

Amplification of pixels in medical image data for segmentation via deep learning object-oriented approach

Ahmad Firdaus Ahmad Fadzil^{1*}, Noor Elaiza Abd Khalid² and Shafaf Ibrahim¹

Universiti Teknologi MARA Cawangan Melaka, Kampus Jasin, Malaysia¹

Universiti Teknologi MARA Shah Alam, Malaysia²

Received: 21-September-2020; Revised: 17-January-2021; Accepted: 19-January-2021

©2021 Ahmad Firdaus Ahmad Fadzil et al. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Medical images serve as a very important tool for medical diagnosis. Medical image segmentation is an area of image processing that segments critical parts of a medical image for diagnosis purposes. The emergence of machine learning approach for medical image segmentation specifically by employing Convolutional Neural Network (CNN) has become a ubiquity as other approaches does not able to compete with its robustness and accuracy. However, this approach is very exhaustive in terms of time and computing resources. The CNN approach mostly emphasizes on the spatial information regarding the image without using much of the individual data contained within the image. Therefore, this paper proposed a method to amplify the pixel data of medical images via Object-oriented Programming (OOP) approach for segmentation using a straightforward sequential deep learning approach. The results indicated that the proposed method allows more than 90 % faster training time with 33.8 seconds average and overall better segmentation performance of 0.744 for balanced-accuracy metric compared to recent state-of-the-art CNN segmentation models such as SegNet and U-Net Models.

Keywords

Pixels, Amplify, Object-oriented programming, Image segmentation, Deep learning.

1.Introduction

Medical image is vital for medical diagnosis. Examples of medical image data includes brain MRI, retinal fundus, and cells histopathology images [1, 2]. One of the most important processes for medical images is segmentation, where part of the medical image is segmented to obtain the lesion required for further analysis [3]. Recent trends in medical image segmentation shows the emergence of machine learning algorithm such as Convolutional Neural Network (CNN) that returns very competent segmentation performance [4]. This renders other segmentation approach to be obsolete due to its robustness and high segmentation performance.

However, this approach suffers in terms of prolonged training time and large computing resource [5] that renders the approach inaccessible for smaller research group or enthusiast.

CNN emphasizes the spatial information [6] contained within an image thus requiring a very in-depth learning regarding the image spatial information. This in turns, disregards that each pixel contain valuable amount of information and can be further amplified to include additional information that can be used for more accurate image segmentation.

Object-oriented Programming (OOP) paradigm is a programming paradigm that able to provide an in-depth description of an object and has been employed universally [7] for various type of problems. This in turns enables more interactions between each object subsequently creating a relationship that helps affluence data manipulation and association [8]. However, the studies on the effort of incorporating such paradigm into crucial areas such as describing medical images are not being given the attention it deserved. Therefore, complementing medical images with object-oriented programming approach would allow each image to encompass a more in-depth relationships between each pixel compare to standard approach of the conventional bitmap representation.

*Author for correspondence

The objectives of the paper are as follows:

1. To propose the employment of OOP for describing the amplified medical pixel data for deep learning segmentation to be fed into a straightforward sequential deep learning neural network.
2. To train the deep learning network using a modestly configured machine with limited amount of memory and processing resources.
3. To evaluate and analyse the training and segmentation performance of the proposed approach and to provide in-depth performance comparison with other state-of-the-art CNN segmentation models.

2.Related works

Medical image segmentation is an area of image processing that has been thoroughly investigated due to its critical importance and the accessibility of different techniques to accomplish it [1]. Conventional image processing approach such as forthright image processing approach is considered as a classic technique. This approach segments the medical image manually by employing various image

processing processes such as enhancement, morphology, and thresholding with different image would necessitating different segmentation parameters [9]. While this approach can return a competent segmentation performance, the approach in the end is impractical for large datasets as it requires great amount of time and effort to individually determine each image's segmentation parameters and to manually segment the images.

Alternatively, the emergence of CNN as the new standard of image processing brought upon the existence of various CNN-based segmentation approaches. CNN approach excels specifically from image dataset as it considers the spatial information contains within an image. Fundamentally, CNN simulates the neurons in the visual cortex to process visual information [10]. This is accomplished by employing layers of "convolutional" networks that interprets image data by convolving the data from one layer to another to finally making sense of the image's content [5].

