

Using Novel Method with Convolutional Neural Network for Colorectal Cancer Classification

Sushama Tanwar¹, S. Vijayalakshmi², Munish Sabharwal³

¹Research Scholar, Galgotias University, Uttar Pradesh, India 201 307

²Professor, Galgotias University, Uttar Pradesh, India 201 307

³Professor, Galgotias University, Uttar Pradesh, India 201 307

tanwar.sushama25@gmail.com¹, s.vijayalakshmi@galgotiasuniversity.edu.in², mscheckmail@yahoo.com³

Abstract—Detection of diseases with the help of computer has helped the doctors in recognizing the colorectal cancer more efficiently, that further helps in treatment and increases the survival rate of patients. This paper shows our work on classification of endoscopic images using convolutional neural network (CNN). The proposed network preserves the spatial details of endoscopic images by changing the dilation factor. It is said that if the dimensionality of an image is reduced, the image may lose spatial details, which may result in confusion among similar looking polyps or may even miss detection of polyps. In our model we also use regularization technique to overcome the problems like noise, artifacts and overfitting. For evaluating over models we have used matrices namely accuracy matrix, precision matrix, recall matrix and F1-score. Our model gives higher accuracy when compared with traditional models for classification of endoscopic images.

Keywords—Colorectal cancer (CRC), Colorectal cancer classification, image classification, CNN.

I. INTRODUCTION

CRC is the third most dangerous cancers diagnosed across the world. The number of estimated death due to colorectal cancer in 2020 in United States is around 53,200 [1]. To provide effective treatment and improve the survival rate of patients, timely detection and diagnosis of colorectal cancer is required. Although, technologies like artificial and computer vision are available these days which help the doctors in detection and classification of colorectal cancer. The research in this domain is going on since decades and has given effective medical imaging technologies [2][3]. The research has also contributed to the automated detection and classification of brain tumors [4][5], skin cancer [6], breast cancer [7], gastric cancer [8], hookworms [9].

The traditional methods available in machine learning for image classification of images is based on hand crafted features, like texture information, shape, color etc. The proposed techniques make use of feature extraction methods as well as classifiers in order to classify CRC images. But with these traditional methods feature extraction is difficult because of some limitations like color insufflations, blurring, variations in viewpoint and lack of illumination.

Motivated with success of deep learning in computer vision [10][11][12], researchers have started applying deep learning for analysis of endoscopic images. But for this the main challenge is to availability of data set, as medical data is not available in large amount. This challenge can be addressed using transfer learning approach, but this also suffers from lot of issues [13][14][15].

Down sampling is used in deep layers of pre-trained networks, this technique works well with classification of natural images using ImageNet dataset [16], but this does not work well with medical images. The presence of small feature maps at higher layers have limited information, which is insufficient for representing features of endoscopic images, hence cannot identify small polyps or similar images. So, it can be argued that if the output is of bigger size then the features can be represented more accurately and hence classification can be improved. Also, there is a possibility of overfitting when the data given to a pre-trained CNN is small in size, as the CNN can even learn artifacts as well as other details from the dataset that may result in inaccurate results when applied to new data. The main limitations for the use of the traditional machine learning methods in medical imaging is high accuracy is a key demand but is difficult in similar looking limited dataset also the pre-trained methods are very much prone to overfitting if proper regularization method is not used. Even a very small error caused in classification of a medical image may result in bad experience. Two of the diseases one is Crohn's disease and the other is ulcerative colitis have features that are same and both of them are differentiated based on the chronic inflammation in the digestive tract. But the mistake in classification of such disease is not acceptable at all. Hence, there is a need of more accurate and effective classification models that can even learn the minute details in endoscopic images.

We propose a method that can increase the classification accuracy by using dilation in convolutional network. The main assumption that we have made in our method is that the model will be able to learn even very minute details from the dataset when dilation method is used. It can also increase the accuracy of classification when feature maps having high resolution are passed to the layers that are responsible for classification. An increase in the dilation factors may miss the required spatial features in similar looking images and small polyps, and hence cannot be applied to images of such classes. In the same manner to address the problem of artifacts, noise and overfitting DropBlock[17] layer is amended after all the dilated convolution layers, that uses regularization method. It helps in regularization of the data by ignoring the adjacent regions of feature map and hence the model is forced to find some other place for fitting the data. The DropBlock used can remove the artifacts that may be preset in the dataset like motion blur, specular reflection, artificial devices etc.

This research paper initially describes the introduction and motivation for the proposed method. The next section i.e. Section II describes about the related work in colorectal

TABLE 1. Comparison and weaknesses of previous approaches.

| Publications | Method | Classification type | Dataset | Accuracy | Weaknesses |
|--|---|---------------------|-----------------|----------|---|
| M. Hafner et al. [21] | Multiscale novel color texture + local binary pattern (LBP) | polyp | private | 85.3% | -Requires handcrafted features |
| D. Mahapatra et al. [26] | Supervised learning + intensity and texture | crohn's | private | 88.9% | -Requires handcrafted features |
| Z. Wei et al. [27] | Gabor filter banks + K-means clustering + histogram | colitis | private | - | -Requires handcrafted features |
| M. Hafner et al. [19] | Local fractal dimension | polyp | private | 88.2% | -Requires handcrafted features |
| S. S. Ahmed et al. [28] | Neuro-fuzzy model | crohn's | private | 97.67% | |
| E. Mossotto et al. [29] | Combined unsupervised ML | crohn's/colitis | private | 83.3% | -Reliance on endoscopic and histological data |
| A. Alammari et al. [30] | Convolutional neural network | colitis | private | - | -Failure in images with similar features |
| G. Wimmer, A. Vecsei, M. Hafner, A. Uhl [31] | CNN + SVM | polyp/colitis | KSAVIR (public) | 92.5% | -Network complexities |
| M. Zaid et al. [32] | Ensemble of texture and CNN features | polyp/colitis | private | 83% | -Requires handcrafted features |
| G. Urban et al. [33] | Deep CNN | polyp | private | 96% | -Use of ensembling techniques |
| R. W. Stidham et al. [34] | Deep CNN | colitis | private | - | -Experiment on less number of class |
| T. Ozawa et al. [35] | CAD + CNN | colitis | private | - | -Lack of diverse dataset |
| R. Zhang et al. [16] | CNN + SVM | polyp | public | 85.9% | -Less reliable, overfitting |
| Y. Shin and L. Balasingham [36] | 3-layer CNN | polyp | public | 91.26% | -Use of imbalanced dataset |
| | | | | | -Less number of images |
| | | | | | -Lack of diverse datasets |
| | | | | | -less number of classes and images |

classification using endoscopic images. A detailed description of our proposed approach for classification of colorectal cancer is given in section III. Section IV and section V describes the data, training procedure, performance, accuracy metrics and results of our proposed model. The last section that is section VI and section VII tells about the significance and contribution of our findings.

II. RELATED WORK

This section of the paper describes about the methods used for feature extraction and classification of colorectal cancer from endoscopic images. It includes the traditional machine learning methods and modern deep learning methods.

