

Automatic *Bacillus anthracis* bacteria detection and segmentation in microscopic images using UNet++



Fatemeh Hooral^a, Hossein Khosravi^{a,*}, Bagher Moradi^b

^a Faculty of Electrical Engineering and Robotics, Shahrood University of Technology, Shahrood, Iran

^b Esfarayen Faculty of Medical Science, Esfarayen, Iran

ARTICLE INFO

Keywords:

Deep learning
Automated detection
Segmentation
Anthrax disease
Bacillus anthracis bacteria
UNet
UNet++

ABSTRACT

Anthrax is one of the important diseases in humans and animals, caused by the gram-positive bacteria spores called *Bacillus anthracis*. The disease is still one of the health problems of developing countries. Due to fatigue and decreased visual acuity, microscopic diagnosis of diseases by humans may not be of good quality. In this paper, for the first time, a system for automatic and rapid diagnosis of anthrax disease simultaneously with detection and segmentation of *B. anthracis* bacteria in microscopic images has been proposed based on artificial intelligence and deep learning techniques. Two important architectures of deep neural networks including UNet and UNet++ have been used for detection and segmentation of the most important component of the image i.e. bacteria. Automated detection and segmentation of *B. anthracis* bacteria offers the same level of accuracy as the human diagnostic specialist and in some cases outperforms it. Experimental results show that these deep architectures especially UNet++ can be used effectively and efficiently to automate *B. anthracis* bacteria segmentation of microscopic images obtained under different conditions. UNet++ produces outstanding results despite the many challenges in this field, such as high image dimension, image artifacts, object crowding, and overlapping. We conducted our experiments on a dataset prepared privately and achieved an accuracy of 97% and the dice score of 0.96 on the patch test images. It also tested on whole raw images and a recall of 98% and accuracy of 97% is achieved, which shows excellent performance in the bacteria segmentation task. The low cost and high speed of diagnosis and no need for a specialist are other benefits of the proposed system.

1. Introduction

Anthrax is a serious infectious disease caused by a bacterium called *B. anthracis*. It is a disease that has a worldwide distribution and mainly affects agricultural workers (Yang et al., 2020). Humans become infected with anthrax when working with animal products such as wool, skin or animal bones. The disease in humans is mainly manifested in four clinical forms: skin, lung, intestinal, and injection. More than 95% of cases of anthrax disease are skin lesions that occur following superficial lesions. The disease is caused by the entry of bacterial spores into wounds or skin scratches. A preliminary diagnosis of anthrax is confirmed by direct observation as well as bacterial culture. Warm staining of the pathology sample and the observation of a warm gram-positive quadrilateral bacilli, sometimes in the form of a long chain, is an important diagnostic key (Bolognia et al., 2012). Microscopic examination of clinical specimens is used to diagnose suspected cases of this disease and in environmental specimens to observe spores of *B. anthracis*. The most common color used is hot and polychrome

methylene blue. In clinical specimens, *B. anthracis* is seen as gram-positive, thick, straight, long, quadrangular, and often single, double, or short chains consisting of three to four bacilli (Doganay et al., 2010).

Indeed, automatic and semi-automatic processing of various medical images, such as computerized tomography (CT) scan, magnetic resonance imaging (MRI), ultrasound, and microscopic images of blood or tissue using machine vision systems and image processing are the most common and important applications in this field. The goal of image processing algorithms is usually to create new images with different features than the original image or to extract features based on the original image (Hosseini et al., 2017).

One of the applications of image processing systems in medicine is the quantitative and qualitative analysis of microscopic images of a patient's blood or tissue sample. In fact, computer-assisted microscopic imaging plays an important role in the diagnosis and prognosis of diseases (Gurcan et al., 2009). Examining a large number of slides containing colored samples is a daunting task, depending on the results of a specialist's visual examination. The results of such examinations by

* Corresponding author.

E-mail address: hosseinkhosravi@shahroodut.ac.ir (H. Khosravi).

several specialists are rarely the same. To overcome such problems, automated analysis of medical images using image processing is considered. The accuracy of the detection and segmentation of cells and bacteria has a great impact on the validity of the diagnosis by pathologists. In general, due to fatigue and decreased visual acuity, microscopic diagnosis of diseases by humans is associated with reduced accuracy and high error.

Deep learning has recently been considered as a leading machine learning tool in almost every scientific field from precision horticulture (Saedi and Khosravi, 2020) to medical image analysis (Mithra and Emmanuel, 2019). Deep learning in a variety of fields, including classification, detection, and segmentation has been instrumental in the analysis of medical images. It is a learning method that directly processes raw data and automatically learns representation and extracts appropriate features. It can be used for object detection, image segmentation and classification. Compared to manually extracted features, the trained representation requires less human intervention and leads to much better results (Goodfellow et al., 2016). Recently, deep learning as a powerful tool has been shown to achieve good results in the analysis of microscopic images such as object or lesion classification (Setio et al., 2016; Nie et al., 2016), organ and substructure segmentation (Moeskops et al., 2016; Korez et al., 2016), organ and region localization (Payer et al., 2016; de Vos et al., 2016; Kong et al., 2016), registration (Simonovsky et al., 2016; Miao et al., 2016) and more.

In digital pathology, many microscopic images are collected for analysis and evaluation. High-dimensional microscopic data contains complex patterns and non-linear dependencies with the type of disease or the label of the images. In addition, due to the preparation of samples from different patients or different methods of data preparation, there are wide differences in the images (Xing and Yang, 2016). These situations significantly challenge classical image analysis and machine learning algorithms for such tasks like core/cell detection, segmentation, classification, and so on.

Deep neural network (DNN) with many convolutional layers and a large number of trainable parameters can learn feature hierarchies and extract input-output relationships very well. Thus, DNNs are better choices for microscopic image analysis, and today, it is the most important method for analyzing microscopic images.

Microscopic imaging has complex unique features compared to other imaging techniques such as MRI, ultrasound, and CT. In digital pathology, image data is usually generated using a specific color scheme and there are many challenges, including the complexity of background, non-uniform intensity, cell/nuclei contact, and overlap (Gurcan et al., 2009; Veta et al., 2014; Irshad et al., 2013).

Recent research works have shown that DNNs can segment microscopic images perfectly (Greenspan et al., 2016; Min et al., 2017). In general, the models offered for segmentation include two categories of semantic segmentation and instance segmentation. Semantic segmentation is a process of classifying an image at the pixel level in which each pixel is assigned a class label. Convolutional neural networks (CNNs), as the first semantic segmentation networks consider this task as a pixel classification problem. For example, to segment the neural membrane in microscopic images in (Ciresan et al., 2012), CNN has been used as a binary pixel classification to label each pixel. Ning et al. (Ning et al., 2005) provided a CNN-based framework as a pixel classifier to segment the image into five categories including the nucleus, membrane, cytoplasm, cell wall, and external culture medium to analyze fetal growth stages.

By considering convolutional layers instead of fully connected layers, the CNN can be trained end-to-end to directly create probabilistic mappings with the same dimensions as the input images. Thus, a fully convolutional neural network (FCNN) (Ronneberger et al., 2015) can be applied effectively to a full input volume or image. However, the computational efficiency of FCNN is very high and cannot be used in practice.

A new architecture based on FCNN was introduced by (Long et al.,

2015). The new architecture, called UNet, was modified and extended to work with fewer training images and to yield more precise segmentation. The main idea is to supplement a usual contracting network by successive layers, where pooling operations are replaced by up-sampling operators. An important modification in UNet is that there are a large number of feature channels in the up-sampling part, which allow the network to propagate context information to higher resolution layers.

In (Çiçek et al., 2016; Drozdal et al., 2016), in order to segment two-dimensional and three-dimensional data, in addition to up-sampling paths, jump connections were also used to directly connect the layers of expansion and contraction convolution. Milletari et al. (Milletari et al., 2016) expanded the UNet to minimize segmentation error using residual blocks similar to those used in ResNet.

UNet has also been used in microscopic imaging to segment nerve membranes (Arganda-Carreras et al., 2015), glioblastoma-astrocytoma and HeLa cells (Maška et al., 2014). Many more progressions have been made to FCNNs to improve segmentation performance in various studies, including SegNet by modifying the number of deconvolution layers using the indicators generated by the pooling layers (Badrinarayanan et al., 2017) and DeepLab which replaces the deconvolution layers with a probabilistic graphical model (Chen et al., 2017).

Recently, RNNs have been used to segment medical images, which have led to satisfactory results. In (Xie et al., 2016) a temporal-spatial RNN consisted of a much smaller number of parameters and a lower computational cost than similar models is presented for the segmentation of the histopathological images. Stollenga et al. presented two different LSTM structures, including three-dimensional LSTM-RNN and PyraMiD-LSTM (Stollenga et al., 2015) for the parallel segmentation of neural structures (Arganda-Carreras et al., 2015) and MR brain images. Andermatt et al. (Andermatt et al., 2016) used a three-dimensional RNN to segment the white and gray area in the MRI database of the brain. In (Chen et al., 2016) to segment the structures in 3D microscopic images, the idea of combining two-sided LSTM-RNNs with two-dimensional UNet-like architecture is presented. In (Poudel et al., 2016) a combination of a two-dimensional UNet architecture with a gated recurrent unit for three-dimensional segmentation is proposed.