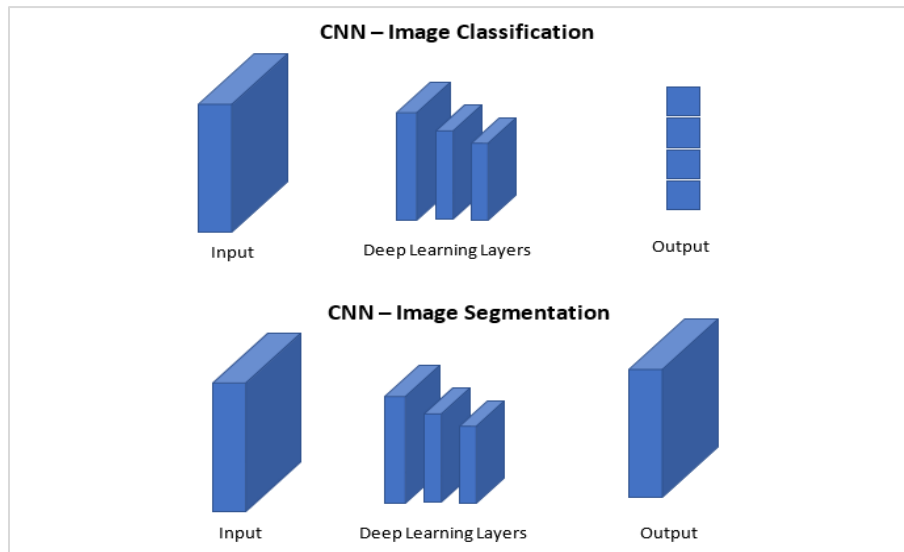


Figure 1 CNN image classification vs CNN image segmentation

Figure 1 explicates the difference between the process overview of image classification via standard CNN and the image segmentation via FCN. While conventional CNN approach primarily is an image classification method, segmentation using CNN is nevertheless demonstrated to also be successfully executed with model such as the Fully Convolutional Neural Network (FCN) [11]. This approach basically opens the fully connected layer in the final layer of CNN into a convolutional layer as the result of

segmentation differs from the standard image classification problem that usually executed via CNN [11]. The final convolutional layer in the FCN basically serves as the map to produce the final output that contain analogous size and dimension to that of the input. FCN approach has demonstrated to produce very competent segmentation results and has since inspired various research to follow suit. This include state-of-the-art segmentation models such as U-Net [12] and SegNet [13] that has shown to be able

to accurately segment various dataset specifically medical images. Nevertheless, CNN-based approaches are indicated to succumb to prolonged training time while necessitating tremendous amount of computing resources [5]. This in turns prevent such approach to be accessible towards smaller research groups that are equipped with modest computing resources.

In this paper, the performance of the proposed segmentation method is compared with U-Net and SegNet segmentation model for in-depth performance comparison in terms of training and segmentation performance which will be further discussed later.

3. Methodology

Generally, an image is represented by an $m \times n$ matrix where m represents the width and n represents the height, respectively. This matrix consists of the information regarding the pixel in the form of red, green, and blue color space or RGB. The RGB color space carries a numerical value of 0-255 for each channel and the combination of all these channels will subsequently generate a color that represents a pixel within the image. In CNN, the approach generates multiple instances of $m \times n$ matrices that convolve from one layer to another, creating a sequence of data that can be interpreted in the final layer of the network.

3.1 Dataset employed

The dataset employed in this paper is a medical image of retinal fundus image via Digital Retinal Images for Vessel Extraction (DRIVE) database [14]. This dataset contains 40 images (20 training and 20 testing) full-colored retinal fundus images for blood vessels extraction. In this paper however, only 10 images are selected for training purpose in order to investigate the robustness of the segmentation approach proposed as CNN-based approach such as U-Net model [12] is demonstrated to be able to produce accurate segmentation with small number of training datasets.