A. CROHN'S DISEASE AND ULCERATIVE COLITIS

Authors of [18] proposed a supervised learning technique that can automatically identify and localize the abdominal areas that are affected by Crohn's Disease. They used features like shape asymmetry of 3D regions, intensity statistics and texture anisotropy in order to differentiate between affected area and unaffected area. Authors in [19] also followed similar approach but used different features like intensity and texture. Authors in [20] detected colitis in tomography scans that were computer by contrast enhancement. They used visual codebook for accurately detecting the tomography scans. Authors in [21] proposed neuro-fuzzy based technique that detects Crohn's disease. They performed tests on various levels of fuzzy partition and for dimensionality reduction they used factor analysis.

Authors in [22] proposed three unsupervised machine learning models. In the first model they used endoscopic data and which gave an accuracy of 71%. In the second model they used histological data, and achieved accuracy of 76.9% and in the third one they used both histological and endoscopic data and achieved an accuracy of 82.7%. Authors in [23] classified Crohn's disease and ulcerative colitis by calculating individualized pathway scores using genes. Authors in [24] used global features, deep convolutions neural network and deep transfer learning to classify different diseases and also made a dataset named as "KVASIR". Authors in [25] proposed a technique to classify the depth of ulcerative colitis. They used deep convolutional network along with the knowledge of endoscopic domain. Authors of [26] showed that the accuracy calculated by a

deep convolutional neural network is equivalent as compared to that of a radiologist when classification of the intensity of ulcerative colitis. Authors in [27] found the GoogLeNet architecture of CNN based on computer aided diagnosis (CAD) systems is very robust for the detection of intensity of ulcerative colitis. Authors in [28] proposed a CAD system for prediction of soreness related to ulcerative colitis.

B. COLONIC POLYPS

Polyps are initial phase of colorectal cancer. Some polyps get converted into cancer cells. Authors of [29], in their study of classification of colorectal polyps proposed a technique called as texture analysis which was based on local fractal dimensions (LED). Their study described three LFD based approaches, this approach was able to discover some extra features from the image like the shape and gradient information which further helped in increasing the accuracy of classification, and it was tested on different set of data. They also proposed a filter bank based technique, which was a texture analysis technique and contributed in colorectal polyp classification [30]. The filter bank had filter mask that classified different polyp from each other. Authors in [31] proposed a color texture operator that made use of local binary pattern variant, this operator was able to automatically classify endoscopic images. They created a color vector field by finding the similarity of neighboring pixels and for further classification they used kNN classifier. Authors in [31] worked on eleven different datasets of endoscopic polyp images and tested wavelet based approaches on those datasets. Authors in [32] proposed three different approaches for feature extraction that were based on wavelet and discovered that those techniques were suitable for automated colonic polyp classification. Authors in [33] proposed a detection system that was based on local features. Authors in [34] proposed an architecture that created a new feature by combining Gabor filter and monogenic local binary pattern. This features generated extracted information about the shape and edgeat multiresolution and kept the color details. In this architecture they used linear discriminant analysis for feature reduction and SVM for classification. Authors in [35] proposed a technique that used two segmentation techniques, through which some features were extracted which can be further used for classification of colonic polyps.

Now a days after the introduction of convolutional neural networks (CNN) the use handcrafted features is reduced for feature extraction and classification and is replaced by CNNs [36]. Authors in [37] worked on detection, multiclass classification and localization of colonic polyps by using various technologies like deep learning, information retrieval and local and global feature analysis. Authors in [38] introduced a transfer learning approach that used deep CNN to learn features from non-medical data and then used low-level features for detection and localization of colonic polyp. Authors in [39] compared the methods that were based on hand crafted features and CNN by using three different databases and found that CNN gave better results. Authors in [40] classified gastrointestinal disease by using deep learning and texture features. Authors in [41] detected polyps by making use of deep CNN and verified the results with human expert. The method proposed by them was able to detect all the polyps that the human expert had detected.

Authors in [42] classified celiac disease and colonic polyps from an endoscopic database. They used three CNN architectures that were pre-trained and SVM as classifier. They performed concatenation and combination on the result obtained from different layers and obtained better

results than CNN based methods.

C. LIMITATIONS OF RELATED WORK

After reviewing the related work some problems in previous methods were found which are listed in Table 1. Also some weakness in the previous methods is listed below:

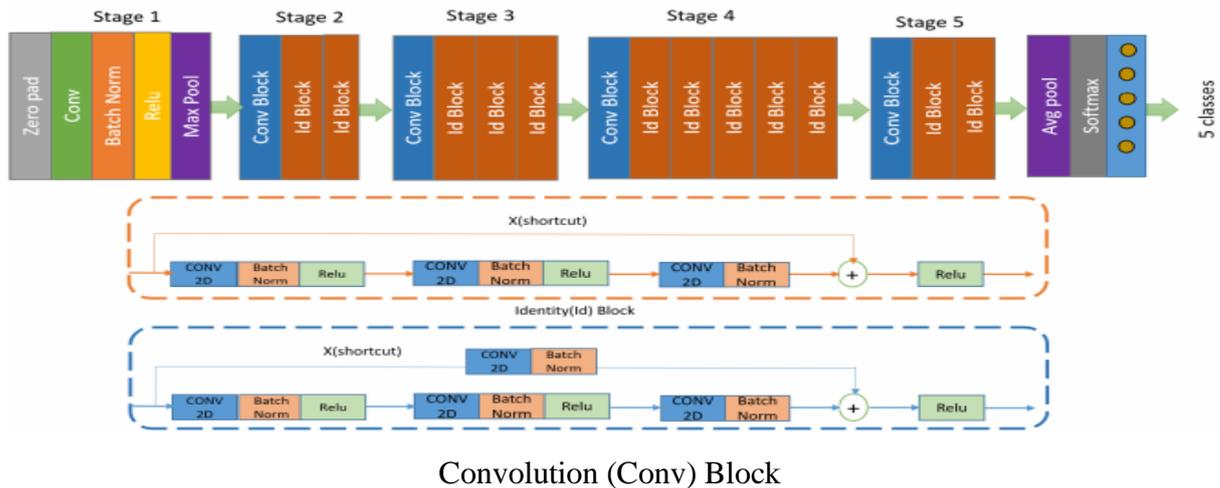


Fig 1. Overview of ResNet50 architecture[45]. At the first stage down-sampling with stride 2 is performed on the feature map and then batch normalization is performed followed by ReLU layer. All the stages need equal number of layer. Convolution block and identity block are present in each stage. The identity block has 3 convolutional layer

a) These approaches had dependency over the handcrafted features. For getting the handcrafted features it is required to get the deep knowledge of image [19][20][29][31]-[35].They needed texture analysis which is computed by the classifier by feeding local descriptors of image in it. Although some of these approaches give high accuracy but still lack in generalization and transfer ability in inter-dataset variability.

b) The dataset used is limited [26], the classes of data is also less [15][41][39].

c) Most of the approaches depend on histological data and endoscopic data, which limits the application of the technique in real scenario as histological images may not be available always.

d) The approaches are not aware of the features that the network learn during the training process [14][25][42][41][26][27][39][24].

