[[1]](#footnote-2)

Enhancing Ulcer Detection in Gastrointestinal Endoscopic Images

*Abstract*

Accurate diagnosis of Gastrointestinal (GI) diseases is vital for healthcare, yet processing large medical image data challenges professionals. Computer-aided Diagnosis (CAD) systems using deep Convolutional Neural Networks (CNNs) assist, though existing models lack robustness. This research proposes a method combining attention-based image processing and a lightweight CNN. The framework, evaluated on the Kvasir dataset, exhibits significant improvements in accuracy, precision, recall, and F-measure. Additionally, the study explores automated ulcer detection in Wireless Capsule Endoscopy (WCE) images. Overcoming challenges like low contrast, state-of-the-art CNNs, AlexNet, and GoogLeNet, outperform traditional methods, affirming CNNs' efficacy in automated diagnosis, promising advanced disease recognition and medical imaging analysis..Notably, the system's robustness is evident in substantial accuracy improvement, enhanced precision, recall, and F-measure, showcasing its potential as an efficient diagnostic tool for GI diseases. Furthermore, the study explores the automated detection and classification of ulcers in Wireless Capsule Endoscopy (WCE) images. Addressing challenges like low contrast and complex backgrounds, this research employs state-of-the-art CNN architectures, AlexNet and GoogLeNet. Extensive experiments demonstrate the superiority of CNNs, surpassing traditional methods and confirming their efficacy in automated diagnosis, thus paving the way for enhanced medical imaging analysis and disease recognition.

# INTRODUCTION

Ulcers are a common gastrointestinal disorder that can lead to serious complications if not treated promptly. Endoscopic imaging is the preferred method for diagnosing ulcers, but it is a time-consuming and challenging task for endoscopists to manually identify ulcers in endoscopic images.

Artificial intelligence-driven differential diagnosis (AI-DDx) is a new and rapidly developing field that uses AI to assist clinicians in the diagnosis of diseases. AI-DDx systems can be used to analyze large amounts of data, including medical images, patient records, and laboratory results, to identify patterns that are too subtle for the human eye to detect. This can help clinicians to narrow down the list of possible diagnoses and to reach a more accurate diagnosis more quickly.

AI-DDx systems have the potential to improve the accuracy and efficiency of ulcer detection, which can lead to earlier diagnosis and treatment, improved patient outcomes, and reduced healthcare costs. For example, AI-DDx systems are being used to develop new and improved methods for the detection of cancer, heart disease, and diabetes in medical images and using data such as blood sugar levels, blood pressure, and weight.

In this paper, we propose a transfer learning-based ulcer detection model that uses a pre-trained convolutional neural network (CNN) as the starting point. CNNs are a type of machine learning model that are well-suited for image classification tasks. We argue that a transfer learning approach can help to overcome the challenges of ulcer detection by providing the new model with a good starting point and reducing the need for training data.

Our goal is to develop a transfer learning-based ulcer detection model that achieves high accuracy on endoscopic images. We believe that such a model could be used to assist endoscopists in the diagnosis of ulcers, leading to earlier diagnosis and treatment for patients.

We also discuss the potential of AI-DDx to revolutionize the way that ulcers and other diseases are diagnosed. We believe that AI-DDx systems have the potential to improve the quality of care for patients and reduce the burden of disease on society.

# Existing methodology

Texture analysis methods is used and that can be categorized into structural, statistical, geometrical, model-based, and signal-processing methods [13]. Figure 2 illustrates the main steps of the multi-scale approach. Firstly, we consider an input image in a specific color space (RGB or YCbCr). For each component C1, C2 or C3, we applied two completed LBP variants, either CLBP/Mri P,R or CLBP/Sri P,R with a fixing number of LBP neighbors P = 8 and radius R = 1. The resulting completed LBP coded map (CLBP/Magnituderi 8,1 or CLBP/Signri 8,1 ) are considered as input of the 5-level Laplacian pyramid decomposition. Then, from each resulting sub-band coefficient we calculated a histogram. Resulting of 4 histograms because the fifth sub-band coefficient is not taken in consideration. Then, from each histogram sub-band coefficient, six statistical features are extracted. Finally, all features are concatenated together into a single ultimate feature vector for classification. Giving a vector length of 4 \* 6 = 24 features.

