Sādhanā (2022)47:234
https://doi.org/10.1007/s12046-022-01970-8

Colorectal polyp detection in colonoscopy videos using image enhancement and discrete orthonormal stockwell transform

J S NISHA and VARUN PALAKUZHIYIL GOPI*

National Institute of Technology, Trichy, Tamilnadu, India
e-mail: nishajs2007@gmail.com; varun@nitt.edu

MS received 9 May 2022; revised 1 August 2022; accepted 9 August 2022

Abstract. Computer-aided detection based on Machine Learning (ML) techniques is increasingly used to detect early-stage colorectal polyps from colonoscopy images. This study presents an efficient ML algorithm that analyses colonoscopy images and accurately detects polyps for reliable diagnosis of the early stages of Colo-Rectal Cancer (CRC). The proposed approach consists of mainly image enhancement, which enhances the low illumination colonoscopy images, followed by feature extraction using Discrete Orthonormal Stockwell Transform (DOST) and classification by a Support Vector Machine (SVM) classifier. We present an efficient image enhancement algorithm that highlights the clinically significant features in the colonoscopy image and the DOST feature extraction method to discriminate between the polyp area and non-polyp region in the colonoscopy data. The proposed method has been trained using the publicly available databases CVC ClinicDB and tested using ETIS Larib and CVC ColonDB. A sliding window with NMS-based post-processing is used in the selection of polyps from the test images. The performance measures are found in terms of precision (93.76%), recall (92.71%), F1 score (93.23%) and F2 score (93.54%) for CVC ColonDB database and precision (80.97%), recall (93.12%), F1 score (86.62%) and F2 score (83.13%) for the ETIS Larib database. Comparison with the existing method shows that the proposed approach surpasses the existing one in terms of precision, recall, F1-score, F2-score in the CVC ColonDB, and in terms of recall, F1-score in the Etis Larib database. This method would help doctors with timely evaluation and analysis of anomalies from colonoscopy data, which would help in the early planning of preventive or therapeutic protocols.

Keywords. Colonoscopy; discrete orthonormal stockwell transform; support vector machine; colorectal cancer; non-maximum suppression algorithm.

1. Introduction

Diseases of the gastrointestinal tract represent a unique set of health concerns for humans. In recent years, Colo-Rectal Cancer (CRC) has become one of the most common forms of malignant tumors, accounting for about one-tenth of all cancer diagnoses and deaths. CRC is the third most common cancer globally among cancer patients and the second most common among cancer-related fatalities [1]. In light of this, timely detection of GI tract anomalies is critical in planning preventive and therapeutic actions. CRC often begins as an abnormal growth known as a polyp inside the colon or rectum. Early diagnosis of polyps and removal of colorectal lesions may dramatically lower CRC incidence and death. Colonoscopy is a useful diagnostic procedure for examining irregularities in the bottom half of the digestive system and determining the location of lesions [2]. It is a beneficial diagnostic tool for finding anomalies in the rectum and colon and preventing them from developing into malignancies [3].

In a traditional colonoscopy, a doctor or specialist visually inspects the images of the colorectal track to detect lesions and polyps. Thus, the detection of polyps and lesions by colonoscopy is limited by the competence and abilities of the physician, who may fail to discern between neoplastic and hyperplastic polyps during colonoscopy screening or perhaps the absence of polyps in the Gastro-Intestinal (GI) tract. The accurate detection and discerning of lesion form, colour, irregularity, size, and texture continues to be a challenge in visual inspection. In addition, a single colonoscopy scan is made up of many images, the analysis of all of which is time-consuming and tedious.

Automatic detection of abnormalities using Artificial Intelligence (AI) and Machine Learning (ML) techniques has been increasingly gaining importance in medical diagnostics to circumvent the problems associated with traditional manual inspection [4]. An AI-based approach can improve the quality of colorectal screening and monitoring based on colonoscopy while reducing needless costs

*For correspondence
Published online: 12 November 2022

[5]. ML techniques are being developed for anomaly detection, segmentation, etc., in various areas of diagnostic imaging. In a colonoscopy, the use of ML to improve the accuracy of detecting different polyp sizes with minimum polyp miss rates can result in enhanced efficacy in diagnostics. This can, in turn, lead to better therapeutic management or prevention of CRC. In recent years, advanced ML algorithms for evaluating polyps based on appearance have been proposed [6]. While ML is promising for polyp detection, detecting and differentiating polyp-like structures in the GI images is challenging. An efficient ML-based network for detecting polyps in colonoscopy images is proposed in this work.

Our contributions can be summarised as

- A novel image enhancement approach has been proposed to highlight the features of the colonoscopic images to detect colorectal polyps in their early stages.
- A 2D-DOST transform is used to extract the time-frequency content in the colonoscopy images and Haralick features from the GLCM method is used for the selection of features and SVM for classification.
- The sliding window-based approach with Non-Maximum Suppression algorithm (NMS) based post-processing significantly reduces the number of FP and false-negative instances in polyp detection.
- Regarding performance measures like precision, F2-score, recall, and F1- score, the suggested technique outperforms existing polyp detection algorithms.

2. Literature review

In recent years, there have been numerous studies to enhance the computer-aided detection of polyps in the GI tract from colonoscopy images. The following are some investigations in this field.

Polyps were identified using different Convolutional Neural Networks (CNN) and hybrid approaches in [7]. CUMED, a CNN-based technique, performed well in polyp identification. An improved rectal polyp classification strategy using CNN was proposed in [8]. Pre-trained networks were evaluated on eight high-definition colonoscopy databases. The "off-the-shelf" CNN properties were found to be relevant for automated rectal colon polyp categorization. The adenoma distribution and detection rate in the lower half of the intestine were established based on colonoscopy indications in [9]. The method indicated that polypoid detection was more effective (98.00%) than fat detection (89.80%). An Inception-ResNet approach was used to recognize colorectal polyps from colonoscopy images. This method outperformed precision 89.70%, recall 84.30%, and F1-score 86.90% in offline learning. In [10], a new CNN architecture was proposed for learning the spatial characteristics of polyps inside the GI tract during colonoscopy. The suggested approach was tested on 17,574 frames from the AsuMayoDB database, which yielded an accuracy of 88.60%.

For polyp identification in colonoscopy videos, hierarchical features were retrieved using CNN [11]. In the CVC-ClinicDB, this technique outperformed other models with an F1-score of 73.28 %, recall of 82.76%, and accuracy of 65.75%. In [12], a DL solution was proposed for colon polyp classification using Global Average Pooling (GAP). In this work, experiments were conducted on the CP-CHILD-A database, and accuracy of 98% was obtained in both models. In [13], the polyp was detected using a combination of contourlet transform and a VGG-19 using 224×224 patch colonoscopy data. In [14], a 2-Dimensional CNN was used to identify polyps. The proposed method detected polyps in static images, and an image style transfer module was introduced to increase the sensitivity. Renji-VideoDB, ETIS-Lirib, and CVC-ClinicDB were used to test the suggested technique. A lightweight CNN approach was suggested in [15] to detect polyp using Mish and ReLU activation functions. A technique for the identification of colon cancer that integrates structural and statistical pattern recognition was put forth in [16]. A comparison of the classification accuracy and error rate of various classifiers, including Multilayer Perception, Sequential Minimal Optimization, Bayesian Logistic Regression, and K-Star using the percentage split approach. The outcome demonstrates that the MLP classifier, which has a kappa statistic of 0.625 and an accuracy of 83.33 percent, is the best method in WEKA.

Although several studies have shown the localization of numerous polyps, their categorization is still in its infancy. In [15], a polyp identification DL system was developed and tested on four different databases, X & Y as training databases, Z as a validation database, and V as a testing database [17]. Polyp detection sensitivity was 96.70 %, 90.20 %, and 87.70 % for databases X, Y, and Z, respectively. The FP rate of 8.30% was obtained when database V was used. Fast Pareto Depth Analysis (FPDA) and Pareto Depth Analysis (PDA) algorithms were used to detect polyps in [18]. FPDA was quicker than PDA, with a 2% decrease in accuracy. Ref [19] provides a comprehensive overview of the important studies that have been conducted using DL for the detection and categorization of GI tract polyps. The ResNet-50 and ResNet-101 models were developed in [20] for extracting features from polyp images. The models were tested with different pertained weights on publicly available databases. A CNN approach with three convolution layers were used for extracting features from colonoscopy images in [21]. The CVC-Clinic database was used for training, while the AsuMayo and ETIS-Larib databases were used for testing. They had 90.82% sensitivity, 91.76% specificity, 91.26% accuracy, and 92.71% precision.

