JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-18, No.-01, January (2023) pp 21-29 ISSN (Print) 0973-8975

UNET MOBILENETV2: MEDICAL IMAGE SEGMENTATION USING DEEP NEURAL NETWORK (DNN)

Bikash Chandra Bag[1], Hirak Kumar Maity[2], Chaitali Koley[3]

[1,2] Department of Electronics and Communication Engineering, College of Engineering and Management, Kolaghat, West Bengal, India

[3]Department of Electronics and Communication Engineering, National Institute of Technology Mizoram, Chaltlang, Aizawl-796012, Mizoram, India

Email: mebikash1986@cemk.ac.in, hmaity@cemk.ac.in, chaitali.ece@nitmz.ac.in

https://doi.org/10.26782/jmcms.2023.01.00002

(Received: September 26, 2022; Accepted: January 6, 2023)

Abstract

In this paper, the framework of polyp image segmentation is developed using a Deep neural network (DNN). Here Unet Mobile NetV2 is considered to evaluate the performance of the image from the CVC-612 dataset for the segmentation method. The proposed model outperformed earlier results. To compare our results two parameters, normally Dice co-efficient and Intersection over Union (IoU) are considered. The proposed model may be used for accurate computer-aided polyp detection and segmentation during colonoscopy examinations to find out abnormal tissue and thereby decrease the chances of polyps growing into cancer. MobileNetV2 significantly outperforms U-Net and MobileNetV2, two key state-of-the-art deep learning architectures, by achieving high evaluation scores with a dice coefficient of 89.71%, and an IoU of 81.64%.

Keywords: Deep Neural network, Semantic segmentation, UNet MobileNetV2.

I. Introduction

The first application of the DNN is based on Image segmentation in healthcare applications. According to PAHO (Pan American Health Organization), colorectal cancer has a significant rate of deaths related to this disease worldwide. The presence of this type of cancer is assessed when there are polyps (a projection of tissue growth from the wall of an empty space, such as the intestine, which can be benign or malignant and can even develop into cancer) therefore, it is extremely important that these polyps are discovered in the early stages by doctors and this investigation occurs through the colonoscopy exam in Lab. According to [IX] many polyps are lost or not observed during the exams, as it depends a lot on the experience of the doctor who performs it, so a method that is able to detect and segment these polyps will automatically help in the diagnosis thus decreasing the chance that some polyps will be forgotten and, consequently, wrong diagnoses will be made.

Colonoscopy is an invasive exam that captures real-time images of the large intestine and part of the terminal ileum (the final portion of the small intestine). To perform the experiment, a device called a colonoscopy is used, which has a thin and flexible tube with a camera at its end capable of filming the inside of the intestine in order to investigate the presence of colorectal cancer, polyps, and inflammatory bowel diseases.

The main objective of this work is to perform the segmentation of polyps present in frames extracted from colonoscopy videos, which have several examples of these anomalies. To solve this problem, the U-Net [XIV] architecture is used with the MobileNetV2 [III] network, with individual tests being carried out for each network. The dataset CVC-612 [x] is used for this proposal.

II. Related Works

The segmentation of an image is the process of classifying each pixel as belonging to a previously defined label. In the last few decades, there has been a lot of research related to the segmentation of different images with different content. The first methods implemented for the segmentation of polyps such as [IV] and [XIII], trained classifiers to distinguish one polyp from the rest of the image, however, these models have a high error rate. Currently, the vast majority of models used to perform segmentation are convolutional neural networks or FCNs (Fully Convolutional Networks) with a pre-trained model to identify and segment polyps [I,X] and this is explained by the significant improvement obtained when these models are implemented. A more current approach presented in the state of the art is the one presented in [VI], in which the authors used a deep neural network for segmentation, valuing the area and limit of the polyps to identify these anomalies, since they have different sizes and formats, analyzing these two characteristics helps a lot in the hit rate.

