International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Intelligent Hemorrhage Identification in Wireless Capsule Endoscopy Pictures Using AI Techniques.

Himanshu Sharma[1]*, Mr. Himanshu Tyagi[2]

[1]*Research Scholar Quantum University, Roorkee
[2]Assistant Professor Quantum University,Roorkee

***Corresponding author:-** Himanshu Sharma
*Research Scholar Quantum University, Roorkee

ABSTRACT
Image segmentation in medical images is performed to extract valuable information from the images by concentrating on the region of interest. Mostly, the number of medical images generated from a diagnosis is large and not ideal to treat with traditional ways of segmentation using machine learning models due to their numerous and complex features. To obtain crucial features from this large set of images, deep learning is a good choice over traditional machine learning algorithms. Wireless capsule endoscopy images comprise normal and sick frames and often suffers with a big data imbalance ratio which is sometimes 1000:1 for normal and sick classes. They are also special type of confounding images due to movement of the (capsule) camera, organs and variations in luminance to capture the site texture inside the body. So, we have proposed an automatic deep learning model based to detect bleeding frames out of the WCE images. The proposed model is based on Convolutional Neural Network (CNN) and its performance is compared with state-of- the-art methods including Logistic Regression, Support Vector Machine, Artificial Neural Network and Random Forest. The proposed model reduces the computational burden by offering the automatic feature extraction. It has promising accuracy with an F1 score of 0.76.

Keywords: Machine learning, Deep learning, Artificial intelligence

1. INTRODUCTION

Gastric cancer is the fifth most common cancer worldwide and seventh most prevalent in accordance to the GLOBOCAN 2018 as shown in Figure 1.1. For males, 10% of the total cancers is Colorectal cancer and that makes it the third most prevalent cancer in males with cases of 663,000 all over the world. Whereas, it is the other most frequent cancer in females with cases of 570,000, which is 9.4% of the total cases globally [1,2].

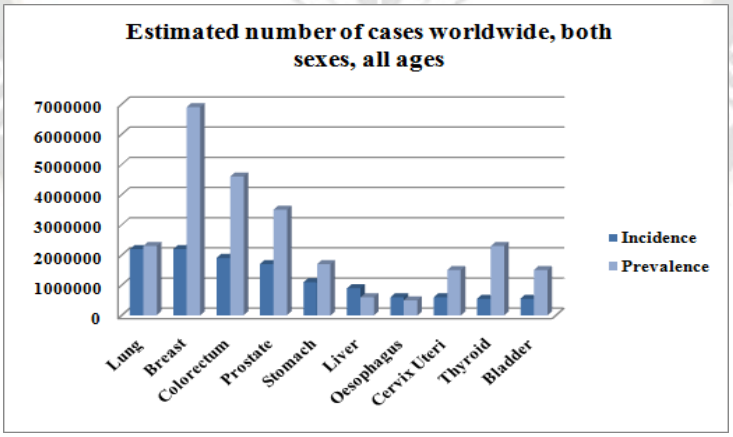


Figure 1.1: The estimated incidence and prevalence of worldwide cancer cases [1,2]

Persistent discomfort in the abdomen is a flag that something is wrong. Other nagging signs are gas, blood in stools, bloating, diarrhea, heartburn, prolonged inflammation in the digestive tract, etc. These are the reasons to go consult a gastroenterologist. Gastroenterology deals with the digestive system and its dysfunctions. The digestive tract or gastrointestinal (GI) tract is an organ system that begins from the mouth and continues up to the anus. GI tract comprises of several organs such as digestive canal, throat or esophagus, liver, duodenum, bile ducts, pancreas, gallbladder, small intestine, and large intestine, colon, rectum. Figure 1.2 shows the GI tract and its major organs.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

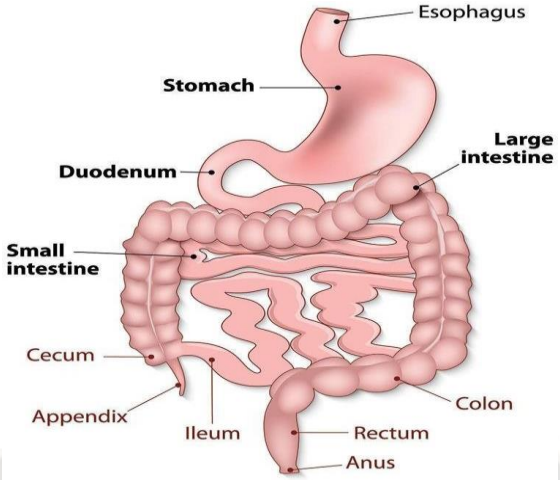


Figure 1.2: Major body parts in the gastrointestinal tract

Gastroenterology is the study of every organ's functionality in the digestive system while concentrating on multiple conditions that may strike them [3]. The study aids an absolute understanding of the digestion method, consumption of nutrient, and excreta. It also addresses the complications that may harm these organs including polyps, ulcers, cancer, and gastroesophageal reflux. Physicians practicing in this domain are addressed as gastroenterologists.

Literature Review

This chapter studies the work done in the field of Capsule Endoscopy to detect plausible abnormalities like polyp, tumor, ulcer, bleeding, etc. The State-of-the- art research done in existing related literature is the motivation for our research work. Related work done by various authors is summarized to analyze different approaches used and compared to analyze the efficiency of the related literature.

Review of literature based on bleeding detection using the Convolutional Neural Network (CNN)