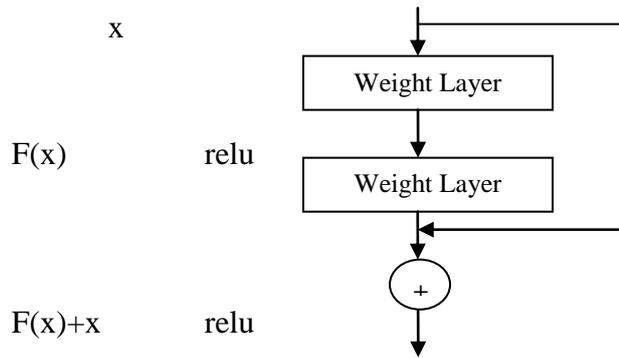


Fig 2. Residual Learning[45]

III. METHODOLOGY

In our model the CNN network uses transfer learning approach instead of training a new CNN, we have used a pre-trained network. The features that are obtained by activating a CNN that is trained in a fully supervised manner for recognition task on large scale can be used again for

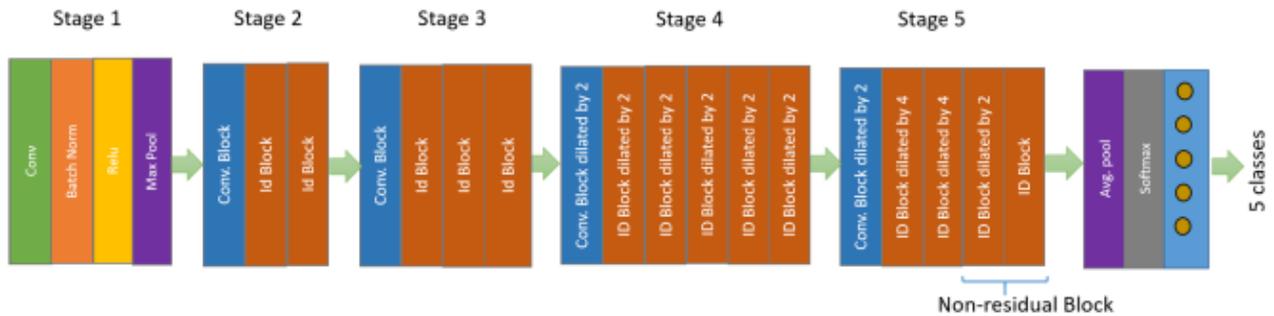


Fig 3. Proposed Architecture. Dilation is added at stage 4 and stage 5. In the end non-residual layer is added. Also strides in stage 4 and 5 are removed and all the other blocks are dilated. DropBlock regularization method is performed after every Convolution layer.

| Layer Name | Conv1 | Conv2_x | Conv3_x | Conv4_x | Conv5_x | Fc |
|-----------------|---|--|---|--|--|---|
| Original ResNet | | | | | | |
| Proposed Method | 7*7,64, stride 2 output size 112 * 112 | 3*3 max pooling, stride 2 $\begin{pmatrix} 1*1,64 \\ 3*3,64 \\ 1*1,256 \end{pmatrix} \times 3$ output size 56*56 | $\begin{pmatrix} 1*1,128 \\ 3*3,128 \\ 1*1,512 \end{pmatrix} \times 4$ output size 28*28 | $\begin{pmatrix} 1*1,256 \\ 3*3,256 \\ 1*1,1024 \end{pmatrix} \times 6$ output size 14*14 | $\begin{pmatrix} 1*1,512 \\ 3*3,512 \\ 1*1,2048 \end{pmatrix} \times 3$ | output size 7*7 |
| | | $\begin{pmatrix} 1*1,64 \\ 3*3,64 \\ 1*1,256 \end{pmatrix} \times 3$ output size 56*56 | $\begin{pmatrix} 1*1,128 \\ 3*3,128 \\ 1*1,512 \end{pmatrix} \times 4$ output size 28*28 | $\begin{pmatrix} 1*1,256 \\ 3*3,256 \text{ dilate } 2 \\ 1*1,1024 \end{pmatrix} \times 6$ output size 28*28 | $\begin{pmatrix} 1*1,512 \\ 3*3,512 \text{ dilate } 2 \\ 1*1,2048 \end{pmatrix} \times 1$ $\begin{pmatrix} 1*1,512 \\ 3*3,512 \text{ dilate } 4 \\ 1*1,2048 \end{pmatrix} \times 2$ $\begin{pmatrix} 1*1,512 \\ 3*3,512 \text{ dilate } 2 \\ 1*1,2048 \end{pmatrix} \times 1$ $\begin{pmatrix} 1*1,512 \\ 3*3,512 \text{ dilate } 1 \\ 1*1,2048 \end{pmatrix} \times 1$ | Average pool, number of class fc, softmax |

Fig 4. Detailed proposed method and baseline model

novel task. Also, the dataset we had contained limited number of images that was not enough for training the CNN as it needs a large number of parameters for getting trained. Authors of

[13][45][46][47][48][49] have shown that a pre trained CNN with proper fine tuning gives better results or in worst case similar results than a CNN trained from scratch. The baseline model used is ResNet50 architecture as by experiments we came to know that for colorectal dataset ResNet50 gave better results as compared to any other architecture.

Fig 2 displays the basic residual network block along with the identity connection. It learns the function below:

$$A(k)=B(k)+k \quad (1)$$

In this equation k is the identity function and $B(x)$ represents the stacked non-linear layers. The ResNet50 architecture used has five stages of blocks as shown in fig 3. In each block small chunks of network are connected to form big network using skip or shortcut connections. Depending on the dimensions of the input and output two blocks are used. If the input activation and output dimension has the same size, then function becomes as follows:

$$y = f(x, W_i) + x \quad (2)$$

in this equation x and y are input and output vectors of individual layers. The residual mapping is represented by the $f(x, W_i) + x$ function. Fig 1 represents the example of an identity block having two paths shortcut path and main path. On the other hand if the size of input and output activation is different, a convolutional layer to the upper path is added as follows:

$$y = f(x, W_i) + W_s x \quad (3)$$

The identity block has three different sets of convolutional layers, and then is the batch normalization followed by ReLU activation function. On the other had convolutional layer also has the same number of layers followed by an additional layer that is the convolutional layer.

We assume that it is required to keep the down sampling approach and the spatial information is preserved in the last layers. This was done by adding dilation and removing the down sampling. The architecture that we have used was made for ImageNet Classification also we have made some changes so as to adjust to our network [43].

