

Deep Learning-Based Semantic Segmentation of Blood Cells from Microscopic Images



S. B. Asha and G. Gopakumar

Abstract Morphological analysis and differential cell counting are important in characterizing many diseases including malaria and leukaemia. The basic building block involved is cell segmentation and is a challenging but beneficial task in cytopathology. Microscopy being the gold standard for cell analysis, approaches have been discovered from traditional image processing operations to deep learning techniques for cell segmentation from microscopy images. In the last few years, classification networks were extended for image segmentation using the pixel-based classification method, known as semantic segmentation. Convolutional neural networks exhibited good performance in image segmentation. However, the networks suffered some limitations due to fully connected layers and pooling layers that restricted the size of images to be given as input and resulted in the loss of spatial context. In this research, experiments were carried out with two popular CNN architectures UNet and SegNet, traditionally used for semantic segmentation of natural images. By identifying the capacity of these networks for cell segmentation on natural images, we have experimented on a custom-built RBC, WBC and platelet cell segmentation dataset based on ALL-IDB. We critically evaluate the performance of both architectures with an intuitive explanation of their difference in performance. The UNet outperformed SegNet that too with limited labelled training data giving a promising Dice score of 0.97. With the experiments and analysis conducted in this work, we propose that the UNet is a very good choice for cell segmentation in cytopathology applications.

Keywords Semantic segmentation · Microscopic image analysis · Biomedical image processing · Deep learning

S. B. Asha (✉) · G. Gopakumar
Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Amritapuri,
India
e-mail: ashasbenny@am.amrita.edu

G. Gopakumar
e-mail: gopakumarg@am.amrita.edu

1 Introduction

The cell is the basic unit in all living organisms. Our body comprises trillions of cells, and analysing the behavioural and physical characteristics of these cells provides clues to the health of the subject being examined. Cytology is the branch of medicine that studies cells in living organisms. Cytopathology is considered as the branch of pathology where usually a microscopic examination of cells and tissues is done for diagnosing diseases. Traditionally, the microscopic examination is done manually, which is time-consuming, costly and often the results depend on the skill of the examiner. As timely detection and treatment are crucial for the patient, this is a good candidate for automation.

Automation in cytology is widely used in the medical field to reduce manual efforts and to get standard results. Image analysis methods are applied to cytology samples like blood smear images for identification, quantification and classification of diseases and abnormal conditions [1]. Microscopic analysis of blood samples is used for differential cell counting, detection of anaemias, presence of malaria parasites [2], tuberculosis [3], different types of leukaemia, eosinophilia, thrombocytosis, thrombocytopenia, etc. [4, 5].

Each cell possesses a standard signature which is contributed by the shape and size of the cell, morphology of the nucleus, presence of granules and the amount of cytoplasm. Depending on the disease, these cell signatures differ and the change in the signature is assessed by the microscopic image analysis. Accurate quantification of the cell signature depends on the detection of accurate spatial locations of cell and cellular structures in the image [6]. Therefore, one of the focus areas in microscopic image analysis is automated detection and segmentation of cellular structures [7]. It is not an easy task due to the challenges like the heterogeneous shape of cells in the image, intracellular variability, occurrence of cells as a cluster, etc. Also, the availability of publicly accessible annotated data that can be used to learn the model is insufficient [6].

In recent decades, several tools and techniques have been developed for the segmentation of cells. However, there is still a great demand for precise, standardized and robust whole cell segmentation algorithms to reliably measure morphological properties and subcellular structures in cell images [8].

To the best of our knowledge, little or no research has been carried out to segment different types of cells with heterogeneous shapes from microscopic images in a single pass using deep learning architectures when there is only a limited amount of data available for training. In this research, we propose UNet architecture for cell segmentation from microscopic images in the above settings by establishing its power to segment out WBCs, RBCs and platelets in a single pass. We also assess the effectiveness of data augmentation strategies in improving the performance of the model.

2 Related Works

An important step in the automation of cytopathology is cell localization and segmentation. Traditionally, cell segmentation was performed by simple thresholding [9]. Later, more advanced techniques based on watersheds [2], morphology [10] and clustering-based techniques [11] were employed. All these techniques demanded expert human knowledge for identifying the morphological features of the cells under study. These methods were not suitable when there is no significant contrast between various objects.

With the advent of deep learning algorithms, data became utmost important and the focus shifted to data-driven models [12]. Image segmentation is now considered as a pixel-level classification problem using labelled pixels.