Authors proposed a way in [26] for detection of bleeding regions in CE images employing a semantic segmentation based on deep neural network, called SegNet. CNN is trained using the SegNet layers. Bleeding regions are detected in CE images by segmenting the test images on the trained CNN. The efficiency achieved is 94.42%. In [27], Authors advised a simplified neural network (NN) for the detection of bleeding frames by performing automatic bleeding regions segmentation on the CE dataset. Fitting color channels are chosen as inputs to the NN. An multi-layer perceptrons and CNN are applied to conduct image classification individually. They decreased the number of computational operations. The performance of the recommended systems is assessed using the DICE score. The area under the receiver operating curve (AUC-ROC) is 0.97. Due to significantly fewer computations, CNN is proved to be more beneficial than multi-layer perceptron. Authors in the paper [28] presented a high-level detection system for the bleeding frames in the WCE image dataset. This system is implemented using a deep CNN that detects both active and inactive frames. Authors have designed CNN to have 8 layers. The network composes3 convolutional layers, 3 pooling layers, and two fully connected layers. The ReLU orthe rectifier function is implemented at the convolutional layers and the first FC layer to increase non-linearity in the network. Images are made of different objects that are not linear to each other. Without applying this activation function, the image classification will be treated as a linear problem, while it is in actual a non-linear one. Pooling layers implement Max-Pooling to preserve the main features while also reducing the size of the image. This helps reduce overfitting. One of the main causes for the overfitting to occur is the too much information fed into CNN. Especially if that information is not relevant in classifying the image. SVM is more beneficial in the case where the user wants to depreciate the entropy loss for prediction. So, CNN is devised by substituting the SoftMax regression function at the secondary FC layer with an SVM classifier. But this proposed system resulted in complex computations and required a large dataset of designated images for training the CNN. So, the authors presented a way in the paper [29] that implements the deep-learning technique of CNN with handcrafted features that are obtained using a k-means clustering technique. This model is focused on detection of frames with active and inactive bleeding. This approach reduces the computational cost incurred in the training of CNN.

2. Review of literature based on bleeding detection other ML techniques

In [30], authors proposed a method for automatic feature extraction and detection of bleeding in endoscopy images. The endoscopy images are segmented using block- based segmentation. The local features are discovered using color characteristics. Different algorithms are investigated in this approach to classify the bleeding and non-bleeding images. The performance parameters are also calculated to test the effectiveness of these algorithms. The accuracy obtained by the proposed approach is 95%. In [31], this approach

presented bleeding detection in WCE images based on region-based feature extraction. This method extracts features from HSI and CIE color spaces. Authors used a uniform library binary pattern to label each region. The secondary set of features can be extracted from the grayscale image. Classification of regions is done by support vector machine (SVM) into three categories, namely, non- bleeding region, bleeding region, and background. GrowCut algorithm is used for the concluding segmentation of CE images. This paper in [32] introduced a three-step algorithm for automated detection of bleeding in endoscopy images. Authors have done Key-frame extraction and edge removal as the first step of preprocessing. In the second step, they separated the bleeding images from the dataset of all the frames using the KNN classifier, applying the concept of principle color spectrum by utilizing the superpixel color histogram feature. In the last step, the segmentation of bleeding regions from the various color spaces is executed by securing a 9-D color feature vector at the superpixel feature. The accuracy attained is 0.9922. This proposed system presented in [33], focused on identifying the multiple blood specks

in the several frames captured from the WCE video. To overcome the performance degradation due to the small size of the region of interest, the authors suggested an unsupervised ML technique. It breaks the principal classification query into many confined classification queries. This technique cluster the training set using fuzzy C- means, then implement improvised KNN on the determined centers rather than the entire training dataset. Authors have also analyzed the performance and results of the proposed algorithm as opposed to the typical KNN and SVM.
In the paper [34], Naive Bayes classifier and superpixel segmentation are used for automated obscure bleeding disclosure in WCE video dataset. They used the color histogram for region discovery and feature extraction. In the final step, they adopted an improvised semi naive bayesian classifier. In [35], authors have classified bleeding and non-bleeding frames in wireless CE images utilizing color histogram of block statistics. Local feature extraction is done using a WCE image block, preferably of an individual pixel. For the contrasting color panels of RGB color space, index values are defined. So, the authors used index values to extract the color histogram. Color histogram is useful in securing distinct color texture characteristics. Feature reduction using color histogram and principal component analysis is adopted to decrease the dimension of these local features. Extracted local features that do not result in any computational strain provides the blocks with bleeding regions. Authors used a public dataset of 2350 images that renders 97.85% accuracy.

In [36], authors have worked on the system that distinguishes the bleeding images and regions from the WCE dataset. WCE images are preprocessed to convert into a color space defined by green to red pixel-ratio. These transformed images are used to obtain various analytical features from the overlying spatial blocks. These various blocks are then clustered into two clusters using K-means based clustering (unsupervised). These clusters are used to obtain cluster-based features. These features combined into a global feature is used along with differential cluster-based features to detect blood zone frames using an SVM (supervised learning classifier). The proposed system achieved 97.05% precision.

Review of literature based on detection of other anomalies of the digestive tract Authors in [37] have worked on a computerized way of finding abnormalities in the endoscopy images. These abnormalities can be erosion, erythema, ulcerations, polyp, bleeding, etc. This proposed method examines images for varied textures so that, it can differentiate abnormal images from the normal ones effectively. The distribution of different textures in an endoscopy image can be captured using a textons histogram. It is done by applying a FB (filter bank) and LBP (local binary patterns). This proposed approach gives 92% recall and 91.8% specificity on WCE images.

In [38], authors have worked on a computer-aided system that uses various approaches to detect the abnormalities in the images obtained from CE. Authors have employed CNN, region recommendation, transfer learning. In the first step, they have used a CascadeProposal to recommend high-recall regions and abnormal frames. In the second step, the authors used a multi-regional combination technique to detect the regions of interest and have also operated a salient region segmentation approach to catch certain region spots. For object boundary filtration, a dense-region fusion algorithm is applied. And lastly, to increase the efficiency of the proposed model, transfer learning tactics are exercised in CNN. Authors of [39] presented a computer- aided look-behind fully CNN (LB-FCN) algorithm to automatically catch the anomalies in CE images. It uses blocks of parallel convolutional layers with varied filter dimensions to derive the multi-scale features from WCE images. All the LB linked features are combined with the features deduced from prior layers. As LB-FCN has fewer free parameter as compared to conventional CNN, it makes it much easier to train the network on smaller datasets. The AUC performance of LB-FCN achieved is 93.5%.

In [40], they have aimed at reducing the analysis time of the WCE video frames by presenting a computer-aided approach that automatically identifies the abnormal frames. They have worked on an ensemble of two SVMs that are based on HSV and RGB color spectrums. Feature selection and parameter tuning are done by using a nested cross-validation approach. For the betterment of performance, exhaustive analysis is carried out to decide the best feature sets. The dataset used comprises 8872 WCE frames. This fusion system renders an accuracy of 95%, specificity of 95.3% and sensitivity of 94%. A CNN is

suggested in [41] for the GI angioectasia detection during small bowel in CE images. Local features are extracted through deep feature extraction using an approach of segmentation of images based on semantics. Authors created a semantic segmentation-based CNN for classification of GI angioectasias. The sensitivity and specificity achieved is 100% and 96% respectively.