In [22], a DL network with four convolution layers, four max-pooling layers, and two FCL was utilized for

extracting the feature. The AsuMayo Test Clinic database with 12872 images with polyps in 3486 of them was used for training, and 4702 frames with polyps in 827 of them were used for testing. The authors reported a recall value of 68.32%, a specificity of 94.97%, a precision of 90.28%, and a sensitivity of 74.34%. A co-occurrence matrix was used for the discrimination of polyps and normal tissue in [23]. A polyp detection approach based on Histogram of Oriented Gradients (HOG) and saliency map formation was introduced in [24]. The clinically important locations of the endoscopic capsule were highlighted using a saliency detector in this study. The polyp detection in colonoscopy images was done with the VGG network, while polyp segmentation was done with the CNN approach in [25]. The suggested network showed an accuracy of 86% and an F2-score of 82% for polyp segmentation on a private database.

A method for detecting polyps in colonoscopy images using DL was proposed in [26]. A pretrained CNN architecture was used as a feature extractor and SVM for classification. The efficiency of the work was proven using the CVC ColonDB database. An automated method for detecting polyps in a colonoscopy video was reported in [27]. An AlexNet architecture was used to train and test CNN to categorize each image into either polyp or non-polyp. To create a more resilient CNN system, each with-polyp patch was rotated, translated, and scaled for invariant. The accuracy and sensitivity of the system were 91.47% and 91.76%, respectively. A precise polyp localization was achieved by following the properties of polyps like shape, colour, texture, and temporal information in [28]. A set of CNNs, each one, was specialized on one sort of feature, and the results were then combined to approve or reject the candidate. The experimental results on an annotated polyp database revealed a considerable increase in the performance, with a significant reduction in False Positives (FPs).

The strategy suggested in [29] combined all conceivable polyp traits by maximizing the benefits of each feature while reducing its flaws. The CAD method relies on the resilience of form features to use the computationally costly features with great discriminative strength to successfully eliminate FPs. Due to its excellent performance, DL [30], notably CNN, has been popular in recent years for extracting highly representative properties from biological images. A CNN employs many layers to extract the features from colonoscopy data and organize the data into numerous classes. The network layers are responsible for filtering, selecting, and using these characteristics in the final FCL for image categorization. However, CNN training for colonoscopy image classification remains difficult because of a lack of big and freely accessible annotated databases. Efficient approaches for properly recognizing polyps from colonoscopy images are necessary to diagnose CRC early.

DL-based approaches have been commonly used for classification and feature extraction from biomedical images due to their high performance. But, the main issues facing colonoscopy image analysis are the availability of datasets and the huge data size. A substantial amount of labelled training data is required for deep CNN models. When large-scale annotated medical datasets are unavailable, this becomes a problematic need. In most of the CNN base research work, augmented data are given for training. Model training is time-consuming and computationally intensive, and it gets even more when dealing with issues like over-fitting and convergence. Also, complicated data models require costly GPUs, which extremely increases the cost of the polyp detector. To address these challenges, this paper provides a method using an efficient image enhancement method followed by a simple feature extractor using Discrete Orthonormal Stockwell (DOST) with Gray Level Co-occurrence Matrix(GLCM) based feature selection and SVM classifier on colonoscopy image analysis that helps the early detection of colorectal cancer during colonoscopy.

3. Proposed method

Different stages in the proposed polyp detection method are represented in figures 1 and 2, respectively. We aim to introduce a novel image enhancement algorithm focusing on illumination and reflection applied to the colonoscopy images to highlight the vital region. The image enhancement stage effectively highlights the clinically significant region in the colonoscopy data. Training patches are collected from the visually salient areas obtained from the image enhancement stage by random sampling.

The Stockwell Transform (ST) [31] is a most modern approach to texture analysis that closely resembles the continuous wavelet transform. ST examines each pixel's local spatial frequency content directly [32]. In the colonoscopy image analysis, textural features of the tissue have deformations when it exhibits abnormalities. The textural analysis is critical in this scenario for diagnosing diseases in the GI tract. Because of its outstanding resolution and preservation of phase content of the image, the ST is suitable for texture research of biological images [33]. A discrete orthonormal basis is employed to eliminate duplication in the frequency-space domain and speed up the computation of the ST. The Discrete Orthonormal ST (DOST), like the discrete wavelet, transform [34], produces a spatial frequency representation. Here we use the DOST feature extractor for extracting the transform-based features from colonoscopy images.

In medical imaging, the Grey Level Co-occurrence Matrix (GLCM) [35] is used for measuring the texture of an image. The GLCM of the input image is used to calculate Haralick features, which are a set of features used to classify and analyze textures [36]. After the DOST feature extractor, the Haralick textural feature selection from the

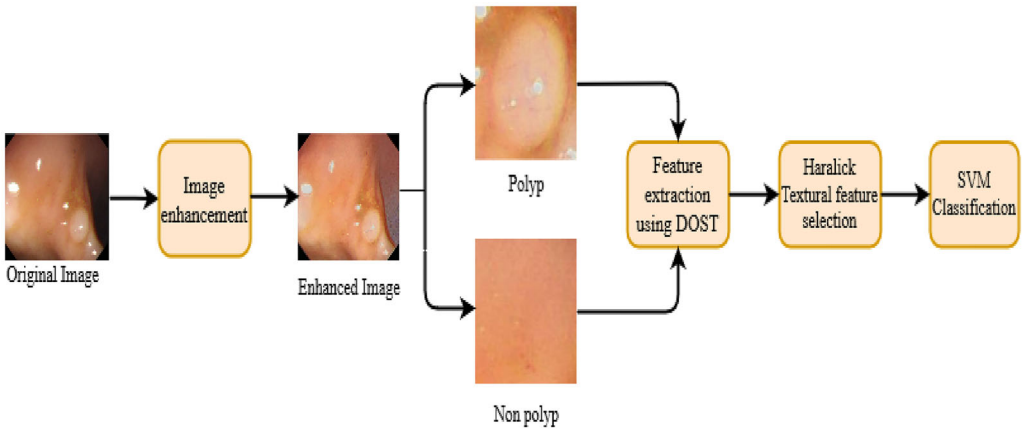

Figure 1. Block diagram representation of the proposed method in training stage.

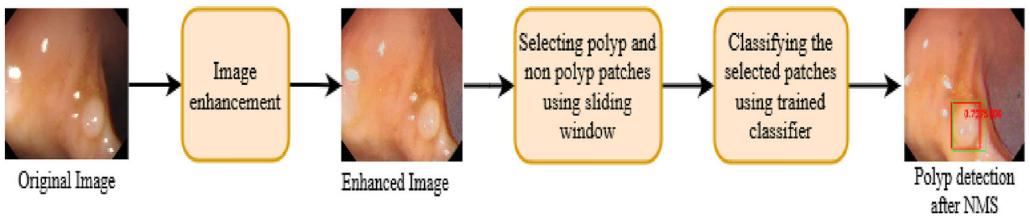

Figure 2. Block diagram representation of the proposed method in testing stage.

GLCM method is used to select relevant features, and these features are trained using an SVM classifier. In the testing stage, the image enhancement stage effectively highlights the clinically significant areas in the CVC ColonDB and Etis Larib databases. After image enhancement, a sliding window approach is used to detect polyps from the colonoscopy data; this method returns a set of detection windows. The trained classifier is used to classify these windows as either polyp or non-polyp. The non-max suppression technique is the final step of the polyp detection algorithms and is used to select the most appropriate bounding box for the polyp from the classifier. The NMS is applied to all detection windows in the image with confidence above a certain threshold. The purpose of NMS is to select the best bounding box for the polyp and suppress all other bounding boxes similarly classified as a polyp.

The steps for the proposed technique are explained in the following sections.

3.1 *The proposed image enhancement method*

The characteristics of low light conditions in colonoscopy images are considered in our suggested image-enhancing approach and perform illumination enhancement in colonoscopy images. The proposed illumination enhancement includes mainly contrast, saturation, and brightness component correction in different color spaces. The block diagram of the proposed image enhancement method is given in figure 3.

In the contrast enhancement stage, the RGB colonoscopy image is converted into L*a*b* space [37]. After the colour conversion, I_{L*}, I_{a*}, I_{b*} represents the light, green-red colour, and a blue-yellow colour component of the image. Contrast Limited Adaptive Histogram Equalization (CLAHE) [38] is performed in the I_{L*} channel. The CLAHE constructively intensifies the image in addition to the noise reduction. It is observed that CLAHE used in the I_{L*} input improved the visual consistency of the polyp in the colonoscopy data. After L* channel enhancement, the image is converted into RGB colour space to obtain better contrast.

However, CLAHE enhancement in the L* component increased the contrast of the RGB image. It does not enhance the darker portions of the colonoscopy data. Hence, we have to enhance the brightness of the colonoscopy image in another colour space. For that, the contrast-enhanced RGB colonoscopy image is converted into HSV space. In HSV color space I_v, I_s, I_h represents value, saturation, and hue component of the image. In HSV colour conversion, a gamma correction is applied in the V channel (I_v) to adjust the brightness. The brightness of an image can be improved using Gamma correction [39], or a nonlinear power transform estimates its diminution.