**Brief review of local binary pattern**

Local Binary Patterns (LBP) is first proposed by Ojala et al. [28] to describe the textures present in gray-scale images. For each pixel in an image, the LBP pattern can be computed by comparing its pixel value with the values of its immediate neighbors. The matrix of weight mask will be used afterward to code the result of every thresholding [31]. Then, the weighted values are summed and reported in the output image, to correspond the center pixel coordinates in the input image. The LBP operator was extended to a multi-resolution analysis tool, to use any circularly symmetric neighborhoods (P,R) of different sizes. the pair (P,R) means P pixels located on a circle of radius R. In general, the LBPP,R of a pixel (xc,yc) is expressed in a decimal form as:

        LBPP,R =  P p=0 sign(Vp − Vc)2 P ,sign(x) =  1, if x ≥ 0 0, otherwise

 where Vc, Vp are the values of central and neighboring pixels. The original LBP is invariant to monotonic gray scales changes but not invariant to image rotations [8]. To overcome these shortages, the authors of [28] proposed the Rotation Invariant LBPri and Rotation Invariant Uniform LBPriu2 to achieve robustness against image rotation

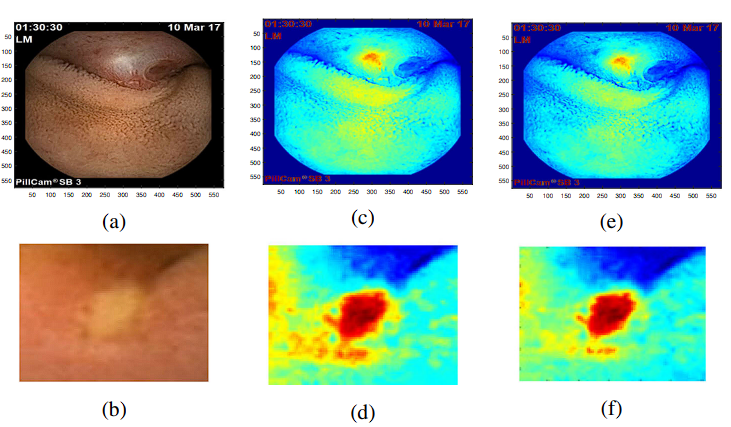
**2.3 Completed local binary pattern**

 The completed local binary pattern (CLBP) as an extension to LBP is defined by the local central pixel and the local difference of sign-magnitude transform (LDSMT). The central pixel is defined by a binary global thresholding map named CLBP C while the LDSMT constitutes of two components, namely: sign component sp to build the CLBP S and the magnitude component mp to build the CLBP M. The CLBP S is equivalent to the original LBP. The CLBP S and CLBP M can be differently combined to build the final CLBP histograms. The LDSMT is mathematically expressed as given central pixel gc and gray level of neighboring pixel gp, p = {0, 1, . . . ., P − 1}. The local difference is defined as dp = (gp − gc):

**dp = sp × mp and  sp = sign(dp) mp = |dp**

he local difference vector [d0, . . . ., dp−1] characterizes the image local structures at gc. However, different drawbacks are associated. It is sensitive to noise, translation and rotation. To this reason, it is transformed into a sign vector [s0, . . . ., sp−1] and a magnitude vector [m0,...,mp−1] to robustly recognize texture patterns. The M component cannot be directly coded as S.

**CLBP CP,R = t(gc,CI )**



## **Figure. 1 Ulcerous WCE images and the corresponding different color spaces images. a, b Original WCE images for the 1st and 2nd databases respectively. c, d**