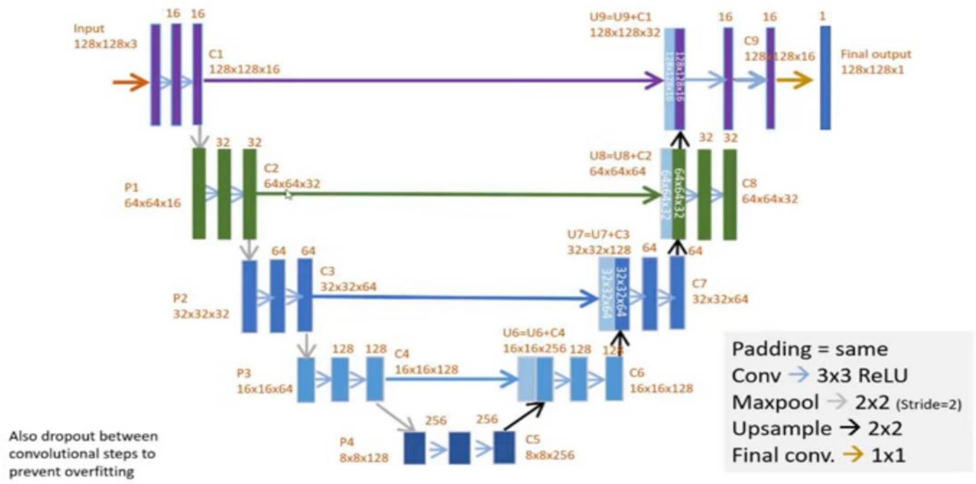


Fig:1. UNet Model

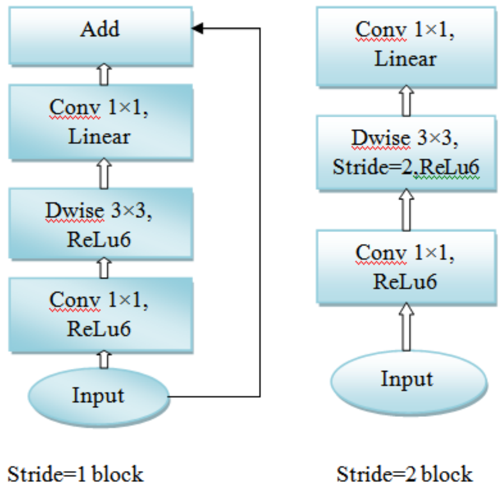


Fig.2. MobileNetV2 Model

With the success of U-Net, proposed by [XIV], when applied to the segmentation of medical images, [V] applied this architecture with some modifications and focused on the segmentation of the entire area polyp and obtained promising results, however, they ignored the restriction of the area limit or size of the anomaly, which may end up depreciating the model's performance in segmentation.

From the original U-Net architecture that consists of an encoder-decoder, it is known that the encoder extracts characteristics (or resources) from the input images and these resources are concatenated with the decoder so that the segmentation task is performed by the network. Using pre-trained networks in large databases in these cases helps a lot when performing segmentation since your weights will already be trained with millions of learned resources and the weights will not start from scratch during training.

III. APPROACH

III.i. Dataset

The dataset used was the CVC-612 [VII] is used, which has 612 images of polyps and their corresponding masks segmented by experienced doctors. The training conducted during this work research, the data will be separated into training, validation, and test sets with a proportion of 80%, 10%, and 10%, respectively.

III.ii. Pre-processing

For the images of the data set to be used, two pre-processing steps were applied. The first was the normalization of the pixels of the images in which these pixels that are in the range of 0 to 255 are reset to a range of 0 to 1, to make the operation faster, since the data was in its raw format. The second was the implementation of data augmentation techniques such as Center Crop, Random Rotate, Horizontal Flip, Vertical Flip and Grid Distortion.

III.iii. U-Net-MobileNetV2

Fig. 3 shows an overview of the proposed U-Net MobileNetV2 network where a pre-trained MobileNetV2 with the weights of the ImageNet database [XII] is applied as the encoder. The proposed model receives the input images with a size of 320x320, which are inserted in the pre-trained encoder, that is based on inverted residual blocks (or structures). Each block includes a combination of spatial convolutions with 3x3 kernels, ReLu activation, and layers of Batch Normalization. In this case, the encoder will use compact depth-to-depth convolution to filter and learn the characteristics of images that feed the network. The use of these inverted residual blocks helps to reduce the number of parameters, making the model easier and faster to train. Another advantage is that the model will perform better and converge faster than if a pre-trained network was not being used. In the decoding path, up-sampling operations are used to increase the size of the feature map back to its original size. During this path, the characteristics are concatenated between the encoder and decoder blocks and also go through a 3x3 convolution layer, followed by Batch Normalization and ReLu activation. Finally, the last block of the network is a block with a 1x1 convolutional layer with a Sigmoid activation, so that the segmentation performed by the network can be generated.