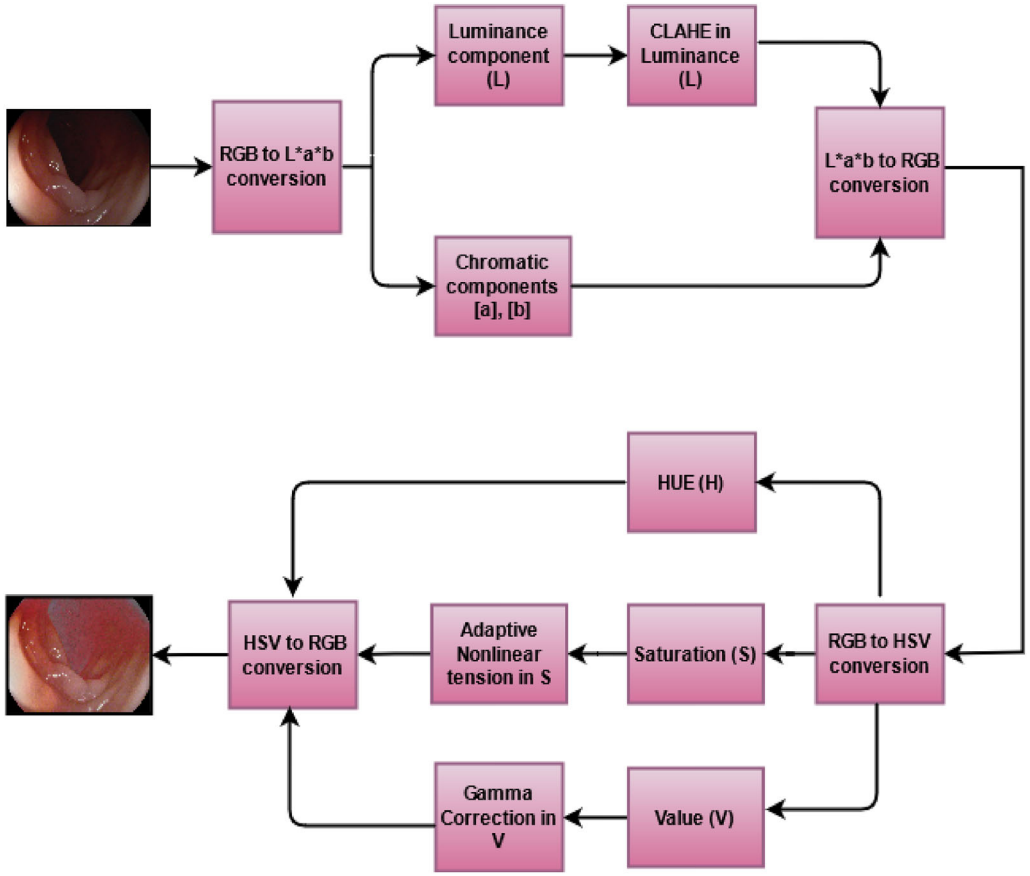

Figure 3. Proposed image enhancement method.

Gamma corrected illumination, I_g is given by,

$$I_g = I_v^{\gamma} \tag{1}$$

In our method, the gamma is set to 1/2.2 [40].

For saturation enhancement [41], a stretching function [42] is utilised in the S channel (I_s). Stretching function in the S channel, I_s is represented as,

$$S' = \frac{1}{2} + \frac{1}{2}\left(\frac{max(H,S,V) + min(H,S,V) + 1}{2 \times mean(H,S,V) + 1}\right)S \tag{2}$$

where S represents before the stretching function and S' represents after the stretching function. Stretching the image in the S component improves the image saturation and gives it a more natural appearance. After performing all these operations in the HSV color space, the image is converted into the RGB color space.

3.2 *Performance analysis of the proposed image enhancement*

Our proposed enhancement technique uses the Average Gradient (AG), Peak signal-to-noise ratio (PSNR), and Information Entropy (IE) factors for the analysis of the proposed enhancement technique. PSNR is the most frequently used image analysis technique, and it focuses on improved image quality. The AG analyses image clarity, whereas the IE measures image information content.

$$AG = \frac{1}{i \times j}\sum_{m=1}^{I}\sum_{k=1}^{J}\sqrt{\frac{(\partial I/\partial x)^2 + (\partial I/\partial y)^2}{2}} \tag{3}$$

$I_e(i,j)$ represents the enhanced image and $i \times j$ represents the size of the enhanced image.

$$IE = -\sum_{m=1}^{L} P(I_m).logP(I_m) \tag{4}$$

I_m represents the set of all symbols in $I_e(i,j)$ and $P(I_m)$ denotes the probability of occurrence of each symbol in $I_e(i,j)$.

$$PSNR = 10.log\left(\frac{(2^n - 1)^2}{MSE}\right) \tag{5}$$

The Mean Square Error of the original and enhanced images is represented by MSE. The number of bits in each

sample value is denoted by n. The value of n is usually set to 256.

$$MSE = \frac{1}{M \times N} \sum_{\alpha=1}^{M} \sum_{\beta=1}^{N} |g(\alpha - \beta) - h(\alpha - \beta)|^2 \qquad (6)$$

The gray value pixels of the input image and the enhanced image is designated by $g(\alpha, \beta)$ and $h(\alpha, \beta)$, where α and β represents corresponding pixels in the row and column.

Tables 1, 2, and 3 represent the quantitative analysis of the proposed image enhancement on the CVC ClinicDB, CVC ColonDB, and ETIS-Larib datasets. AG, IE, and PSNR values are increased by a large margin by the proposed method. Histogram Equalization (HE) [43] and CLAHE [44] are commonly used enhancement techniques on biomedical images. To justify our proposed method, we have compared three evaluation parameters: AG, IE, and PSNR, with the existing HE and CLAHE techniques. From the evaluation, it was found that the suggested approach performs better than others regarding colonoscopy image enhancement.

3.3 *Feature extraction and classification*

A feature extractor is a simplified image description that contains only the most important information about the image. Structural analysis has been used to analyze the characteristics of an image because the human visual system identifies an object by elucidating its structures. Edges, corners, and textures are the commonly used structures heightened by the pixel spatial relationship. Features obtained in this section are finally fed to the classifier that maps the input data to a specific category.

3.4 *Discrete Orthonormal Stockwell Transform (DOST)*

After performing preprocessing, the computer-aided system needs some relevant characteristics of normal and polyp images to differentiate them. Several methods, including statistical, transformation, and structural-based approaches, can extract texture characteristics. The statistical technique is employed to estimate the spatial distribution of grey values. The texture of the image also determines the frequency content of an image. An effective image transformation approach may be used to retrieve local spatial frequency information from an image.

DOST is a powerful feature extraction method because it counts gradient alignment in each localized portion of an image. This feature extraction method is effectively used to detect and diagnose various abnormalities in biomedical images. Because of its robustness to illumination conditions and computational efficiency, the DOST feature extraction approach has been useful in medical image analysis. In [45] 2D DOST was used to extract the time-frequency features from brain MRIs. In [46] based on DOST and statistical features were used to distinguish normal and diseased retinal fundus images. The Two-Dimensional (2D) DOST may give pixel-by-pixel textural information of a biomedical image by producing local spectra, including frequency information collected from the Fourier Transform (FT) of an image. The DOST is a fast and simple frequency-domain method for analyzing every pixel in a large image quickly. Because of its superior performance in discriminating between polyp and non-polyp patches, the DOST is utilized to detect polyps in colonoscopy images in this context.

The development of the Short-Time FT (STFT) [47] produced a high impact in the field of time-frequency analysis. The windowing capacity enhances the FT and time localization for the frequency components analysis; thus, STFT is also referred to as windowed FT. Due to the usage of a fixed size time window, the major disadvantage of STFT is the poor resolution in time-frequency. To overcome this drawback, Wavelet Transform (WT) was introduced [48]. The main advantage of WT is it uses a flexible size of windows. The basis function used in the WT allows us to expand and compress the frequency information. The main limitation of WT is that the information regarding the phase was not retained [31]. To overcome the limitations of WT and STFT, a time-frequency distribution-based ST was proposed in 1996, which combines the promising features of both the transforms. ST employs a window based on a Gaussian function that keeps information about the phase.

Table 1. Quantitative analysis of Average Gradient in the databases.

Name of the database	*AG* (WOE)	*AG*(HE) [43]	*AG*(CLAHE) [44]	*AG* (PEM)
CVC ClinicDB	2.18	3.15	3.53	4.43
CVC ColonDB	2.26	3.08	3.87	5.17
Etis Larib	5.73	7.16	7.25	9.22

*WOE- without enhancement, *PEM-proposed Enhancement Method

Table 2. Quantitative analysis of Entropy in the databases.