A. APPROACH

This section gives the details of the approach used to learn and represent the endoscopic features from the colon disease images by our model. In order to attain this we have used dilation at the end of each layer. The proposed model contains five layers as described earlier in section III. We assume L^i as the group of layer where $i=1 \dots 5$. The j^{th} layer in group I is represented as L^i_j . The filter corresponding to L^i_j is represented by F^i_j . The output of L^i_j calculated by our model is

$$(L^i_j + F^i_j) (p) = \sum_{a+b=p} L^i_j(a) F^i_j(b) \quad (4)$$

To attain the desired result dilated convolutions is used in the last two groups so in stage 4 dilated operators are used with rate 2 for each layer in the block.

$$(L^4j + 2F^4j) (p) = \sum_{a+b=p} L^4j(a)F^4j(b) \quad (5)$$

In L^5_1 that is the first layer of stage 5 following transformations are performed

$$(L^5j + 2F^5j) (p) = \sum_{a+b=p} L^5j(a)F^5j(b) \quad (6)$$

And in other blocks of stage 5 a dilated factor of 4 is used by analogy

$$(L^5j + 4F^5j) (p) = \sum_{a+b=p} L^5j(a)F^5j(b) \quad (7)$$

for $j=1, \dots, 4$. In L^5_4 that is the forth block of stage 5 a dilation factor of 2 is used

$$(L^5j + 2F^5j) (p) = \sum_{a+b=p} L^5j(a)F^5j(b) \quad (8)$$

for $j=3,4,5$. In the end a non-residual block with normal convolution is added and a global average pooling layer is also added after the non-residual layer, which is same as in the original architecture. The global average pooling layer is used to limit down the feature maps of the output to a vector. This vector is mapped to another vector having the prediction scores of all the classes by 1×1 convolution as shown in Fig 3 and Algorithm 1.

Fig4.gives the details of each layer in the original architecture and modified architecture. The model proposed by us has a total of fifty-seven layers. The earlier layers have same shape as well as structure but the structure of the later layers is changed because of the addition of dilation. A feature map of size $112 \times 112 \times 64$ is generated by the first convolutional layer. The size of the input image is kept as 224×224 and 64 filters of $7 \times 7 \times 3$ size are applied on the image. Further a max pooling layer is used that generates output feature map of size $56 \times 56 \times 64$ by processing the input by a filter of size 3×3 . Downsampling is performed in the original model by applying 1×1 convolution layer. Also stride of size 2 is applied in layers of stage 3 to stage 5. On the other hand, in the proposed model the stride applied is of size 1 and also the 3×3 dilation layer is used in stage 3 instead of 3×3 convolution layer. Finally, the global average pooling layer is used to generate optimal feature vector of size $1 \times 1 \times 2048$

Algorithm 1 Overall Proposed Method

```

1: Initialize the ResNet50 Networks
2: ResNet ← load()
3: features, labels = get_batch(dataset)
4: model ← create_model()
5: if mode == training then
6:     fitted-model ← model.fit(features, labels)
7:     for e = 1 : epochs do
8:         for i = 1: Gi do ▷ Gi is different stages of layers.
9:             if Gji == Gj4 then ▷ Gji is jth layer in group i
10:                Execute equation number (5),
11:                Add DropBlock
12:            end if
13:            if Gji == Gj5 then
14:                Execute equation number (6,7,8),
15:                Add DropBlock
16:            end if
17:        end for
18:    end for
19: model ← fitted-model()
20: end if
    
```

Algorithm 2 DropBlock Regularization Method

```

1: Input: output activations of a layer (A), block_size, γ
2: Sample mask matrix M randomly, where Mi,j ~ Bernoulli(γ);
3: For each zero position Mi,j, expand a spatial square mask with the center being M : Mi,j, the width, height being block_size and set all the values of M in the square to be zero.
4: Apply the mask: A = A * M
5: Normalize the features: A = A * count(M) / count_ones(M)
    
```

When the frequency content of feature map is higher than the dilated convolution sampling rate the problem of gridding artifacts arise. In order to avoid gridding artifacts, two non-residual blocks are added in the proposed model having decreased dilation. Also this prevents the propagation of gridding artifacts from one layer to the other. The modified network gives an output of 28x28 after G⁵ layers, which in turn supports global average pooling layer to work on large number of values. It also supports the classifier to identify the features that cover a small part in any image.

The problem of overfitting or poor local minima may arise due to addition of more non-residual blocks at the last layer which in turn increases the network size. The limited dataset creates issues in such case. Also, the endoscopic images have a lot background noise and artifacts that is one of the limitations in classification. To handle these problems DropBlock method is used [17][52][53][54] which regularize the convolutional network. The DropBlock method decreases adjacent region of the feature map, whereas the Dropout method decreases features randomly [44][50][51]. DropBlock is applied to all the blocks of stage 4 and stage 5 after convolutional layers. All these details are given in algorithm 2. It contains two parameters α and β . The parameter α contains the information about the continuous regions that are required to be decreased whereas β tells the number of units to be decreased. By keeping the size of α as 7*7 the value of β is calculated as

$$\beta = ((1 - k)(s^2)) / \beta^2 (s - \beta + 1)^2 \quad (9)$$

here k denotes the probability of keeping an activation unit. The initial binary mask is sampled using Bernoulli distribution having (1-k) as mean. s denotes the feature map size and (s - β + 1)

represents the valid seed region. For computing β the value of k is kept 0.9.

TABLE 2. Details of colorectal dataset.

| Actual Class | Training set | | Validation set | Testing set |
|--------------------|---------------------|--------------------|----------------|-------------|
| | Before Augmentation | After Augmentation | | |
| Adenocarcinoma | 474 | 7380 | 80 | 80 |
| Adenoma | 615 | 7380 | 80 | 80 |
| Crohn's | 403 | 7380 | 80 | 80 |
| Ulcerative colitis | 613 | 7380 | 80 | 80 |
| Normal | 610 | 7380 | 80 | 80 |

IV. EXPERIMENT

A. DATA COLLECTION

1) KVASIR DATASET

KVASIR is an open dataset that is available for free. It has around 8000 images, 8 classes each having 1000 images. It contains a large set of images divided into two categories namely anatomical landmarks that shows Z-line,pylorus and cecum and pathological finding that shows esophagitis, polyps and ulcerative colitis. Some sets also describe about the removal of lesions. They include dyed, lifted polyps and dyed resection margins. It has images of various resolutions ranging from 720x576 to 1920x1072pixels. Pogorelov et al performed evaluation of these databases with three methods: classification using global features, deep convolution neural networks and deep transfer learning. For the experiments the data is divided in the ratio of 50:50 in order to compare with the original paper. This paper shows the comparison of the proposed method with other approaches.

2) COLORECTAL DATA

The dataset is provided by Mahaveer Cancer Hospital, Jaipur. The dataset is divided into five classes of 3515

TABLE 3. F-score (F1), precision, recall for the evaluated different CNNs architecture on the colorectal dataset.