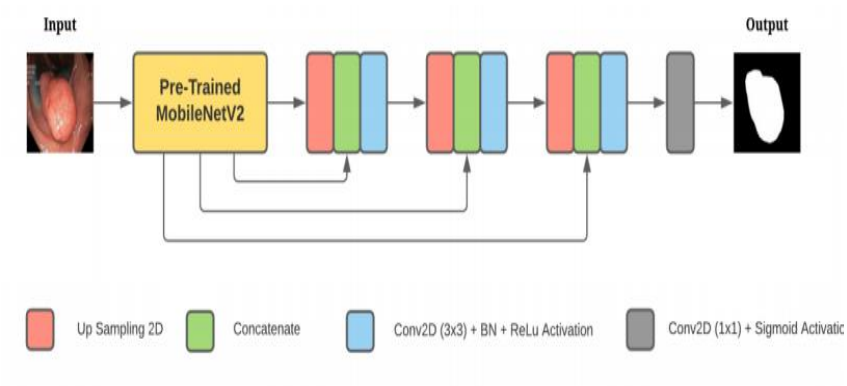
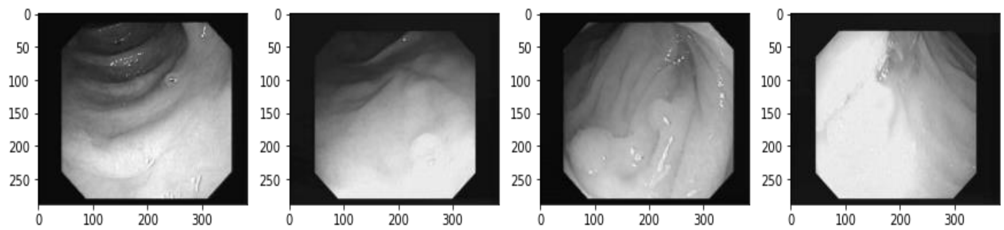


Fig.3. U-Net-Mobile-NetV2 Architecture

Input Images



Output Images

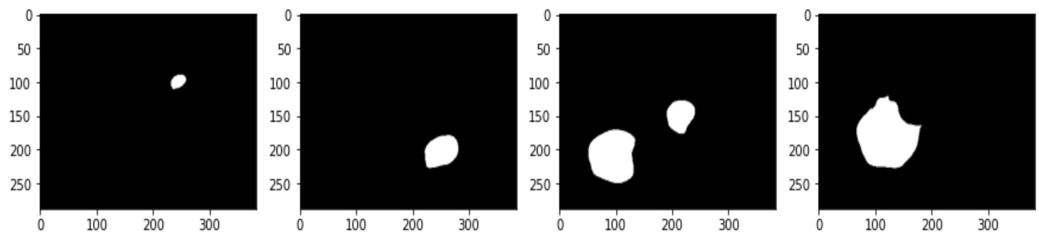


Fig:4. Input and Output Images

III.iv. Implementation details

The proposed model was implemented with the Keras [I] framework and Tensor Flow [II] as a backend. The data set was divided into subsets of training, validation, and testing with a proportion of 80%, 10%, and 10%, respectively. The training was carried out with a learning rate of 0.0001, and a batch size of 16 since a lower or very small value could lead to overfitting, the chosen optimizer was Adadelta and the model was trained for 100 epochs and generally converged with 76 epochs. Early-stoping regularization was also implemented in the validation subset, specifically in the validation loss as another way to avoid overfitting. The metrics chosen to assess the model's performance were the Dice Coefficient(DSC), and the Intersection over Union(IoU). As a loss function, Dice Loss was implemented, as it generates better results in segmentation tasks, even though some cases in the state of the art still carry out experiments with loss functions such as binary cross entropy.

Dice Coefficient (DSC) is the area of overlap divided by the total no of pixels in both images.

$$\text{DSC}= \frac{|X \cap Y|}{|X|+|Y|} \tag{1}$$

$$\text{IOU}=\frac{Area\ overlap}{Area\ of\ Union}=\frac{X \cap Y}{X \cup Y} \tag{2}$$

Overlap between prediction and ground truth /Union between prediction and ground truth. where

X indicates the input image, and Y indicates image masks.