Name of the database	*IE* (WOE)	*IE*(HE) [43]	*IE*(CLAHE) [44]	*IE* (PEM)
CVC ClinicDB	6.29	7.16	7.69	7.85
CVC ColonDB	6.84	7.31	7.65	7.86
Etis Larib	6.89	7.30	7.66	7.88

*WOE- without enhancement, *PEM-proposed Enhancement Method

Table 3. Quantitative analysis of PSNR in the databases.

Name of the database	*PSNR* (HE) [43]	*PSNR*(CLAHE) [44]	*PSNR* (PEM)
CVC ClinicDB	14.76	19.54	22.59
CVC ColonDB	16.38	21.29	23.31
Etis Larib	17.10	22.64	25.57

*WOE- without enhancement, *PEM-proposed Enhancement Method

The ST function represented for an 1D signal s(t) is given by

$$S(f,\tau) = \int_{-\infty}^{\infty} s(t)g(t-\tau)e^{-j2\pi.f.t}dt \tag{7}$$

where g(t) represents the Gaussian filter and the width of the Gaussian function is represented by σ.

The amount of redundancy in the time-frequency representation is high in ST. To overcome this problem, the DOST is proposed, which has a spatial frequency representation similar to the DWT. A dyadic sampling scheme [33] in the frequency domain is used to calculate the DOST of an image. The DOST of a signal is calculated by finding the basis functions from the FT and applying appropriate frequency and phase changes.

The two dimensional FT of the image $i(g, h)$ is given by

$$I(c,d) = \sum_{g=-0}^{M-1}\sum_{h=-0}^{N-1} i(g,h)exp^{-j2\Pi(cg/M+dh/N)} \tag{8}$$

where $M \times N$ represents the size of the input image.

Its corresponding 2D ST is given by

$$S_t(\hat{u},\hat{v},f_{\hat{u}},f_{\hat{v}}) = \frac{|f_{\hat{u}}||f_{\hat{v}}|}{2\Pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} i(g,h)exp\left(\frac{(\hat{u}-g)^2)(\hat{v}-h)^2}{2}\right) exp(-j2\Pi(f_{\hat{u}}g+f_{\hat{v}}h))dgdh \tag{9}$$

$f_{\hat{u}}$ and $f_{\hat{v}}$ are the frequencies in two directions and $\hat{u}$ and $\hat{v}$ represents shift parameters.

The main steps in calculating the 2D-DOST of an image are: initially partitioning the 2D-FT $I(c, d)$, in the second step multiply the partitioned FT with the square root of the number of points in that division. In the last step, take the inverse of 2D-FT.

The 2D-DOST of an $M \times N$ image $i(g, h)$ is represented as

$$S_d(\hat{u},\hat{v},f_{\hat{u}},f_{\hat{v}}) = \frac{1}{\sqrt{2^{p_g+p_h-2}}}\sum_{c=-2^{p_g-2}-1}^{2^{p_g-2}}\sum_{d=-2^{p_h-2}-1}^{2^{p_h-2}} I(c+\hat{v}_g, d+\hat{v}_h) exp\left(j2\Pi\left(\frac{c\hat{u}}{2^{p_g-1}}+\frac{d\hat{v}}{2^{p_h-1}}\right)\right) \tag{10}$$

The 2D-DOST operated image has the same number of coefficients as the original image. As a result, Haralick features from the GLCM method are used to reduce dimensionality.

3.5 *Haralick textural feature selection*

Haralick [36] introduced textural feature extraction based on a co-occurrence matrix for the statistical analysis. Based on the relationships between pixels in the co-occurrence matrix, GLCM and textural features are computed. This method may be used in many image analysis applications, from biomedical to sensor array processing. It offers information on image area properties, including homogeneity, contrast, boundaries, and complexity. The paper by Haralick [35] included a list of other factors that may be computed using the recommended method. The suggested method for reducing the dimensionality of DOST coefficients employs Haralick textural feature selection.

A total of 20 features (Energy, Entropy, Inverse difference, Correlation, Dissimilarity, Auto-correlation, Contrast, Homogeneity, Sum of average, Sum of entropy, Sum of variance, Maximum probability, Sum of squares,

Difference entropy, Maximal correlation coefficient, Inverse difference normalized, Difference variance, Cluster prominence, Inverse difference moment normalized, Cluster shade) based on image transform (DOST) were extracted for each image. Because features have diverse value ranges, a normalization approach is required to transform them into a consistent range. All characteristics were normalized to have a mean of zero and a standard deviation of one.

3.6 *Support vector machine (SVM)*

The choice of the classifier is an essential factor in anomaly detection in medical image processing. Here, a linear SVM classifier is used for the classification of features extracted from the DOST. SVM [49] and related learning algorithms are supervised learning methods. In a higher-dimensional space, if the vectors are non-linearly separable, the SVM can make them linearly separable. Incoming data is classified into binary classes in the most popular applications. For SVM models, each data point is represented as an m-dimensional vector by fitting the data point inside a hyperplane. For classification, each input is turned into a point in n-dimensional space, and the value of each attribute is described as the value of a single hyperplane coordinate. Classification is established by optimizing the hyperplane that most differentiates the two groups (polyp and non-polyp).

3.7 *Non-maximum suppression (NMS)*

The object detection approach highly depends on NMS. It starts by sorting all detection windows by their scores. All detection boxes with a large overlap with D are muted, and the detection window D with the highest score is picked. This procedure is repeated for the remaining windows. The NMS algorithm can suppress the detection of the same polyp in the same image to detect only once. Considering so many detection windows, a polyp may be included in several adjacent windows, so all these detections will predict a polyp. Concretely, what the NMS does is that it will look at the predicted probability from these windows and pick the window that has the highest predicted probability for each category.

The steps in the NMS algorithm:

Input: A list of proposal windows E, corresponding confidence scores G and overlap threshold H.

Output: A list of filtered windows F.

Algorithm:

1. Remove the proposal with the greatest confidence score from E and place it on F, the final proposal list. (F is initially empty).
2. Now compare this proposal to all of the others by calculating the Intersection Over Union (IOU) of this proposal with all of the others. Remove that proposal from E if the IOU exceeds the threshold H.
3. From the remaining suggestions in E, select the proposal with the highest level of confidence and transfer it to F.
4. Compute the IOU of this proposal from the IOU of all the proposals in E and delete the windows with IOUs greater than the threshold. In our proposed method, the threshold is set to 0.5.
5. This process is continued until E is emptied of suggestions.

4. Results and discussion

The suggested approach is implemented using a TensorFlow backend in the Keras framework in Python using the Spyder platform.. The proposed model assessments are performed on a computer with a Dual Nvidia Geforce RTX 2080 Ti 11 GB GPU with 64 GB RAM.

4.1 *Dataset description*

Publicly available databases such as the ETIS-Larib [52], CVC ClinicDB [50] and CVC ColonDB [51] shown in table 4 are used to conduct detailed analyses. The proposed method is trained using CVC-ClinicDB. It consists of a total of 612 frames. Publicly available databases such as ETIS-Larib and CVC-ColonDB are used to test the proposed model. The CVC-ColonDB consists of a total of 379 frames, and the Etis-Larib database consists of a total of 196 frames extracted from colonoscopy videos. Several polyps can be seen in these frames. All databases consist of the ground truth image and the frames.

4.2 *Performance measures*

The evaluation matrices like precision, recall, F1-score, and F2-score are used to evaluate the performance of the proposed approach. The precision determines the positive prediction value (P). This figure indicates the success of our system in preventing FP. Recall (R), also called sensitivity, describes how the model successfully limits FN. The accuracy defines the total performance of the method. The F1-score indicates the harmonic mean of precision and recall value. The F2-score reduces the relevance of

Table 4. Dataset description.

Dataset	Resolution	Content
CVC-ClinicDB [50]	384 × 288	612 Polyps
CVC-ColonDB [51]	574 × 500	379 Polyps
ETIS-Larib [52]	1225 × 966	196 Polyps

accuracy while raising the significance of recall. While accuracy maximization reduces the number of FP and recall maximization reduces FN, the F2-score is more concerned with reducing FN than with reducing FP.

$$Precision, P = \frac{TP}{TP + FP} \tag{11}$$

$$Recall, R = \frac{TP}{TP + FN} \tag{12}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Recall + Precision} \tag{13}$$

$$F2 - score = \frac{5 \times Precision \times Recall}{4 \times Recall + Precision} \tag{14}$$

The IOU is a measure to evaluate the accuracy of a detector. We need the ground-truth bounding boxes and the predicted bounding boxes of the proposed model to calculate the IOU for the evaluation. The red bounding box in figures 4 and 5 reflects the ground truth, whereas the green bounding box represents the predicted output. If the IOU of the predicted window falls inside the ground truth area, the polyp will be examined to be detected. If more than one window with IOUs is identified within the ground truth zone, in that case, it will be considered as one TP. If the sum of IOUs is greater than 0.5, or select the detection window with the highest score, which must be greater than 0.5. In figure 5, we can see that two windows with IOUs of 0.33 and 0.34 are identified within the ground truth zone; in that case, the sum of IOUs is taken as the performance measure. FP is defined as a detection window with an IOU of less than 0.5 or an IOU that is outside the ground truth area. A polyp that is undetectable by the search window is referred to as an FN.