Initially, classification networks with fully connected layers were only used for pixel-based classification. This was done using a patch of the image around the pixel due to the fixed size restriction of fully connected layers [13]. Long et al. [14] proposed a fully convolutional network (FCN) in which only convolutional layers are present to allow input images of any size. Probabilistic graphical models like conditional random fields (CRFs) and Markov random fields are used along with FCN to integrate more semantic contexts [15, 16].

The pooling layers in convolutional neural networks discard the spatial context. To solve this problem, architectures were developed that gradually recover spatial information. One of them is the convolutional encoder-decoder-based architecture where shortcut connections are provided between the encoder and the decoder. The encoder-decoder based on the convolutional neural network was used for image segmentation in [17]. Badrinarayanan et al. [18] came up with an encoder-decoder framework with a final pixel-wise classification layer, known as SegNet, and Ronneberger et al. [17] came up with UNet architecture for the segmentation of images. It consists of contracting and expansive paths similar to the encoder-decoder architecture. Segmentation of neural structures was done by combining residual blocks in ResNet architecture with UNet, called residual deconvolutional network [19]. The dense mechanism was incorporated into UNet to reduce gradient vanishing problems [20]. To account for the loss of features during downsampling and upsampling operations, dilated convolutions were introduced along with UNet [21]. Several other architectures are being developed to better address the need for semantic segmentation [22].

Though UNet is experimented within diverse applications [23] including biomedical segmentation, the power of these networks is seldom explored for the segmentation of multiple cellular structures at once using fewer data. We establish the efficacy of UNet in similar settings by showing improved accuracy on the cell segmentation on the benchmarked dataset ALL-IDB [24] using Dice's coefficient and Intersection over Union (IoU) metrics. The model is applied to segment three types of cells, namely red blood cells, white blood cells and platelets present in microscopic images of blood samples in a single pass.

Table 1 Distribution of pixels among the classes

| Class | Total no. of pixels | Percentage of pixels (%) |
|-----------|---------------------|--------------------------|
| RBC | 208946235 | 46.82 |
| WBC | 13882586 | 3.11 |
| Platelets | 549005 | 0.12 |

3 Dataset

We used the dataset developed by Shahzad et al. [25] for semantic segmentation. It is a manually generated dataset consisting of blood cell images and is an extension of the ALL-IDB dataset [24].

ALL-IDB dataset is a public dataset containing 108 whole-slide microscopic images of blood samples. The total count of blood elements is about 39000. The images were captured with 300–500 magnification rate microscopes. The dataset was originally released for the development of algorithms in detecting Acute Lymphoblastic Leukaemia (ALL). There were 59 images from healthy individuals and 49 images from ALL patients.

Blood images of humans contain at least three components, RBCs, WBCs and platelets. RBCs appear round in shape and are 7–8 μm in diameter. Variation in size and shape may indicate some abnormal conditions. WBCs are the largest among the three with a diameter from 10 to 20 μm . The dataset developed by Shahzad et al. [25] includes individual masks for WBCs, RBCs and platelets in the image. The dataset contains ground truth masks for 108 images of the ALL-IDB dataset. We have only used 106 masks out of this in our study, due to a size mismatch between the original images and masks for the remaining two images. The distribution of pixels across different classes for the dataset is shown in Table 1.

4 Methodology

The overall methodology for semantic segmentation is to design a network that extracts features through successive convolution operations and produces a segmentation map. We have implemented the segmentation framework based on SegNet and UNet architectures for the segmentation of RBCs, WBCs and platelets from microscopic images. The overview of the framework is shown in Fig. 1.

4.1 Preprocessing

The dataset contains whole-slide microscopic images and the corresponding individual masks for RBCs, WBCs and platelets as shown in Fig. 2. To perform multiclass segmentation, ground truth masks corresponding to RBC, WBC and platelets need

