Journal of Human Centered Technology

https://www.humentech.utm.my E-ISSN: 2821-3467

Polyp Segmentation from Colonoscopy Image Using DeepLabV3+

Nur Anis Jasmin Sufri[1], Biragathis Babu[1], Muhammad Amir As'ari[1*,2]

[1]Department of Biomedical Engineering and Health Sciences, Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia
[2]Sport Innovation and Technology Centre (SITC), Universiti Teknologi Malaysia, Skudai, 81310, Malaysia

*Corresponding Author amir-asari@utm.my

Cite: https://doi.org/10.11113/humentech.v3n2.81 Research Article

Abstract:
Colorectal cancer is one of the most common cancers in the world, and it is also one of the leading causes of mortality from cancer. Over time, polyp development on the inner lining of the colon or rectum may progress to cancer. Due to the differences in size, shape, and color, this is a difficult undertaking polyp texture and differences between several forms of hard polyps mimics. Colonoscopy is the most effective way to detect polyps. Screening and detection are possible, but they are operator-dependent, time-consuming, and prone to errors. Polyp segmentation method based on the convolutional neural network is developed and to enhance the performance of the method. In this study, polyp segmentation algorithm was developed using Python as the main programming language which was applied with TensorFlow and OpenCV on Kvasir-SEG datasets. DeepLabv3+ model was used to train the filters and pooling operations were applied to images to segment the polyps. The data obtained was further analysed and evaluated the performance analysis. As a result, the developed algorithm was tested and shows an average accuracy 0.7215 and average IoU score 0.7240 for 20 epochs. As a conclusion, the algorithm developed is able to segment the polyps and it can be further developed with large number of data to improve the accuracy.
Keywords: Colonoscopy; Polyp; Segmentation; Deep Learning; DeepLabv3+

1. INTRODUCTION

Colon or rectal cancer develops when cells proliferate unchecked and take over the body. Colon cancer is another name for it. The large intestine and gut are other terms for the colon. The rectum connects the colon with the anus through a tube. Colonic polyps, or abnormal growths, may occur. It's possible that some polyps may become cancerous over time. Screening tests may detect polyps, which can then be surgically removed before they develop into cancer. When the condition is discovered in its earliest stages, treatment is more successful.

The incidence of colorectal cancer was quite low until a few decades ago. Around 10% of cancer-related deaths occur as a result of this malignancy in Western countries. An increasing number of people are developing colorectal cancer as a result of an ageing population, poor modern eating habits, and an increase in risk factors including smoking, lack of physical exercise, and obesity. Laparoscopic surgery for main disease, more aggressive excision of metastatic disease (such as liver and lung metastases), radiation for rectal cancer, and neoadjuvant and palliative chemotherapies have emerged as new therapeutic options for primary and metastatic colorectal cancer. As for cure rates and long-term survival, innovative therapy techniques have had just a minimal impact. Colorectal cancer is preceded by polypoid precursors, and screening programmes have gained momentum as a consequence of these features. This primer provides an overview of the current understanding of colorectal cancer's epidemiology and causes, as well as diagnosis and treatment. Colorectal cancer (CRC) is one of the most frequently diagnosed malignancies worldwide (1). CRC is the third most frequent cancer, accounting for 10.2 % of all cancer diagnoses in 2018, behind lung cancer 11.6 %, and breast cancer 11.6 % (2). It is the second most deadly malignancy in terms of mortality, accounting for 9.2% of all cancer fatalities. According to data (3), both males and females are almost equally impacted. Nonetheless, despite the high incidence and mortality rates, CRC-related fatalities have been declining at an increasing pace for both men and women since 1980 (4). This increase is mostly due to advancements in early diagnosis and treatment.