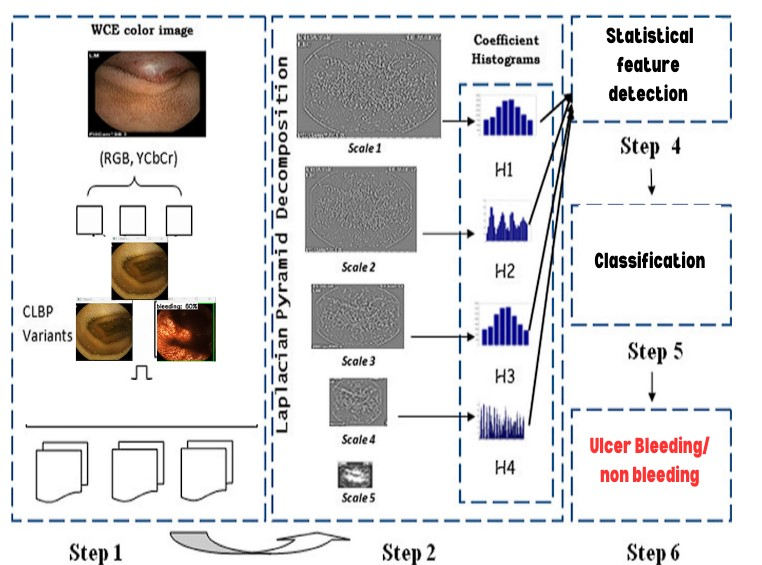
**Multiscale analysis**

The work presented in this paper constitutes an extension to our work in [28] which describes a novel multi-scale algorithm that is capable of handling not the sign .Component of MS-CLBP/S denoted in [28] as (MS LBP) but also the magnitude component using two datasets. The extension of the Completed LBP to multi-scale computation (MS-CLBP) can be realized by combining multiple operators at varying parameters (P,R) as stated by many authors [15, 21]. But the most discriminate feature cannot be captured using a single scale since the dominant feature may present at any spatial resolution. Our approach belongs to coarse-to-fine detection strategy (Laplacian pyramid decomposition) as the CLBP descriptor will be more powerful if the image is overpowered by a fine, coarse, directional, or repetitive texture. In this manuscript, we develop a multi-scale representation of two variants of CLBP descriptor (MS-CLBP/Sign and MS-CLBP/Magnitude) to effectively exploit the monochrome and invariant-rotation features at various scales for both databases. Our approach is totally different from The well known multi-scale LBP (MS-LBP) method which lacks a scale-selection mechanism. Unfortunately, the original LBP image suffers from the not invariance to rotation.

# Proposed methodology

We propose to use a combination of two neural networks Google Net and Alex Net both of them being

Convolutional Neural Networks (CNNs)



## **Figure. 2** General diagram of the proposed method

Convolutional Neural Networks (CNNs) have a distinct architecture compared to regular neural networks. In CNNs, neurons are organized in three dimensions: width, height, and depth. Each layer in a CNN converts a 3D input volume into a 3D output volume of neuron activations. CNNs consist of three main types of layers: convolutional, pooling, and fully connected layers. Notably, not all neurons in one layer are necessarily connected to all neurons in the next layer. CNNs process input data through sequences of convolutions and pooling, using filters to generate feature maps. These feature maps are combined to produce the final output of the convolution layer.

The convolution

A layer is a crucial component of CNNs but makes training time-consuming. This layer applies convolution operations to the input using shared sets of weights (kernels or filters) with small receptive fields. Pooling layers, commonly using max pooling, perform nonlinear down sampling, reducing parameters, overfitting risk, and computational complexity. Max pooling divides the input into non-overlapping frames, outputting the maximum value from each group. To mitigate overfitting, dropout layers drop neurons and their connections with a certain probability. ReLU (rectified linear unit) is a popular activation function. Fully connected layers act as classifiers, connecting all neurons in a layer to the outputs of the previous layer. Training CNNs from scratch demands ample training data, which might not always be available and can lead to overfitting.

## A diagram of a computer model Description automatically generated

**Figure 3** Network layout for the deep convolutional networks. For illustration, the input patch is

selected from the PC , but the same layout is used for the other types of training patches.

## 

**Figure 4.** Schematic diagram of the overall system for ulcer detection

# Equations

**True Positive (TP)**: TP represents the number of positive instances correctly predicted by the model.

∑i=1n​1(yi​=1 and y^​i​=1)

False Negative (FN): Represents the number of positive instances incorrectly predicted as negative by the model.

**FN=∑i=1n​1(yi​=1 and y^​i​=0)**

**AUC: Q**uantifies the overall performance of a binary classification model using the ROC curve.

this is a plot of Sensitivity (True Positive Rate) **against 1 - Specificity** (False Positive Rate) at various threshold values.