3.2 Amplified pixel description

Instead of focusing on learning the spatial information from one layer to another analogous to that of CNN approach, this paper suggests that each data in the pixels can be further amplified to generates even more profound information. To accomplish this, the individual pixel in the image will

carry additional information such as additional color space information, neighborhood statistical data, and class information. To accomplish this, the pixel data is encoded in the form of Object-oriented Programming (OOP) classes that would allow structured and accessible description regarding each individual pixel.

Figure 2 depicts the overview of Object-oriented Pixel Descriptor (OOPD). This descriptor stores additional color space information such as RGB, XYZ, YUV, LAB, CMYK and HSL of each individual pixel compared to that of the conventional approach of only storing the RGB color space information. This is especially useful as different color spaces describe color differently. While RGB color space predominantly describe the color in terms of the combination of red, green, and blue channel, color space such as LAB otherwise describe color in terms of lightness from black to white, green to red, and blue to yellow respectively. This generates a stark difference from the way RGB color space describe a pixel thus giving the approach extra information that can be utilized for segmenting the medical images.

The neighborhood statistical information on the other hand stores the statistical information of each color space mentioned, with a 3×3 neighborhood statistical information that contain the standard mean, mode, median, and range of every color space implemented as depicted in *Figure 3*. This would generate large number of dependent variables as the statistic will be calculated based on all the color spaces employed for each of the individual color channels described in each color space. Consequently, a single pixel in a medical image can hold large number of information that can be fed into a deep learning network to produce a thorough segmentation.

Additionally, the amplified pixel would also contain class information that individually labels which pixel belongs to which classes of the image as illustrated in *Figure 4*. This is especially crucial for medical images where image such as retinal fundus image contains different features that requires classification. This also mean that each individual pixel may contain its own class information that specifically described which features information within the medical image that the pixel currently belongs to.

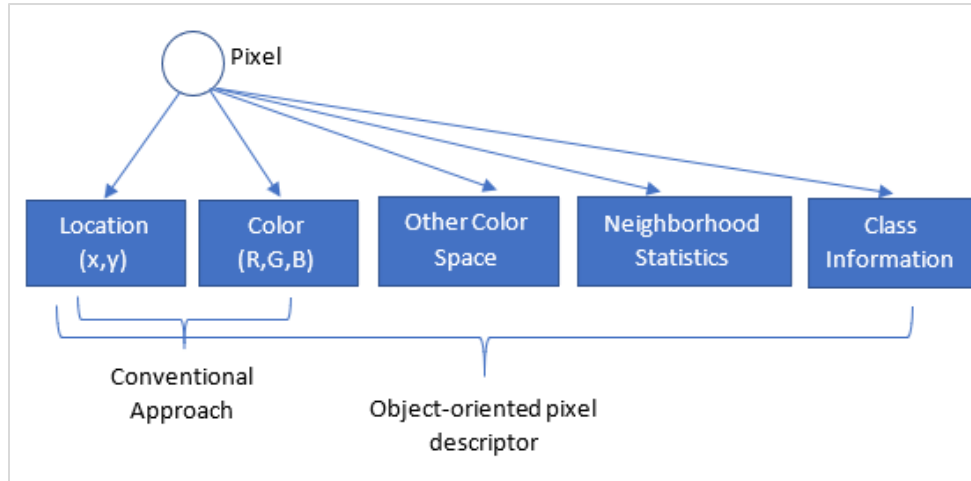


Figure 2 Object-oriented pixel descriptor (OOPD)