Colorectal cancer is one of the most common cancers in the world, and it is also one of the leading causes of mortality from cancer. Over time, polyp development on the inner lining of the colon or rectum may progress to cancer. Early identification and diagnosis of polyps are critical for the patient's life. Because of the differences in size, form, and texture of polyps, as well as differences between various kinds of hard mimics, detecting and segmenting polyps in colonoscopy is a difficult process. Colonoscopy is the most effective procedure for polyp identification and screening, although it is operator-dependent, time-consuming, and prone to errors. Polyps are sometimes overlooked or difficult to detect during a colonoscopy. There are several deep learning-based segmentation have been established in recent years which beneficial

Received 24 April 2024; Accepted 24 July 2024
Available online 09 August 2024

in producing an automated segmentation for medical imaging and helping in reducing human error. DeepLab v3+ is one of the common and prominent deep based segmentations. Therefore, this study is focused on formulating the DeepLabv3+ for segmenting polyp from colonoscopy image and evaluate the developed model with several segmentation performance metrics.

2. BACKGROUND STUDY

2.1 Colonoscopy Image

A colonoscopy is a diagnostic procedure used to find anomalies or alterations in the colon and rectum. Colonoscopy is a procedure where the colon's lining is examined using a flexible camera called an endoscope. Colonoscopy is a biopsy-based invasive technique. A long, flexible tube called a colonoscope is introduced into the rectum during a colonoscopy. The entire colon can be seen by the doctor thanks to a tiny video camera at the end of the tube. During the colonoscopy, polyps or other abnormal tissue may be removed using the endoscope if necessary. During the endoscope, tissue samples (biopsy samples) could also be taken. To avoid colorectal cancer and recurring adenomas, a high-quality colonoscopy is essential. The goal of recent technological developments has been to increase the adenoma segmentation rate (ADR). Image-enhanced endoscopy (IEE) facilitates the detection and characterization of polyps, especially nonpolypoid colorectal neoplasms. Due to its ability to deposit in areas of depression, which facilitates the identification of flat and depressed lesions, indigo carmine is the dye most frequently employed in colonoscopies. With the development of virtual chromoendoscopy, a practical contrast enhancement method has emerged that does not require the preparation or administration of dyes through the colonoscope operating channel. Using optical filters, narrow-band imaging (NBI) enhances the mucosa's surface and capillary pattern. Second-generation NBI yields a picture that is twice as bright as the first system's, which is optimistic for ADR results. Additionally, it has been demonstrated that LASEREO, a second-generation blue laser imaging system, increases polyp identification and ADR, making it a feasible IEE modality (5).

2.2 Convolution Neural Network based Semantic Segmentation

Deep learning is a branch of machine learning that focuses on the study of artificial neural networks with multiple hidden layers, along with associated machine learning techniques (6-7). To achieve accurate outcomes from deep learning algorithms, a substantial amount of data must be input into the system. Current deep learning models mostly utilize artificial neural networks, namely convolutional neural networks (CNNs). However, these models may also integrate propositional equations or latent variables organized in layers within deep generative models. Deep learning plays a significant role in various aspects of artificial intelligence, including image processing, language processing, and face detection. Several examples of deep models are Convolutional Neural Networks (CNN), Deep Belief Networks (DBNs), and encoder-decoder architectures (8). Neural networks with numerous concealed layers are commonly employed as deep learning models. Manual feature extraction is unnecessary when training deep learning models with large amounts of labeled data and neural network topologies (8).

Semantic Segmentation is a computer vision problem that entails the categorization of visually comparable components in an image that share the same class. Semantic segmentation is the process of assigning a specific category to each individual pixel in an image. This task is crucial for picture comprehension and the development of self-driving vehicles (9). The problem of pixel-wise labelling, which is addressed by semantic segmentation, is typically formulated using Convolutional Neural Networks (CNNs) (10). Semantic segmentation involves the assignment of a specific class to each individual pixel inside a picture. Several semantic segmentation techniques have been introduced.