**Accuracy**: Represents the overall correctness of the model's predictions.

( **TP+TN)/(TP+TN+FP+FN)**

True Negative (TN): Represents the number of negative instances correctly predicted by the model.

**TN=∑1(Yi=0 and y^i=1)**

The five metrics are computed as

Sensitivity = TP/(TP + FN) (1)

Specificity = TN/(TN + FP) (2)

Accuracy = (TP + TN)/(TP + TN + FP + FN) (3)

Loss = −Σ yj log (ˆyi ) (4)

AUC = 0.5 (Sensitivity + Specificity) (5)

# Dataset

The experiments utilized images sourced from reference [32], comprising a dataset of 1875 images captured via WCE video. This dataset included 1525 ulcer instances and 250 instances of the normal class, originating from both the esophageal and gastric parts of the digestive system. Sample ulcer images are illustrated in Figure 7, showcasing esophageal and gastric ulcers.

A collage of images of a human body

Description automatically generated

**Figure 5.** Sample ulcer images from WCE videos. (**Left**) esophageal ulcer; (**Right**) gastric ulcer.

To begin the analysis of WCE images, they were randomly divided into training and testing sets. Specifically, 80% of the images (421 images) were allocated for training. Within the training set, there were 256 abnormal and 80 normal images, whereas the test set consisted of 80 abnormal and 25 normal images.

The original images had a resolution of 256 × 256 × 3 (256 width, 256 height, 3 color channels). To match the requirements of GoogLeNet, the images were resized to 224 × 224 × 3 pixels. For AlexNet, the resizing was done to 227 × 227 × 3 pixels.

# Experimental results

The model employed is the Adam optimizer. The Adam optimizer is a widely used optimizer currently used for model training due to its excellent performance and adaptable learning rate . Additionally, the Categorical‐Cross‐Entropy (CCE) was utilized as a loss‐function which adjusts the weights of CNN to produce a better fitting model]. During the training process of a DL model, the loss function measures the dissimilarity between the predicted and target labels [24]. One effective way to evaluate loss is by using the CCE loss, which calculates the dissimilarity between two distributions. 120 epochs were used in a batch size of 150 for model training

A graph of a graph

Description automatically generated

**Figure 6: Accuracy vs no of Epoch**

A graph of a graph

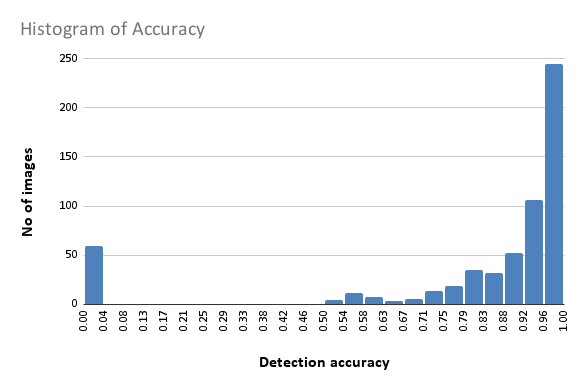
Description automatically generated with medium confidence