In 2018, with the changes in the UNet structure, a new and more powerful architecture called UNet++ was introduced. This new architecture of deep networks has been adapted and developed especially for the segmentation of medical images, in which the encoder and decoder sections are connected through dense skip paths in such a way that the semantic distance between the feature maps of these two sections decreased. Using this idea made network training easier. The network is used to segment various medical images, including CT scans of the chest for node segmentation, CT scans of the abdomen for liver segmentation, microscopic images for core segmentation, and colonoscopy videos for polyp segmentation and in all cases, produced better results compared to UNet (Zhou et al., 2018).

In this paper, two important deep architectures including UNet and UNet++ have been used for the automatic diagnosis of anthrax disease and the segmentation of the anthrax bacterium. This study which has been presented for the first time shows that these two networks, especially UNet++ can be highly effective in the diagnosis of anthrax disease and bacterial detection and segmentation. In this study, the microscopic image of the tissue sample of anthrax patients are obtained under different conditions, and despite the many challenges such as high image dimension, image artifacts, object crowding, and overlapping, the proposed method performed very well.

2. Materials and methods

2.1. The dataset

In this research, we prepared a dataset of images taken from slides containing *B. anthracis* bacteria. These slides belong to the tissues of



Fig. 1. Image acquisition system used in the proposed method.

patients suffering from the cutaneous anthrax disease.

The method of preparing the tissues on the slides was that we first, using a scalpel blade, picked up some sanies from the wound tissues and placed them on clean slides. Then drying was performed at room temperature and the smear fixation was conducted using Methanol. Fixed tissue smears were painted by Giemsa reagent. After 15 min, tissue smears were washed by water. In the next stage, dried slides were examined by an Olympus microscope and we took specific diagnostic images using a camera mounted on a stand.

Slide preparation and database provision was done in collaboration with the vice-chancellor of Esfarayen faculty of medical sciences.

Using the imaging system shown in Fig. 1, the images of the samples were acquired in RGB space using a Canon camera mounted on the microscope. Several microscopic images of different parts (On average, about 5 images per slide) are taken from each slide.

In this paper, 200 slides containing tissue samples related to anthrax patients are used as input data. Samples are fixed on slides using the Giemsa staining method. 956 images were obtained by imaging different areas of the slides. 80% of these images are used for training and 20% for test. Then, 1021 patches containing bacteria were extracted from the training images and 281 patches from the test images. To evaluate the proposed detection and segmentation models, the extracted patches along with the raw images are used. Some examples of prepared patch images are shown in Fig. 2. As can be seen in the different patches which are extracted from microscopic images, *B. anthracis* bacteria can appear in any shape, angle, and color intensity in the microscopic images. These are the challenges we face in detecting and segmenting bacteria in this research.

In this paper, two DNN architectures are used to segment microscopic images of the tissue sample of anthrax patients. One is UNet architecture, which has been used for segmentation in various fields such as natural and medical images, and the other is UNet++ which has been expanded and adapted to segment medical images by making changes in the UNet architecture.

2.2. UNet

In 2015, UNet was developed to segment medical images. It achieved the best results in two ISBI challenges in the fields of segmentation of neural structures and cell tracking (Ronneberger et al., 2015). This network has also been adapted for new medical applications, including volumetric segmentation of three-dimensional structures and sparse computed tomography reconstructions (Çiçek et al., 2016; Han and Ye, 2018). In this paper, we adapted this architecture for bacteria segmentation in microscopic images for anthrax disease diagnosis.

UNet is a CNN architecture that is similar to a 'U' and is used to quickly and accurately segment images. As shown in Fig. 3, the architecture of this network consists of three main sections: 1- The contraction section or encoder, which consists of a large number of contraction blocks, and each block consists of two 3×3 convolution layers followed by a 2×2 max pooling. The number of filters doubles after each block so that it can learn complex structures of the image effectively. 2- Bottleneck section, which is the interface between the contraction and expansion sections. 3- Expansion section or decoder which includes several expansion blocks and each block includes two convolution layers and one up-sampling layer. The final mapping passes through another layer of convolution with the number of filters equal to the number of desired segments.

2.3. UNet++

Segmentation of medical images requires higher accuracy than the segmentation of natural images, and a small error in the segmentation of medical images leads to a significant reduction in the quality of diagnosis and the stage of the disease. Thus, it is important to use more appropriate architecture for this purpose. Most of the deep network architectures for image segmentation are derived from the UNet. In all of these models, skip connections are used to connect the encoder and

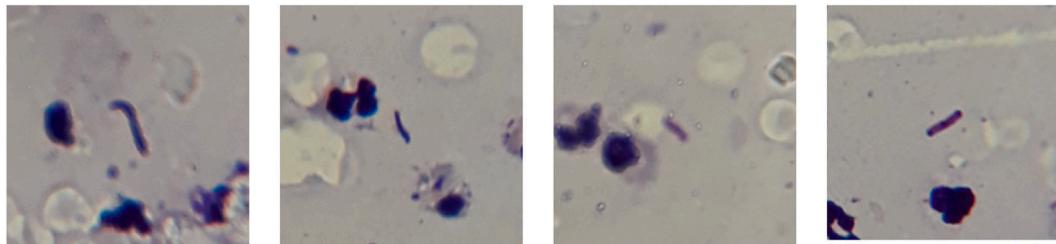


Fig. 2. Some images of the prepared dataset containing *B. anthracis* bacteria.

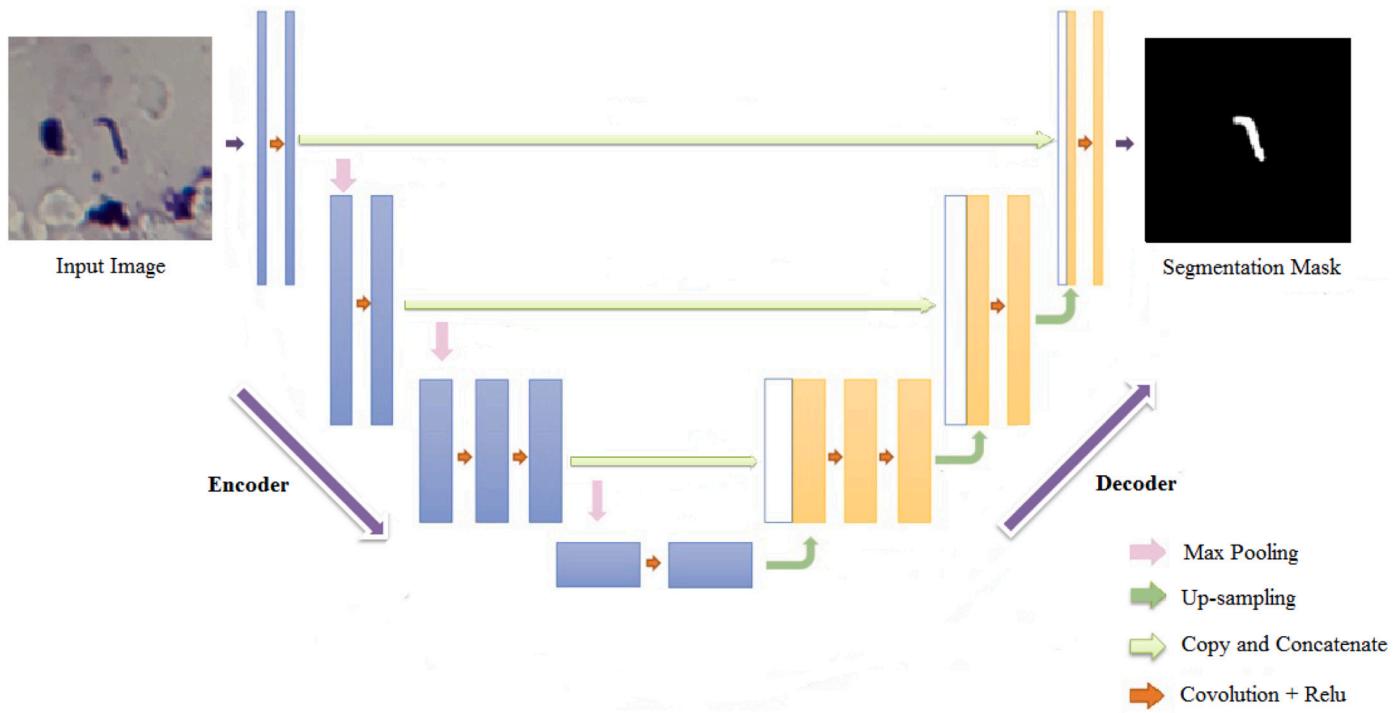


Fig. 3. UNet architecture used for bacteria segmentation.

the decoder. These connections play an important role in retrieving object details and creating complete and accurate segmentation mask.

UNet++ is one of these architectures that we used in this paper for detecting and segmenting bacteria. It is provided with changes in the skip connections section of UNet. The idea behind the UNet++ model is that using dense skip connections, the feature maps of the encoder section will be more semantically similar to the feature maps of the decoder section. In fact, these dense skip connections make it possible to maintain more details of the foreground objects by strengthening the feature maps of the encoder before combining them with the feature maps of the decoder. The advantage of using this architecture is that, unlike UNet, which could only combine feature maps with the same scale, using skip connections, it is possible to receive feature maps with different scales in the decoder and allows the aggregation layer to decide how to integrate feature maps carried by skip connections with

decoder feature maps. Fig. 4 shows the UNet++ architecture in which dense skip connections are shown in black and green and deep supervision are shown in red. For more details on Unet++ refer to (Zhou et al., 2018).