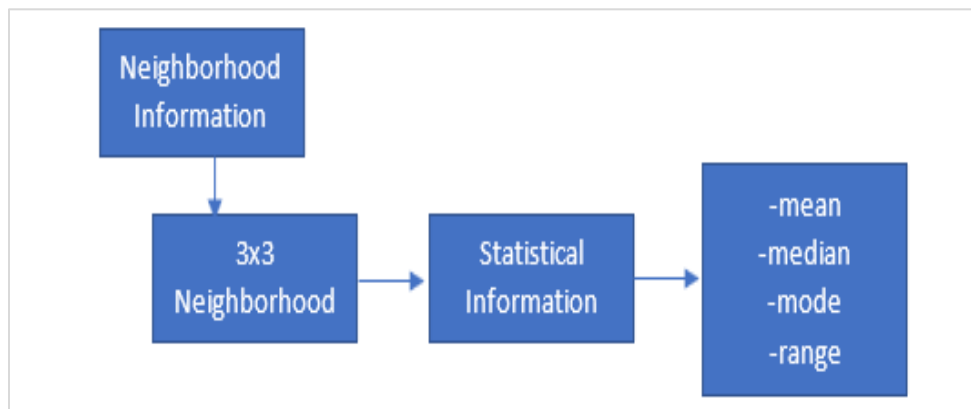


Figure 3 Pixel neighborhood information

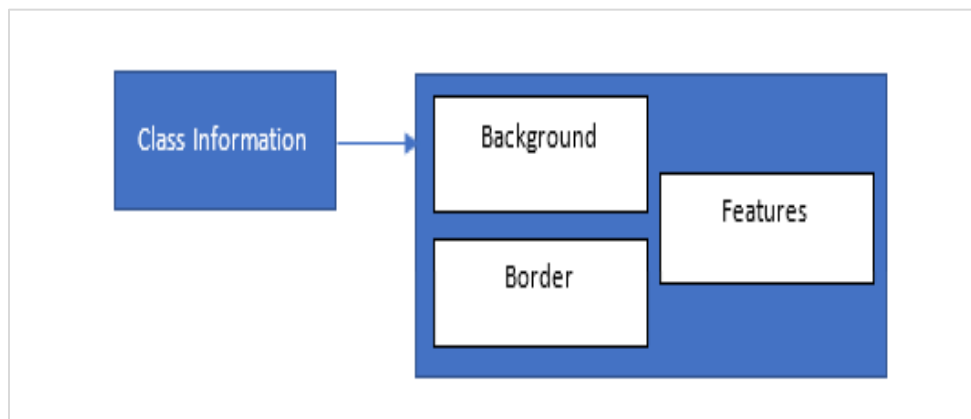


Figure 4 Pixel class information

3.3Deep learning neural network

The amplified pixel information via the OOPD will finally be ready for the segmentation training process. This process basically employs sequential deep-learning neural network with large number of

neurons due to the large number of information that is stored via the OOPD.

Figure 5 depicts the layers of network employed to train the segmentation model. Compared to that of

CNN approach [15, 16], this approach is significantly more abridged and upfront. This is due to the approach does not employ convolutional layers in CNN as it can make use of the amplified individual pixel information to create an in-depth relationship in the neural networks employed. Rectified Linear Unit (ReLU) activation layers are employed in 3 layers of

the neural network as this each layer employs dense network with large number of neurons. Subsequently, the classification process where the pixel will be either labelled with “segmentation” or “non-segmentation” is accomplished via the Sigmoid activation function.

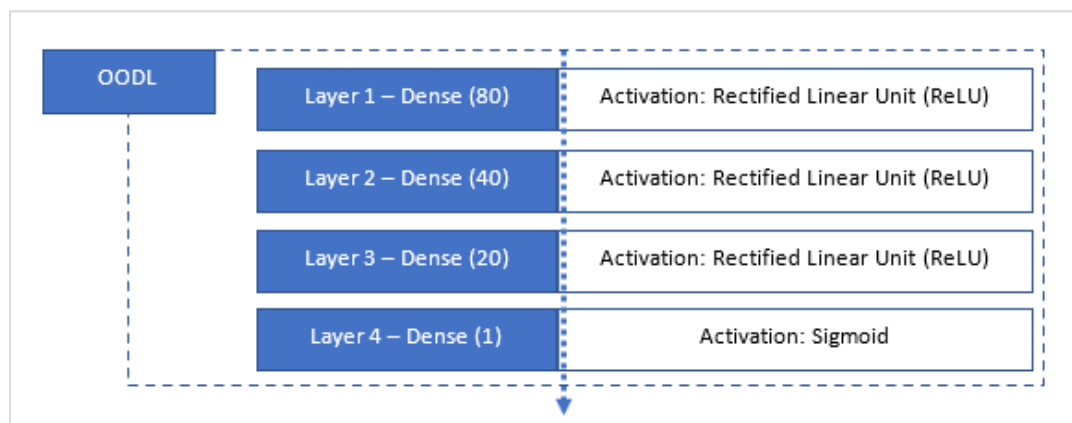


Figure 5 Object-oriented deep learning networks (OODL)