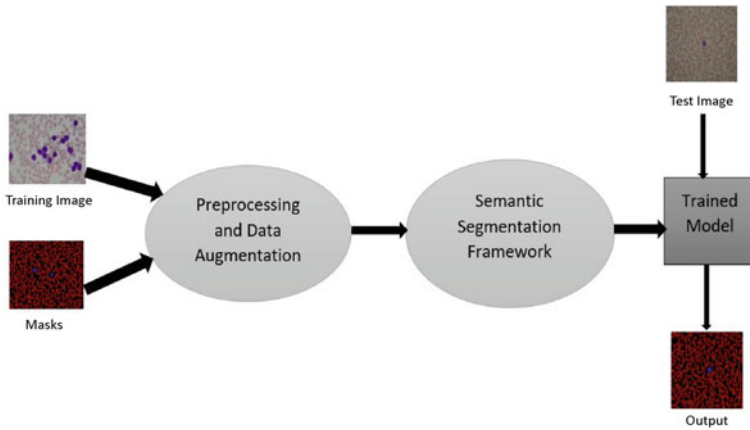


Fig. 1 Overall view of the segmentation framework

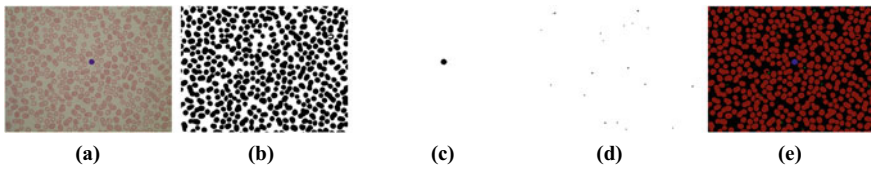


Fig. 2 a Input image. b RBC mask. c WBC mask. d Platelet mask. e Pixel mask

to be combined for each image. Then pixel IDs are assigned for each pixel as per the class to which it belongs. The resulting pixel labelled masks, as shown in Fig. 2e, are used in the framework.

The images are also subjected to area opening operation to remove small objects due to random noise or artefacts. Since the dataset is small, images are augmented by flipping the image from left to right and top to bottom directions. Also, the images are rotated by 90°, 180° and 270°. These types of augmentations aid the transformation invariant mechanism of deep neural networks. The rotation angle is in multiples of 90°, and hence the augmentation does not affect the shape, texture, symmetry and size of cells present. Flip augmentation operations will only reverse rows or columns of the image without affecting the cells present. Therefore, these simple augmentation strategies will not bias the segmentation of cells. Using the said augmentation strategies, the model is expected to learn enough information that can segment new cell images.

4.2 Semantic Segmentation Framework

Semantic segmentation networks are designed to make dense predictions for the image. Each pixel of the image is provided with a label of the class to which it belongs. This helps in the identification of objects and their boundaries in the image.

In cell segmentation, a microscopic image is split into segments to capture the relevant morphological information provided by cellular structures. Variations in shape, size, texture and contrast among cellular structures and lack of global applicability of existing approaches led to the usage of deep learning techniques for the cell segmentation problem [26]. The latest development in semantic segmentation is the encoder-decoder architecture. The structure of encoder-decoder networks helps in capturing semantic information efficiently. It consists of an encoder network followed by a decoder network. The encoder module takes the input and produces intermediate states which are given as input to the decoder module and produces the output. It was initially used for machine translation applications and later used in sequence-to-sequence prediction. The basic encoder-decoder architecture is shown in Fig. 3.

In the case of image segmentation, the encoder module gradually reduces the feature maps and captures higher semantic information. Then the decoder module gradually recovers the spatial information [27].

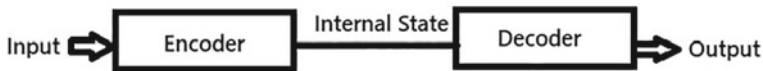


Fig. 3 Encoder-decoder architecture

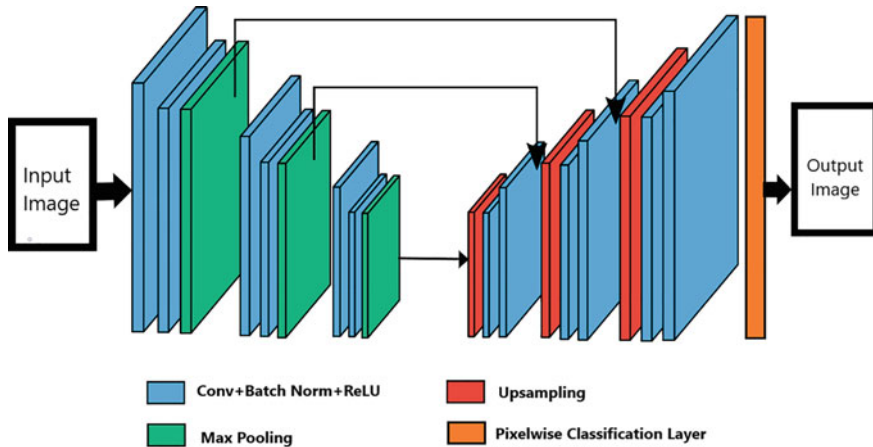


Fig. 4 SegNet architecture

We have tried to segment RBCs, WBCs and platelets from the ALL-IDB dataset using the concepts of SegNet as well as UNet. A brief description of the architectures is given in the next two paragraphs.