2.4. *B. anthracis* bacteria detection and segmentation

Effective and automatic segmentation of the nucleus/cell or bacterium in microscopic images, as a basic prerequisite for computer-assisted image-based diagnosis, is the basis of many image analyses such as disease diagnosis and progression, diagnosis the type and morphology of the cell, the type of bacteria, etc. (Xing and Yang, 2016).

The purpose of this paper is to diagnose anthrax from microscopic images of tissue samples from patients. So far, no research has been conducted on the automatic diagnosis of anthrax disease and the

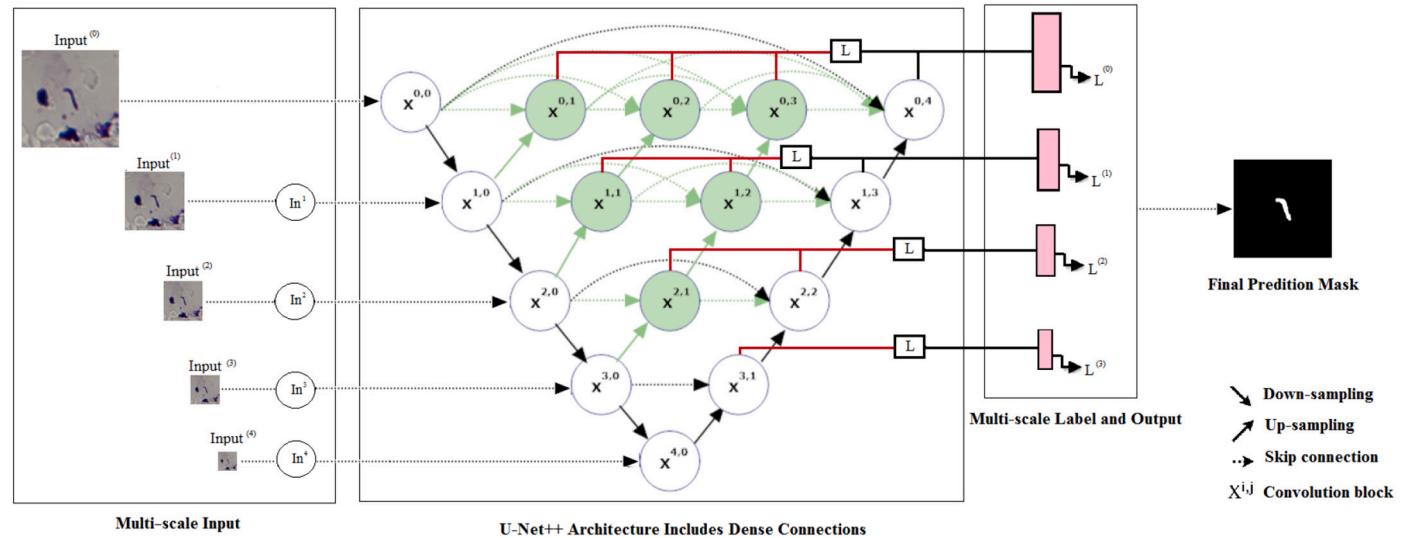


Fig. 4. UNet++ architecture used for bacteria segmentation.

accurate detection and segmentation of the bacterium *B. anthracis* in microscopic images. Previously UNet and Mask-RCNN are used for nuclei segmentation and achieved the best results in Kaggle's 2018 Data Science Bowl (Caicedo et al., 2019).

In this paper, we use UNet and UNet++ to detect and segment *B. anthracis* bacteria of the microscopic images of the tissue sample related to anthrax patients. A challenging dataset has been prepared (Sec 2.1) and all experiments are conducted on it.

2.5. Implementation

2.5.1. Data preparation

Usually, the microscopic images have a high resolution, and due to the limited memory of RAM or GPU, they must be divided into small patches and/or resized to smaller dimensions. Therefore, to perform our experiments on bacteria detection and segmentation, we extracted patches of size 512×512 pixels from the microscopic images containing bacteria. As mentioned earlier in Sec. 2.1, 956 images were obtained by imaging different areas of the slides. 80% of these images are used for training and 20% for test. Then, 1021 patches containing bacteria were extracted from the training images and 281 patches from the test images. The results of applying different network structures on raw images are given in the experimental results section. Due to the high resolution of the raw images (2696×2696), these images are resized to 1536×1536 and then the sliding window technique with a window size of 512×512 and overlap of 50% is used to sweep and test the whole raw image. The window size is selected based on the size of the largest bacteria in all resized images.

To train the DNN for segmentation, we need to annotate data and create a training dataset. A microbiologist conducted this task by detecting and segmenting the area of the *B. anthracis* bacterium using OCLAVI software which provides high-quality medical annotation data.

2.5.2. Pre-processing

Before training the network, pre-processing is applied to the input images in order to increase the generalization power and speed up the training process of the model.

First, we apply the feature normalization to raw microscopic images, which is a common approach in training DNNs and accelerates the convergence of the gradient descent algorithm. Image normalization is performed according to Eq. 1.

$$\tilde{X} = \frac{X - \mu}{\sigma} \quad (1)$$

Where μ and σ are the mean intensity value and standard deviation across all training data. This process applies to both training data and test data.

A common problem in medical image segmentation is the small number of training samples which leads to overfitting on training images and bad segmentation results in unseen data. Therefore, data augmentation is used to increase the size of the training data, and improve the generalization capability of the network. Using data augmentation we can generate more training data without the need of annotation, and this increases the network accuracy.

In this paper, data augmentation includes rotation at 8 different angles and color amplification. Because, the bacteria may appear in the microscopic image with any angle, rotation and horizontal flip (rotation with angle 180°) are used. Rotation is applied with multiples of 45 degrees such as 45, 90, ..., 315 as shown in Fig. 5. As can be seen in the patch images, *B. anthracis* bacteria can appear in any shape and at any angle and any color intensity in the microscopic images. These are the challenges we face in detecting and segmenting bacteria, which can be overcome to a large extent by proper data pre-processing, increasing the size of the database using data augmentation, selecting and adapting appropriate models and precise adjustment of model parameters.

One of the major challenges in microscopic image analysis is the different staining conditions and the different thickness of the slides, which lead to images that look very different. In this paper, to deal with these problems, color augmentation is used to create data with different color patterns. Color augmentation is performed by adding a random mean (between -0.08 and $+0.08$) and multiplying random coefficients (between 0.92 and 1.08) in each color channel of the image and adding a blurring factor. Three different sets of random values are generated for each image. These values must not be too large because the color is a key feature for the classification of the bacteria and large changes may have an adverse effect.

Therefore, using rotation and color augmentation operations for each image, we will obtain 11 images (8 by rotation and 3 by color augmentation) for a single input image. So the total number of training images will be increased 11 times.

2.5.3. Parameter settings

In this section, more details are provided about the parameters of the two deep architectures used for microscopic image segmentation. The initial learning rate is 0.001 which is reduced by 10% per 1000 iterations. The batch size is 5. Stochastic gradient descent with a momentum of 0.9 is used as the optimization method. To prevent overfitting, the correct number of epochs during network training is determined based on validation data and is not determined only by the accuracy or loss of the training data. In this way, the training continues while validation loss is decreasing and when it starts to increase, network training stops.

We tried two structures for each model (UNet and UNet++). Details of these structures, called A and B, are shown in Table 1. Several structures with different depths have been examined also, but these two structures have led to better results. The size of the filters used in all layers is 3×3 and down-sampling is conducted through 2×2 pooling with stride 2. We compared three different pooling techniques: max pooling, average pooling and stochastic pooling (Song et al., 2018).

The number and arrangement of layers in the two structures are similar, but the number of filters used in each layer and the number of feature maps are different. In both models, the dropout technique is used to improve the generalization of the networks.

The second structure, St. B, has 19,533,217 trainable parameters which is much higher than the first structure (St. A) with 5,306,273 parameters. Thus St. B is more computationally complex and has a longer execution time than St. A.

2.6. Evaluation criterion

To evaluate the models, five metrics are employed for verifying the quality of the segmentation results, including the Dice Similarity Coefficient (DSC), Accuracy, Specificity, Precision and Recall criteria according to the Eq. 2 to Eq. 6:

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

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

$$Specificity = \frac{TN}{FP + TN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

TP, FP, TN and FN are the numbers of true positive, false positive, true negative and false negative detections, respectively which are defined as follow:

True Positive (TP): pixels correctly segmented as bacteria.