3.4 Performance evaluation

To investigate the performance of the proposed approach, the performance evaluation is divided into two; training performance and segmentation performance. The training performance basically records the amount of time required for the approach to train per epoch of the segmentation model compared to other popular CNN approaches such as U-Net and SegNet model. The loss and accuracy figure for each training epochs are also recorded for training performance. The training parameters of U-Net and SegNet segmentation models are configured with 256×256 input data and training process is limited with only 5 epochs as that is the maximum configuration that can be executed via the device

employed. The system configuration employed in this paper are as follows:

Table 1 illustrates the system configuration employed in this paper. The configuration employed can be considered as modest with conventional memory size and discrete laptop GPU. Another performance that is investigated is the segmentation performance where the segmentation produced from the dataset mentioned earlier will be investigated via the Receiver Operating Characteristic (ROC). The metrics include sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), and balanced accuracy which will be further elaborated later.

Table 1 System configuration

Hardware	Configuration
CPU	Intel i-7 8565U
Memory	8GB LPDDR3-RAM
GPU	NVIDIA GeForce MX 150

4. Results and discussion

4.1 Training performance

In terms of the training performance, the approach is evaluated based on the total training time (in seconds) for all 5 epochs and the training loss and accuracy figure produced while training the segmentation model. The time lapse in seconds(s) is recorded while training each segmentation model.

The metrics employed while training the models are loss calculation via cross entropy and straightforward accuracy calculation.

In *Figure 6*, it has been suggested that OODL able to produce almost analogous training accuracy figure compared to that of U-Net and SegNet segmentation model. However, the loss figure of OODL is

noticeably higher compared to the other approach, suggesting that the approach can be further improved in terms of improving the loss figure. However, this does not entirely reflect the segmentation performance which will be further deliberated later in this section.

In terms of the training time, the OODL is demonstrated to be significantly faster for each epoch compared to that of the other CNN approaches as shown in *Figure 7*. This transpires due to the number

of layers employed and structure of the OODL neural network to be significantly simpler compared to the CNN-based neural network. Both U-Net and SeNnet shows similar figure in terms of training time required for each epoch, with OODL shown to be more than 90% faster with the average of 33.8 seconds. This shows that OODL require significantly less time and computing resources compared to CNN-based approaches.

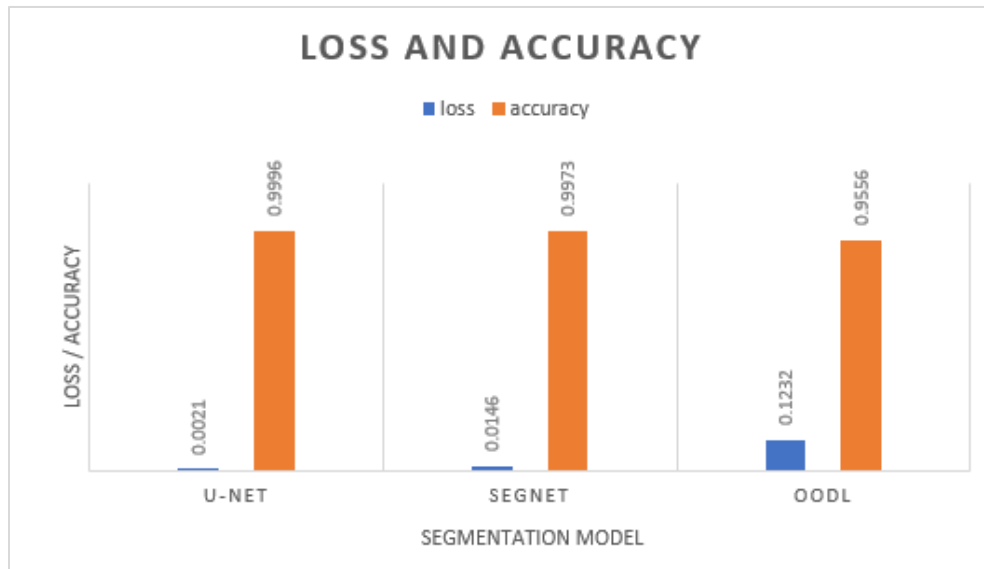


Figure 6 Training loss and accuracy



Figure 7 Training time (Seconds)