4.3 *Implementation*

We used DOST and an SVM classifier to develop an efficient polyp detector. First, we collect polyp and non-polyp image patches from the CVC ClinicDB dataset listed in table 4. As shown in figure 6, positive patches are extracted with polyps using the ground truth, whereas non-polyp patches are randomly collected from the images. To obtain a total of 1224 polyp patches, we have extracted 612 polyp patches from ground truth, and the remaining 612 polyps are collected by the process of random sampling. 1224 non-polyp patches were extracted from the random sampling. Thus, a total of 1224 polyp patches and 1224 non-polyp patches before and after image enhancement were accumulated from the CVC-ClinicDB database and were used for training. We resize the different-sized image patches to 128×128. The 128×128 patch size is considered as the aspect ratio according to the image size. This can improve the classification performance and provide more reliable detection in the analysis of polyp characteristics in colonoscopy imaging. This patch size substantially improves the problems of differentiating between polyps and false positives using the DOST function.

The 2D-DOST transformed patch has the same coefficients as the original patch, i.e., 128×128. GLCM with Haralick feature extraction is used to reduce the dimensionality of a DOST-based feature vector. The features obtained from the haralick textural feature selection and the corresponding labels indicating a polyp or a non-polyp are used to learn the SVM classifier. The proposed method takes 324.60 seconds to get trained on the CVC ClinicDB dataset. In the testing stage, the CVC-ColonDB and ETIS Larib databases pass through the image enhancement stage to get the uninformative frames in the detection. Our work aims to estimate the area that is likely to contain polyps within an image. In the proposed method, a sliding window-based approach is used for the detection. The sliding window size (128×128) is set the same as the patch size. At each window, we compute DOST descriptors, and GLCM with Haralick feature selection is performed and applied to the trained classifier. If the classifier detects a polyp with a sufficiently large probability, it records the bounding box of the detected window. One of the problems

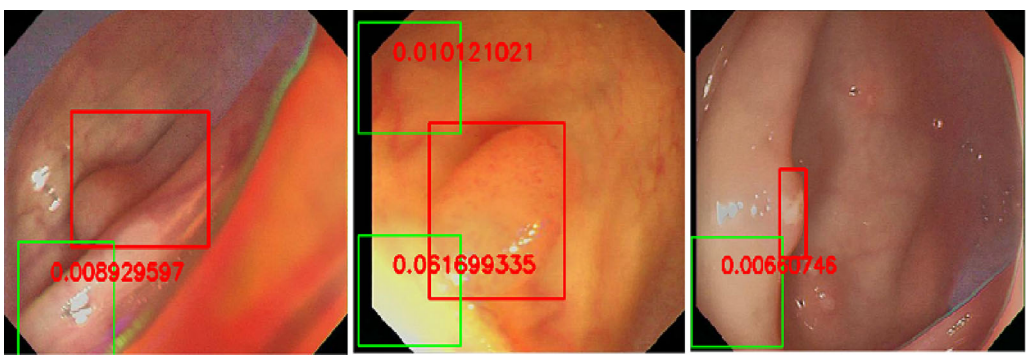

Figure 4. False positive obtained after NMS algorithm.

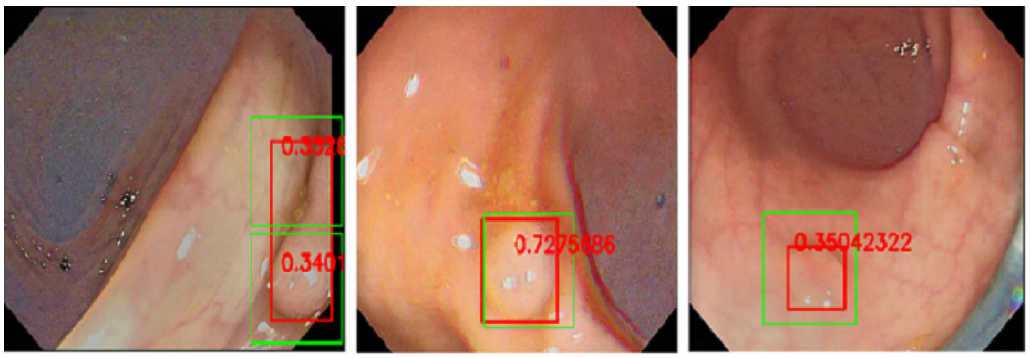

Figure 5. True positive obtained after NMS algorithm.

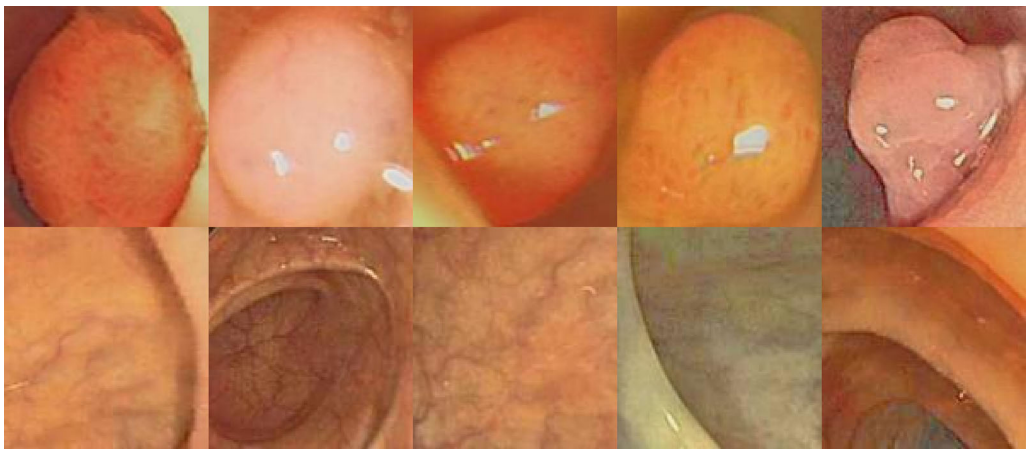

Figure 6. Samples of training data.

with the sliding window-based approach is that it may detect one polyp in an image more than once. The use of the NMS algorithm overcomes this problem. After finishing the scan of the image, apply NMS to remove redundant and overlapping bounding boxes. To evaluate the performance of our proposed enhancement, testing is also done without enhanced data. The inference time on CVC ColonDB is 124 ms and the processing speed of 8.06 frames per second, and on the ETIS Larib database, the inference time is 153 ms and the processing speed of 6.53 frames per second, respectively.

4.4 *Performance evaluation on databases*

We have selected the standard metrics like recall, precision, F1-score, and F2-score to evaluate the performance of the proposed model. Performance metrics evaluation on both databases is described in table 5. In CVC ColonDB, out of 379 polyps, 331 polyps are correctly detected. The number of FN instances is 26, and the number of FP instances is 22. Thus, the CVC ColonDB obtained a precision value of 93.76%, recall value of 92.71%, F1-score 93.23%, and F2-score 93.54%. In the ETIS Larib database with patch size, 149 polyps are correctly detected. The number of polyp patches not detected is 11, and the number of FN instances is 35. Thus, the ETIS Larib database obtained a precision value of 80.97%, a recall value of 93.12

To evaluate the performance of our proposed model, the performance evaluation in the three different databases without image enhancement is also analyzed. In the ETIS Larib database, only 79 polyps are correctly detected, in which FN instances are 129 and FP instances are 151. Without enhancement, the ETIS Larib database obtained only a precision value of 34.34%, recall value of 37.98%, an F1-score of 36.06%, and an F2-score of 35.01%, and the CVC ColonDB obtained only a precision value of 31.02%, recall value of 37.58%, an F1-score of 33.98%, and an F2-score of 32.14%. In the CVC ColonDB, only 112 polyps are correctly detected, in which FN instances are 249, and FP instances are 186. From these analyses, it has been observed that the proposed image enhancement method effectively highlights the polyp in the image and thereby

Table 5. Performance analysis of the proposed method.