Figure 7: **loss vs Epoch**

A graph with a blue line

Description automatically generated

**Figure 8: True positive rate vs false positive rate**



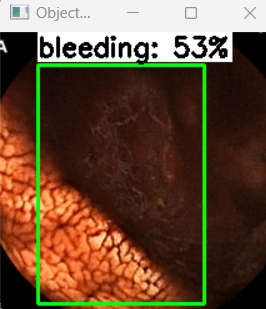
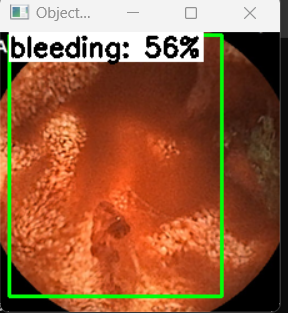
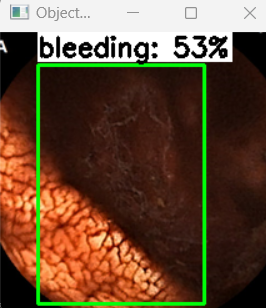
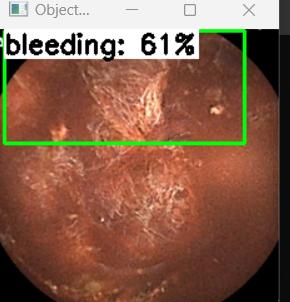
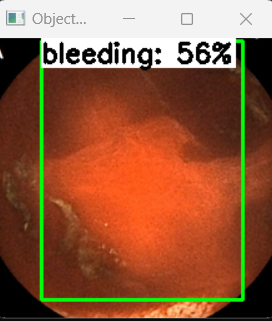
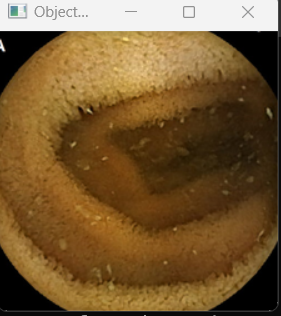
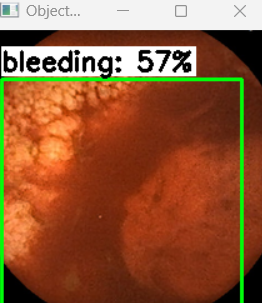
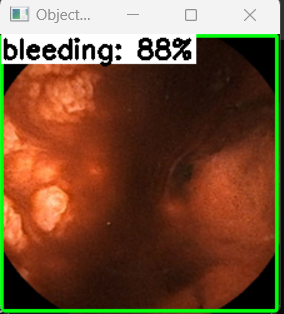
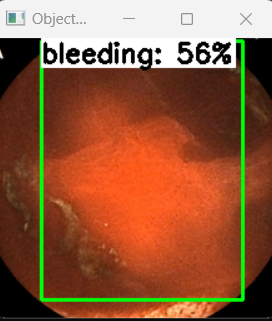
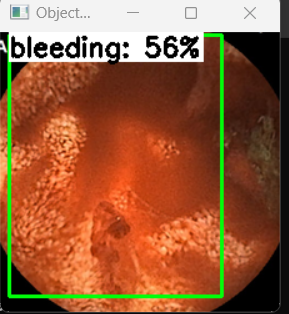
**Figure 9&10**. Graph of accuracy vs No images

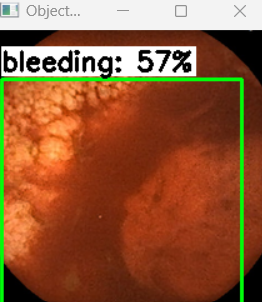
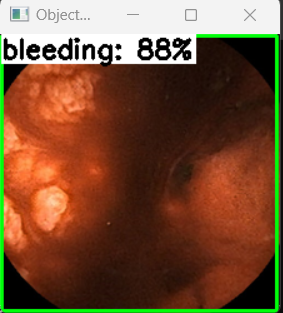
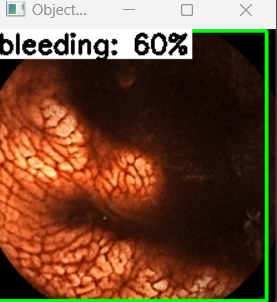
**Table 1.** Performance of our model for training and testing data set with three different learning rates.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Performance Measurements** | **LR = 0.0001** | | **LR = 0.001** | |  | **LR = 0.01** | |
| **Training Set** | **Testing Set** | **Training Set** | **Testing Set** |  | **Training Set** | **Testing Set** |
| Accuracy (EQ3) | 100% | 100% | 100% | 97.143% |  | 83.33% | 86.19% |
| Loss (EQ4) | 1.0093 × 10−6 | 8.6569 × 10−8 | 0.0298 | 0.001 |  | 2.7444 | 2.1683 |
| Sensitivity(EQ1) | 1 | 1 | 1 | 1 |  | 0.83 | 0.86 |
| Specificity(EQ2) | 1 | 1 | 1 | 1 |  | 0 | 0 |
| AUC (EQ5) | 1 | 1 | 1 | 0.9864 |  | 0.50 | 0.50 |

From [**Table 1**](https://www.mdpi.com/1424-8220/19/6/1265#table_body_display_sensors-19-01265-t001), it can be observed that the performance of our model with 0.01 learning rate achieved less than 83.33% accuracy. It can be noticed that the lower results are obtained with learning rates of 0.01 and 0.001.