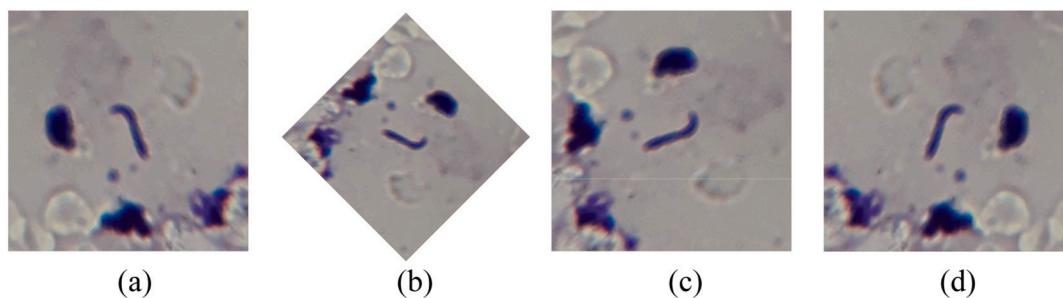


Fig. 5. Data augmentation using image rotation and flipping. (a) Original patch image, (b) Patch Image with 45-degree rotation, (c) Patch image with 90-degree rotation, (d) Patch image with flipping.

False Positive (FP): pixels falsely segmented as bacteria.

True Negative (TN): pixels correctly detected as background.

False Negative (FN): pixels falsely detected as background.

In order to evaluate the models, the segmentation results are compared with the ground truth that is prepared by a specialist with annotating the regions related to the bacteria.

3. Results

3.1. Experimental results

The two proposed models, UNet and UNet++, are implemented in Python using Keras and Tensorflow libraries. The total number of microscopic images was 954. From the total images, 80% was selected for training and 20% for testing. Then, from the training and test images, 1021 and 281 patches of size 512×512 pixels containing bacteria have been extracted, respectively. Furthermore, 18% of the training data

were randomly considered as validation data in order to prevent overfitting. To have better performance, data augmentation was applied as described in Sec. 2.5.B. As said before, we used two structures for both UNet and UNet++ (The details of them are found in Sec. 2.5.C).

In this section, two experiments and their results are given. In the first experiment, the networks are evaluated on the extracted patches (512×512) from test images containing bacteria. In the second experiment, the networks are examined to detect and segment bacteria on the whole raw images.

The segmentation results of a patch of the microscopic image containing the bacterium are depicted in Fig. 6 during network training. These results are for UNet++B. As the number of epochs increases the network performs better and the quality of the segmentation improves. After 250 epochs, the bacterium is detected and segmented with high accuracy.

The segmentation quality of the applied models (using the max-

Table 1

Structures A and B for UNet and UNet++.

Layer	Type	Shape		# Param		Connected to
		St.A	St.B	St.A	St.B	
0	input	(512,512,3)	(512,512,3)	0	0	-
1	lambda	(512,512,3)	(512,512,3)	0	0	input
2	Conv.1 + Drop1	(512,512,16)	(512,512,32)	448	896	lambda
3	Conv.2	(512,512,16)	(512,512,32)	2320	9248	Drop1
4	pool.1	(256,256,16)	(256,256,32)	0	0	Conv.2
5	Conv.3 + Drop2	(256,256,32)	(256,256,64)	4640	18,496	pool.1
6	Conv.4	(256,256,32)	(256,256,64)	9248	36,928	Drop2
7	Pool.2	(128,128,32)	(128,128,64)	0	0	Conv.4
8	Conv.5 + Drop3	(128,128,64)	(128,128,64)	18,496	36,928	pool.2
9	Conv.6	(128,128,64)	(128,128,64)	36,928	36,928	Drop3
10	Pool.3	(64,64,64)	(64,64,64)	0	0	Conv.6
11	Conv.7 + Drop4	(64,64,128)	(64,64,128)	73,856	73,856	pool.3
12	Conv.8	(64,64,128)	(64,64,128)	147,584	147,584	Drop4
13	Pool.4	(32,32,128)	(32,32,128)	0	0	Conv.8
14	Conv.9 + Drop5	(32,32,256)	(32,32,256)	295,168	295,168	pool.4
15	Conv.10	(32,32,256)	(32,32,256)	590,080	590,080	Dropout5
16	ConvTrans.1	(64,64,128)	(64,64,128)	131,200	131,200	Conv.10
17	Concatenate1	(64,64,256)	(64,64,256)	0	0	ConvTrans.1, Conv.8
18	Conv.11 + Drop6	(64,64,128)	(64,64,128)	295,040	295,040	Concatenate1
19	Conv.12	(64,64,128)	(64,64,128)	147,584	147,584	Dropout6
20	ConvTrans.2	(128,128,64)	(128,128,64)	32,832	32,832	Conv.12
21	Concatenate2	(128,128,64)	(128,128,64)	0	0	ConvTrans.2, Conv.6
22	Conv.13 + Drop7	(128,128,64)	(128,128,64)	73,792	73,792	Concatenate2
23	Conv.14	(128,128,64)	(128,128,64)	36,928	36,928	Dropout7
24	ConvTrans.3	(256,256,32)	(256,256,32)	8224	8224	Conv.14
25	Concatenate3	(256,256,64)	(256,256,96)	0	0	ConvTrans3, Conv.4
26	Conv.15 + Drop8	(256,256,32)	(256,256,64)	18,464	55,360	Concatenate3
27	Conv.16	(256,256,32)	(256,256,64)	9248	36,928	Dropout8
28	ConvTrans.4	(512,512,16)	(512,512,16)	2064	4112	Conv.16
29	Concatenate4	(512,512,32)	(512,512,48)	0	0	ConvTrans.4, Conv.2
30	Conv.17 + Drop9	(512,512,16)	(512,512,32)	624	13,856	Concatenate4
31	Conv.18	(512,512,16)	(512,512,32)	2320	9248	Dropout9
32	Conv.19	(512,512,1)	(512,512,1)	17	33	Conv.18

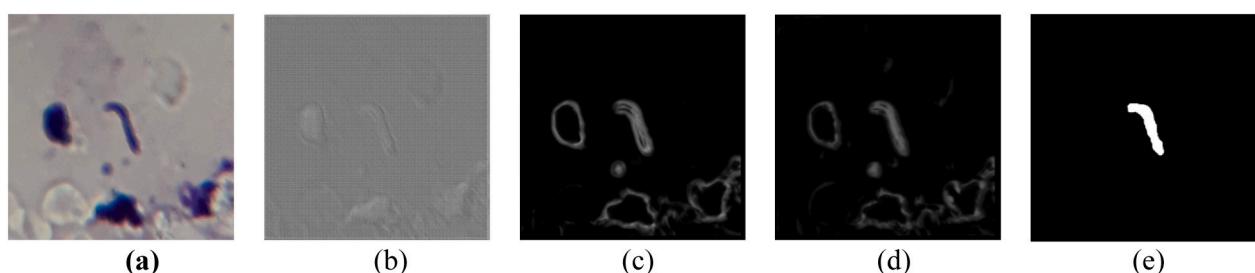


Fig. 6. (a) Original image patch, The segmentation results of a patch of the microscopic image containing the bacterium after (b) one epoch, (c) 25 epochs, (d) 50 epochs and (e) 250 epochs.

Table 2

Evaluation of the bacteria segmentation on the microscopic test patches of size 512×512 with two proposed configurations of UNet and UNet++ using the max-pooling method.

	Structure A		Structure B	
	UNet	UNet++	UNet	UNet++
Recall	0.88	0.90	0.93	0.95
Precision	0.90	0.90	0.91	0.96
Specificity	0.95	0.96	0.96	0.98
DSC	0.89	0.90	0.92	0.96
Accuracy	0.93	0.94	0.96	0.97

pooling method) on the microscopic patch test images based on the mentioned criteria (Recall, Precision, Specificity, DSC, and Accuracy) is evaluated and results are reported in [Table 2](#).

Observing the results in [Table 2](#), we see that both architectures produced good results, but UNet++B performed better in most evaluation criteria and in some respects has been similar to those of UNetB. The UNet++B architecture has the best quality of segmentation, and despite the many challenges in the microscopic images; it has led to excellent results.

The average segmentation time per each patch image is also shown in [Table 3](#) for different structures on CPU (Core i7 7700HQ @ 2.8GHz) and GPU (Nvidia GTX 1070 8GB). This time was obtained by averaging the execution time for the total patch test images, including 281 images. The results show that the lowest run time is for UNetA with 0.023 s and the highest run time is for UNet++B with 0.078 s (on GPU). This is because the number of filters used in the convolution layers of the UNetB and UNet++B structures is more than those in UNetA and UNet++A. Also, in general, the execution time for all structures on GPU is significantly shorter than the CPU.

A comparison of different configuration results based on three evaluation criteria (recall, precision, and specificity) is showed in [Fig. 7](#). The recall in this configuration was 0.95 while the corresponding value for UNetB is 0.93. Based on precision criteria, the UNet++B with a precision of 0.96 has been much better than the UNetB (with a precision of 0.91). According to the specificity criterion, the UNet++B model has achieved the best result among all the criteria. It performed better than other configurations with a specificity of 0.98. Dice score and accuracy are shown in [Fig. 8](#). The results showed that based on the Dice score and accuracy evaluation criteria, the UNet++B with DSC

and accuracy values of 0.96 and 0.97 has performed best in the bacteria segmentation task.