The work done in [42] aims for an automated way for the detection of angioectasias in WCE image dataset. This approach depends on the automatic separation of a region of importance. That region is chosen by applying a module for the task of image segmentation based on the approach of Maximum a Posteriori where a new hastened variant of the Expectation-Maximization is also advised. This proposed method attained sensitivity and specificity values of 96% and 94.08% respectively with 95.58% accuracy in a database comprising 800 WCE frames designated by two gastroenterologists.

In this paper [43], they have conducted ulcer and lesion detection and classification in WCE dataset employing two pre-trained CNN, GoogleNet and AlexNet. These two networks perform object classification to obtain ulcer and non-ulcer frames. Due to a huge number of layers in GoogleNet, AlexNet resulted in double the efficiency of GoogleNet for training. The efficiency of both networks is enhanced by tuning the parameters. It is also found in this study that higher the learning rate of the network, higher is the resulting accuracy. The learning rate of 0.0001 renders adequate results for both the networks. AlexNet attained 100% accuracy with the rate of 0.001.

Authors of [44] trained a deep CNN to distinguish ulcers and erosions in small bowel CE images automatically. This CNN is based on an single shot multiBox detector that holds 16 layers. The CNN is trained using SSD on the 5360 images. For the testing phase, 10,440 WCE images are fed to CNN, out of which, 440 are of erosions or ulcers. This system renders an accuracy of 91.5%. Whereas, in the paper [45], they have focused on decreasing the misidentification rate for a polyp in CE images. This will support the professionals in finding the most significant regions to pay consideration. Features are deduced using color wavelet and CNN. These extracted features are then fed to a train an SVM. SVM will classify the CE frames into the polyp region and normal frames classes. They achieved 98.34% accuracy.

The study performed by [46] has focused on detecting a hookworm abnormality in wireless capsule endoscopy images. They have adopted the deep learning algorithm to recognize the tube-like pattern of hookworm. For the better activity of the classification, two neural networks are employed, edge extraction CNN and hookworm classification CNN. Both the CNNs are seamlessly integrated into the recommended system to evade edge feature caching. The edge extraction CNN provides the tubular regions and the hookworm CNN gives the feature maps. Both the results are integrated into the pooling layers to produce an intensified feature map accentuating tubular regions. They achieved 88.5% accuracy.

Comparative Analysis

Most of the investigated modern techniques based on this comparison have worked on the automated detection of abnormalities seen in the GI Tract Endoscopy. Our research concentrates on the automatic bleeding detection in capsule endoscopy videos using a convolutional neural network. Literature review organization in terms of techniques has been shown in Figure 2.1 while highlighting the CNN based approaches. We are using CNN in our research for the detection of the bleeding frames in WCE images as it offers a wide range of benefits over other ML techniques. CNN is fast, efficient, and it needs limited preprocessing of the images [39, 43, 44]. CNN derives the features from images automatically, which results in a reduced computational burden. Features are learned during the training of the network on the image dataset. One more benefit of using CNN is that only the number of filters and the filter size is required to be defined, whereas the values of the filters are determined by CNN automatically during the training phase. Unlike most of the other ML techniques, object detection in images is carried out by CNN regardless of the location of the object to be recognized. Pooling feature of CNN also prevents overfitting of the network.

Methodology

A computerized system for bleeding detection in WCE images is proposed to catch the presence of a threat in the digestive tract that caused the bleeding. A dataset of WCE videos with frames holding both bleeding and non-bleeding frames is collected (from PSRI hospital, New Delhi) and used for the proposed approach. At very first, we extracted images from the collected WCE videos using a VLC media player at a rate of 2 frames per second. We have used VLC software for the extraction as it is very simple easy to use. One can set the properties of the images to be extracted like dimensions, bit depth, frame extraction rate, etc. in the preferences tab of the VLC software.

At an initial phase, image preprocessing steps are done to enhance the performance of the algorithm to which the processed images are being fed. The first step of the pre-processing is to *resize* all the images to the same size, in case the acquired images have different sizes. It is done as the deep learning algorithms need the images to be of the same size to ensure the efficient implementation of the operations being carried out. In the second step, the resized images are treated to a *de-noising* algorithm to soften the images, to remove any noise and, unwanted distortions present in the images. The noise present in the images degrade the algorithm's performance and increase complexity and computations [51]. The essential quality of a reliable image de-noising technique is that it should effectively eliminate noise as much as plausible as well as conserve edges and information [52]. We have used the stationary discrete wavelet transform (SWT) for de-noising process in our work. The stationary discrete wavelet transform is a wavelet transform tool that enables to examine the statistically non-predictable signals, particularly at those regions that have discontinuities [53]. Since images have discontinuities at the edges, they can be

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

presented spatially in a multi-resolution manner by using the SWT technique [54]. These RGB de-noised images are then classified using supervised deep learning.

As detecting bleeding frames is a binary classification problem. Supervised learning has been chosen because labeled dataset has been obtained from known gastroenterologists for model training. Thus, we can also evaluate its performance based on the correct classifications and misdetection rate as given by the labeled dataset. It is conducted by employing a Convolutional Neural Network (CNN) architecture to classify the frames into bleeding and non-bleeding to detect the presence of an anomaly in the digestive system as it offers a wide range of benefits over other ML techniques. It is fast, efficient, and it needs very limited preprocessing of the images [39, 43, 44] as it derives the features from images automatically, which results in a reduced human intervention. Features are learned during the training of the network on the image dataset. Another benefit of using CNN is that only the number of filters and the filter size is required to be defined, whereas the values of the filters are determined by CNN automatically during the training phase. Unlike most of the other ML techniques, object detection in images is carried out by CNN regardless of the location of the object to be recognized. Pooling feature of CNN also prevents overfitting of the network. CNN carry out automatic segmentation of the images for classification.

3. Proposed approach

The proposed method as shown in figure 4.2 works on preprocessing of WCE images by applying a de-noising algorithm and color segmentation on them and implementing supervised deep learning for classification.

- WCE video dataset is collected from PSRI hospital (New Delhi) and fed to the VLC software to extract images. Figure 4.1 shows the snapshot of the extracted image dataset.