For instance, the Fully Convolutional Network (FCN) is primarily designed for dense prediction and has demonstrated performance in a wide variety of visual tasks. FCN utilizes a convolutional neural network (CNN) to convert the pixels of an image into different categories. Fully Convolutional Networks (FCNs) differ from convolutional neural networks (CNNs) in that they use transposed convolution layers to restore the spatial dimensions (height and breadth) of intermediate layer feature maps to match those of the input image. This results in predictions that have a one-to-one correspondence with the input image in terms of spatial dimensions. The FCN is a type of network that does not include any "Dense" layers, which are commonly seen in regular CNNs. Instead, it utilizes 1x1 convolutions to fulfill the role of fully connected layers, also known as Dense layers. In order to achieve polyp segmentation instead of basic detection, deconvolution layers were incorporated into the pre-trained CNNs to generate a pixel classification map as the output (11).

Other than that, ParseNet is an efficient neural network for semantic segmentation that relies on fully convolutional networks. The incorporation of global context into the categorization of areas or objects in an image has demonstrated advantages in detection and segmentation tasks, which shares many characteristics with patch-based algorithms. In (12), research was conducted on semantic segmentation. Our approach involves incorporating global context into a fully convolutional network that operates end-to-end, rather than using a patch-based method. Instead of segmenting the image into regions or objects, the network simultaneously predicts the values of all pixels. During the process, the feature map of a layer is subjected to pooling across the entire image, resulting in a context vector.

U-Net is another semantic segmentation designed for biomedical image based on FCN architecture (13). The UNet model is simple with few parameters. It is suitable for the classification of medical images with small dataset. In UNet analysis, two steps of feature extraction, sub-sample and up-sample are needed. In the U-Net, an encoding network captures the visual context and a decoding network, which allows the localization of important parts, is built on top of it. UNet++ reduces the semantic gap between the encoder and decoder networks before fusion by inserting a sequence of layered skip connections into the U-Net. The ResUNet++ combines residual units with the spatial attention-based Atrous Spatial Pyramidal Pooling (ASPP) and channel attention-based squeeze-and-excitation block. An image contains objects of sizes ranging from small area to large area in different regions.

Another segmentation model is Pyramid Scene Parsing Network (PSP Net) where it has highest segmentation accuracy where study in Yan *et al*. (14) proved that compare with FCN and U-Net. Pyramid Scene Parsing Network utilises a pyramid parsing module that exploits global context information by different region-based context aggregation. The pyramid pooling module fuses features under four different pyramid scales. The coarsest level highlighted in red is global pooling to generate a single bin output. The following pyramid level separates the feature map into different sub-regions and forms pooled representation for different locations (15). PSPNet solves the common problems in FCN and performs well in image segmentation trained data. Some research articles show that, PSP-Net used to design an automatic segmentation of brain tumor from three-dimensional MR datasets (16).

The initial implementation of the Deeplab architecture was constructed using Deep Convolutional Neural Networks (DCNNs) and fully connected Conditional Random Fields (CRFs) approach (17). CRFs are commonly used in segmentation maps with noise to reduce noise and highlight boundaries. The application of CRFs in DeepLabv1 is based on the same rationale. Repeatedly applying this strategy to the output of the DCNN results in a smoother segmentation map with sharper boundaries between different regions of the picture. A precise segmentation map can be obtained in this DeepLabv1 model with a minimal number of repetitions. Following that, DeepLabv2, which is the successor to DeepLab, introduced a post-processing technique that incorporated the existing DeepLabv1 architecture (a combination of DCNN and CRFs) to enhance the smoothness of the edges while keeping them unaltered. The implementation of Atrous Spatial Pyramid Pooling (ASSP) has been incorporated into the convolutional layers. The ASPP method is employed to segment objects at various scales using filters with multiple sample rates and effective field-of-views. Because CRFs are also present in DeepLabv2, the architecture is still unable to achieve end-to-end learning. The problem was addressed by the introduction of DeepLabv3 (18). DeepLabv3 was introduced as an upgrade to the DeepLab framework, aiming to further improve the previous methodologies. However, DeepLabv3 has eliminated the final CRF step from the network to enable end-to-end learning, which is different from DeepLabv2. Although the smoothing layer was removed, the network's performance surpassed that of previous designs. A network architecture known as an encoder-decoder is employed to enhance the accuracy of segmentation predictions. DeepLabv3+ was introduced by enhancing the encoding component using Separable Atrous Convolution instead of traditional atrous convolutions. Upsampling techniques are employed to increase the resolution of the encoder's output in the decoder's structure. This study focuses on formulating DeepLabv3+ for the purpose of segmenting polyps from microscopic images.