Dataset	Methods	TP	FP	FN	P	R	F1	F2
CVC-ColonDB [51]	DOST-SVM (WE)	331	22	26	93.76	92.71	93.23	93.54
	DOST-SVM (WOE)	112	249	186	31.02	37.58	33.98	32.14
ETIS-Larib [52]	DOST-SVM (WE)	149	35	11	80.97	93.12	86.62	83.13
	DOST-SVM (WOE)	79	151	129	34.34	37.98	36.06	35.01

increases the performance of the polyp detector. Our suggested technique detects the highest number of polyps, i.e., 331 in CVC ColonDB and 149 in ETIS Larib, and a reduced number in FP and FN, surpassing the existing state-of-the-art methods significantly.

4.5 *Comparison of the proposed network with existing CNN architectures*

CNN-based approaches have been commonly used for classification and feature extraction from biomedical images due to their high performance. But, the main issues facing colonoscopy image analysis are the availability of datasets and the huge data size. A substantial amount of labelled training data is required for deep CNN models. When large-scale annotated medical datasets are unavailable, this becomes a problematic need. In most of the CNN base research work, augmented data are given for training.

Tables 6, 7, 8 and 9 compare polyp detection on CVC ColonDB and ETIS Larib data with well-known CNN architectures like VGG16, VGG19, ResNet152, and InceptionV3 after increasing the size of the dataset through data augmentation and GANs. CVC ClinicDB dataset consist of 612 colonoscopy images. Using data augmentation and GAN the size of the CVC ClinicDB dataset is increased to ten thousand, this dataset is used for training the pretrained network. After increasing the size of the data for training the number of false positives and false negative instance were significantly reduced in the test data. CVC ColonDB and Etis Larib datasets are used for testing. The pretrained network is connected to three dense layers: the last dense layer with the softmax activation function. Batch normalization is used to avoid overfitting. From the results, it is observed that ResNet152 and VGG16 gives higher reacall on CVC ColonDB and on the Etis Lari dataset VGG19, ResNet152 and InceptionV3 gives slightlty higher precision values.

While considering computational complexity analysis of the pretained network after the performance evaluation using VGG16 has 15043137 trainable parameters, number of floating point operations(FLOPS) are 0.1323 BFLOPs, GPU Memory requirement is 1.6186 GB and memory required by model weights 57.38 MB. VGG19 has 20,352,833 trainable parameters, number of floating point operations are 0.1676 BFLOPs, GPU Memory requirement is 1.7565 GB and memory required by model weights 77.63 MB. RESNET152 has 59,334,401 trainable parameters, number of floating point operations are 0.1125 BFLOPs, GPU Memory requirement is 9.2258 GB and memory required by model weights 226.92MB. InceptionV3 has 22,883,233 trainable parameters, number of floating point operations are 0.0214 BFLOPs, GPU Memory requirement is 1.6034GB and memory required by model weights 87.42MB. Since the pretrained network consist of numerous number of layers, it which increases the number of trainable parameters and thereby the complexity of the model. In deep learning, FLOPS are a unit of measurement for the number of operations required to execute a network model. From the analysis it has been seen that number of FLOPS of the pretrained networks are high. The GPU memory requirement and memory required by model weight are also high in pretrained networks.

Pretrained models are usually trained on large amounts of data and using resources that are no't usually available to everyone. Take ImageNet, for example. It contains over 14 million images, with 1.2 million assigned to one of a thousand categories. From the analyses, the proposed

Table 6. Performance analysis on CVC ColonDB: after augmentation.

Model	Precision (%)	Recall (%)	F1-Score (%)	F2-Score (%)
VGG16 [53]	68.37	100.00	81.21	72.98
VGG19 [54]	89.33	80.92	84.91	87.51
ResNet152 [55]	72.17	99.20	83.56	76.33
InceptionV3 [56]	79.80	87.89	83.65	81.29
Proposed Method	93.76	92.71	93.23	93.54

Table 7. Performance analysis on ETIS LARIB database: after augmentation.

Model	Precision (%)	Recall (%)	F1-Score (%)	F2-Score (%)
VGG16 [53]	78.08	71.64	74.73	76.71
VGG19 [54]	86.58	69.60	77.17	82.55
ResNet152 [55]	93.79	69.74	80.00	87.74
InceptionV3 [56]	82.05	75.73	78.76	80.70
Proposed Method	80.97	93.12	86.62	83.13

Table 8. Performance analysis on CVC ColonDB: GAN.

Model	Precision (%)	Recall (%)	F1-Score (%)	F2-Score (%)
VGG16 [53]	70.30	99.61	82.43	74.70
VGG19 [54]	86.32	82.34	84.28	85.50
ResNet152 [55]	70.77	98.11	82.22	74.94
InceptionV3 [56]	82.35	89.28	85.67	83.65
Proposed Method	93.76	92.71	93.23	93.54

Table 9. Performance analysis on ETIS LARIB database: GAN.

Model	Precision (%)	Recall (%)	F1-Score (%)	F2-Score (%)
VGG16 [53]	74.85	69.71	72.19	73.76
VGG19 [54]	87.50	68.29	76.71	82.84
ResNet152 [55]	90.06	63.04	74.17	82.95
InceptionV3 [56]	83.33	77.38	80.25	82.07
Proposed Method	80.97	93.12	86.62	83.13

method has better outcomes than the existing pre-trained networks after augmented data in terms of precision, F1-score, and F2-score on the CVC ColonDB and recall and F1-score on the ETIS-Larib data set. The proposed method has better outcomes than the existing pre-trained networks with GAN data in terms of precision, F1-score, and F2-score on the CVC ColonDB and recall, F1-score, and F2-score on the ETIS-Larib data set. While using pretrained model, model training is time-consuming and computationally intensive, and it gets even more when dealing with issues like over-fitting and convergence. Also, complicated data models require costly GPUs, which extremely increases the cost of the polyp detector. To address these challenges, this paper provides a method using an efficient image enhancement method followed by a simple feature extractor using Discrete Orthonormal Stockwell (DOST) with Gray Level Co-occurrence Matrix(GLCM) based feature selection and SVM classifier on colonoscopy image analysis that helps the early detection of colorectal cancer during colonoscopy.

4.6 *Comparison with existing works*

A comparative study of our method and the other state-of-the-art methods recently developed in polyp detection on the CVC ColonDB and ETIS Larib databases is detailed in table 10 and table 11. From the tables, we can see that our proposed approach outperforms by a wide margin in performance measures than methods in [7, 12, 15, 20–25, 51, 57, 58] and [14]. Our suggested methodology is better than the other methods in terms of precision, recall, F1 and F2-score in the CVC ColonDB database, surpassing methods in [57] based on single-shot feed-forward fully CNN approach, [24] based on enhancement and SVM classification, [21] based on CNN, [12] based on modifications in VGGNet, [20] based on ResNet architecture, [23] based on GLCM-SVM, [58] based on DFL, [51] based on valley information and [22] based on CNN by a large margin. On the ETIS Larib database, our method surpasses [57] based on single-shot feed-forward fully CNN approach, [24] based on HOG feature extraction, [12] based on VGG, [20] based on ResNet, [25] based on VGG network, [15, 23] based on GLCM-SVM, [7] based on CUMED and [14] based on 2D CNN in terms of recall and F1-score.

The higher the precision, the more confident the classifier is in predicting the presence of polyps. The precision value achieved for the suggested technique for the ETIS-Larib database is 80.97 percent and precision value achieved for the suggested technique for the CVC ColonDB database is 93.76%. Our method attained a much higher precision value in CVC ColonDB, surpassing Qadir *et al* [57] by 5.41% Farah *et al* [24] by 43.47%, Wang *et al* [12] by

Table 10. Comparison with existing works in CVC ColonDB Dataset.

Method	Description	P (%)	R (%)	F1 (%)	F2 (%)
Proposed Method	DOST-SVM	93.76	92.71	93.23	93.54
Qadir *et al* [57]	MDeNet plus	88.35	91.00	89.65	88.86
Farah *et al* [24]	WE-SVM	50.29	86.33	63.56	75.51
Wang *et al* [12]	VGGNets-GAP	92.70	90.80	91.74	92.31
Sornapudi *et al* [20]	ResNet	89.94	91.64	90.73	91.27
Shin *et al* [21]	CNN	91.76	92.71	92.22	91.94
Akbari *et al* [22]	CNN	92.02	85.90	88.90	90.72
Cheng *et al* [23]	GLCM - SVM	65.29	77.35	67.39	70.81
Yoon *et al* [58]	DFL	70.67	70.67	70.67	70.67
Bernal *et al* [51]	Valley information	47.15	71.67	56.88	64.92

*P-Precision,*R- Recall

Table 11. Comparison with existing works in ETIS-Larib database.