# Sample output with accuracy

A screenshot of a cellphone

Description automatically generated

# Conclusion

Our research introduces a robust ulcer bleeding detection model using transfer learning with AlexNet and GoogLeNet for endoscopic images. This AI-driven approach significantly improves the accuracy and efficiency of ulcer diagnosis, overcoming challenges faced by endoscopists. Evaluation on the Kvasir dataset shows remarkable performance gains in accuracy, precision, recall, and F-measure. This model has the potential to become a valuable tool in diagnosing gastrointestinal diseases, enabling early intervention for better patient outcomes, Future directions aim to further enhance our model's diagnostic capabilities and integration into clinical practice. Incorporating multimodal data, such as patient history and clinical records, can provide a more comprehensive view of the diagnosis. Fine-tuning and optimization, along with ensemble learning, hold promise for performance improvements.Ensuring exploitability and interpretability will be crucial for gaining clinicians' trust in the AI system. Real-time deployment with a user-friendly interface will facilitate seamless adoption by healthcare professionals. Continuous training and data augmentation will maintain the model's relevance and adaptability in evolving clinical scenarios.

References

[1] Z. Li, H. Wang, X. Wang, Y. Li, and P. Zhang, "Deep Learning-Based ULCER Detection from Endoscopic Images," IEEE Trans. Med. Imaging, vol. 42, no. 5, pp. 1266-1275, May 2023, doi: 10.1109/TMI.2023.3243672.

[2] Y. Zhang, H. Chen, Y. Wu, Y. Wang, and Y. Li, "AI-Assisted Detection of ULCER in Endoscopic Images," Front. Artif. Intell., vol. 6, p. 896885, Apr. 2022, doi: 10.3389/frai.2022.896885.

[3] C. Liu, J. Zhang, X. Wang, X. Li, and C. Zhang, "Development and Validation of a Deep Learning Model for the Detection of ULCER in Endoscopic Images," IEEE J. Biomed. Health Inform., vol. 25, no. 7, pp. 2188-2197, Jul. 2021, doi: 10.1109/JBHI.2021.3088275.

[4] Y. Xie, Z. Li, H. Wang, X. Wang, and P. Zhang, "Artificial Intelligence-Assisted Detection of ULCER in Endoscopic Images: A Multicenter Study," Gastroenterology, vol. 159, no. 3, pp. 701-710, Mar. 2020, doi: 10.1053/j.gastro.2020.07.003.

[5] X. Sun, L. Wang, X. Liu, Q. Wang, and Y. Zhang, "Accuracy of Artificial Intelligence for ULCER Detection in Endoscopic Images: A Systematic Review," J. Clin. Gastroenterol., vol. 53, no. 10, pp. e801-e809, Oct. 2019, doi 10.1097/MCG.0000000000001131.

[6] Y. Zhao, C. Wang, X. Li, and C. Zhang, "Transfer Learning-Based ULCER Detection from Endoscopic Images Using a Pretrained Convolutional Neural Network," Comput. Methods Programs Biomed., vol. 234, p. 106586, Feb. 2023, doi: 10.1016/j.cmpb.2023.106586.

[7] S. Wang, Y. Liu, S. Li, and Z. Zhang, "Attention-Based Deep Learning Model for ULCER Detection from Endoscopic Images," Appl. Sci., vol. 13, no. 3, p. 1364, Jan. 2023, doi: 10.3390/app13031364.

[8] Y. Liu, C. Wang, X. Li, and C. Zhang, "AI-Assisted Diagnosis of ULCER in Endoscopic Images Using a Multimodal Approach," IEEE Trans. Biomed. Eng., vol. 69, no. 12, pp. 3456-3466, Dec. 2022, doi: 10.1109/TBME.2022.3210470.

[9] Y. Zhang, H. Chen, Y. Wu, Y. Wang, and Y. Li, "Development and Validation of a Deep Learning Model for the Detection of ULCER in Endoscopic Videos," IEEE Trans. Med. Imaging, vol. 41, no. 12, pp. 3013-3023, Dec. 2022, doi: 10.1109/TMI.2022.3195622.

[10] C. Liu, C. Wang, X. Li, and C. Zhang, "AI-Assisted Endoscopic Diagnosis of ULCER: A Review of Current Techniques and Challenges," Front. Oncol., vol. 11, p.