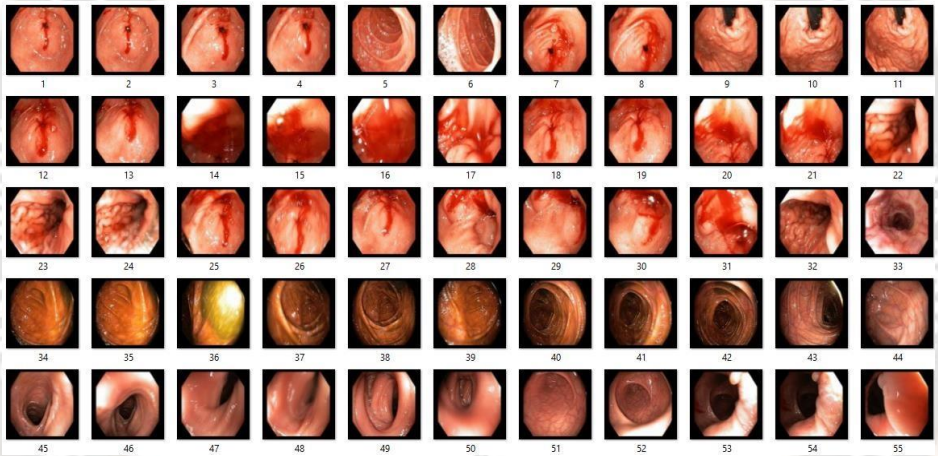


Figure 4.1: Snapshot of the WCE dataset collected from PSRI hospital [6]

- The number of WCE images extracted for the dataset is 2,621 with 505bleeding and 2,116 normal frames.
- All the images extracted are resized to prepare them to feed to our proposed model. High-resolution images involve more computations and higher memory specifications. If the input is a scaled-down variant of the bigger images, then determining key features in the initial layers will be easier for the network. For higher resolution images, these principal features might be discovered at the subsequent layers of the network [55]. So, we have resized our WCE images to a size of 100×100 pixels for a scaled-down CNN.
- Stationary Discrete Wavelet Transform tool in MATLAB (SWT) is used as a de-noising algorithm to smoothen the images, to remove any noise and, undesired distortions present in the images.
- These processed images are then fed to the convolutional neural network (CNN). The complexity of the model depends upon the complexity of the data. We can start by adding only one hidden layer in the network with neurons and then check the test accuracy of the
- network using cross-validation. And then deepen the network by adding more layers and neurons. The proper way to decide the number of hidden layers and neuron is discussed in section 4.4 in detail.
- CNN works by automatically extracting features from the image using the training set of WCE images.
- CNN parameters and options are tuned to obtain better performance. The methods for tuning the options of the CNN is discussed in section 4.4 in detail.
- The trained network is applied to the test set of images for the classification after it met the criteria for validation.
- The accuracy, sensitivity, and specificity of the network are measured from confusion matrix, Precision-Recall graph curve is also plotted.

- Performance of the proposed model is compared with state-of-the-art methods such as Linear, SVM, ANN and random forest algorithms.

4. Dataset preparation

The WCE dataset used in this research is collected from the PSRI institute of gastroenterology (Pushpawati Singhania Research Institute, Delhi, India). A set of 2621 WCE images is extracted from the video using VLC software at a rate of 2 frames per second. This dataset comprises of 505 bleeding frames and 2,116 normal frames. The WCE images have color homogeneity problem as the color to be detectedhas a wide range of shades. The blood shade may fluctuate extensively from bright red to deep red to brownish, also containing redness of the normal skin tissues.

Pre-processing on the images

After acquiring the WCE image dataset, image preprocessing steps are done to enhance the performance of the model to which the processed images are being fed. Pre-processing steps of image processing techniques are applied to the acquired image dataset for getting more enhanced images which are then fed to the network for binary classification. The steps for preprocessing are as follows:

Resizing to a global size

The first step of the pre-processing is to resize all the images to the same size, in case the acquired images have different sizes. It is done as the deep learning algorithms need the images to be of the same size to ensure the efficient implementation of the network. High-resolution images involve more computations and higher memory specifications. If the input is a scaled-down variant of the bigger images, then determining key features in the initial layers will be easier for the network. For higher resolution images, these principal features might be discovered at the subsequent layers of the network [55]. So, beforehand performing the next steps of the proposed system, the CE images are resized into a size of 100×100 pixels to make all the images of equal dimensions. We have chosen 100×100 pixels size for a scaled-down CNN.

De-noising using SWT

In our thesis, we have implemented stationary wavelet transform (SWT) for noise removal. The stationary discrete wavelet transform is a wavelet transform tool that enables to examine the statistically non-predictable signals, particularly at the region of discontinuities [53]. Since images have discontinuities at the edges, they can be presented spatially in a multi-resolution manner by using the SWT technique [54]. SWT is a wavelet transform that is used to transform signals or images to derive valuable information for the analysis as well as saves the computational cost and reduce required memory space. As we are working with medical images, the loss in information can adversely affect the result. So, SWT has been chosen for de-noising to secure the medical information in the images while undergoing the decomposition and de-noising. SWT de-noising requires decomposition, level thresholding, and inverse transforming for reconstructing the image. For the step of decomposition, SWT algorithm decomposes an image into coefficients to get details about the image like contrast, correlation, energy, homogeneity, entropy, etc. It is done by a choosing a wavelet that decomposes an image to get desired results. There are many wavelets that can be applied for decomposition like Haar, Daubechies, SymN, etc. We have chosen Daubechies for our work as it best proven in paper [56] for the image decomposition and de-noising. Daubechies wavelet being an orthogonal wavelet does not unnecessarily color the white noise, preserves the energy and relatively offer longer support [57]. The SWT decomposition of the image results into four sub- images to get the coefficients. These sub-images are obtained from different applications of vertical and horizontal filters that are Low-pass filter (LPF) and High- pass filter (HPF). The resultant four sub-images of varying contrast, orientations, sharpness, and resolutions are called as approximations (average component) and detail components (horizontal, vertical and diagonal). The approximation coefficient is the low-resolution components and is the result of the LPF (averaging filter) of the SWT. Whereas, the detail coefficients are the output of the HPF (difference filter) of the SWT and are considered as the high-resolution components. The SWT involves up-sampling performed by of the LPF and HPF instead of down-sampling as done in discrete wavelet transform. Up-sampling involves transforming the filters by embedding zeros between the coefficients of the filters which enables the SWT to preserve the coefficients at each level of transformation. In this way, the SWT provides us with coefficients at each level over the similar frequency ranges and same as original in length. It means that the produced coefficients are half in resolution but are same in size as the input image. Whereas, in discrete wavelet transform, the coefficients are halved in size with each successive level as it involves down- sampling [58]. During the SWT decomposition of the images, some of the obtained coefficients contain important details (high-resolution sub-images). Very small details are disregarded without essentially altering the key features of the images. The purpose of thresholding is to replace all high-resolution sub-image coefficients that are smaller than a selective threshold with zero. These coefficients are utilized in the reconstruction of the image by using an inverse wavelet transformation. The effectiveness of SWT de-noising depends on the threshold value [59].