IV. Results

To evaluate the performance of the proposed model all, the experiment is conducted with the CVC-612 dataset. The obtained results with the proposed model are compared with the other three state-of-the-art models that were also proposed for this polyp segmentation task. The values of the Dice Co-efficient and the IoU are considered for evaluation and comparison between the proposed model with state of art works. Table 1 presents the results of the models U-Net-MobileNetV2, ResUNet [x], ResUNet++ [VIII], and the model proposed in [XV] based on CVC-612.

Table 1: Quantitative results of all models

Method	Dice	IoU
Proposed Model	0.8971	0.8164
Tomar et. All[15]	0.8411	0.7565
ResUNet++[14]	0.8133	0.7927
ResUNet[10]	0.7877	0.7777

Table 1 shows that the U-Net-MobileNetV2 model obtained the highest Dice Coefficient and also the highest IoU among the models analyzed for the CVC-612 dataset. Based on the models proposed in [VII], [VIII], and [XV] with the CVC-612 dataset, U-Net-Mobile NetV2 achieved better results with percentage averages of 13.888%,10.303% and 6.657% using Dice Coefficient and of 4.976%, 2.989% and 7.918% using IoU, respectively.

The proposed model surpassed the baseline architecture[x] in terms of all metrics, mainly in terms of the Dice Coefficient, which was surpassed by a large margin. This may be said that this margin of difference between the U-NetMobileNetV2 and the other state-of-the-art models indicates that the use of a pre-trained network as a U-Net encoder can optimize the performance of the model in the segmentation task and can reduce computational cost and training time. Fig 5 presents some samples of the segmentations obtained by the proposed model, which we refer to as qualitative results. In Fig 2 I have samples of original images, and their corresponding masks beside, the mask predicted by U-Net-MobileNetV2 are shown in the diagram.

V. Discussions

The proposed U-Net-MobileNetV2 obtained satisfactory results in the CVC-612 dataset (see Table 1). From Fig 3, it is clear that the segmentation generated by the proposed model is better than those generated by the other state-of-the-art models used for comparison, as it was able to better identify the shape and location of the polyps in the images. In general, our proposed method generated a segmentation mask that was closer to the ground truth provided b the dataset than the other models.

The proposed model learns the information from depth-to-depth convolutions and filters this information to each block and stage of the network, allowing the creation of more refined resource maps, where the most relevant resources of the images are extracted, which ended up helping to increase the efficiency of the segmentation.

U-Net-MobileNetV2 can and should also be applied to segmentation tasks of other types of images and not only to medical images since it will probably also perform well even if for other images the pixel classification needs more detailed validations, as well as apply to different class targets and not just binary pixel classification.