4.2 Segmentation performance



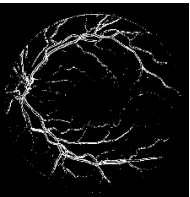



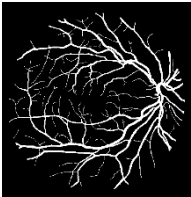
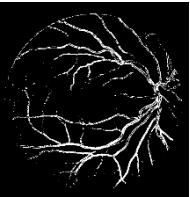
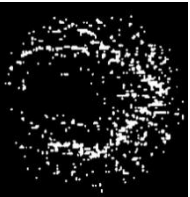
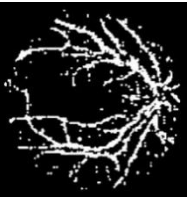

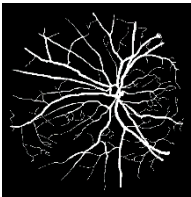
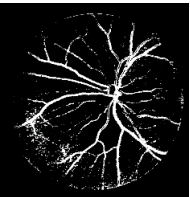
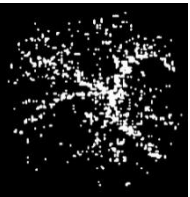
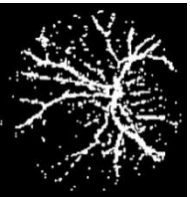
In terms of segmentation performance evaluation, the segmentation results of segmenting all the 20 testing datasets mentioned in earlier section is initially collected. Table below depicts the example results obtained from all 3 segmentation models. The segmentation performance is evaluated by overlapping the segmented image and the ground truth image.

Table 2 depicts the example of segmentation results obtained from all 3 segmentation models. This in turns will enable calculation of values such as true positive / TP (image segmentation correspond to the ground truth image), false positive / FP (image segmentation does not correspond to the ground truth image), false negative / FN (image unsegmented does not correspond to the ground truth image) and finally true negative / TN (image unsegmented correspond to the ground truth image). These values are then employed to calculate ROC metrics such as sensitivity (true positive rate), specificity (true negative rate), PPV, and NPV, and Balanced

Accuracy (B-ACC). In terms of appearance, ODDL's segmentation looks to be the sharpest as this approach took the input of full-sized images compared to SegNet and U-Net model of only 256x256 input which is the highest configuration supported by the hardware employed.

Figure 8 illustrates the segmentation performance for all 20 testing datasets employed. It is indicated that ODDL produces the best sensitivity, PPV, NPV, and overall B-ACC for the dataset under the training parameters and device employed mentioned in the earlier section of the paper. This result demonstrates that while ODDL require significantly less amount of processing and training time, it can produce results analogous if not better than the state-of-the-art CNN-based segmentation model such as U-Net and SegNet. This also suggested that amplifying the pixel data in medical image such as retinal fundus images would allow be able to generate accurate image segmentation when complemented with deep learning network.