3. METHODOLOGY

3.1 Research Methodology Flow

Figure 1 shows the workflow along the project. The project begins with data collection from existing datasets of images of polyp from the datasets. Then, the images were labelled and converted into .jpg format. Next, the CNN models were developed in Google Collab and their performance was evaluated accordingly. In this project, the aim is to develop automated polyp segmentation system and to model the deep CNN in recognizing polyps from colonoscopy image. Additionally, the performance of the developed model in segmentation polyp will be evaluated after the DeepLabv3+ is developed. For the software implementation part, it is based on Python in which programming language.

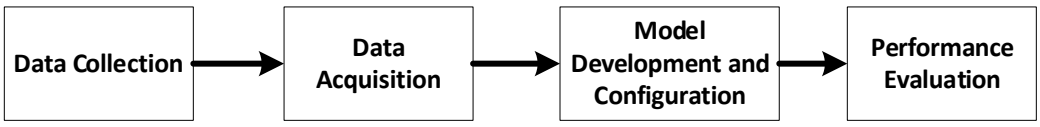

Figure 1. Project's workflow

3.2 Data Collection

Several earlier datasets from Kaggle and Kvasir-SEG datasets were used. The Kvasir dataset contains 8000 photos of the gastrointestinal (GI) tract, divided into 1000 classes (19). 1000 polyp pictures and the accompanying ground truth from the Kvasir Dataset v2 are included in the 46.2 MB Kvasir-SEG dataset. The pixels in Kvasir-SEG range from 332×487 to 1920×1072. Two other folders with the same filename contain the photographs and their matching masks. The dataset is categorized into eight categories, including pathological results, endoscopic operations, and anatomical landmarks. Each class's images are stored in a different folder with the name of the class to which they correspond. The Kvasir-SEG dataset (size 46.2 MB) contains 1000 polyp images along with the ground truth associated with them from the Kvasir Dataset v2. The resolution of the images in Kvasir-SEG varies from 332×487 to 1920×1072 pixels. Two separate files with identical names are saved containing the images and masks. Picture files are compressed using JPEG compression, which facilitates web browsing.

3.3 Model Development and Configuration

In this project, the DeepLabv3+ with pretrained ResNet50 was formulated by training on Kvasir-Seg datasets. Moreover, in this project, import pickle module where pickle dump method to save the train, test and validation into file <filename>: this will save the object in this file in byte format. Lists, dictionaries, tuples, and other Python objects may be serialised and deserialized with the help of Python's pickle modules.The function tf.cast() used to cast a specified Tensor to new data type. The model was trained with TensorFlow 1.10.0 framework and Keras as a backend. It supports CPU and GPU computation for model training and testing. Tensorflow open-source library developed by Google for running machine learning models and deep learning neural networks in the browser.The optimized Deeplabv3+ network is accurate for

Overall, the development of the models training, testing was done on Google Colab using NVIDIA Tesla T4-16GB GPU. Pre-trained weights from the Segmentation Models package were adapted to the baseline trials. The encoder for each model was SE-ResNeXt-50-32x4D with weights taken from ImageNet. All the images and masks were reduced in size using bilinear interpolation before this were restored back to their original size. Passing the result through a two-dimensional softmax function generated the final forecast. The Adam optimizer used a starting learning rate of 0.0001 for all models, and after 20 epochs, the learning rate was lowered to 0.00001. Table 1 summarizes the hyperparameter setup performed in developing the proposed segmentation model.