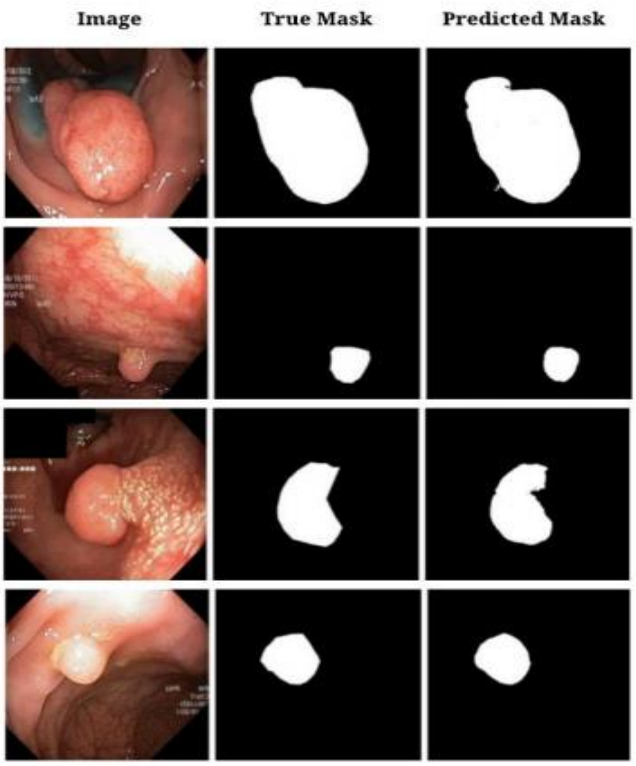


Fig. 5. Quantitative results of the proposed model

VI. Conclusions

In our discussion, the U-Net-MobileNetV2 model is proposed for the automatic segmentation of polyps in colonoscopy exam images. Our model and the successive test results that were carried out demonstrate that the proposed approach outperformed the models present in the state-of-the-art that were cited and used for comparison. Very representative and competitive values were achieved for the segmentation task with the CVC-612 dataset as we reached a Dice Coefficient of 0.8971 and an IoU of 0.8164. A great advantage of this model that help a lot in training was the implementation of inverted residual blocks, which make training faster and more efficient, as well as making the model converge more quickly.

Conflict of Interest:

The authors declare that no conflict of interest to report the present study.

References

I. A. Arezzo, A. Koulaouzidis, A. Menciassi, D. Stoyanov, E. B. Mazomenos, F. Bianchi, G. Ciuti, P. Brandao, P. Dario, R. Caliò, Fully convolutional neural networks for polyp segmentation in colonoscopy. In: International Society for Optics and Photonics. Medical Imaging 2017: Computer-Aided Diagnosis, v. 10134, p. 101340F, 2017.

II. A. Howard, M. Zhu, A. Zhmoginov, L-C. Chen, M. Sandler, MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, p. 4510-4520, 2018.

III. A. Howard, M. Zhu, A. Zhmoginov, L-C. Chen, M. Sandler, MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, p. 4510-4520, 2018.

IV. A. V. Mamonov, I. N. Figueiredo, P. N. Figueiredo Y-H. R. Tsai, Automated polyp detection in colon capsule endoscopy. IEEE transactions on medical imaging, v. 33, n. 7, p. 1488–1502, 2014.

V. B. Paul, S. A. Fattah, T. Mahmud, Polypsegnet: A modified encoderdecoder architecture for automated polyp segmentation from colonoscopy images. Computers in Biology and Medicine, p. 104119, 2020.

VI. D-P. Fan, G. Chen, G-P. Ji, H. Fu, J. Shen, L. Pranet Shao, T. Zhou, : Parallel reverse attention network for polyp segmentation. In: SPRINGER.International Conference on Medical Image Computing and ComputerAssisted Intervention, p. 263–273, 2020.

VII. D. Jha, D. Johansen, H. D. CVC-612, Johansen, M. A. Riegler, P. Halvorsen, P. H. Smedsrud, T. Lange : A Segmented Polyp Dataset. In Proc. of International Conference on Multimedia Modeling (MMM), p. 451-462, 2019.

VIII. D. Jha, H. D. Johansen, M. A. Riegler, P. Halvorsen, P. H. Smedsrud, T. Lange, ResUNet++: An Advanced Architecture for Medical Image Segmentation. 2019 IEEE International Symposium on Multimedia (ISM),2019.

IX. E. Dekker, J. C. V. Rijn, J. B. Reitsma, J. Stoker, P. M. Bossuyt, S. J. V. Deventer, Polyp miss rate determined by tandem colonoscopy: a systematic review. The American journal of gastroenterology, v. 101, p. 343, 2006.

X. E. Nasr-Esfahani, K. Najarian, M. Akbari, M. Mohrekesh, N. Karimi, S. M. R. Soroushmehr, S. Samavi, Polyp segmentation in colonoscopy images using fully convolutional network. In: IEEE. 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), p. 69–72, 2018.

XI. F. Chollet, Keras. https://github.com/fchollet/keras, 2015.

XII. J. Deng, K. Li, L. Fei-Fei, L-J. Li, R. Socher, W. Dong, Imagenet: A large-scale hierarchical image database. In: IEEE conference on computer vision and pattern recognition, p. 248–255, 2009.

XIII. J. Liang, N. Tajbakhsh, S. R. Gurudu, Automated polyp detection in colonoscopy videos using shape and context information. IEEE transactions on medical imaging, v. 35, n. 2, p. 630–644, 2015.

XIV. O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks forbiomedical image segmentation. In: SPRINGER. International Conference on Medical image computing and computer-assisted intervention, p. 234–241, 2015.

XV. N. K. Tomar, Automatic Polyp Segmentation using Fully Convolutional Neural Network, 2021.