| Model | Class | precision | recall | F1 |
|-----------------------|--------------------|-------------|-------------|-------------|
| VGG16 [12] | adenocarcinoma | 0.88 | 0.91 | 0.90 |
| | adenoma | 0.94 | 0.91 | 0.92 |
| | crohn's disease | 0.85 | 0.69 | 0.76 |
| | normal | 0.95 | 0.94 | 0.94 |
| | ulcerative colitis | 0.79 | 0.94 | 0.86 |
| | | 0.88 | 0.87 | 0.87 |
| InceptionResnet2 [43] | adenocarcinoma | 0.83 | 0.96 | 0.89 |
| | adenoma | 0.93 | 0.86 | 0.90 |
| | crohn's disease | 0.93 | 0.78 | 0.84 |
| | normal | 0.99 | 0.90 | 0.95 |
| | ulcerative colitis | 0.82 | 0.96 | 0.89 |
| | | 0.90 | 0.89 | 0.89 |
| Xception [44] | adenocarcinoma | 0.89 | 0.89 | 0.89 |
| | adenoma | 0.90 | 0.86 | 0.88 |
| | crohn's disease | 0.88 | 0.72 | 0.79 |
| | normal | 0.97 | 0.93 | 0.95 |
| | ulcerative colitis | 0.74 | 0.94 | 0.83 |
| | | 0.87 | 0.86 | 0.86 |
| ResNet [42] | adenocarcinoma | 0.90 | 0.94 | 0.92 |
| | adenoma | 0.94 | 0.91 | 0.92 |
| | crohn's disease | 0.91 | 0.79 | 0.85 |
| | normal | 0.96 | 0.97 | 0.97 |
| | Ulcerative colitis | 0.83 | 0.93 | 0.88 |
| | | 0.91 | 0.91 | 0.91 |
| DenseNet [45] | adenocarcinoma | 0.90 | 0.91 | 0.91 |
| | adenoma | 0.94 | 0.93 | 0.93 |
| | crohn's disease | 0.90 | 0.76 | 0.82 |
| | normal | 0.97 | 0.95 | 0.96 |
| | ulcerative colitis | 0.81 | 0.95 | 0.87 |
| | | 0.90 | 0.90 | 0.90 |
| NasNet [46] | adenocarcinoma | 0.84 | 0.96 | 0.90 |
| | adenoma | 0.93 | 0.88 | 0.90 |
| | crohn's disease | 0.93 | 0.79 | 0.85 |
| | normal | 0.99 | 0.89 | 0.94 |
| | ulcerative colitis | 0.82 | 0.96 | 0.89 |
| | | 0.90 | 0.89 | 0.90 |

TABLE 4. F-score(F1), precision, recall for the evaluated different existing approaches on colorectal dataset.

| Model | Class | precision | recall | F1 |
|---------------------------------|--------------------|--------------|--------------|-------------|
| R. Zhang et al. [16] | adenocarcinoma | 0.78 | 0.81 | 0.80 |
| | adenoma | 0.82 | 0.94 | 0.88 |
| | crohn's disease | 0.79 | 0.79 | 0.79 |
| | normal | 0.99 | 0.98 | 0.99 |
| | ulcerative colitis | 0.79 | 0.84 | 0.81 |
| | | 0.83 | 0.87 | 0.85 |
| Y. Shin and I. Balasingham [36] | adenocarcinoma | 0.79 | 0.90 | 0.84 |
| | adenoma | 0.89 | 0.95 | 0.92 |
| | crohn's disease | 0.93 | 0.54 | 0.73 |
| | normal | 0.93 | 0.95 | 0.94 |
| | ulcerative colitis | 0.64 | 0.90 | 0.75 |
| | | 0.83 | 0.84 | 0.83 |
| W. Ryan et al.[34] | adenocarcinoma | 0.89 | 0.91 | 0.90 |
| | adenoma | 0.94 | 0.91 | 0.92 |
| | crohn's disease | 0.86 | 0.75 | 0.80 |
| | normal | 0.97 | 0.94 | 0.96 |
| | ulcerative colitis | 0.81 | 0.94 | 0.87 |
| | | 0.894 | 0.89 | 0.89 |
| Proposed Method | adenocarcinoma | 0.96 | 0.94 | 0.95 |
| | adenoma | 0.95 | 0.97 | 0.96 |
| | crohn's disease | 0.92 | 0.82 | 0.87 |
| | normal | 0.99 | 0.97 | 0.98 |
| | ulcerative colitis | 0.84 | 0.94 | 0.89 |
| | | 0.932 | 0.928 | 0.93 |

TABLE 5. Accuracy (Acc), f-score (F1), recall, precision for the kvasir dataset.

| Method | Acc | Precision | Recall | F1 |
|--------------------------|--------------|--------------|--------------|-------------|
| 6 Layer CNN | 0.914 | 0.661 | 0.640 | 0.651 |
| 3 Layer CNN | 0.959 | 0.589 | 0.408 | 0.453 |
| Inception v3 TFL | 0.924 | 0.698 | 0.689 | 0.693 |
| 2 GF Random Forest | 0.928 | 0.713 | 0.715 | 0.711 |
| 2 GF Logistic Model | 0.926 | 0.706 | 0.707 | 0.705 |
| 6 GF Random Forest | 0.933 | 0.732 | 0.732 | 0.727 |
| 6 GF Logistic Model Tree | 0.937 | 0.748 | 0.748 | 0.747 |
| Proposed Method | 0.957 | 0.868 | 0.922 | 0.88 |

images that include 634 images of adenocarcinoma, 775 images adenoma, 563 images of Crohn's disease, 773 images of ulcerative colitis and 770 normal images. Initially the images were of different size ranging from 400x400 to 2000x2000 pixels. So, the images were first processed to make their size according to the architecture. Normalization was performed on images with default features as needed by the architecture. As the data set was small augmentation was performed on the dataset so as to increase the size of the dataset. The dataset available was not balanced so while augmenting it was taken care that the class having less data was augmented more so that the data becomes balanced and hence reducing the problem of overfitting. For augmentation various techniques like scaling, flipping, zooming, rotating, shearing, contrast normalization were used. The images were first rotated, then flipped in both horizontal and vertical direction and finally zoomed. The complete dataset was first divided into two groups one was used for training purpose and the other for testing and validation. The description of the dataset is given in Table 2.

B. PERFORMANCE METRICES

For checking the performance we made four evaluation matrices accuracy, precision, recall and F1 score. The matrices were calculated as follows:

$$\text{Accuracy} = \frac{\text{true positive} + \text{true negative}}{\text{true positive} + \text{true negative} + \text{false positive} + \text{false negative}} \quad (10)$$

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \quad (11)$$

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \quad (12)$$

$$F1 \text{ score} = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (13)$$

Here Accuracy is the ratio of correctly classified images to total images. Recall is the ratio of true positives to that of true positives and false negatives. Precision is the ratio of true negatives to that of true negatives and false positives. F1 score is the average of recall and precision.

C. NETWORK TRAINING

The methodology used is Keras on the front end and Tensorflow at the backend. The model was trained and parameters were learned using the training dataset. To optimize and correct the learning rate validation dataset is used. To evaluate the model for recognition and generalization test dataset is used.

The weights of ResNet50 are initialized and stochastic gradient descent with batch size 16 is used. The learning rate starts from 0.001 is divided by 10 when patience level becomes more than 8. A momentum of 0.9 and weight decay 0.0001 is used.