SegNet is a symmetric convolutional architecture comprising of an encoder and a decoder. The schematic architecture as in the original paper is shown in Fig. 4 [18]. In the encoder, layers identical to convolutional layers of VGG16 [28] are present. Each encoder consists of convolutional layers with batch normalization and ReLU non-linearity, followed by max-pooling and sub-sampling layers. Convolution with a filter bank is performed on the input image resulting in a set of feature maps which are passed to the batch normalization layer and the ReLU layer. Then the max-pooling operation is done using 2×2 windows with stride 2. Only the max-pooling indices are stored. This means that only the locations of maximum feature value in each pooling window are stored for each encoder feature map. For a 2×2 pooling window, this is done using 2 bits. The fully connected layers are removed thereby making the network small in size. For each encoder, there exists a corresponding decoder. The max-pooling indices from the corresponding encoder are used for upsampling the feature map to produce a sparse feature map. These are then convolved with filters to produce a dense feature map. The key advantages of this technique are retaining the boundary information and the reduction in the number of parameters required for end-to-end training. The output from the final decoder is fed into the multiclass softmax classifier to predict the class probabilities for each pixel [18].

The UNet architecture comprises a contraction path called an encoder and an expansion path called a decoder. It is a U-shaped architecture since the encoder is more or less symmetric to the decoder. Schematic architecture as in the original paper is shown in Fig. 5 [17]. In the contraction path, repeated convolutions are applied followed by ReLU activation and max-pooling operations. In the expansion path, the feature map is upsampled and then convolved by a 2×2 convolution. This is then concatenated with the corresponding cropped feature map from the contraction path. The concatenated feature map is then convolved by 3×3 maps and is then followed by ReLU activation. The feature maps from the contraction path are cropped because border pixels may be lost during convolution operation. The feature map is subjected to 1×1 convolution in the final layer to map it to the corresponding classes [17].

UNet concatenates the full feature map from the corresponding encoder in the contraction path to the feature map from the decoder. Due to this, UNet makes use of more features for recovering the spatial context. Therefore, UNet can work efficiently with augmented datasets with a small number of images. As in SegNet, it doesn't reuse the max-pooling indices. UNet requires more memory compared to SegNet as the entire feature maps from the encoder are stored and used in the decoder. SegNet only stores the max-pooling indices from the encoder [29].

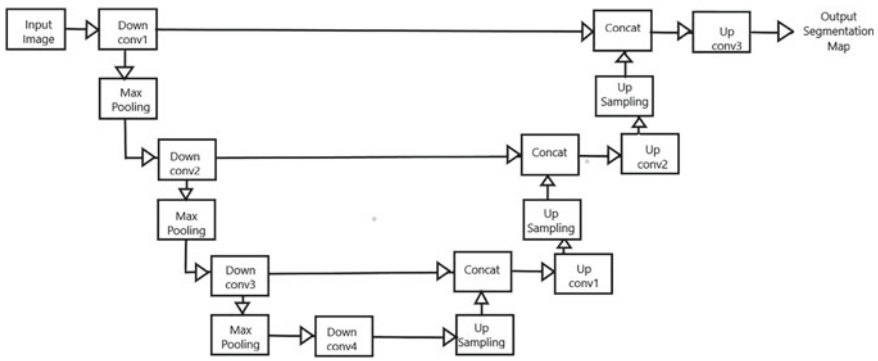


Fig. 5 UNet architecture

5 Experimentation and Results

To train and compare the above presented deep architectures, namely SegNet and UNet, we have implemented the architectures and have experimented on the ALL-IDB dataset for RBCs, WBCs and platelets segmentation. The models are trained on 106 images of the ALL-IDB dataset. The implementation was done using Keras and TensorFlow frameworks on a single GPU machine with 16 GB RAM and NVIDIA GEFORCE GTX-1050 Ti. To quantify the performance of the models, Dice's coefficient, IoU and Pixel Accuracy (PA) measures are used as evaluation metrics.