Table 2 Example of segmentation results

Image	Original	Ground Truth	ODDL	SegNet	U-Net
Image 1					
Image 2					
Image 3					

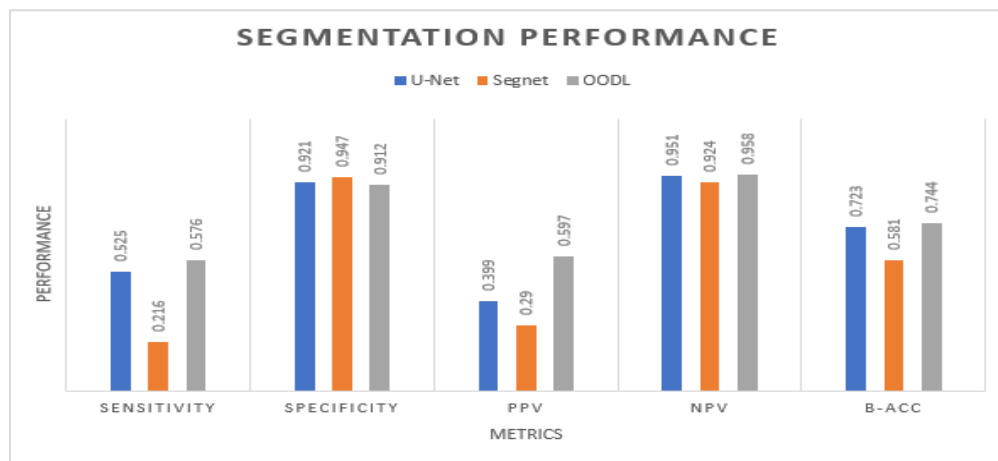
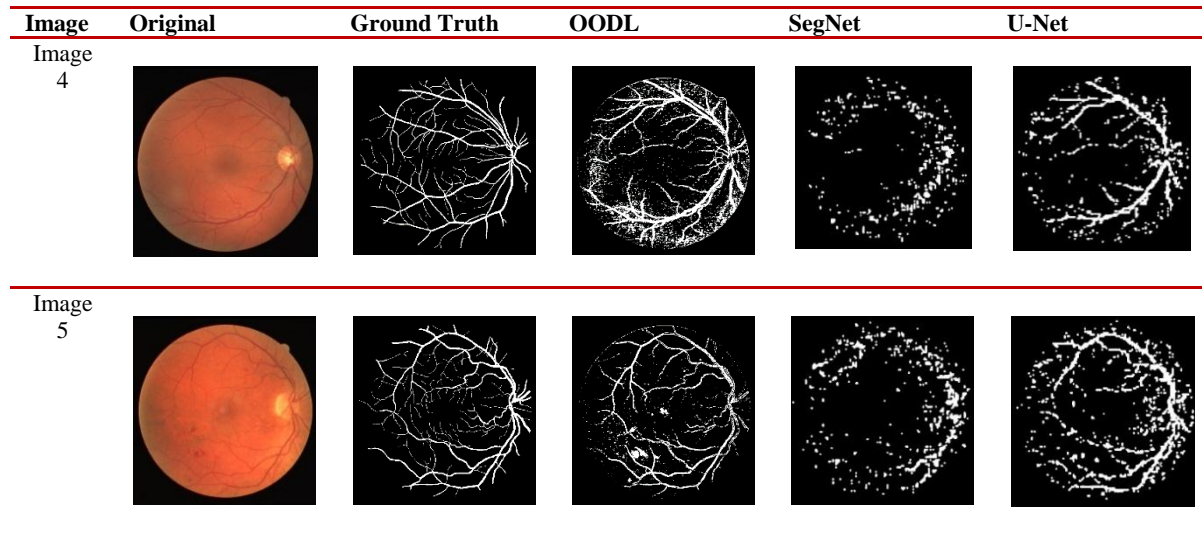


Figure 8 Segmentation performance

5. Conclusion and recommendation

In this paper, the approach of amplifying the individual pixel data in a medical image is demonstrated to be able to provide good segmentation performance even when being compared to other state-of-the-art CNN segmentation model such as U-Net and SegNet. However, the experiment and training process executed in this paper only limits to the amount of processing resources currently available to the authors. The CNN-based segmentation models are demonstrated to be able to produce better results if more training time and processing resources are being made available. Nevertheless, OODL is demonstrated to be very robust despite using modest computing resources and training time. Employing more time and computing resources towards the OODL approach and configuring the deep learning layer towards its

optimal configuration is also a noteworthy route to further improve this approach.

Acknowledgment

The research was supported by Ministry of Education Malaysia (MoE), and Universiti Teknologi MARA through the Fundamental Research Grant Scheme (FRGS) (600-IRMI/FRGS 5/3 (215/2019)).

Conflicts of interest

The authors have no conflicts of interest to declare.