TABLE 6. Ablation studies for the dilated rate at stage 4 and stage 5. We evaluated the proposed method on the given values.

| Method | Dilation rate at stage 4 | Dilation rate at stage 5 | DropBlock | precision | recall | F1 |
|---------------------------|--------------------------|--------------------------|-----------|-----------|--------|-------|
| Baseline model (ResNet50) | 1 | 1 | - | 0.91 | 0.91 | 0.91 |
| | 1 | 2 | - | 0.91 | 0.92 | 0.91 |
| | 2 | 2,4 | - | 0.88 | 0.91 | 0.89 |
| | 2 | 2,4 | ✓ | 0.88 | 0.92 | 0.90 |
| | 2 | 2,4,2 | - | 0.91 | 0.90 | 0.90 |
| | 2 | 2,4,2 | ✓ | 0.90 | 0.93 | 0.91 |
| ResNet50 | 2 | 2,4,2,1 | - | 0.93 | 0.92 | 0.92 |
| | 2 | 2,4,2,1 | ✓ | 0.932 | 0.928 | 0.93 |
| | Proposed Method | 2 | 2,4,2,1 | ✓ | 0.932 | 0.928 |

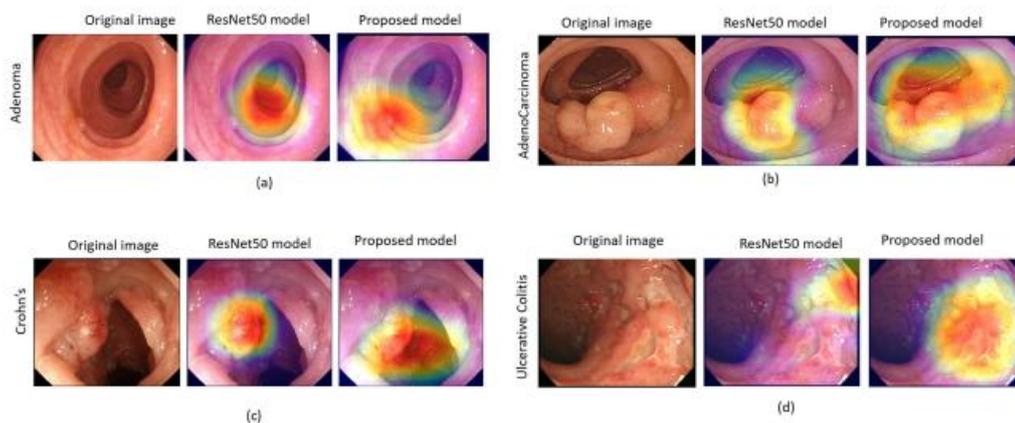


FIGURE 5. Comparison of CAMs generated from the proposed and the baseline ResNet50 method. The proposed method highlighted the specific regions which were misclassified by the baseline model.

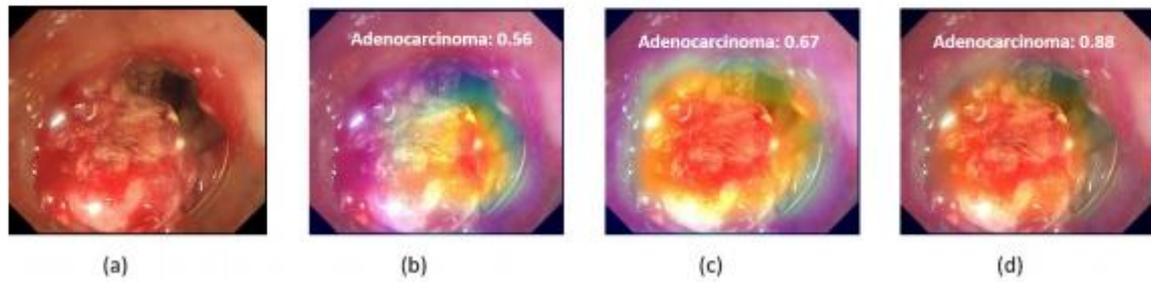


FIGURE 6. Class activation mapping (CAM): (a) original image (b) original ResNet50 model, (c) Dilated ResNet50 trained without DropBlock, (d) Dilated ResNet50 trained with DropBlock (Proposed Method). A proposed method tends to focus on the affected part only and less sensitive to noise, unlike (c).

V. RESULTS

A. PERFORMANCE OF KVASIR DATASET

The results using KVASIR dataset are represented in Table 5 using accuracy, recall, precision and F1 score matrices. The results are compared with the proposed model in accordance with the baseline model given in the original research. The baseline models include classification using global features, deep convolution neural networks and deep transfer learning. The proposed model achieved an accuracy of 95.7% and F1 score 0.88 which overcomes the performance using 2 layer CNN and 3% better than model using transfer learning in deep learning.

B. PERFORMANCE OF COLORECTAL DATASET

This section compares the performance of the proposed model with the earlier models used for classification of endoscopic images. The work done on colorectal cancer using deep learning is limited so we have compared with other similar tasks. The comparison results are presented in Table 3 when different models were trained using colorectal image dataset. The evaluation was performed keeping the number of parameters, augmented dataset and validation dataset same. The comparison showed same results but ResNet50 outperformed all the others. Table 4 compares the output of each model using F1 score.

The proposed model gives better results than all the other models giving F1 score as 0.93. Authors of [38] were able to achieve similar results in normal class but lagged behind in similar looking disease class. Authors in [39] obtained F1 score 0.836 using three layers which shows that it cannot learn the complex features in the image as it is not that deep so it gave poor performance. Authors in [26] were able to obtain F1 score of 0.89 with 159 layers, it shows that its discriminative capability is powerful but with other CNN methods accuracy obtained was similar.

C. ABLATION STUDY

To evaluate the performance of proposed model, it is compared with the baseline models with same experimental settings. For the experimentation the dilation at stage 4 and stage 5 is removed. Table 6 shows the results obtained by recall, F1 score, precision. From the results recorded in Table 6 following conclusions can be drawn:

- 1) If dilation is added only at the end layer, the classification will not give good result infact it may even degrade the performance. Also if a convolution block with dilation

rate 2 is added at stage 4 does not have any impact.

2) The F1 score gets down from 0.91 to 0.89 and 0.90 when dilation rate changes are made at 4th and 6th row of stage 4 and stage 5 as gridding artifacts effect the system. The DropBlock technique used for regularization increases the F1 score at 5th and 7th branch which is higher as compared 4th and 6th row.

3) Also there is an improvement in the F1 score by 0.91 to 0.92 because of addition of increasing and decreasing dilation rates at the last layers. Also, the performance increases by 0.92 to 0.93 by the use of DropBlock regularization. Hence it can be concluded that the proposed model is effective in recognizing small polyp and similar looking images.

VI. DISCUSSION

The proposed method has obtained the best results on the given dataset in terms of classification as shown by Table 4. It can be seen that applying downsampling and preserving features at the end blocks boosts up the recall rate to 92.8% and precision rate by 93.2%. From table 3 it can be concluded that in medical domain transfer learning may not give good results. Due to progressive down-sampling approach used by CNN it does not give good results with the dataset that has high interclass similarity and interclass variation.

ResNet50 fails in classification of endoscopic images at some place as shown in Fig5. If the endoscopic image has very small polyp than classification is difficult as CNN uses progressive down-sampling approach. As it can be seen Fig 5a adenoma is classified as normal image. This kind of misclassification occurs because of the inappropriate spatial information because of reduction in image resolution that represents the tiny feature maps of size 7x7 in the end. At deep layers the learned features are more class specific, the classification becomes more difficult if similar features occur in different classes. To quote an example some adenoma that is polyp and may grow as adenocarcinoma have similar shape to that of continuous inflammation. So the model may get confused and misclassify adenocarcinoma as polyp as shown in Fig5b. In Fig 5 as shown the model is able to detect only some patterns in each class of last two sets of images. These features are further used for classification process. But the proposed method preserves the learned information till the end layers. So, it can be concluded that if dilation is added at the end layers classifying process becomes more effective as compared to the fine tuning of CNN and other techniques.