Dice's coefficient measures the overlap accuracy of segmentation images with the ground truth images. The value of Dice's coefficient can range from 0 to 1, where a value of 1 indicates perfect segmentation. Dice's coefficient is given by Eq. 1 in terms of True Positives (TP), False Positives (FP) and False Negatives (FN):

$$\text{Dice's Coefficient} = \frac{2TP}{2TP + FP + FN} \quad (1)$$

IoU (Eq.2) is defined by the ratio of the area of overlap and the area of union between the predicted image and the ground truth mask. The value of the IoU metric can range from 0 to 1, where a value of 1 indicates perfect segmentation:

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (2)$$

Pixel accuracy is given by the percentage of pixels classified correctly as shown in Eq.3. It provides the overall accuracy of the cell segmentation:

$$\text{Pixel Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

We tried to assess semantic segmentation accuracy using pixel accuracy. It has to be noted that accuracy measure suffers from class imbalance problems since the images contain background class as the majority class. Hence, Dice's coefficient and

IoU measures were used along with it in the current model. They are both positively correlated metrics and are used particularly for class imbalanced problems as they measure relative overlap between predictions and ground truth. The IoU metric weighs false positives more than Dice's coefficient. Dice's coefficient measures the average performance of the segmentation, whereas IoU measures worst-case performance. Therefore, the models are optimized such that Dice's coefficient and IoU are maximized.

Since we are segmenting RBCs, WBCs and platelets, this is a multiclass pixel-wise classification problem. To train both networks, we have used the Adam optimizer [30]. Adam is a gradient-based optimization algorithm using an adaptive learning rate. Training has been done on a batch size of 5 for SegNet and 8 for UNet for 100 epochs. Batch size is kept to a minimum due to memory limitations. The models are trained from scratch without any transfer learning. The initial learning rate is fixed at 0.001 and is reduced by a factor of 0.1 with respect to the reduction in the validation of Dice's coefficient. Since the problem deals with multiclass segmentation, categorical cross-entropy is taken as the loss function, and the ground truth segmentation mask is provided after one-hot encoding.

Initially, both the models are trained using original images without applying augmentation techniques to the dataset. Later, different combinations of data augmentations are applied, and the impact is analysed on the performance of the models. The effect of applying data augmentation before and after splitting into train and validation sets is also analysed. In addition to this, we try to measure the performance of models when the train and validation sets are both augmented after splitting. The results and analysis are discussed below.

The dataset consisting of 106 images are split into train and validation sets without any augmentation in an 80:20 ratio in the first case. When they are used with UNet and SegNet models, 0.95 and 0.82 are obtained as Dice's coefficient values respectively which is listed as Case I in Table 2. To improve the performance of the models, the number of training images is increased by applying augmentation techniques. The training images are first rotated by 90°, 180°, 270° and added to the training set. This resulted in 0.964 and 0.85 as Dice's coefficient values for UNet and SegNet (Case

Table 2 Performance metrics

| Model | Case | Dice's coefficient | IoU | Pixel accuracy (%) |
|--------|------------|--------------------|-------------|--------------------|
| UNet | I | 0.95 | 0.83 | 94 |
| | II | 0.96 | 0.87 | 96 |
| | III | 0.97 | 0.88 | 96 |
| | IV | 0.98 | 0.93 | 96 |
| | V | 0.98 | 0.91 | 97 |
| SegNet | I | 0.82 | 0.54 | 78 |
| | II | 0.85 | 0.59 | 81 |
| | III | 0.88 | 0.68 | 85 |
| | IV | 0.92 | 0.78 | 90 |
| | V | 0.89 | 0.72 | 87 |

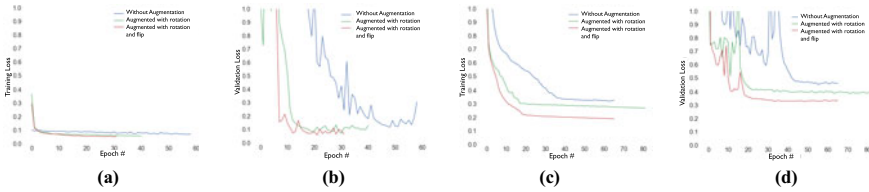


Fig. 6 Effect of applying augmentation techniques. **a** Training loss—UNet. **b** Validation loss—UNet. **c** Training loss—SegNet. **d** Validation loss—SegNet