We have chosen global thresholding by employing a singular threshold value to every detail coefficient in all wavelet levels. The values higher than the threshold value are considered as data or details whereas smaller values are deemed as noise. The threshold limit must be implemented in one of the two modes, a soft or hard thresholding mode. As stated in [60], Soft thresholding provides a more visually pleasing image and decreases the sudden sharp

transitions that happen in hard thresholding. Therefore, soft thresholding is chosen over hard thresholding. After de-noising the images with these threshold limits, we apply an inverse transform to reconstruct the original image from the sub-images but without noise. The reconstruction is the reversed course of decomposition or inverse wavelet transform. The modified approximation and detail coefficients at all levels are up sampled and moved through the low pass reconstruction filter (LPRF) and high pass reconstruction filters (HPRFs) and then added to obtain a level reconstructed image. To obtain the original image, this manner is maintained through the corresponding number of levels as in the process of decomposition of the original image [61].

Performance metrics from the confusion matrix

Several performance parameters are derived from the confusion matrix such as accuracy, sensitivity, specificity, and others.

- **Accuracy**

Accuracy refers to the total number of correct classifications done by the network out of the total number of classifications made. As shown in equation (1), accuracy is the rate of true predictions by all the true and false predictions combined.

$$\textbf{Accuracy} = (TP + TN) / (TP + FP + FN + TN) \quad (1)$$

- **Sensitivity or Recall**

Sensitivity is the rate of accurately predicted positives to genuine positives. Recall gives us an idea about a model's performance proportionate to false negatives. As shown in equation (2), recall focuses on catching all the frames that have "bleeding" with the prediction as "bleeding", not just concerning catching frames correctly. For medical image classification problems, high sensitivity is preferred as it indicated high true positive value and low number of false negatives.

$$\textbf{Sensitivity or Recall} = TP / (TP + FN) \quad (2)$$

- **Specificity or True Negative rate (TNR)**

Specificity is the rate of accurately predicted negatives to the actual negatives. Specificity is the exact reverse of sensitivity. Equation (3) shows the specificity derived from a confusion matrix.

$$\textbf{Specificity} = TN / (TN + FP) \quad (3)$$

- **Precision**

Precision is the rate of accurately predicted positives to all the predicted positives. In equation (4), precision shows the proportion of the frames that are detected as having the presence of bleeding, actually had bleeding.

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

Recall provides us an idea about a network's performance concerning false negatives, the frames that the network missed, and the Precision provides us with the idea of its performance concerning false positives, the frames that were predicted. Precision is about predicting frames correctly, whereas Recall is about prediction all the positive frames correctly [67]. So, for minimizing false negatives, we have to focus on getting Recall as best as possible with a decent and acceptable Precision value. The values of both Precision and Recall can be monitored by a single value performance metric called as F1 score.

- **F1 Score**

To consider the role of both precision and recall, the F1 score is computed. As presented in equation (5), F1 score is simply the harmonic mean of precision and recall. In the case of unbalanced class distribution in the dataset, F1 score is a better evaluation metric than accuracy. Low value of F1 score indicates a problem when one of the Precision and Recall has a low value. In that case, F1 score is closer to the smaller value than the bigger value out of these two.

$$\textbf{F1 Score} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (5)$$

- **Precision-Recall curve**

To demonstrate the trade-off in precision and recall, the precision-recall (PR) curve is plotted. In an image dataset with imbalanced class distribution, the PR curve is a better evaluation metric than accuracy. Thus, PR curve gives a more informational depiction of the performance of the network with unbalanced dataset [68,69]. The area under the PR curve varies from 0 to 1 and also gives an idea about the network's performance. The closer AUC is to 1, the better the network [70]. The closer the curve is to the top-right edge, the more reliable the system. Henceforth, a greater area under the curve (AUC) symbolizes that the system has higher precision and higher recall [69].

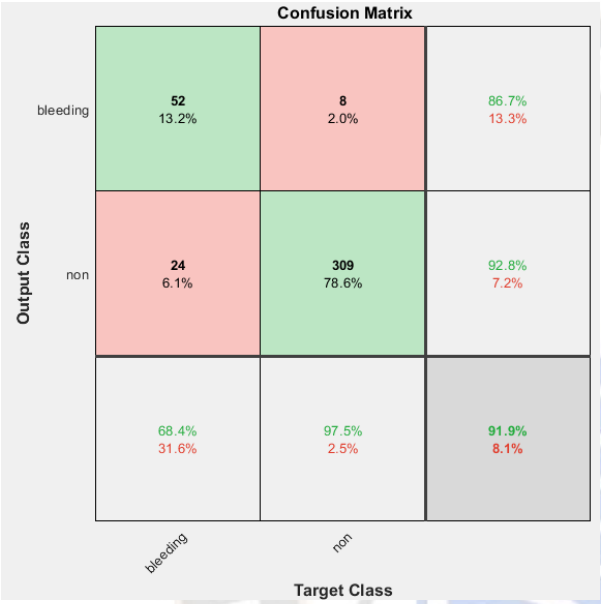
Comparison with state-of-the-art methods

For the evaluation of the proposed model, we have compared the performance of our model with state-of-the-art methods like Logistic Regression model, SVM, artificial neural network (ANN) and Random Forest. These models accept the data of color and texture features as input datasets with the labels for all WCE images obtained from the PSRI hospital, New Delhi. These color and texture features can be extracted from the color histogram and co-occurrence matrix of an individual image respectively. The models are trained, validated and tested on the data to get the evaluation metrics for the respective model. The ratio of train, validation and test data is set as same as our model that is 70%, 15%, and 15% respectively. The performances of these models are then compared with the proposed model.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

5. Results of WCE

The video obtained by the pill-sized camera used in WCE is carefully observed by thedoctor for irregularities inside the digestive system. It may need a few days or even weeks to get the outcomes of the wireless capsule endoscopy. While some issues can be immediately discerned, like bleeding or contractions, some may be obscure like inactive bleeding, blood specks that hard to detect.