Method	Description	P (%)	R (%)	F1 (%)	F2 (%)
Proposed Method	DOST-SVM	80.97	93.12	86.62	83.13
Qadir *et al* [57]	MDeNet plus	86.12	86.54	86.33	86.20
Farah *et al* [24]	WE- SVM	38.08	47.60	42.31	45.33
Rahim *et al* [15]	CNN	72.00	63.82	67.66	65.30
Wang *et al* [12]	VGGNets-GAP	89.70	88.30	88.99	89.41
Xu *et al* [14]	2D CNN	83.24	71.63	77.00	73.69
Sornapudi *et al* [20]	ResNet	72.93	80.29	76.43	78.70
Haj *et al* [25]	VGG	–	84.00	–	–
Cheng *et al* [23]	GLCM – SVM	62.56	57.35	59.84	61.44
Bernal *et al* [7]	CUMED	72.36	69.23	70.76	69.83

*P-Precision,*R- Recall

1.06% , Sornapudi et. al. [20] by 3.82%, Shin *et al* [21] by 2%, Akbari *et al* [22] by 1.74%, Cheng *et al* [23] by 28.47%, Yoon *et al* [58] by 23.09% and Bernal *et al* [51] by 46.61%. On the ETIS Larib database, our method surpass Farah *et al* [24] by 42.89%, Sornapudi *et al* [20] by 8.04%, Rahim *et al* [15] 8.97%, Cheng *et al* [23] by 18.41%, Yoon *et al* [58] by 65.08% and Bernal *et al* [7] by 8.61% in terms of precision.

A high recall value lowers the risk of misclassification by reducing the missing of clinically significant characteristics from the images. Our proposed network shows highest recall value on both the databases. Proposed method exceed recall value than the methods proposed in Qadir *et al* [57] by 1.71%, Farah *et al* [24] by 6.38%, Wang *et al* [12] by 1.91%, Sornapudi *et al* [20] by 1.07% , Akbari *et al* [22] by 6.81%, Cheng *et al* [23] by 15.36% , Yoon *et al* [58] by 22.04% and Bernal *et al* [51] by 21.04% in CVC ColonDB database. For the ETIS-Larib database, it perform better in recall value than the other detection methods in [57] by 6.58% , [24] by 45.52%, [12] by 4.82% , [20] by 12.83%, [25] by 9.12%, [15] by 29.30%, [14] by 21.49%, Cheng *et al* [23] by % 35.77, Yoon *et al* [58] by 35.91% and Bernal *et al* [7] by 23.89%.

The high precision result shows that our model's FPs are reduced. On the other hand, it aids in the early detection of polyps and avoids CRC development. The F1 and F2 scores are important performance indicators to emphasize recall and precision measures in computer-aided detection systems. Averaging the precision and recall numbers yields the F1 and F2 scores. The F1-score is determined by giving precision and recall equal weight, whereas the F2-score is determined by giving precision less weight and recall more weight. Lower F1 and F2-score values result in the absence of clinically important characteristics, preventing early detection of polyps. On the CVC ColonDB database, our technique can get the highest F1 and F2-score. The achieved F1-score is 93.23% and the F2-score is 93.54%. For the ETIS-Larib database, our technique can get an F1-score of 80.97% and F2-score of 83.13%.

4.7 *Statistical Evaluation using ANOVA Test*

The statistical evaluation of the proposed approach is carried out by the Analysis of Variance (ANOVA) Test. We employ the indicated strategy and the more accurate techniques described in tables 12 and 13 for the analysis. We

Table 12. Analysis on CVC ColonDB.

Source of Variation	Mean Squares	Sum of Squares	df	F crit	F	P-value
Between Methods	89.9767	269.9303	3	2.8662	2.8662	0.3925×10^{-8}
Within Methods	184.2464	6632.8716	36			
Total		6902.8019	39			

Table 13. Analysis on ETIS LARIB Dataset.

Source of Variation	**Mean Squares**	**Sum of Squares**	**df**	**F crit**	**F**	**P-value**
Between Methods	317.1112	951.3337	3	2.8387	0.4819	0.4819×10^{-8}
Within Methods	658.0390	26321.5600	40			
Total		27272.8938	43			

use the proposed method as well as existing detection methods shown in tables 10 and 11. The P-value is the probability of getting at least as bad as the one observed, suggesting that the null hypothesis is valid. A lower p-value means that the other hypothesis is better supported by evidence. Statistical significance is greater when the P-value is modest. In this situation, the P-value for the CVC ColonDB is 0.3925×10^{-8} and the Etis-Larib database is anticipated to be 0.4296×10^{-11}. The low P-value shows the statistical significance of the recommended technique. In ANOVA, the F-test is performed to see the variability of the individual data. On the ETIS-Larib database, the F crit value is 2.8387, and CVC ColonDB is 2.8662.

5. Conclusion

This paper proposed a method for detecting intestinal polyps from colonoscopy images using machine learning techniques. The novelty of the proposed automated system lies in selecting image enhancement methods. After patch extraction, we have employed a simple feature extraction scheme based on DOST and SVM-based classification for polyp detection. The post-processing algorithm NMS significantly reduces the number of FPs and FNs in the polyp detector. The experimental results validate the effectiveness of the proposed image enhancement model, improving the detector performance in terms of the recall value by 55.13%, precision value by around 62.74%, and the F1-score value by around 59.25% and the F2-score value by around 61.40% for the CVC Colon DB database. For the ETIS Larib database, an improvement of 55.14%, 46.63%, 50.56%, and 48.12% can be achieved in the recall, precision, F1-score, and F2-score, respectively. Based on the results obtained, it is clear that the proposed method outperforms the existing one. Dr. Sudip Chakraborty, Chakraborty hospital, Port Blair, Andaman & Nicobar, has verified the colonoscopy dataset and performance measures obtained on both the datasets. By experimenting with different feature extraction and classification methods, the proposed approach can be extended to detect anomalies in various medical image schemes. We intend to expand this research in the future by segmenting the detected polyps from the colonoscopy images.

Acknowledgements

The authors would like to thank Dr. Omkar Singh, Director (ANIIMS), Port Blair, Andaman & Nicobar and Dr. Sudip Chakraborty, Chakraborty hospital, Port Blair, Andaman & Nicobar for the valuable guidance throughout the research.

References

[1] Siegel R L, Miller K D, Goding Sauer A, Fedewa S A, Butterly L F, Anderson J C, Cercek A, Smith R A and Jemal A 2020 Colorectal cancer statistics, 2020. *CA: a cancer journal for clinicians* 70(3): 145–164

[2] Lieberman D 2005 Quality and colonoscopy: a new imperative. *Gastrointestinal endoscopy* 61(3): 392–394

[3] Hazewinkel Y and Dekker E 2011 Colonoscopy: basic principles and novel techniques. *Nature reviews Gastroenterology & hepatology* 8(10): 554

[4] Tajbakhsh N, Gurudu S R and Liang J 2015 A comprehensive computer-aided polyp detection system for colonoscopy videos. In: *International Conference on Information Processing in Medical Imaging*, Springer, pp 327–338

[5] Goyal H, Mann R, Gandhi Z, Perisetti A, Ali A, Aman Ali K, Sharma N, Saligram S, Tharian B and Inamdar S 2020 Scope of artificial intelligence in screening and diagnosis of colorectal cancer. *Journal of clinical medicine* 9(10): 3313

[6] Behera B, Kumaravelan G and Kumar P 2019 Performance evaluation of deep learning algorithms in biomedical

document classification. In: 2019 11th *International Conference on Advanced Computing* (ICoAC), IEEE, pp 220–224

[7] Bernal J, Tajkbaksh N, Sánchez F J, Matuszewski B J, Chen H, Yu L, Angermann Q, Romain O, Rustad B and Balasingham I *et al* 2017 Comparative validation of polyp detection methods in video colonoscopy: results from the miccai 2015 endoscopic vision challenge. *IEEE transactions on medical imaging* 36(6): 1231–1249

[8] Ribeiro E, Uhl A, Wimmer G and Häfner M 2016 Exploring deep learning and transfer learning for colonic polyp classification. *Computational and mathematical methods in medicine* 2016

[9] Boroff E S, Disbrow M, Crowell M D and Ramirez F C 2017 Adenoma and polyp detection rates in colonoscopy according to indication. *Gastroenterology Research and Practice* 2017

[10] Zhang R, Zheng Y, Poon C C, Shen D and Lau J Y 2018 Polyp detection during colonoscopy using a regression-based convolutional neural network with a tracker. *Pattern recognition* 83: 209–219

[11] Park S, Lee M and Kwak N 2015 Polyp detection in colonoscopy videos using deeply-learned hierarchical features. Seoul National University

[12] Wang W, Tian J, Zhang C, Luo Y, Wang X, Li J (2020) An improved deep learning approach and its applications on colonic polyp images detection. *BMC Medical Imaging* 20(1): 1–14

[13] Hasan M M, Islam N and Rahman M M 2020 Gastrointestinal polyp detection through a fusion of contourlet transform and neural features. *Journal of King Saud University-Computer and Information Sciences*