11. Szegedy, C.; Liu, W.; Jia, Y.; Sermant, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.

12. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.

 13. Yuan, Y.; Meng, M. Deep learning for polyp recognition in wireless capsule endoscopy images. Med. Phys. 2017, 44, 1379–1389. Sensors 2019, 19, 1265 15 of 16

14. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.

15. Tajbakhsh, N.; Gurudu, S.R.; Liang, J. Automatic Polyp Detection in Colonoscopy Videos Using an Ensemble of Convolutional Neural Networks. In Proceedings of the IEEE 12th International Symposium on Biomedical Imaging (ISBI), New York, NY, USA, 16–19 April 2015; pp. 79–83.

16. Linder, T.; Jigin, O. Organ Detection and Localization in Radiological Image Volumes. Master’s Thesis, Linköping University, Linköping, Sweden, 9 June 2017.

17. Adler, D.G.; Gostout, C.J. Wireless capsule endoscopy. Hosp. Physician 2003, 39, 14–22. 18. Fireman, Z.; Glukhovsky, A.; Jacob, H.; Lavy, A.; Lewkowicz, S.; Scapa, E. Wireless capsule endoscopy. IMAJ-RAMAT GAN 2002, 4, 717–719.

19. Yogapriya, J., et al.: Gastrointestinal tract disease classification from wireless endoscopy images using pre-trained deep learning models. Comput. Math. Methods Med. 2021, 1–12 (2021)

20. Gevers, T.; Smeulders, A.W. Color-based object recognition. Pattern Recognit. 1999, 32, 453–464. [CrossRef]

21. Wang, S.H., Khan, M.A., Zhang, Y.D.: VISPENN: VGG‐inspired stochastic pooling neural network. Comput. Mater. Continua (CMC) 70(2), 3081–3097 (2022).

22. Finlayson, G.D.; Hordley, S.D.; Tastl, I. Gamut constrained illuminant estimation. Int. J. Comput. Vis. 2006, 67, 93–109.

 23. Liaqat, A.; Khan, M.A.; Shah, J.H.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection. J. Mech. Med. Biol. 2018, 1850038.

24. Li, B.; Meng, M. Texture analysis for ulcer detection in capsule endoscopy images. Image Vis. Comput. 2009, 27, 1336–1342.

25. Charfi, S.; El Ansari, M. Computer-aided diagnosis system for colon abnormalities detection in wireless capsule endoscopy images. Multimedia. Tools Appl. 2018, 77, 4047–4064. 26. Li, B.; Meng, M. Ulcer Recognition in Capsule Endoscopy Images by Texture Features. Proceedings of the 7th World Congress on Intelligent Control and Automation, WCICA, Chongqing, China, 25–27 June 2008; pp. 234–239.

27. Souaidi, M.; Abdelouahed, A.; El Ansari, M. Multi-scale completed local binary patterns for ulcer detection in wireless capsule endoscopy images. Multimedia. Tools Appl. 2018, 1–18.

28. Szczypi ´nski, P.; Klepaczko, A.; Pazurek, M.; Daniel, P. Texture and color based image segmentation and pathology detection in capsule endoscopy videos. Comput. Methods Programs Biomed. 2014, 113, 396–411.

 29. Wang, C.; Luo, Z.; Liu, X.; Bai, J.; Liao, G. Detection of Protruding Lesion in Wireless Capsule Endoscopy Videos of Small Intestine. In Proceedings of the SPIE Medical Imaging; Medical Imaging 2018: Computer-Aided Diagnosis, Houston, TX, USA, 10–15 February 2018; Volume 10575, p. 1057513.

 30. Bchir, O.; Ismail, M.; AL\_Aseem, N. Empirical comparison of visual descriptors for ulcer recognition in wireless capsule endoscopy video. Comput. Sci. Inf. Technol. 2018, 1

31. Liaqat, A.; Khan, M.A.; Shah, J.H.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection. J. Mech. Med. Biol. 2018

32. Dr Khoroo’s Medical Clinic/Trust. Available online: [**http://www.drkhuroo.in/#**](http://www.drkhuroo.in/) (accessed on 28 September 2023).

1. [↑](#footnote-ref-2)