The loss values of the two configurations of the UNet and UNet++ models and the training curve of the UNet++B architecture are depicted in [Fig. 9](#). As shown, the loss value of the UNet++ model in both configurations, A and B, is less than the value in the corresponding configurations of the UNet model. The loss value of UNet++B is the lowest value and this configuration has led to the best result with a loss value of 0.02 ([Fig. 9 \(a\)](#)). Also, as the number of epochs increases, and the network trains more, the amount of loss decreases significantly ([Fig. 9 \(b\)](#)).

To have a qualitative evaluation, some segmentation results for patch test images are shown in [Fig. 10](#). Examining the qualitative results of the various architectures applied, it can be seen that the predictions and ground truth for the configuration B of UNet and UNet++, look pretty similar. The UNetA had the worst performance compared to others.

In the second experiment, we evaluate the performance of the networks to detect and segment bacteria on the whole raw test images. We examined three popular pooling techniques, including average pooling, max-pooling, and stochastic pooling. Due to the proximity of the results of max-pooling and stochastic pooling and for brevity, only the results of average pooling and stochastic pooling are mentioned in this section.

As mentioned in data preparation section, due to the high resolution of the whole raw test images (2696×2696), these images are resized to 1536×1536 and then the sliding window technique with window size of 512×512 and 50% overlap is used to sweep and segment the whole image as shown in [Fig. 11](#). The window size is selected based on the size of the largest bacteria in all resized images.

The segmentation quality of the applied models on the whole microscopic raw images based on the mentioned criteria (Recall, Precision, Specificity, DSC, and Accuracy) is evaluated and results are reported in [Tables 4 and 5](#) for two different pooling methods. Due to the evaluation of the networks on the whole images, the quality of the segmentation is somewhat reduced compared to the first experiment in which the networks were tested only on the extracted patches. However, the segmentation results are still acceptable, especially the second structure of UNet++ using a stochastic pooling method, which produced good results.

[Table 4](#) shows the results of evaluating the structure A on the whole images of both networks based on two different pooling methods. As shown in this table, again the UNet++ performed better than UNet. Furthermore, the stochastic pooling method compared to average pooling leads to better results.

[Table 5](#) displays the results of evaluating the second structure of both networks on the whole raw test images based on stochastic and average pooling methods. Again, the second structure of UNet++ with a stochastic pooling produced the best results among all other structures. Based on specificity criteria, the UNet++B with a specificity of 0.97 has been much better than the UNetB with a specificity of 0.93. According to the recall criterion, the UNet++B model has achieved the best result among all the criteria. It performed better than other

Table 3

Average segmentation time of the proposed structures for test patches of size 512×512 pixels.

	Structure A		Structure B	
	UNet	UNet++	UNet	UNet++
Test-Run Time (CPU)	0.479 s	0.565 s	1.315 s	1.578 s
Test-Run Time (GPU)	0.023 s	0.028 s	0.065 s	0.078 s

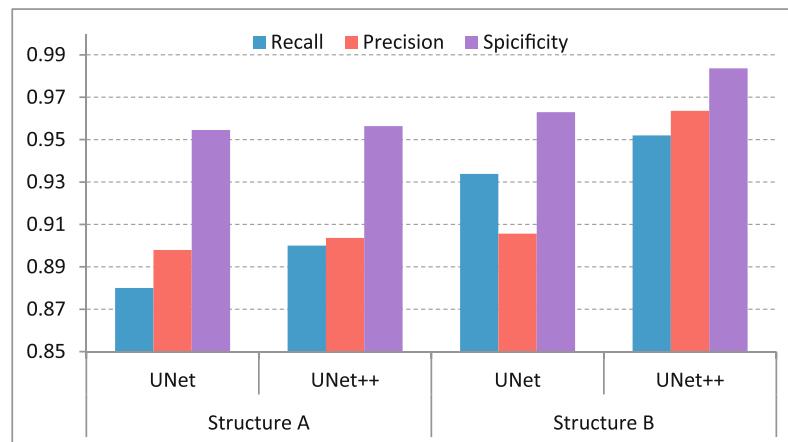


Fig. 7. Comparison of different configuration results for microscopic patch images based on evaluation criteria (recall, precision and specificity).

configurations with a recall of 0.98. Based on the Dice score and accuracy evaluation criteria, the UNet + + B with DSC and accuracy values of 0.92 and 0.97 has performed best in the bacteria segmentation task. These comparisons have been made for the stochastic pooling method, which showed better performance.

The average segmentation time per whole raw image for different structures on CPU and GPU is shown in Fig. 12. Because in the second experiment, the whole raw images are evaluated using the sliding window technique with a 50% overlap, the average test time is increased compared to the first experiment. Also, in general, the execution time for all structures on GPU is significantly shorter than the CPU. The results show that UNetA with 0.7 s has the lowest run time and the highest run time is for UNet + + B with 1.95 s (on GPU).

The loss values of the two configurations of the UNet and UNet + + models are shown in Fig. 13. As shown, stochastic pooling leads to lower loss values compared to average pooling, meanwhile the loss value of the UNet + + model in both configurations, A and B, is less than the value in the corresponding configurations of the UNet model. The loss value of UNet + + B with stochastic pooling is 0.029 which is the lowest value among other models.

In Fig. 14 we have a qualitative evaluation of the models on the whole raw images, especially in more challenging images, such as crowded images or images with overlapping bacteria. Considering that in the previous experiments, the structure B of both networks led to better segmentation results, in this figure, only the segmentation results of the UNetB and UNet + + B are shown, in which two different pooling methods (average pooling and stochastic pooling) are used. Examining the qualitative results of the various architectures, it can be seen that the predictions and ground truth for the UNet + + B with stochastic pooling, look pretty similar. The UNetB with average pooling had the worst performance compared to others.

The 3rd and 4th columns of the figure are related to the crowded images with overlapping bacteria. Nevertheless, the UNetB and UNet + + B with stochastic pooling, have been able to well segment the region related to the two overlapping bacteria. One of the cases in which the UNetB with average pooling failed to detect and segment the bacteria, but the UNet + + B succeeded, is shown in the second column. For the raw image in the 1st column, all the structures have performed well.

In some microscopic images, filaments of tissue may appear elongated and thin, or there may be image artifacts that look like bacteria in that the algorithm mistakenly identifies as bacteria. Fig. 15 shows two misclassified images with these features in which even the best structure (UNet + + B with stochastic pooling) failed to correctly segment and mistakenly identified two non-bacteria areas as bacteria.

3.2. Discussion

In this study, we proposed a novel UNet and UNet + + based method for detection and segmentation of *B. anthracis* bacteria from patients' tissue microscopic images for anthrax disease diagnosis. Despite the challenges, such as the high resolution of input images, object crowding, and overlapping, our method achieved very good results. Based on the Dice score and accuracy criteria, the UNet + + B showed excellent performance in the bacteria segmentation task. So, this method could be considered as an automated anthrax diagnosis. In addition, because the number of bacteria is inherently low in the wound and microscopic vision, the automatic localization of *B. anthracis* would facilitate the difficulty of microscopic diagnosis. In the present study, we achieved 95% recall and 96% precision in the segmentation of *B. anthracis* in patch test images. Also when dealing with the whole raw test images we achieved a recall of 98% and a precision of 87%. The lower precision in whole raw images is because the number of *B.*

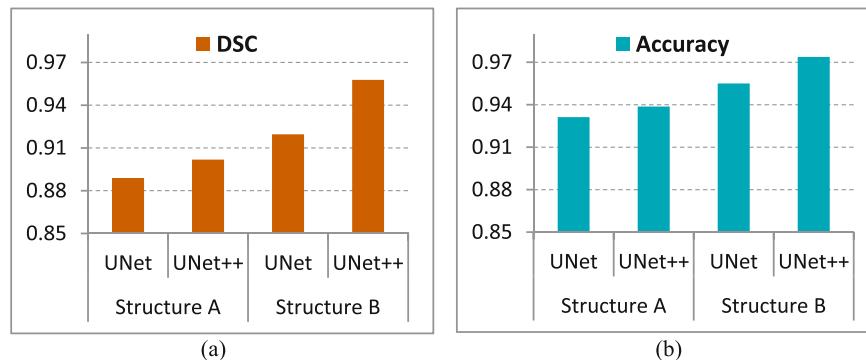


Fig. 8. Comparison of different configuration results for microscopic patches based on a) Dice Score and b) Accuracy.

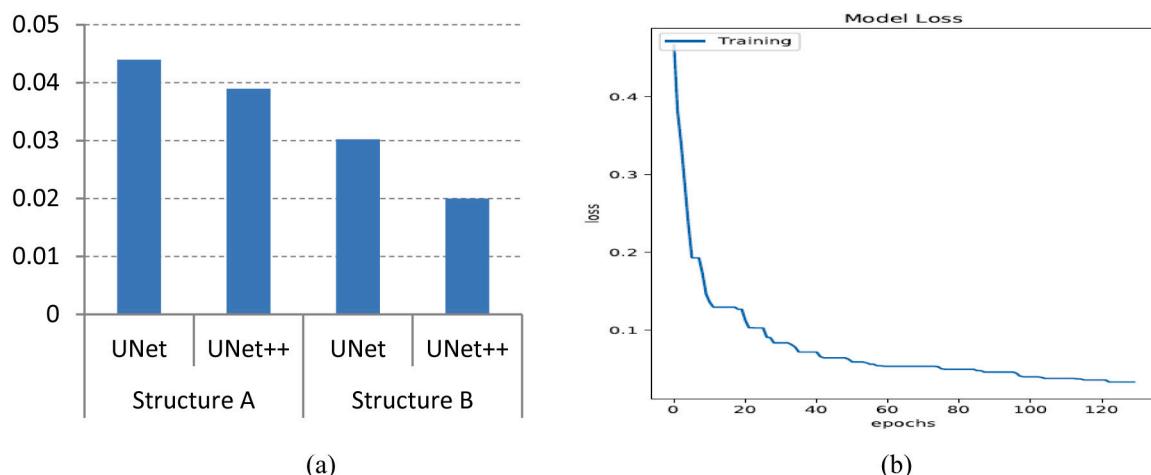


Fig. 9. a) Loss values of the two proposed models b) Loss curve of the UNet + + B during the training phase.

anthracis segments are very low concerning the whole image.

