," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 4, pp. 3040–3048, 2021. https://doi.org/10.11591/ijece.v11i4.pp3040-3048

[16] A. C. P. Coelho, A. Pereira, A. Leite, and M. Salgado, "A deep learning approach for red endoscopies," in *Proc. Int. Conf. Image Anal. Recognit.*, 2018, vol. 1, no. June, pp. 553–561. http://dx.doi.org/10.1007/978-3-319-93000-8_103

[17] M. M. Saraiva *et al.*, "Artificial intelligence and colon capsule endoscopy: Automatic detection of blood in colon capsule endoscopy using a convolutional neural network," *Endosc. Int. Open*, vol. 09, no. 08, pp. E1264–E1268, 2021. doi: 10.1055/a-1490-8960

[18] G. Sreejesh, "Bleeding frame and region detection in the wireless capsule endoscopy video," *IEEE J. Biomed. Heal. Informatics*, vol. 2, no. 1, 2023.

[19] G. Meena, K. K. Mohbey, M. Acharya, and K. Lokesh, "Original research article an improved convolutional neural network-based model for detecting brain tumors from augmented MRI images. journal of autonomous intelligence," vol. 6, no. 1, 2023. https://jai.front-sci.com/index.php/jai/article/view/561

[20] A. Sarraf, A. Jalali, and J. Ghaffari, "Recent applications of deep learning algorithms in medical image analysis," *Am. Sci. Res. J. Eng. Technol. Sci.*, vol. 72, pp. 58–66, 2020.

[21] J. Bernal *et al.*, "Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: A review," *Artif.*

Intell. Med., vol. 95, no. 8, pp. 64–81, 2019. https://doi.org/10.1016/j.artmed.2018.08.008
[22] H. S. Pannu, S. Ahuja, N. Dang, S. Soni, and A. K. Malhi, "Deep learning based image classification for intestinal hemorrhage," *Multimed. Tools Appl.*, vol. 79, no. 29–30, pp. 21941–21966, 2020. https://doi.org/10.1007/s11042-020-08905-7
[23] A. Nakada, R. Niikura, K. Otani, Y. Kurose, Y. Hayashi, K. Kitamura, and K. Hasatani, "Improved object detection artificial intelligence using the revised RetinaNet model for the automatic detection of ulcerations, vascular lesions, and tumors in wireless capsule endoscopy," *Biomedicines*, vol. 11, P. 942, 2023.
[24] Y. Chu, F. Huang, M. Gao, D. W. Zou, J. Zhong, W. Wu, and L. F. Wang, "Convolutional neural network-based segmentation network applied to image recognition of angiodysplasias lesion under capsule endoscopy," *World Journal of Gastroenterology*, vol. 29, no. 5, P. 879, 2023.
[25] A. Sahafi, A. Koulaouzidis, and M. Lalinia "Polypoid lesion segmentation using YOLO-V8 network in wireless video capsule endoscopy images," *Diagnostics*, vol. 14, no. 5, P. 474, 2024.
[26] S. Al-Ali, J. Chaussard, S. Li-Thiao-Té, É. Ogier-Denis, A. Percy-du-Sert, X. Treton, and H. Zaag, "Detection of ulcerative colitis lesions from weakly annotated colonoscopy videos using bounding boxes," *Gastrointestinal Disorders*, vol. 6, no. 1, pp. 292–307, 2024.