References

- [1] Abuzaraida MA, Elrajubi OM. Improving laser mark detection for retinal images based on the AlexNet model. *International Journal of Advanced Trends in Computer Science and Engineering*. 2020; 9(4): 4597-603.
- [2] Devani D., Suryadibrata A., Young J. C. Implementation of Siamese convolutional neural

- network from cell images for malaria disease identification. *International Journal of Advanced Trends in Computer Science and Engineering*. 2020; 9(4):4195–4200.
- [3] Wadhwa A, Bhardwaj A, Verma VS. A review on brain tumor segmentation of MRI images. *Magnetic Resonance Imaging*. 2019; 61:247-59.
- [4] Kayalibay B, Jensen G, Van Der Smagt P. CNN-based segmentation of medical imaging data. *arXiv preprint arXiv:1701.03056*. 2017.
- [5] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*. 2017; 60(6):84-90.
- [6] Zhang M, Li W, Du Q. Diverse region-based CNN for hyperspectral image classification. *IEEE Transactions on Image Processing*. 2018; 27(6):2623-34.
- [7] Fadzil AF, Ahmad Z, Abd Khalid NE, Ibrahim S. Retinal fundus image blood vessels segmentation via object-oriented metadata structures. *International Journal of Engineering & Technology*. 2018; 7(4.33):110-3.
- [8] Mehtre VV, Nigam Y. Review on concepts related to object oriented programming system. *IRE Journals*. 2019; 3(6):56-8.
- [9] Abd Khalid NE, Ismail MF, Manaf MA, Fadzil AF, Ibrahim S. MRI brain tumor segmentation: a forthright image processing approach. *Bulletin of Electrical Engineering and Informatics*. 2020; 9(3):1024-31.
- [10] Jogin M, Madhulika MS, Divya GD, Meghana RK, Apoorva S. Feature extraction using convolution neural networks (CNN) and deep learning. In *international conference on recent trends in electronics, information & communication technology 2018* (pp. 2319-23). IEEE.
- [11] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In *proceedings of the conference on computer vision and pattern recognition 2015* (pp. 3431-40).
- [12] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In *international conference on medical image computing and computer-assisted intervention 2015* (pp. 234-41). Springer, Cham.
- [13] Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017; 39(12):2481-95.
- [14] <https://drive.grand-challenge.org/> . Accessed 26 August 2020.
- [15] Manikandan N, Priyanka M, Sasikumar R. Approximation computing techniques to accelerate CNN based image processing applications—a survey in hardware/software perspective. *International Journal of Advanced Trends in Computer Science and Engineering*. 2020; 9(3):3828-46.
- [16] Larrazabal AJ, Martinez C, Ferrante E. Anatomical priors for image segmentation via post-processing with denoising autoencoders. In *international conference on medical image computing and computer-assisted intervention 2019* (pp. 585-93). Springer, Cham.



Ahmad Firdaus Ahmad Fadzil is a lecturer at Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Cawangan Melaka, Kampus Jasin. He is a proud graduate of Universiti Teknologi MARA where he has received his Diploma, Bachelor's Degree, and Masters (Computer Science) from. He is currently pursuing his PhD also in UiTM with his current research interests include Evolutionary Computing, Parallel Processing, Image Processing, and Machine Learning.
Email: firdausfadzil@uitm.edu.my



Noor Elaiza Abd Khalid is an associate professor at Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Shah Alam, Selangor. She has received BSc (Comp Sc) from USM, Malaysia in 1985, MSc (Comp Sc) in University of Wales in 1992 and PhD (Comp.Sc) from Universiti Teknologi MARA, Shah Alam in 2010. Currently hold position as senior lecturer in Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam. Her current research interest is in Image Processing, Biomedical, Artificial Intelligence and Parallel Computing
Email: elaiza@uitm.edu.my



Shafaf Ibrahim is a senior lecturer in Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Cawangan Melaka, Kampus Jasin. Have a qualification of Diploma, Bachelor's Degree, Masters and PhD in Computer Science. Her research interests are Artificial Intelligence, Evolutionary Algorithms, Image Processing, and Optimization.
Email: shafaf2429@uitm.edu.my