There is no similar study in the automation of anthrax microscopic diagnosis but we can compare our results with the other studies based on other microbial automated diagnoses. For example, Kant et al. achieved the recall and precision of 83.78% and 67.55% respectively for tuberculosis (TB) bacteria or *Mycobacterium tuberculosis* detection from microscopy images (Kant and Srivastava, 2018).

Mengying et. al represented the first study on automatically detecting *M. tuberculosis* in sputum smear scans obtained by a high-resolution slide scanning system. In this study, three CNN models (Inception v3, ResNet, and DenseNet) were used for training and testing on a dataset of 2630 sputum smear microscopic images, and the

accuracy and reliability of the models were verified. The metrics of Accuracy, Precision, Sensitivity, F1 score, ROC curve and AUC were utilized to evaluate the performance of the models systematically. The experimental results got the best performances on Inception v3, where all indicators were 98.4%. This study also has some limitations. Considering the characteristics of the sputum smear scans, they used subgraph classification instead of object detection, which greatly reduced the workload. However, when there are more than one bacteria in a subgraph, the predicted amount of *M. tuberculosis* may be less than the actual value. This error rate can be decreased by choosing a proper subgraph size (Hu et al., 2019).

In another research, Panicker et al. achieved the recall and precision

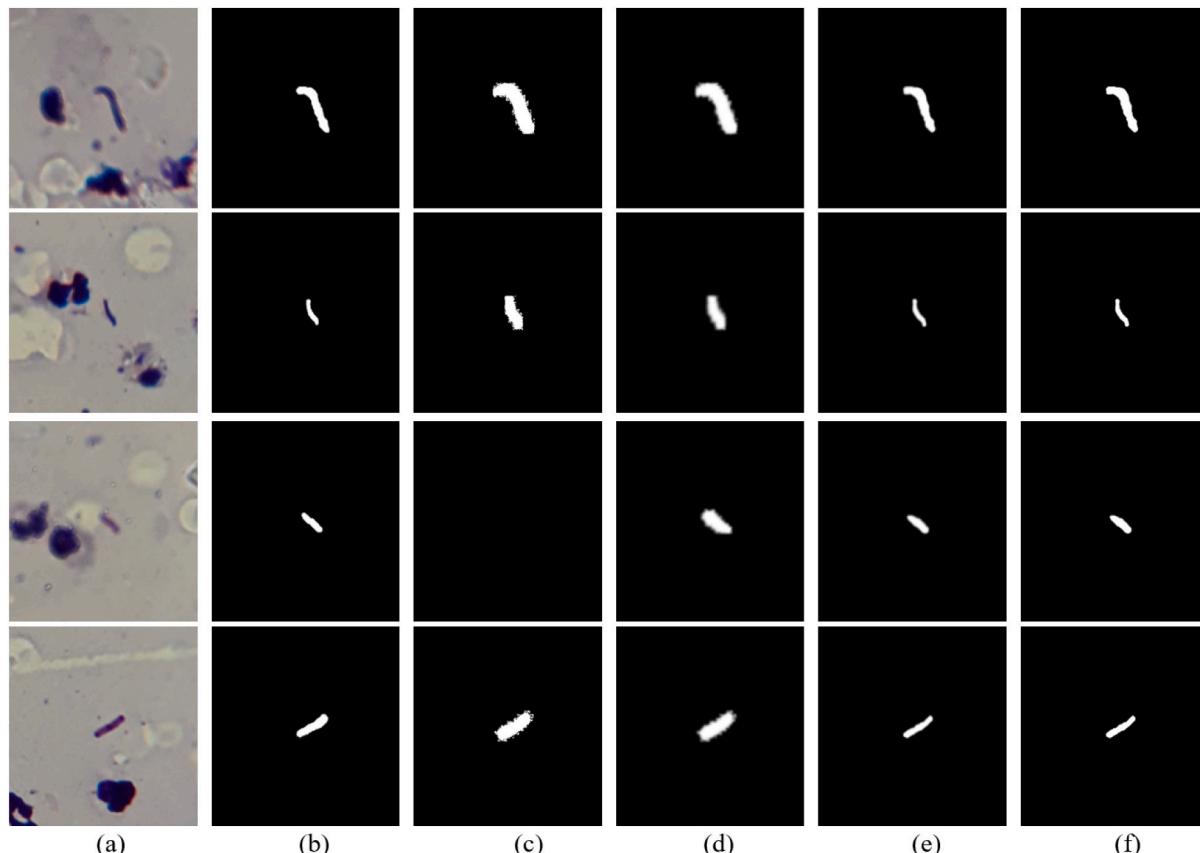


Fig. 10. Qualitative results for patch test images. a) Input patch images, b) Ground truth annotations, c) UNetA outputs, d) UNet + + A outputs, e) UNetB outputs, f) UNet + + B outputs.

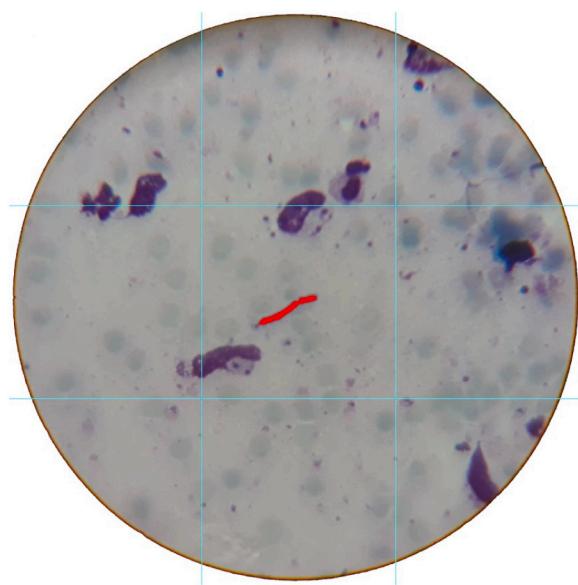


Fig. 11. An example of a tiled image to evaluate different architectures on the whole raw images.

Table 4

Evaluation of the bacteria segmentation of the first configuration of UNet and UNet++ (structure A) on the whole test images in terms of recall, precision, specificity, dice score, and accuracy.

UNet A		UNet++ A	
Average Pooling	Stochastic pooling	Average Pooling	Stochastic pooling
Recall	0.84	0.85	0.87
Precision	0.46	0.49	0.52
Specificity	0.80	0.82	0.84
DSC	0.59	0.62	0.65
Accuracy	0.81	0.83	0.84

Table 5

Evaluation of the bacteria segmentation of the second configuration of UNet and UNet++ (structure B) on the whole test images in terms of recall, precision, specificity, Dice score, and accuracy.

UNet B		UNet++ B	
Average Pooling	Stochastic pooling	Average Pooling	Stochastic pooling
Recall	0.94	0.95	0.97
Precision	0.71	0.74	0.82
Specificity	0.92	0.93	0.96
DSC	0.80	0.83	0.89
Accuracy	0.92	0.94	0.96

of 97.13% and 78.4% for tuberculosis detection from microscopy images (Panicker et al., 2018). In (Górriz et al., 2018) a UNet-based method for automatic segmentation of Leishmania parasites into three classes including promastigotes, amastigotes, and adhered parasites was presented. Evaluation of this method in terms of Dice score, precision, recall, and F1-score at the best state (for the amastigote class) resulted in 77.7%, 75.7%, 82.3%, and 77.7% respectively.