The proposed method is also capable of handling with noises and artifacts in the endoscopic image. The dilation convolution used with DropBlock regularization gives better results as can be seen in Fig 6. The proposed method gives a probability score from 56% to 88%, so it can be concluded that the proposed model can better classify more specific and essential regions and does not get affected by noise and artifacts.

The proposed method is also helpful for extracting useful features from endoscopic images. The experiments also show that when the dataset had images that were hard to distinguish CNN was able to learn the features from last layers using class activation map approach [45] and similar methods were not able to attain such accuracy rates. Also, the proposed method was able to get a F1 score 0.88 with 92% recall rates with KVASIR dataset, that shows the high capability of recognizing in disease class. From the experiments and their results it can be seen that the proposed method is better and more stable than the other traditional methods for classification of

endoscopic images.

VII. CONCLUSION

This work investigates the use of deep learning techniques for classification of endoscopic images. It can be seen that the features shown by the layers before global average pooling were not enough because of large use of down-sampling because of which spatial information was lost. So, to preserve the information at last layer the dilation convolution is used in increasing and decreasing order. Also the use of DropBlock at deep layer was able to recognize specific regions without getting effected by noise and artifacts. The classification task was also done attaining a higher accuracy which shows that the proposed model was able to capture very small and detailed information among similar images. From the experiments and comparison with KVASIR dataset it can be seen that the proposed method gives better performance in endoscopic image classification. In future this technique can be used for classification of other endoscopic images. It can also be extended by using earlier feature layers and deep features with dilation to handle classification problems in other domain.

REFERENCES

- [1] The American Cancer Society medical and editorial content team, Accessed on: Aug 8, 2020. [Online]. Available: <https://www.cancer.org/content/dam/CRC/PDF/Public/8604.00.pdf>
- [2] M. Owais, M. Arsalan, J. Choi, and K. R. Park, "Effective diagnosis and treatment through content-based medical image retrieval (CBMIR) by using artificial intelligence," *J. Clin. Med.*, vol. 8, no. 4, p. 462, Apr. 2019.
- [3] F. Amato, A. López, E. M. Peña-Méndez, P. Vañhara, A. Hampl, and J. Havel, "Artificial neural networks in medical diagnosis," *J. Appl. Biomed.*, vol. 11, no. 2, pp. 47–58, 2013. 99236 VOLUME 8, 2020 S. Poudel et al.: Colorectal Disease Classification Using Efficiently Scaled Dilation in CNN
- [4] B. Li and M. Q.-H. Meng, "Tumor recognition in wireless capsule endoscopy images using textural features and SVM-based feature selection," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 323–329, May 2012
- [5] S. Sawant and M. Deshpande, "Tumor recognition in wireless capsule endoscopy images," *Int. J. Comput. Sci. Netw. Secur.*, vol. 15, no. 4, p. 85, 2015.
- [6] H. Takiyama, T. Ozawa, S. Ishihara, M. Fujishiro, S. Shichijo, S. Nomura, M. Miura, and T. Tada, "Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks," *Sci. Rep.*, vol. 8, no. 1, p. 7497, Dec. 2018.
- [7] D. M. Vo, N.-Q. Nguyen, and S.-W. Lee, "Classification of breast cancer histology images using incremental boosting convolution networks," *Inf. Sci.*, vol. 482, pp. 123–138, May 2019.
- [8] T. Hirasawa, K. Aoyama, T. Tanimoto, S. Ishihara, S. Shichijo, T. Ozawa, T. Ohnishi, M. Fujishiro, K. Matsuo, J. Fujisaki, and T. Tada, "Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images," *Gastric Cancer*, vol. 21, no. 4, pp. 653–660, Jul. 2018.
- [9] J.-Y. He, X. Wu, Y.-G. Jiang, Q. Peng, and R. Jain, "Hookworm detection in wireless capsule endoscopy images with deep learning," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2379–2392, May 2018.

- [10] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, “ImageNet large scale visual recognition challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [11] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, arXiv:1409.1556. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [12] K. Yun, J. Park, and J. Cho, “Robust human pose estimation for rotation via self-supervised learning,” *IEEE Access*, vol. 8, pp. 32502–32517, 2020.
- [13] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, “Convolutional neural networks for medical image analysis: Full training or fine tuning?” *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
- [14] K. Pogorelov, M. Riegler, S. L. Eskeland, T. de Lange, D. Johansen, C. Griwodz, P. T. Schmidt, and P. Halvorsen, “Efficient disease detection in gastrointestinal videos—Global features versus neural networks,” *Multimedia Tools Appl.*, vol. 76, no. 21, pp. 22493–22525, Nov. 2017.
- [15] R. Zhang, Y. Zheng, T. W. C. Mak, R. Yu, S. H. Wong, J. Y. W. Lau, and C. C. Y. Poon, “Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain,” *IEEE J. Biomed. Health Informat.*, vol. 21, no. 1, pp. 41–47, Jan. 2017.
- [16] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [17] G. Ghiasi, T.-Y. Lin, and Q. V. Le, “DropBlock: A regularization method for convolutional networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 10727–10737.
- [18] D. Mahapatra, P. Schueffler, J. A. Tielbeek, J. M. Buhmann, and F. M. Vos, “A supervised learning based approach to detect Crohn’s disease in abdominal MR volumes,” in *Proc. Int. MICCAI Workshop Comput. Clin. Challenges Abdominal Imag.* Berlin, Germany: Springer, 2012, pp. 97–106.
- [19] D. Mahapatra, P. Schueffler, J. A. W. Tielbeek, J. M. Buhmann, and F. M. Vos, “A supervised learning approach for Crohn’s disease detection using higher-order image statistics and a novel shape asymmetry measure,” *J. Digit. Imag.*, vol. 26, no. 5, pp. 920–931, Oct. 2013.
- [20] Z. Wei, W. Zhang, J. Liu, S. Wang, J. Yao, and R. M. Summers, “Computer-aided detection of colitis on computed tomography using a visual codebook,” in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 141–144.
- [21] S. S. Ahmed, N. Dey, A. S. Ashour, D. Sifaki-Pistolla, D. Balas-Timar, V. E. Balas, and J. M. R. S. Tavares, “Effect of fuzzy partitioning in Crohn’s disease classification: A neuro-fuzzy-based approach,” *Med. Biol. Eng. Comput.*, vol. 55, no. 1, pp. 101–115, Jan. 2017.
- [22] E. Mossotto, J. J. Ashton, T. Coelho, R. M. Beattie, B. D. MacArthur, and S. Ennis, “Classification of paediatric inflammatory bowel disease using machine learning,” *Sci. Rep.*, vol. 7, no. 1, p. 2427, Dec. 2017.
- [23] L. Han, M. Maciejewski, C. Brockel, W. Gordon, S. B. Snapper, J. R. Korzenik, L. Afzelius, and R. B. Altman, “A probabilistic pathway score (PROPS) for classification with