From the table 1, the training parameters have the base learning rates which are values for the training process may fail if the learning rate is set too low, while learning a sub-optimal set of weights too quickly or an unstable training process may ensue if the rate is set too high. When setting up a neural network, the learning rate is a crucial hyperparameter to optimise. For this learning the base learning rate was set 0.0001. Overshooting is possible if we set the learning rate to a value near the upper limit of this range. It happens when we make a huge leap in the direction of the minimised loss function and end up missing by a little margin.

Table 1. Hyperparameter.

Training Parameters	Values
Base Learning Rate	0.0001
Training number of steps	16
Learning power	0.9
Batch size	20
Atrous rates	1,6,12,18
No of epochs trained	20
Number of filters	256
Kernal size	3

3.4 Performance Evaluation

The performance of the architectures relied on different metrics perhaps, most commonly metrics are Intersection over Union (IoU) and Dice coefficient. For the image segmentation, every pixel in the picture corresponds to either a polyp or a non-polyp area. Based on this idea, the Dice coefficient and mean IoU were calculated using Equation (1) and Equation (2) respectively. The symbol 't' is the threshold where it was calculated by predicting object to all ground truth objects. For further research parameter will be refer from Kvasir datasets (19). The value of IoU ranged between 0.0000 and 1.0000.

$$Dice\,Coeffiecient(A,B) = \frac{2\times|A\cap B|}{|A|+|B|} = \frac{2\times TP}{2\times TP+FP+F} \quad (1)$$

$$IoU(A,B) = \frac{A\cap B}{A\cup B} = \frac{TP(t)}{TP(t)+FP(t)+FN(t)} \quad (2)$$

The accuracy of object segmentation is conceptualized using measurements of the Jaccard Index. When a computer vision system is tasked with identifying faces in an image, the Jaccard index can be used to gauge how well the computer's face recognition matches the training set. A statistic for figuring out how similar two sample sets are called the Jaccard Index, or Jaccard Similarity Coefficient. Formally defined as the size of the intersection divided by the size of the union of the sample sets, the measurement highlights similarity between finite sample sets. Equation (3) is the mathematical expression for the index.

$$Jaccard\,Coefficient(A,B) = \frac{2\times|A\cap B|}{|A|+|B|} \quad (3)$$

4. RESULTS AND DISCUSSION

Using the test dataset, the suggested DeepLabV3+ model produced an overall mean IoU of 0.7216 and average dice loss of 0.1686 (see Table 2). Figure 2 displays examples of qualitative result comparisons between the DeepLabV3+ model and the algorithm on the Kvasir-SEG dataset. Based on the quantitative and qualitative findings (refer to Table 2 and Figure 2, 3, 4), the research demonstrates that the DeepLabV3+ model outperforms the algorithm in polyp pixel segmentation. As a result of the algorithm's lack of a learning mechanism or learning parameters, the DeepLabV3+ fully utilizes data augmentation techniques. Because clustering relied on color features as a significant due to their comparable appearance, it was unsuccessful. In this project, the DeepLabv3+ model functioned admirably.

Table 2. Comparison of training, validation and test performance metric for proposed DeepLab v3+.

DeepLabV3+ Model	Average Dice Loss	Mean IoU
Train	0.0905	0.8388
Validation	0.1952	0.724
Test	0.1686	0.7216

Table 3. Performance of proposed deeplabv3+ based on different performance metrics.

Model	Jaccard Coefficient	Dice Coefficient Score	Test Accuracy
DeepLabV3+	0.8039	0.45244	0.7215

Table 3 presents the quantitative results of proposed DeepLabv3 in several different performance metrics. It is suggested that the proposed model achieves excellent result in segmenting polyp this model achieves Jaccard Coefficient of 0.8039. The low values of Jaccard coefficient for the all the layers shows 0.2282 with the dice coefficient score 0.5913. The test accuracy for this model is 0.7215. The test accuracy can be higher than train accuracy due to the difference in distribution. Figures 2, 3, and 4 represent the example of tested sample images, the actual mask and the predicted mask produced by the developed DeepLabv3+ model. From Figure 4, it can be interpreted that the proposed approach manages to detect almost most of the pixel belonging to polyp when compared to actual mask in Figure 3.