[14] Xu J, Zhao R, Yu Y, Zhang Q, Bian X, Wang J, Ge Z and Qian D 2021 Real-time automatic polyp detection in colonoscopy using feature enhancement module and spatiotemporal similarity correlation unit. *Biomedical Signal Processing and Control* 66: 102503

[15] Rahim T, Hassan S A and Shin S Y 2021 A deep convolutional neural network for the detection of polyps in colonoscopy images. Biomedical Signal Processing and Control 68: 102654

[16] Akbar B, Gopi V P and Babu V S 2015 Colon cancer detection based on structural and statistical pattern recognition. In: 2015 2nd *International Conference on Electronics and Communication Systems* (ICECS), IEEE, pp 1735–1739

[17] Lee J Y, Jeong J, Song E M, Ha C, Lee H J, Koo J E, Yang D H, Kim N and Byeon J S 2020 Real-time detection of colon polyps during colonoscopy using deep learning: systematic validation with four independent datasets. *Scientific Reports* 10(1): 1–9

[18] Figueiredo I N, Dodangeh M, Pinto L, Figueiredo P N and Tsai R 2020 Fast colonic polyp detection using a hamilton–jacobi approach to non-dominated sorting. *Biomedical Signal Processing and Control* 61: 102035

[19] Nogueira-Rodríguez A, Domínguez-Carbajales R, López-Fernández H, Iglesias Á, Cubiella J, Fdez-Riverola F, Reboiro-Jato M and Glez-Peña D 2020 Deep neural networks approaches for detecting and classifying colorectal polyps. *Neurocomputing*

[20] Sornapudi S, Meng F and Yi S 2019 Region-based automated localization of colonoscopy and wireless capsule endoscopy polyps. *Applied Sciences* 9(12): 2404

[21] Shin Y and Balasingham I 2017 Comparison of hand-craft feature based svm and cnn based deep learning framework for automatic polyp classification. In: 2017 39th *annual international conference of the IEEE engineering in medicine and biology society* (EMBC), IEEE, pp 3277–3280

[22] Akbari M, Mohrekesh M, Rafiei S, Soroushmehr S R, Karimi N, Samavi S and Najarian K 2018 Classification of informative frames in colonoscopy videos using convolutional neural networks with binarized weights. In: 2018 40th *annual international conference of the IEEE engineering in medicine and biology society* (EMBC), IEEE, pp 65–68

[23] Cheng D C, Ting W C, Chen Y F, Jiang X 2011 Automatic detection of colorectal polyps in static images. *Biomed Eng: Appl. Basis Commun.* **23**: 357

[24] Deeba F, Bui F M, Wahid K A 2020 Computer-aided polyp detection based on image enhancement and saliency-based selection. *Biomedical Signal Processing and Control* 55: 101530

[25] Haj-Manouchehri A and Mohammadi H M 2020 Polyp detection using cnns in colonoscopy video. IET Computer Vision 14(5): 241–247

[26] Taha B, Dias J and Werghi N 2017 Convolutional neural networkasa feature extractor for automatic polyp detection. In: 2017 *IEEE International Conference on Image Processing (ICIP)*, IEEE, pp 2060–2064

[27] Yuan Y, Li D and Meng M Q H 2017 Automatic polyp detection via a novel unified bottom-up and top-down saliency approach. *IEEE journal of biomedical and health informatics* 22(4): 1250–1260

[28] Tajbakhsh N, Gurudu S R and Liang J 2015 Automatic polyp detection in colonoscopy videos using an ensemble of convolutional neu*ral networks. In: 2015* IEEE 12th International Symposium on Biomedical Imaging (ISBI), IEEE, pp 79–83

[29] Tajbakhsh N, Shin J Y, Gurudu S R, Hurst R T, Kendall C B, Gotway M B and Liang J 2017 On the necessity of fine-tuned convolutional neural networks for medical imaging. In: *Deep Learning and Convolutional Neural Networks for Medical Image Computing*, Springer, pp 181–193

[30] Kim M, Yan C, Yang D, Wang Q, Ma J and Wu G 2020 Deep learning in biomedical image analysis. In: *Biomedical information technology*, Elsevier, pp 239–263

[31] Drabycz S, Stockwell R G, Mitchell J R 2009 Image texture characterization using the discrete orthonormal s-transform. *Journal of digital imaging* 22(6): 696–708

[32] Stockwell R G 2007 A basis for efficient representation of the s-transform. *Digital Signal Processin*g 17(1): 371–393

[33] Gibson P C, Lamoureux M P and Margrave G F 2006 Letter to the editor: Stockwell and wavelet transforms. *Journal of Fourier analysis and applications* 12(6): 713–721

[34] Hariharan M, Vijean V, Sindhu R, Divakar P, Saidatul A and Yaacob S 2014 Classificati*on of mental tasks using stockwell transform.* Computers & Electrical Engineering 40(5): 1741–1749

[35] Haralick R M 1979 Statistical and structural approaches to texture. *Proceedings of the IEEE* 67(5): 786–804

[36] Haralick R M, Shanmugam K, Dinstein I H 1973 Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* (6): 610–621

[37] Leon K, Mery D, Pedreschi F and Leon J 2006 Color measurement in lab units from rgb digital images. *Food research international* 39(10): 1084–1091

[38] Reza A M 2004 Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology* 38(1): 35–44

[39] Rahman S, Rahman M M, Abdullah-Al-Wadud M, Al-Quaderi G D and Shoyaib M 2016 An adaptive gamma correction for image enhancement. *EURASIP Journal on Image and Video Processing* 2016(1): 1–13

[40] Palanisamy G, Ponnusamy P and Gopi V P 2019 An improved luminosity and contrast enhancement framework for feature preservation in color fundus images. Signal, *Image and Video Processing* 13(4): 719–726

[41] Chan T F, Kang S H and Shen J 2001 Total variation denoising and enhancement of color images based on the cb and hsv color models. *Journal of Visual Communication and Image Representation* 12(4): 422–435

[42] Dian-Wei W, Peng-Fei H, Jiu-Lun F, Ying L, Zhi-Jie X and Jing W 2018 Multispectral image enhancement based on illuminance-reflection imaging model and morphology operation. *Acta Physica Sinica* 67(21)

[43] Sahu S 2012 Comparative analysis of image enhancement techniques for ultrasound liver image. *International Journal of Electrical and Computer Engineering* 2(6): 792

[44] Salem N, Malik H and Shams A 2019 Medical image enhancement based on histogram algorithms. *Procedia Computer Science* 163: 300–311

[45] Soleimani M, Vahidi A and Vaseghi B 2020 Two-dimensional stockwell transform and deep convolutional neural network for multi-class diagnosis of pathological brain. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29: 163–172

[46] Deepa V, Kumar C S and Andrews S S 2019 Automated detection of microaneurysms using stockwell transform and statistical features. *IET Image Processing* 13(8): 1341–1348

[47] Durak L and Arikan O 2003 Short-time fourier transform: two fundamental properties and an optimal implementation. *IEEE Transactions on Signal Processing* 51(5): 1231–1242

[48] Bentley P M and McDonnell J 1994 Wavelet transforms: an introduction. *Electronics & communication engineering journal* 6(4): 175–186

[49] Daqi G and Tao Z 2007 Support vector machine classifiers using rbf kernels with clustering-based centers and widths. In: 2007 *International Joint Conference on Neural Networks*, pp 2971–2976, 10.1109/IJCNN.2007.4371433

[50] Bernal J, Sánchez F J, Fernández-Esparrach G, Gil D, Rodríguez C, and Vilariño F 2015 Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics* 43: 99–111

[51] Bernal J, Sánchez J and Vilarino F 2012 Towards automatic polyp detection with a polyp appearance model. *Pattern Recognition* 45(9): 3166–3182

[52] Silva J, Histace A, Romain O, Dray X and 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

[53] Kaur T and Gandhi T K 2019 Automated brain image classification based on vgg-16 and transfer learning. In: 2019 *International Conference on Information Technology* (ICIT), IEEE, pp 94–98

[54] Mateen M, Wen J, Song S and Huang Z *et al* 2019 *Fundus image classification using vgg-19 architecture with pca and svd. Symmetry* 11(1):1

[55] McAllister P, Zheng H, Bond R and Moorhead A 2018 Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets. *Computers in biology and medicine* 95: 217–233

[56] Wang C, Chen D, Hao L, Liu X, Zeng Y, Chen J and Zhang G 2019 Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access* 7: 146533–146541

[57] Qadir H A, Shin Y, Solhusvik J, Bergsland J, Aabakken L and Balasingham I 2021 Toward real-time polyp detection using fully cnns for 2d gaussian shapes prediction. *Medical Image Analysis* 68: 101897

[58] Bae S H and Yoon K J 2015 Polyp detection via imbalanced learning and discriminative feature learning. *IEEE transactions on medical imaging* 34(11): 2379–2393