S. No.	Evaluation metric	Value
1	Validation Accuracy	90.84%
2	Test Accuracy	91.92%
3	Sensitivity/Recall	68.42%
4	Specificity	97.48%
5	Precision	86.67%
6	F1 Score	0.7647

Types of disorders diagnosed with WCE

Wireless capsule endoscopy progresses to grow better as it has reformed diagnosis completely by providing a fine-tuned, comfortable, and non-invasive ways of monitoring the core of the small intestine. Some anomalies of the small intestine that are diagnosed by WCE include:

- Angiodysplasias: severe bleeding from the small blood vessels foundunderneath the internal intestinal wall causing blood loss resulting in anemia.
- Small intestinal polyps and tumors such as carcinoid tumor, and lymphoma,and colon cancer.
- Angioectasias: bleeding from tiny vascular tumors or angioectasias, lesions.
- Inflammatory bowel disease or Crohn's disease.

Accuracy

The exactness of WCE can be differed by the purpose of the examination and the equipment used. CE is deemed superior to diagnose the IBD in the small intestine or Crohn's disease, as it detects inflammatory tumors earlier as contrasted to all other methods. Furthermore, CE is more specific in identifying celiac disease, even thougha biopsy is nevertheless required for an absolute diagnosis.

Capsule camera

There are several pill cameras available in the market such as PillCam SB3 (originally known as M2A) by Given Imaging Ltd., MiroCam by IntroMedic company, EndoCapsule by Olympus America, OMOM@Jishan by Science and Technology, CapsoCam Plus by CapsoVision, etc. [14]. M2A was the first capsule-sized camera ensemble to introduced into the capsule endoscopy field [7]. Table 1.1 shows the comparison between these pill-sized cameras available in the market for evaluating the performance, affordability and approvals. This table presents the difference in these pill cameras, how they differ in size, weight, durability, resolution.

Table 1.1: Comparison of pill cameras available in the market [15]

Capsule	PillCam SB 3 Given Imaging	EndoCapsuleOlympus America	MicroCamIntromedicCompany	OMOM Jinshan Science and Technology
Size	Length: 26.2 mm Diameter:	Length: 26 mm Diameter: 11mm	Length: 24.5 mmDiameter: 10.8 mm	Length: 27.9 mm Diameter:

	11.4 mm			13 mm
Weight	3.00 g	3.50 g	3.25-4.70 g	6.00 g
Battery Life	8 h or longer	8 h or longer	11 h or longer	6-8 h or longer
Resolution	340 x 340	512 x 512	320 x 320	640 x 640

6. Conclusions

The proposed bleeding detection model checks for the bleeding frames in the wireless capsule endoscopy dataset. We have proposed a convolutional neural network for the detection to be done as the network is being fed with the de-noised WCE images. The main tasks accomplished in the research are listed below:

- For the dataset to be used in the classification, the WCE video dataset is collected from PSRI hospital (New Delhi, India).
- WCE images are extracted from the collected video using VLC software and are pre- processed to get the equal-sized, de-noised images to make it easier for the model to work on them.
- A convolutional neural network is implemented for the classification of WCE frames into the bleeding and non-bleeding categories.
- The performance of the proposed model is compared with other state-of-the-art models on the basis of evaluation parameters.

7. Challenges and Limitations

The proposed approach involves multiple challenges and assumptions. Challenges involve (a) movements of the camera and organs (b) limited quality of luminance to capture the site texture for diagnose inside the body. Assumptions involve (a) labeled data must be available to train the model (b) de-noising must be done for image clarity (c) data class ratio should be balanced for better classifier accuracy (d) parameter tuning for CNN is also an open question (e) dataset sample used for the case study is quite small to capture most of the image attributes for real world applications (f)system setup for this automatic method and its maintained is also an overhead.

8. Future Scope

The future scope of the proposed model of CNN is listed below.

- Bleeding in the digestive tract is a huge threat to the health of a person as the disorders in the digestive system can be fatal. For example, ulcers, colon cancers, etc.

References

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *68*(6), 394-424.
2. Rawla, P., & Barsouk, A. (2019). Epidemiology of gastric cancer: global trends, risk factors and prevention. *Przeglad gastroenterologiczny*, *14*(1), 26.
3. Burridge, A. L. (2003). Ancient Egyptian medicine. *JAMA*, *290*(6), 826-827.
4. Gillespie, M. P. (n.d.). The Language of Medicine - ppt download. Retrieved from https://slideplayer.com/slide/2374097/
5. Chitra, S., Ashok, L., Anand, L., Srinivasan, V., & Jayanthi, V. (2004). Risk factors for esophageal cancer in Coimbatore, southern India: a hospital-based case- control study. *Indian journal of gastroenterology*, *23*(1), 19-21.(n.d.). Retrieved from https://www.psrihospital.com/
6. Iddan, G., Meron, G., Glukhovsky, A., & Swain, P. (2000). Wireless capsule endoscopy. *Nature*, *405*(6785), 417.
7. Delvaux, M., & Gay, G. 2008).Capsule endoscopy: technique and indications. *Best Practice & Research Clinical Gastroenterology*, *22*(5), 813-837.
8. Anderson, J., & Roberts, E. (2019, June 24). How Capsule Endoscopy Is Used to Diagnose Digestive Disease. Retrievedfrom https://www.verywellhealth.com/capsule-endoscopy-for-celiac-disease- diagnosis-562692
9. McAlindon, M. E., Ching, H. L., Yung, D., Sidhu, R., & Koulaouzidis, A. (2016). Capsule endoscopy of the small bowel. *Annals of translational medicine*, *4*(19).
10. Lapalus, M. G., Dumortier, J., Fumex, F., Roman, S., Lot, M., Prost, B., ...& Ponchon, T. (2006). Esophageal capsule endoscopy versus
11. esophagogastroduodenoscopy for evaluating portal hypertension: a prospective comparative study of performance and tolerance. *Endoscopy*, *38*(01), 36-41.
12. Irvine, A. J., Sanders, D. S., Hopper, A., Kurien, M., & Sidhu, R. (2016). How does tolerability of double balloon enteroscopy compare to other forms of endoscopy?. *Frontline Gastroenterology*, *7*(1), 41-46.
13. (2019, July 16). Capsule endoscopy. Retrievedfrom
14. https://www.mayoclinic.org/tests-procedures/capsule-endoscopy/about/pac- 20393366.
15. Cave, D. (2002). Wireless video capsule endoscopy. *Clin Perspect Gastroenterol*, *5*(4), 203.
16. Van de Bruaene, C., De Looze, D., & Hindryckx, P. (2015). Small bowel capsule endoscopy: Where are we after almost 15 years of use?. *World journal of*