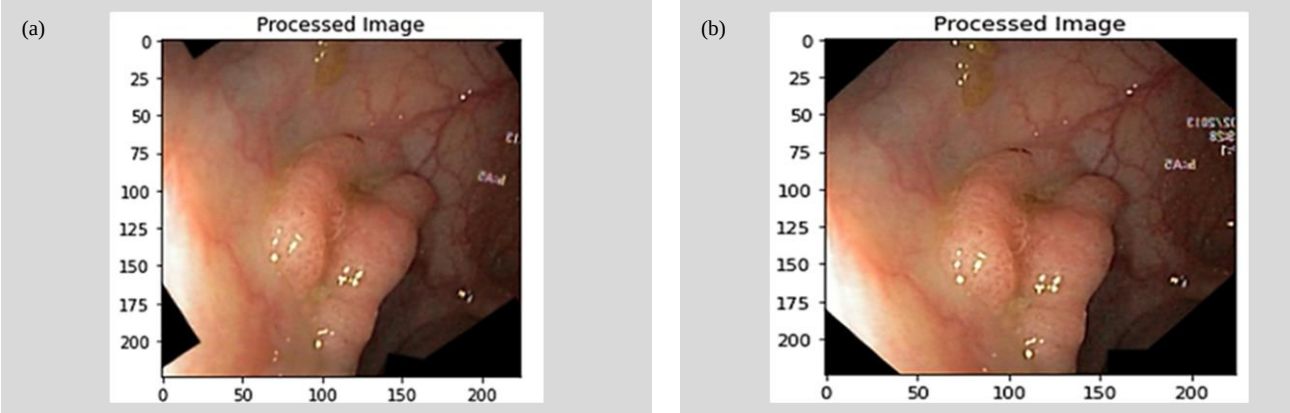

Figure 2. Results of processed image from two sample images (a) and (b).

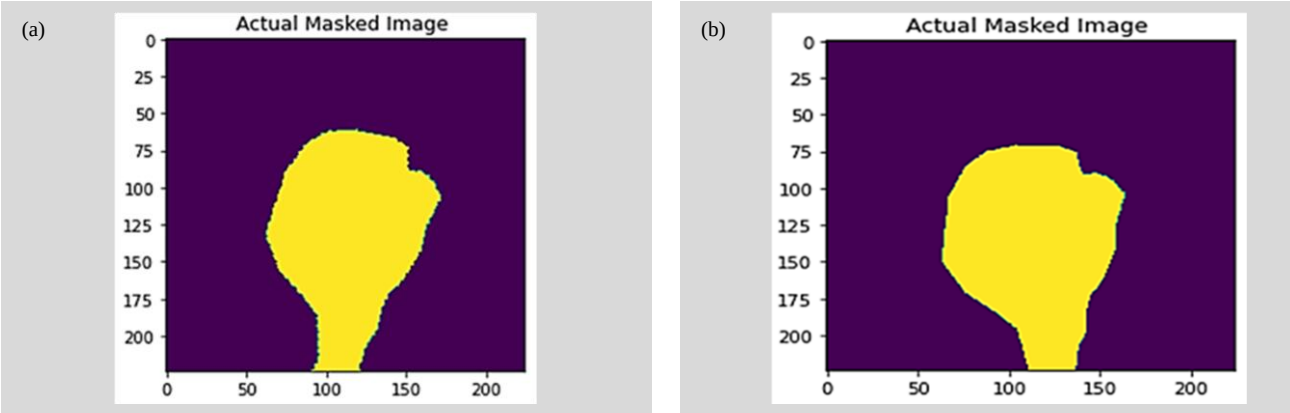

Figure 3. Actual masked images which yellow shows the polyps (segment) and purple unwanted parts (background).