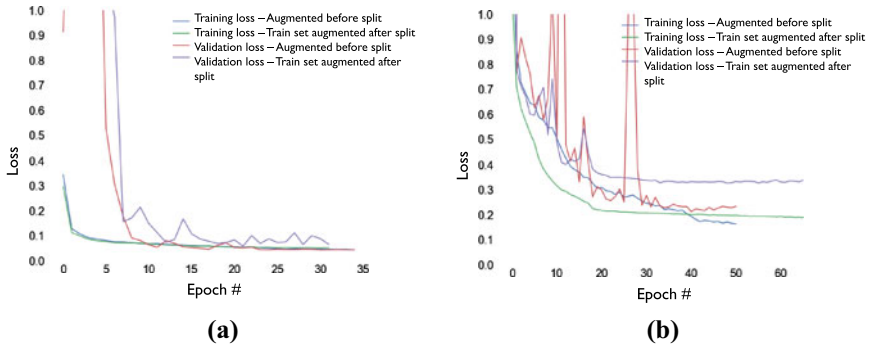


Fig. 7 Effect of splitting data to train and validation sets before and after applying augmentation for **a** UNet and **b** SegNet models

II in Table 2). Further, the training images are flipped along both axis directions separately and added to the training set (Case III in Table 2). In this case, values increased to 0.97 and 0.88. The values for IoU and accuracy are also increased with augmentation. The training and validation loss learning curves for UNet and SegNet models applied to the dataset without augmentation and with different combinations of augmentation are shown in Fig. 6. The models show minimum loss when used with the dataset augmented by rotation and flipping operations. However, training and validation loss for SegNet doesn't even fall below 0.1.

The training and validation loss learning curves for UNet and SegNet models in which the dataset is split before and after applying augmentation are shown in Fig. 7, and metrics are summarized in Case IV of Table 2. The 106 images in the dataset are first augmented and then split into train and validation sets. The performances using these augmented sets are evaluated and compared with the dataset consisting of augmented train images and original validation images. Minimum validation loss, maximum Dice's coefficient and maximum IoU are obtained in Case IV, as rotated training images are present in the validation set and this might have lead to overfitting.

Further, the effect of augmentation on the validation set is also analysed. As given in Case V of Table 2, we found that a slightly higher value for Dice's coefficient is obtained with the augmented validation set for UNet and SegNet. However, this rise