gastrointestinal endoscopy, *7*(1), 13.
17. Swain, P., & Fritscher-Ravens, A. (2004). Role of video endoscopy in managing small bowel disease. *Gut*, *53*(12), 1866-1875.
18. Liangpunsakul, S., Mays, L., & Rex, D. K. (2003). Performance of Given suspected blood indicator. *The American journal of gastroenterology*, *98*(12), 2676.
19. D'Halluin, P. N., Delvaux, M., Lapalus, M. G., Sacher-Huvelin, S., Soussan, E. B., Heyries, L., ... & Heresbach, D. (2005). Does the "Suspected Blood Indicator" improve the detection of bleeding lesions by capsule endoscopy?. *Gastrointestinal endoscopy*, *61*(2), 243-249.
20. Signorelli, C., Villa, F., Rondonotti, E., Abbiati, C., Beccari, G., & de Franchis,
21. R. (2005). Sensitivity and specificity of the suspected blood identification system in video capsule enteroscopy. *Endoscopy*, *37*(12), 1170-1173.
22. Buscaglia, J. M., Giday, S. A., Kantsevoy, S. V., Clarke, J. O., Magno, P., Yong, E., & Mullin, G. E. (2008). Performance characteristics of the suspected blood indicator feature in capsule endoscopy according to indication for study. *Clinical gastroenterology and hepatology*, *6*(3), 298-301.
23. Westerhof, J., Koornstra, J. J., & Weersma, R. K. (2009). Can we reduce capsule endoscopy reading times?. *Gastrointestinal endoscopy*, *69*(3), 497-502.
24. Mahapatra, S. (2019, January 22). Why Deep Learning over Traditional Machine Learning? Retrieved from
25. https://towardsdatascience.com/why-deep-learning-is- needed-over-traditional-machine-learning-1b6a99177063.
26. Rubeaux, M. (n.d.). AI in Medical Imaging. Retrieved from https://blog.keosys.com/ai-in-medical-imaging.
27. Johnson, J., &Karpathy, A. (n.d.). Convolutional Neural Networks for Visual Recognition. Retrieved from http://cs231n.github.io/convolutional-networks/.
28. Brownlee, J. (2019, August 6). Use Early Stopping to Halt the Training of NeuralNetworks At the Right Time. Retrieved from https://machinelearningmastery.com/how-to-stop-training-deep-neural- networks-at-the-right-time-using-early-stopping/.
29. Ghosh, T., Li, L., &Chakareski, J. (2018, October). Effective Deep Learning for Semantic Segmentation Based Bleeding Zone Detection in Capsule Endoscopy Images. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (pp. 3034-3038). IEEE.
30. Hajabdollahi, M., Esfandiarpoor, R., Soroushmehr, S. M., Karimi, N., Samavi, S., &Najarian, K. (2018). Segmentation of bleeding regions in wireless capsule endoscopy images an approach for inside capsule video summarization. *arXiv preprint arXiv:1802.07788*.
31. Jia, X., & Meng, M. Q. H. (2016, August). A deep convolutional neural network for bleeding detection in wireless capsule endoscopy images. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 639-642). IEEE.
32. Jia, X., & Meng, M. Q. H. (2017, July). Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 3154-3157). IEEE.
33. Obukhova, N., Motyko, A., Timofeev, B., & Pozdeev, A. (2019, April). Method of Endoscopic Images Analysis for Automatic Bleeding Detection and Segmentation. In *2019 24th Conference of Open Innovations Association (FRUCT)* (pp. 285-290). IEEE.
34. Tuba, E., Tomic, S., Beko, M., Zivkovic, D., & Tuba, M. (2018, November). Bleeding Detection in Wireless Capsule Endoscopy Images Using Texture and Color Features. In *2018 26th Telecommunications Forum (TELFOR)* (pp. 1-4). IEEE.
35. Xing, X., Jia, X., & Meng, M. H. (2018, July). Bleeding Detection in Wireless Capsule Endoscopy Image Video Using Superpixel-Color Histogram and a Subspace KNN Classifier. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1-4). IEEE.
36. Bchir, O., Ismail, M. M. B., & AlZahrani, N. (2019). Multiple bleeding detection in wireless capsule endoscopy. *Signal, Image and Video Processing*, *13*(1), 121-126.
37. Sivakumar, P., & Kumar, B. M. (2018). A novel method to detect bleeding frame and region in wireless capsule endoscopy video. *Cluster Computing*, 1-7.
38. Ghosh, T., Fattah, S. A., & Wahid, K. A. (2018). CHOBS: Color Histogram of Block Statistics for Automatic Bleeding Detection in Wireless Capsule Endoscopy Video. *IEEE journal of translational engineering in health and medicine*, *6*, 1-12.
39. Ghosh, T., Fattah, S. A., Wahid, K. A., Zhu, W. P., & Ahmad, M. O. (2018). Cluster based statistical feature extraction method for automatic bleeding detection in wireless capsule endoscopy video. *Computers in biology and medicine*, *94*, 41-54.
40. Nawarathna, R., Oh, J., Muthukudage, J., Tavanapong, W., Wong, J., De Groen,
41. P. C., & Tang, S. J. (2014). Abnormal image detection in endoscopy videos using a filter bank and local binary patterns. *Neurocomputing*, *144*, 70-91.
42. Lan, L., Ye, C., Wang, C., & Zhou, S. (2019). Deep Convolutional Neural Networks for WCE Abnormality Detection: CNN Architecture, Region Proposal and Transfer Learning. *IEEE Access*, *7*, 30017-30032.