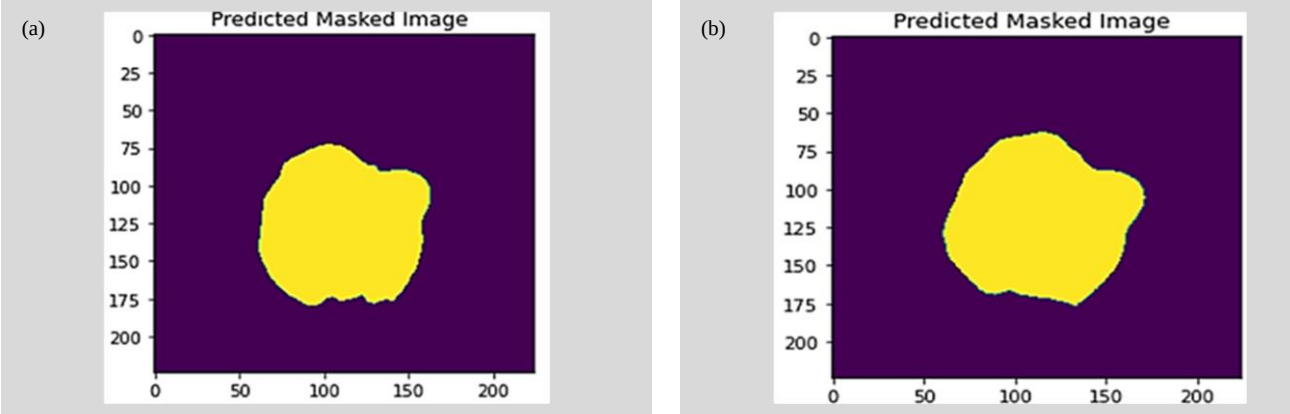

Figure 4. Results of predicated masked images which yellow shows the polyps (segment) and purple unwanted parts (background).

5. CONCLUSION

This technique differs from earlier ones in that it does not depend on just one or a few physical qualities; rather, it makes full use of all available picture features, including colour, texture, shape, and temporal information. In practically all operating points, our analyses based on the biggest annotated polyp database show better performance to the state of the art, considerably lowering polyp segmentation latency and the frequency of false positives. Our novel 3-way visual display and excellent utilisation of convolutional neural networks are responsible for such a significant performance boost. A system for detecting polyps based on a model of polyp appearance in the context of colonoscopy picture processing. This model is based on the appearance of valleys around polyps caused by the colonoscope's light and the camera being pointed in the same direction, resulting in the appearance of shadows around conspicuous surfaces such as polyps. Additionally, we gave the depth of valleys picture, which combines ridges and valleys detector information with the morphological gradient. This approximation seems to be appropriate for the photographs we're working with, since it emphasises both the valleys and edges that define the borders of the buildings shown on the images. DeepLabv3+ has an overall accuracy of 72.2% in its testing results. To sum it all up, the created model is dependable and has fulfilled its goal of segmenting polyps using deep CNN modelling, as well as evaluating the performance of other CNN models. There are, however, a few flaws in this study that might be addressed in future research.

ACKNOWLEDGMENT

The authors would like to express their appreciation to Universiti Teknologi Malaysia (UTM) for endowing this research and the Ministry of Higher Education (MOHE) Fundamental Research Grant Scheme (FRGS/1/2023/ICT02/UTM/02/1) for funding this research.