To identify the characteristics of Acute Lymphoblastic Leukemia (ALL), Alaa et al used a set of images of microscope blood samples that have been obtained from the “Oncology Center-Faculty of Medicine-Mansoura University Hospital-Egypt”. It is made up of 50 infected microscopic samples and 50 normal microscopic samples of the blood. The microscope blood images are exposed to a series of pre-processing steps

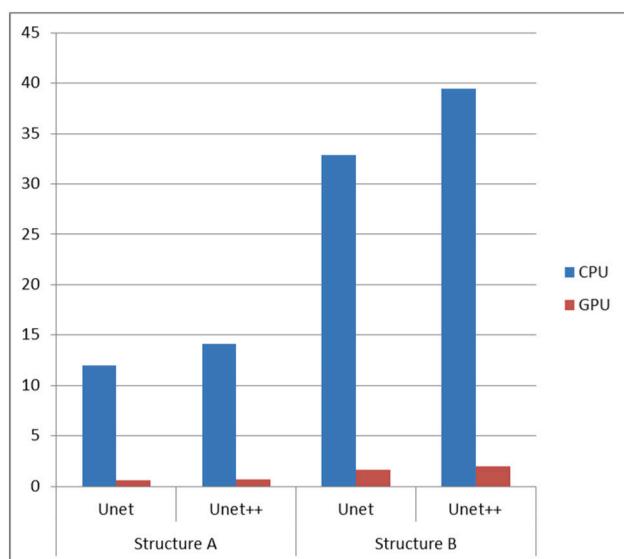


Fig. 12. Average test time of the bacteria segmentation on the whole raw test images.

which include image resizing to 512×512 and 256×256 , and contrast enhancement. By executing K-means clustering on the resultant images, the nuclei of cells under consideration are obtained. Shape features, texture features, color features, and Hausdorff dimension are then extracted for classification. A backpropagation neural network was employed for classification. The system was evaluated using a dataset of 45 infected images and 45 non-infected images and achieved an accuracy of 99.74%. The system has a low computational requirement that makes it well suited for low-cost hardware implementation (Ahmed et al., 2016).

Fatin et al proposed an intelligence identification and counting system to detect the presence of *M. tuberculosis* in the ZN-stained smear sputum image. This system is designed to identify the presence of *M. tuberculosis* and count the number of *M. tuberculosis* by applying digital image processing and artificial intelligence techniques. The image processing technique consists of contrast enhancement, segmentation, and feature extraction. The contrast of the original image was enhanced by the combination of global enhancement, local enhancement, and Contrast Limited Adaptive Histogram Equalization (CLAHE). Then, the enhanced image was segmented using color thresholding. The features they used, consists of 18 color features, 15 shape features, and 5 texture features. Afterward, the features underwent feature selection to select the relevant features by using Neighborhood Component Analysis (NCA) and ReliefF Analysis. The study showed that there more relevant features chosen by ReliefF include 8 color features, 11 shape features, and 3 texture features. For the classification process, Multilayer Perceptron (MLP) trained by Scaled Conjugate Gradient (SCG), k-Nearest Neighborhood (k -NN), and Support Vector Machine (SVM) are used with 6 folds cross-validation. It was found that MLP has the highest accuracy, sensitivity, and specificity with 93.8%, 93.4%, and 94.1% respectively (Rosli et al., 2019).

Another research (Mohamed et al., 2020) aims to automatically identify and classify WBCs in a microscopic image into four types with high accuracy. This model combines the advantage of deep models in automatically extracting features with the higher classification accuracy of traditional machine learning classifiers. The proposed model in this paper consists of two main stages: a shallow tuning pre-trained model and a traditional machine learning classifier on top of it. In this study, ten different pre-trained models with six types of machine learning are used. Moreover, the FCNN of pre-trained models is used as a baseline classifier for comparison. The evaluation process shows that the hybrid of MobileNet-224 as a feature extractor and logistic

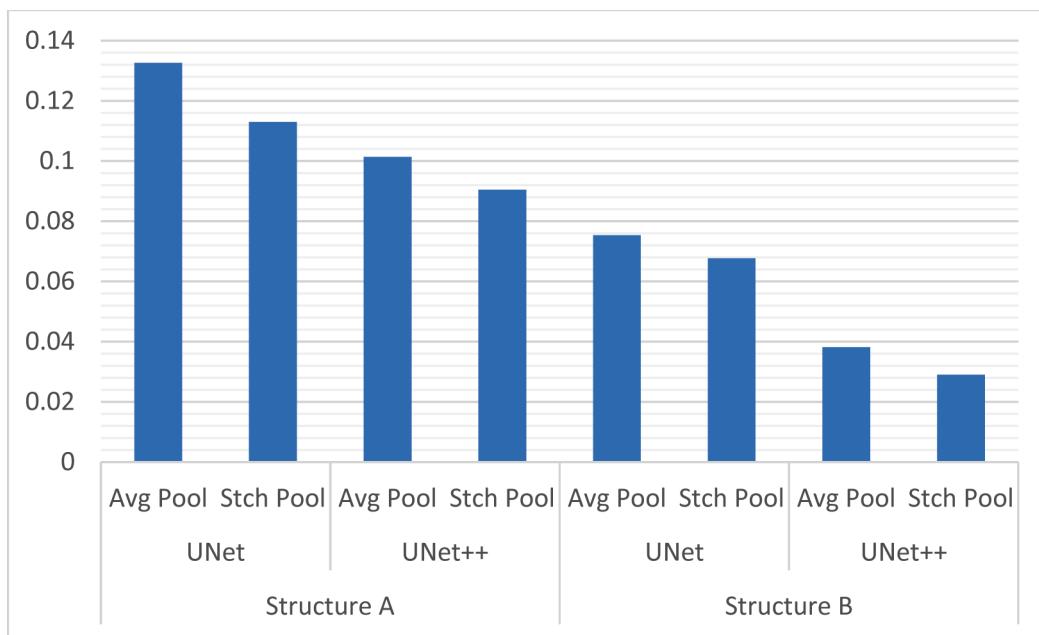


Fig. 13. Loss values of the two configurations of the UNet and UNet + + models using two different pooling methods.

regression as classifier has a higher rank-1 accuracy of 97.03%. Besides, the proposed hybrid model outperformed the baseline FCNN by 25.78% on average.

The motive of Swetha et al. study is to categorize the sputum images based on *M. tuberculosis* present in it and to lessen the dependency of technicians and pathologists through automated detection. In this work, the pre-processing is performed on ZN based sputum images by using noise reduction and intensity modification. Segmentation is based totally on the Channel Area Thresholding (CAT) that is the multilevel-thresholding method. Furthermore, features inclusive of Histogram of Oriented Gradients (HOG) and Speed up Robust Feature (SURF) are extracted. Classification is performed using CNN classifier. It classifies the *M. tuberculosis* as mild, moderate, and severe primarily based on the wide variety of pixels with 99.5% accuracy, 94.7% sensitivity, and 99% specificity (Swetha et al., 2020).

Song et al. proposed an automatic method to diagnose bacterial vaginosis in microscopic images and achieved a sensitivity, specificity, and accuracy of 58.3%, 87.1% and 79.1% (Song et al., 2016). Besides, the average running time per microscope image on CPU was around 30s, while the average execution time for test data in our research for UNet + + B is much less (1.578 s) on CPU. These results confirm that the current study has an excellent performance in the bacteria detection and segmentation task and outperforms previous works with higher speed.

Furthermore, in Mithra et. al study the accuracy of 97.5% was achieved in automated identification of *M. tuberculosis* (Mithra and Emmanuel, 2019). According to this and other studies, our accuracy result that is 97% could be considered as an acceptable index in automated microscopic infection diagnosis.

Also, our results compared with the other related studies confirmed that in a very short time (0.078 s for patch images and 0.7 s for the whole raw images on GPU) with the automated method we would be able to diagnose microscopic infectious agents, dozens of bacteria, causing important diseases with the difficult diagnosis microscopically.

4. Conclusion

Given that microscopic diagnosis of diseases by humans due to fatigue and decreased visual acuity may lead to decreased accuracy and increased error rate. In this paper, a novel UNet and UNet + + based

method for detection and segmentation of *B. anthracis* bacteria in microscopic images for anthrax disease diagnosis is proposed. We start by data preparation and pre-processing stages consisting of patch normalization and artificially data augmentation by rotating, flipping, and color amplification. The reason for using data augmentation is to increase the size of training data without the cost of annotating the images and simultaneously reducing the network overfitting and increasing the generalization power of the model. The dropout technique has also been used to increase the generalizability of the system.

We used two architectures to segment microscopic images of the tissue sample of anthrax patients. One is UNet architecture, which has been used for segmentation in various fields such as natural and medical images, etc., and the other is UNet + + , which has been expanded and adapted to segment medical images by making changes in UNet architecture.

We evaluated the adapted models on a database consisting of 956 microscopic images (1302 image patches) which are obtained by imaging from the tissue sample of anthrax patients on slides. The images obtained under different conditions.

Based on experiment results (on patch images and whole raw images in two different experiments), it has been shown that these two DNNs, especially UNet + + can be highly effective in the diagnosis of anthrax disease and bacterial detection and segmentation in the microscopic images of the tissue sample of anthrax patients which obtained under different conditions, with many challenges in this field. *B. anthracis* bacterial automated detection and segmentation system offers almost the same level of accuracy as the human diagnostic specialist and in some cases outperforms it. The low cost and high speed of diagnosis and no need for a specialist doctor are other benefits of the proposed system.