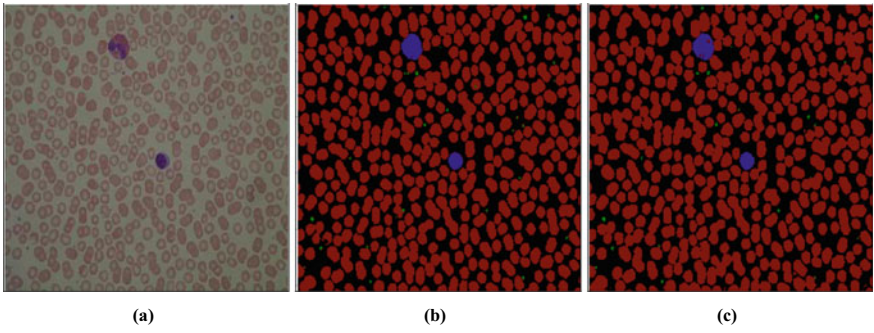


Fig. 8 a Original image. b Ground truth mask. c UNet segmentation result

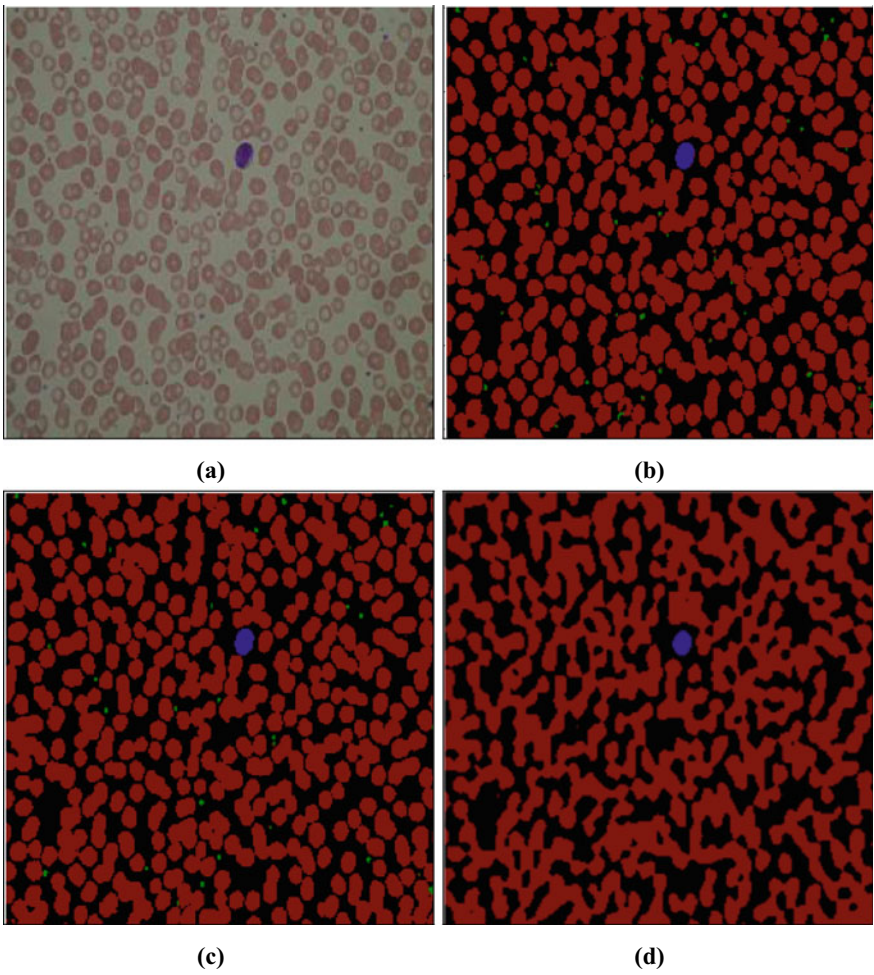


Fig. 9 a Original image. b Ground truth. c UNet result. d SegNet result

is caused by the augmented images in the validation set and may not be good for unseen images.

The performance metrics are evaluated on the validation set after each epoch, and the values for the best trained model with respect to maximum Dice's coefficient in all the above-discussed cases are summarized in Table 2.

With respect to the various cases, the best performance was achieved by UNet using a training set augmented with rotation and flipping operations.

Semantic segmentation on a validation image is shown in Fig. 8. UNet can identify pixels belonging to RBCs, WBCs and platelets from the microscopic blood images. Postprocessing techniques are required to separate the overlapped RBCs and WBCs in predictions. We observed that the predictions made using the SegNet model using this small dataset are not accurate enough as indicated by the metrics. The result of segmentation on a sample image by both models is shown in Fig. 9.

SegNet was proposed for the segmentation of road or indoor scenes, and it focuses on reducing memory requirements at the cost of accuracy [31]. Therefore, in the case of microscopic images, the encoding-decoding process in SegNet may result in the loss of relevant features to identify the cellular structures. This may be the reason behind the poor performance of SegNet on our dataset, whereas UNet was originally designed to work with limited images. UNet concatenates the entire low-level features extracted in the encoder to the decoder which helps in identifying the blood cells correctly. This leads to the better performance of the UNet model and makes UNet an ideal candidate for biomedical cell segmentation.

6 Conclusion

The effectiveness of applying deep learning architectures for the multiclass cell segmentation problem in microscopic images was assessed. When the traditional CNN demands at least a few thousand images to get decent segmentation results, a very good Dice score of 0.97 was obtained when experimented with UNet architecture even when trained from scratch with very limited images from the ALL-IDB dataset. Also when compared with the SegNet architecture, multiple cell types that differ in size and shape are segmented correctly using UNet in a single pass even without applying augmentation. Further, the effect of augmentation techniques for improving the model was evaluated, and the results of applying augmentation before and after splitting the dataset into training and validation sets were analysed. The test images were drawn at random and the validation set offered a Dice score of 0.97 making it an ideal candidate for the cell segmentation task.