CONFLICT OF INTEREST

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

REFERENCES

(1) Eddy DM. Screening for colorectal cancer. Ann Intern Med. 1990; 113(5):373–384. https://doi.org/10.7326/0003-4819-113-5-373.
(2) Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA Cancer J Clin. 2009; 59(6):366–378. https://doi.org/10.3322/caac.20038.
(3) Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RG, Barzi A, Jemal A. (2017). Colorectal cancer statistics. CA Cancer J Clin. 2017; 67(3):177–193. https://doi.org/10.3322/caac.21395.
(4) Siegel RL, Miller KD, Goding SA, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA, Jemal, A. Colorectal cancer statistics. CA Cancer J Clin. 2020; 70(3):145–164. https://doi.org/10.3322/caac.21601.
(5) Rex DK, Schoenfeld PS, Cohen J, Pike IM, Adler DG, Fennerty MB, Lieb JG, Park WG, Rizk MK, Sawhney MS. Quality indicators for colonoscopy. Gastrointest Endosc. 2015; 81(1):31–53. http://doi.org/10.1038/ajg.2014.385.
(6) Deng L, Yu D. Deep learning: methods and applications. Found Trends Signal Process. 2014; 7(3–4):197-387.
(7) Ongsulee P. Artificial intelligence, machine learning and deep learning. 15th Int Conf ICT Knowl Eng (ICT&KE), Bangkok, Thailand. 2017; 1–6. https://doi.org/10.1109/ICTKE.2017.8259629.
(8) Ravì D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, Yang GZ. Deep learning for health informatics. IEEE J Biomed Health Inform. 2016; 21(1):4–21. http://doi.org/10.1109/JBHI.2016.2636665.
(9) Wang P, Chen P, Yuan Y, Liu D, Huang Z, Hou X, Cottrell G. Understanding convolution for semantic segmentation. 2018 IEEE Winter Conf Appl Comput Vis (WACV). 2018. https://doi.org/10.48550/arXiv.1702.08502.
(10) Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Semantic image segmentation with deep convolutional nets and fully connected CRFs. arXiV. 2014. https://doi.org/10.48550/arXiv.1412.7062.

(11) Brandao P, Zisimopoulos O, Mazomenos E, Ciuti G, Bernal J, Visentini-Scarzanella M, Menciassi A, Dario P, Koulaouzidis A, Arezzo A. Towards a computed-aided diagnosis system in colonoscopy: Automatic polyp segmentation using convolution neural networks. J Med Robot Res. 2018; 3(02):1840002. https://doi.org/10.1142/S2424905X18400020.
(12) Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, & Rabinovich A. Going deeper with convolutions. Proc IEEE Conf Comput Vis Pattern Recognit. 2015. http://doi.org/10.1109/CVPR.2015.7298594.
(13) Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Int Conf Med Image Comput Comput Assist Interv. 2015. http://doi.org/10.1007/978-3-319-24574-4_28.
(14) Yan L, Liu D, Xiang Q, Luo Y, Wang T, Wu D, Chen H, Zhang Y, Li Q. PSP net-based automatic segmentation network model for prostate magnetic resonance imaging. Comput Methods Programs Biomed. 2021; 207:106211. https://doi.org/10.1016/j.cmpb.2021.106211.
(15) Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. 2017; 6230–6239. http://doi.org/10.1109/CVPR.2017.660.
(16) Deng L, Yu D. Deep learning: methods and applications. Found Trends Signal Process. 2014; 7(3–4):197–387. http://doi.org/10.1561/2000000039.
(17) Chen LC, Papandreou G, Kokkinos I, Murphy K, & Yuille AL. Semantic image segmentation with deep convolutional nets and fully connected CRFs. arXiv Preprint arXiv:1412.7062. 2014; 40(4):834–848. https://doi.org/10.48550/arXiv.1412.7062.
(18) Tsang S-H. Review: Deeplabv3—atrous convolution (semantic segmentation) [Online]. Medium. 2019 [cited 2024 Jul 26]. Available from: https://medium.com/@sh.tsang/review-deeplabv3-atrous-separable-convolution-semantic-segmentation-a625f6e83b90.
(19) Jha D, Ali S, Tomar NK, Johansen HD, Johansen D, Rittscher J, Riegler MA, Halvorsen P. Real-time polyp detection, localization and segmentation in colonoscopy using deep learning. IEEE Access. 2021; 9:40496–40510. http://doi.org/ 10.1109/ACCESS.2021.3063716.