- applications to inflammatory bowel disease,” *Bioinformatics*, vol. 34, no. 6, pp. 985–993, Mar. 2018.
- [24] K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen, C. Spampinato, D.-T. Dang-Nguyen, M. Lux, P. T. Schmidt, M. Riegler, and P. Halvorsen, “KVASIR: A multi-class image dataset for computer aided gastrointestinal disease detection,” in *Proc. 8th ACM Multimedia Syst. Conf.*, Jun. 2017, pp. 164–169.
- [25] A. Alammari, A. R. Islam, J. Oh, W. Tavanapong, J. Wong, and P. C. de Groen, “Classification of ulcerative colitis severity in colonoscopy videos using CNN,” in *Proc. 9th Int. Conf. Inf. Manage. Eng.*, 2017, pp. 139–144.
- [26] R. W. Stidham, W. Liu, S. Bishu, M. D. Rice, P. D. R. Higgins, J. Zhu, B. K. Nallamotheu, and A. K. Waljee, “Performance of a deep learning model vs human reviewers in grading endoscopic disease severity of patients with ulcerative colitis,” *JAMA Netw. Open*, vol. 2, no. 5, May 2019, Art. no. e193963.
- [27] T. Ozawa, S. Ishihara, M. Fujishiro, H. Saito, Y. Kumagai, S. Shichijo, K. Aoyama, and T. Tada, “Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis,” *Gastrointestinal Endoscopy*, vol. 89, no. 2, pp. 416–421, 2019.
- [28] Y. Maeda, S.-E. Kudo, Y. Mori, M. Misawa, N. Ogata, S. Sasanuma, K. Wakamura, M. Oda, K. Mori, and K. Ohtsuka, “Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video),” *Gastrointestinal Endoscopy*, vol. 89, no. 2, pp. 408–415, Feb. 2019.
- [29] M. Häfner, T. Tamaki, S. Tanaka, A. Uhl, G. Wimmer, and S. Yoshida, “Local fractal dimension based approaches for colonic polyp classification,” *Med. Image Anal.*, vol. 26, no. 1, pp. 92–107, Dec. 2015.
- [30] G. Wimmer, A. Uhl, and M. Hafner, “A novel filterbank especially designed for the classification of colonic polyps,” in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 2150–2155.
- [31] M. Häfner, M. Liedlgruber, A. Uhl, A. Vécsei, and F. Wrba, “Color treatment in endoscopic image classification using multi-scale local color vector patterns,” *Med. Image Anal.*, vol. 16, no. 1, pp. 75–86, Jan. 2012.
- [32] G. Wimmer, T. Tamaki, J. J. W. Tischendorf, M. Häfner, S. Yoshida, S. Tanaka, and A. Uhl, “Directional wavelet based features for colonic polyp classification,” *Med. Image Anal.*, vol. 31, pp. 16–36, Jul. 2016.
- [33] T. Tamaki, J. Yoshimuta, M. Kawakami, B. Raytchev, K. Kaneda, S. Yoshida, Y. Takemura, K. Onji, R. Miyaki, and S. Tanaka, “Computeraided colorectal tumor classification in NBI endoscopy using local features,” *Med. Image Anal.*, vol. 17, no. 1, pp. 78–100, Jan. 2013.
- [34] Y. Yuan and M. Q.-H. Meng, “A novel feature for polyp detection in wireless capsule endoscopy images,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 5010–5015.
- [35] T. Stehle, R. Auer, S. Gross, A. Behrens, J. Wulff, T. Aach, R. Winograd, C. Trautwein, and J. Tischendorf, “Classification of colon polyps in NBI endoscopy using vascularization features,” *Proc. SPIE*, vol. 7260, Feb. 2009, Art. no. 72602S.

- [36] E. Ribeiro, A. Uhl, G. Wimmer, and M. Häfner, “Exploring deep learning and transfer learning for colonic polyp classification,” *Comput. Math. Methods Med.*, vol. 2016, Oct. 2016, Art. no. 6584725.
- [37] Poudel, S., Kim, Y. J., Vo, D. M., & Lee, S. W. (2020). Colorectal Disease Classification Using Efficiently Scaled Dilation in Convolutional Neural Network. *IEEE Access*, 8, 99227-99238.
- [38] Y. Shin and I. Balasingham, “Comparison of hand-craft feature based SVM and CNN based deep learning framework for automatic polyp classification,” in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 3277–3280.
- [39] S. Nadeem, M. A. Tahir, S. S. A. Naqvi, and M. Zaid, “Ensemble of texture and deep learning features for finding abnormalities in the gastro-intestinal tract,” in *Proc. Int. Conf. Comput. Collective Intell. Cham, Switzerland: Springer*, 2018, pp. 469–478.
- [40] G. Urban, P. Tripathi, T. Alkayali, M. Mittal, F. Jalali, W. Karnes, and P. Baldi, “Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy,” *Gastroenterology*, vol. 155, no. 4, pp. 1069–1078, 2018.
- [41] G. Wimmer, A. Vécsei, M. Häfner, and A. Uhl, “Fisher encoding of convolutional neural network features for endoscopic image classification,” *J. Med. Imag.*, vol. 5, no. 3, 2018, Art. no. 034504.
- [42] F. Yu, V. Koltun, and T. Funkhouser, “Dilated residual networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 472–480.
- [43] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [44] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2921–2929.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [46] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 4278–4284.
- [47] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258. VOLUME 8, 2020 99237 S. Poudel et al.: Colorectal Disease Classification Using Efficiently Scaled Dilation in CNN
- [48] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
- [49] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, “Learning transferable architectures for scalable image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8697–8710. [47] [48]
- [50] K.Venkatachalam, A.Devipriya, J.Maniraj, M.Sivaram, A.Ambikapathy, Iraj S Amiri, “A

Novel Method of motor imagery classification using eeg signal”, Journal Artificial Intelligence in Medicine Elsevier, Volume 103, March 2020, 101787

- [51] Yasoda, K., Ponmagal, R.S., Bhuvaneshwari, K.S. K Venkatachalam, “ Automatic detection and classification of EEG artifacts using fuzzy kernel SVM and wavelet ICA (WICA)” Soft Computing Journal (2020).
- [52] P. Prabu, Ahmed Najat Ahmed, K. Venkatachalam, S. Nalini, R. Manikandan, Energy efficient data collection in sparse sensor networks using multiple Mobile Data Patrons, Computers & Electrical Engineering, Volume 87, 2020,
- [53] V.R. Balaji, Maheswaran S, M. Rajesh Babu, M. Kowsigan, Prabhu E., Venkatachalam K, Combining statistical models using modified spectral subtraction method for embedded system, Microprocessors and Microsystems, Volume 73, 2020.
- [54] Malar, A.C.J., Kowsigan, M., Krishnamoorthy, N. S. Karthick, E. Prabhu & K. Venkatachalam (2020). Multi constraints applied energy efficient routing technique based on ant colony optimization used for disaster resilient location detection in mobile ad-hoc network. Journal of Ambient Intelligence and Humanized Computing, 01767-9.