43. Diamantis, D. E., Iakovidis, D. K., & Koulaouzidis, A. (2019). Look-behind fully convolutional neural network for computer-aided endoscopy. *Biomedical Signal Processing and Control*, *49*, 192-201.
44. Deeba, F., Islam, M., Bui, F. M., & Wahid, K. A. (2018). Performance assessment of a bleeding detection algorithm for endoscopic video based on classifier fusion method and exhaustive feature selection. *Biomedical Signal Processing and Control*, *40*, 415-424.
45. Leenhardt, R., Vasseur, P., Li, C., Saurin, J. C., Rahmi, G., Cholet, F., ... & Sacher-Huvelin, S. (2019). A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointestinal endoscopy*, *89*(1), 189-194.
46. Vieira, P. M., Silva, C. P., Costa, D., Vaz, I. F., Rolanda, C., & Lima, C. S. (2019). Automatic Segmentation and Detection of Small Bowel Angioectasias in WCE Images. *Annals of biomedical engineering*, *47*(6), 1446-1462.
47. Alaskar, H., Hussain, A., Al-Aseem, N., Liatsis, P., & Al-Jumeily, D. (2019). Application of Convolutional Neural Networks for Automated Ulcer Detection in Wireless Capsule Endoscopy Images. *Sensors*, *19*(6), 1265.
48. Aoki, T., Yamada, A., Aoyama, K., Saito, H., Tsuboi, A., Nakada, A., ... & Matsuda, T. (2019). Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointestinal endoscopy*, *89*(2), 357-363.
49. Billah, M., & Waheed, S. (2018). Gastrointestinal polyp detection in endoscopic images using an improved feature extraction method. *Biomedical engineering letters*, *8*(1), 69-75.
50. He, J. Y., Wu, X., Jiang, Y. G., Peng, Q., & Jain, R. (2018). Hookworm detection in wireless capsule endoscopy images with deep learning. *IEEE Transactions on Image Processing*, *27*(5), 2379-2392.
51. Byrne, M. F., & Donnellan, F. (2019). Artificial intelligence and capsule endoscopy: Is the truly "smart" capsule nearly here?. *Gastrointestinal endoscopy*, *89*(1), 195-197.
52. Novozámský, A., Flusser, J., Tachecí, I., Sulík, L., Bureš, J., & Krejcar, O. (2016). Automatic blood detection in capsule endoscopy video. *Journal of biomedicaloptics*, *21*(12), 126007.
53. Tuba, E., Tuba, M., & Jovanovic, R. (2017, May). An algorithm for automated segmentation for bleeding detection in endoscopic images. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 4579-4586). IEEE.
54. Al-Rahayfeh, A. A., & Abuzneid, A. A. (2010). Detection of bleeding in wireless capsule endoscopy images using range ratio color. *arXiv preprint arXiv:1005.5439*.
55. Kumar, V., & Samadhiya, A. (2019). Comparative performance analysis of image de-noising techniques. *arXiv preprint arXiv:1901.06529*.
56. Anutam, R. (2014). Performance analysis of image de-noising with wavelet thresholding methods for different levels of decomposition. *Int. J. Multimed. Its Appl*, *6*(3).
57. Acharya, T., & Ray, A. K. (2005). *Image processing: principles and applications*. John Wiley & Sons.
58. Pandey, V. (2014). Analysis of image compression using wavelets. *International Journal of Computer Applications*, *103*(17).
59. Van Noord, N., & Postma, E. (2017). Learning scale-variant and scale-invariant features for deep image classification. *Pattern Recognition*, *61*, 583-592.
60. Luisier, F., Blu, T., Forster, B., & Unser, M. (2005, September). Which wavelet bases are the best for image denoising?. In *Wavelets XI* (Vol. 5914, p. 59140E). International Society for Optics and Photonics.
61. Masumdar, R. E., &Karandikar, R. G. (2016). Comparative Study of Different Wavelet Transforms in Fusion of Multimodal Medical Images. *International Journal of Computer Applications*, *146*(11).
62. Alwan, I. M. (2012). Color Image Denoising Using Stationary Wavelet Transform and Adaptive Wiener Filter. *Al-Khwarizmi Engineering Journal*, *8*(1), 18-26.
63. Bitenc, M., Kieffer, D. S., &Khoshelham, K. (2015). EVALUATION OF WAVELET DENOISING METHODS FOR SMALL-SCALE JOINT ROUGHNESS ESTIMATION USING TERRESTRIAL LASER SCANNING. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, *2*.
64. Anutam, R. (2014). Performance analysis of image de-noising with wavelet thresholding methods for different levels of decomposition. *Int. J. Multimed. Its Appl*, *6*(3).
65. Mortazavi, S. H., &Shahrtash, S. M. (2008, September). Comparing de-noising performance of DWT, WPT, SWT and DT-CWT for partial discharge signals. In *2008 43rd International Universities Power Engineering Conference* (pp. 1-6).IEEE.
66. Heaton, J. (2015). *Artificial Intelligence for Humans, Volume 3: Deep Learning and Neural Networks*. s.l.: Createspace Independent Publishing Platform.
67. Zoumpourlis, G., Doumanoglou, A., Vretos, N., &Daras, P. (2017). Non-linear convolution filters for cnn-based learning. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4761-4769).
68. (n.d.). Retrieved from https://jhui.github.io/2018/02/11/How-to-start-a-deep- learning-project/.
69. Yu, F. (n.d.). A Comprehensive guide to Fine-

tuning Deep Learning Models in Keras. Retrieved from https://flyyufelix.github.io/2016/10/03/fine-tuning-in-keras- part1.html.

70. Silva, J., Histace, A., Romain, O., Dray, X., &Granado, B. (2014). Toward embedded detection of polyps in wce images for early diagnosis of colorectal cancer. *International Journal of Computer Assisted Radiology and Surgery*, *9*(2), 283-293.
71. Drakos, G. (2018, September 12). How to select the Right Evaluation Metric for Machine Learning Models: Part 3 Classification Metrics. Retrieved from https://towardsdatascience.com/how-to-select-the-right-evaluation-metric-for- machine-learning-models-part-3-classification-3eac420ec991.
72. Davis, J., & Goadrich, M. (2006, June). The relationship between Precision- Recall and ROC curves. In *Proceedings of the 23rd international conference on Machine learning* (pp. 233-240). ACM.
73. Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS one*, *10*(3), e0118432.
74. Brownlee, J. (2019, September 25). How to Use ROC Curves and Precision- Recall Curves for Classification in Python. Retrieved from https://machinelearningmastery.com/roc-curves-and-precision-recall-curves-for- classification-in-python/.
75. Ahmmad Musha(2023,Aug). Computer-Aided Bleeding Detection Algorithms for Capsule Endoscopy: A Systematic Review from https://www.mdpi.com/1424-8220/23/16/7170
76. Abhinav Patel(Aug 2021), Automated bleeding detection in wireless capsule endoscopy images based on sparse coding from https://www.researchgate.net/publication/344446498 –