References

1. Bui MM, Pantanowitz L (2020) Image analysis in cytopathology. *Monogr Clin Cytol Basel Karger* 25(9):91–98
2. Gopakumar GP et al (2018) Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner. *J Biophotonics* 11(3)
3. Simon A, Vinayakumar R, Sowmya V, Soman KP (2019) Shallow CNN with LSTM layer for tuberculosis detection in microscopic image. *Int J Recent Technol Eng* 7:56–60
4. Thanmayi AVL, Reddy SD, Kochuvila S (2021) Detection of leukemia using K-means clustering and machine learning. In: *Ubiquitous communications and network computing. UBIUNET 2021. Lecture notes of the institute for computer sciences, Social informatics and telecommunications engineering*, vol 383. Springer, Cham
5. Kalmady KS, Gopakumar G et al (2017) Improved transfer learning through shallow network embedding for classification of leukemia cells. In: *2017 ninth international conference on advances in pattern recognition (ICAPR)*, pp 1–6
6. Dimopoulos S, Mayer CE, Rudolf F (2014) Accurate cell segmentation in microscopy images using membrane patterns. *Bioinformatics* 30(18):2644–2651
7. Song J, Xiao L, Lian Z (2018) Contour-seed pairs learning-based framework for simultaneously detecting and segmenting various overlapping cells/nuclei in microscopy images. *IEEE Trans Image Process* 27(12):5759–5774
8. Al-Kofahi Y, Zaltsman A, Graves et al (2018) A deep learning-based algorithm for 2-D cell segmentation in microscopy images. *BMC Bioinf* 19(365)
9. Kong J et al (2015) Automated cell segmentation with 3d fluorescence microscopy images. In: *Proceedings of the IEEE international symposium on biomedical imaging*, pp 1212–1215
10. Gopakumar G, Jagannadh VK et al (2016) Framework for morphometric classification of cells in imaging flow cytometry. *J Microsc* 261(3):307–319
11. Kothari S, Chaudry Q, Wang MD (2009) Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques. In: *IEEE international symposium on biomedical imaging: from nano to macro*, pp 795–798
12. Gopakumar G, Sai Subrahmanyam GRK (2019) Deep learning applications to cytopathology: a study on the detection of malaria and on the classification of Leukaemia cell-lines. In: *Handbook of deep learning applications*. Springer International Publishing, pp 219–257
13. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J (2012) Deep neural networks segment neuronal membranes in electron microscopy images. In: *Proceedings of the 25th international conference on NIPS 2*, pp 2843–2851
14. Long J, Shelhamer E, Darrell T (2017) Fully convolutional networks for semantic segmentation. *IEEE TPAMI* 39(4):640–651
15. Chen L et al (2014) Semantic image segmentation with deep convolutional nets and fully connected CRFs. [arxiv:1412.7062](https://arxiv.org/abs/1412.7062)
16. Liu Z, Li X, Luo P, Loy C, Tang X (2015) Semantic image segmentation via deep parsing network. *ICCV*, pp 1377–1385
17. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. *MICCAI*, pp 234–241
18. Badrinarayanan V et al (2017) SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE TPAMI* 39(12):2481–2495
19. Fakhry A, Zeng T, Ji S (2017) Residual deconvolutional networks for brain electron microscopy image segmentation. *IEEE TMI* 36(2):447–456
20. Li X et al (2018) H-DenseUNet: hybrid densely connected UNET for liver and tumor segmentation from CT volumes. *IEEE TMI* 37(12):2663–2674
21. Tureckova A, Rodríguez-Sánchez AJ (2019) ISLES challenge: U-shaped convolution neural network with dilated convolution for 3d stroke lesion segmentation. *Brainlesion MICCAI* 11383:319–327

22. Sachin TKS, Sowmya V, Soman KP (2021) Performance analysis of segmentor adversarial network (SegAN) on bio-medical images for image segmentation. In: *Advances in automation, signal processing, instrumentation, and control*. Springer, pp 751–758
23. Darshik AS, Dev A, Bharath M, Nair BA, Gopakumar G (2020) Semantic segmentation of spectral images: a comparative study using FCN8s and U-NET on RIT-18 dataset. In: *2020 11th international conference on computing, communication and networking technologies*, pp. 1–6
24. Labati RD, Piuri V, Scotti F (2011) All-IDB: the acute lymphoblastic leukemia image database for image processing. In: *18th IEEE international conference on image processing*, pp 2045–2048
25. Shahzad M et al (2020) Robust method for semantic segmentation of whole-slide blood cell microscopic images. *Comput Math Methods Med* 2020:4015323
26. Haque IR, Neubert J (2020) Deep learning approaches to biomedical image segmentation. *Inf Med Unlock* 18:100297
27. Zhu Y, Chen LC, Papandreou G, Schroff F, Adam H (2018) Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Computer vision-ECCV*, pp 833–851
28. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. [arxiv:1409.1556](https://arxiv.org/abs/1409.1556)
29. Laves MH, Bicker J, Kahrs LA, Ortmaier T (2019) A dataset of laryngeal endoscopic images with comparative study on convolution neural network-based semantic segmentation. *Inter J of Comp Assist Radiol Surg* 14:483–492
30. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. [arxiv:1412.6980](https://arxiv.org/abs/1412.6980)
31. Zhang L et al (2020) Automated deep learning method for whole-breast segmentation in diffusion-weighted breast MRI. *J Magn Reson Imaging* 51(2):635–643