Authorship statement

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript. Furthermore, each author certifies that this material or similar material has not been and will not be submitted to or published in any other publication before its appearance in the Hong Kong Journal of

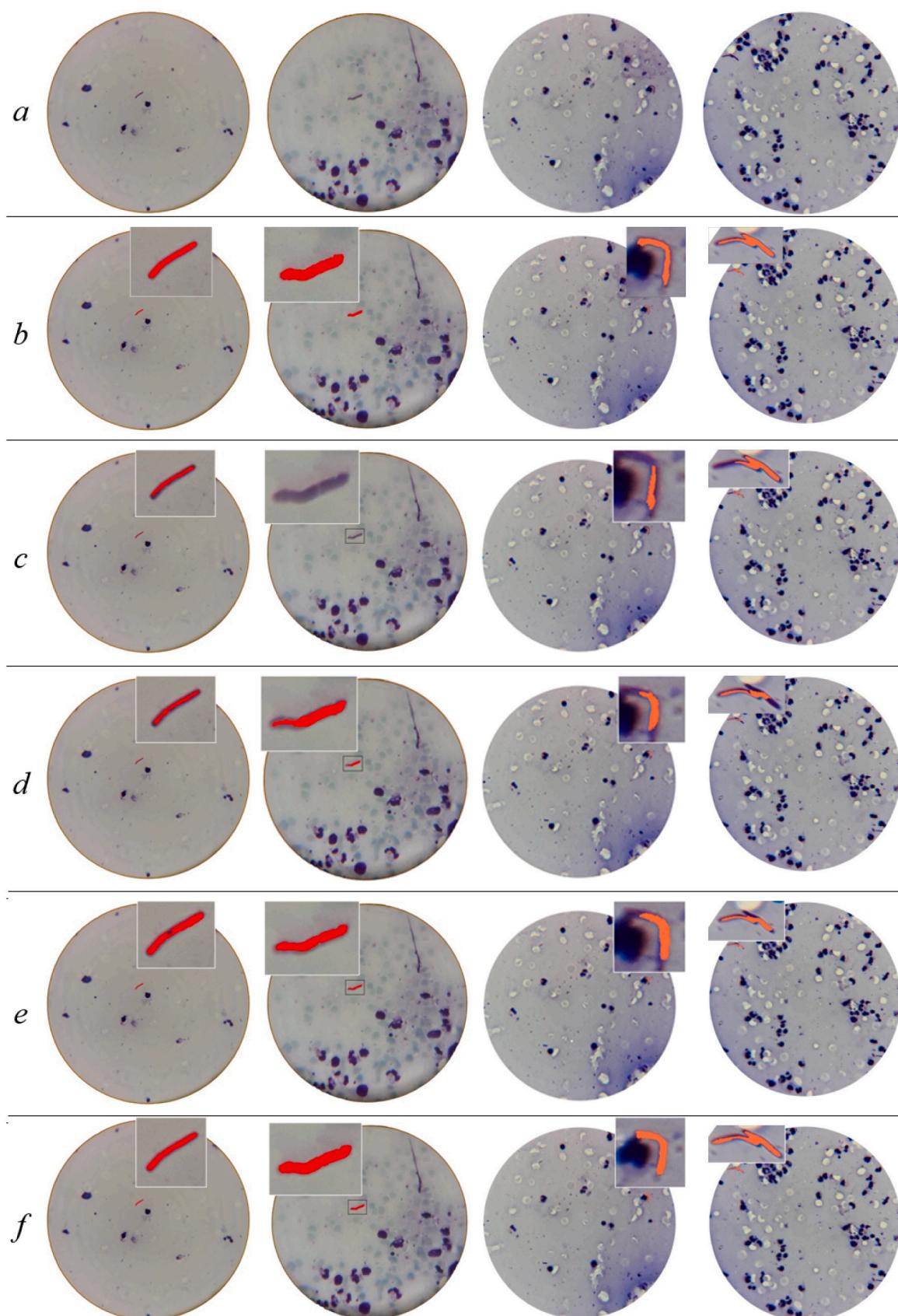


Fig. 14. Qualitative results for the whole raw test images. a) Input raw images, b) Ground truth annotations, c) UNetB outputs with average pooling, d) UNet++B outputs with average pooling, e) UNetB outputs with stochastic pooling, f) UNet++B outputs with stochastic pooling.

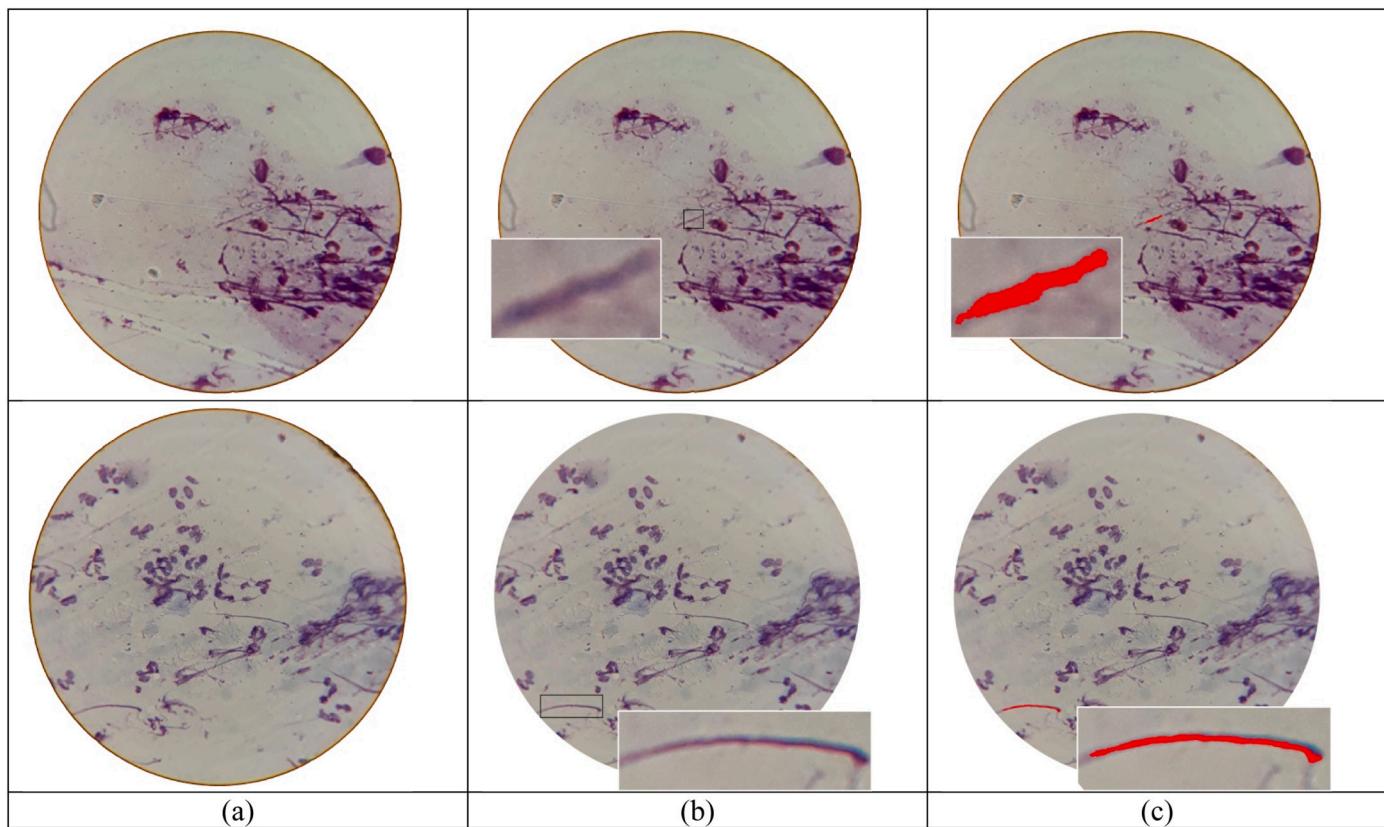


Fig. 15. Two examples of misclassified images. a) Input raw images, b) Magnified view of the part of the image that most closely resembles bacteria, c) UNet++B outputs (using stochastic pooling).

Occupational Therapy.

Authorship contributions

Category 1.

Conception and design of study: Fatemeh Hoorali, Hossein Khosravi. acquisition of data: Bagher Moradi, Fatemeh Hoorali. analysis and/or interpretation of data: Fatemeh Hoorali, Hossein Khosravi, Bagher Moradi.

Category 2.

Drafting the manuscript: Fatemeh Hoorali. revising the manuscript critically for important intellectual content: Hossein Khosravi, Bagher Moradi.

Category 3.

Approval of the version of the manuscript to be published (the names of all authors must be listed):

Fatemeh Hoorali, Hossein Khosravi, Bagher Moradi.

Declaration of Competing Interest

I, Dr. Hossein Khosravi, on behalf of other authors of the paper entitled “Automatic *Bacillus anthracis* Bacteria Detection and Segmentation in Microscopic Images Using UNet++”, declare that we have no financial or personal relationships with other people or organizations that could inappropriately influence our work.

Acknowledgements

All persons who have made substantial contributions to the work reported in the manuscript (e.g., technical help, writing and editing assistance, general support), but who do not meet the criteria for authorship, are named in the Acknowledgements and have given us their

written permission to be named. If we have not included an Acknowledgements, then that indicates that we have not received substantial contributions from non-authors.

References

- Ahmed, A.S., Morsy, M., Abo-Elsoud, M.E.A., 2016. Microscopic digital image processing of acute leukemia. *Int. J. Eng. Res. Technol.* 5 (01).
- Andermatt, S., Pezold, S., Cattin, P., 2016. Multi-dimensional gated recurrent units for the segmentation of biomedical 3D-data. In: Deep Learning and Data Labeling for Medical Applications. Springer, pp. 142–151.
- Arganda-Carreras, I., et al., 2015. Crowdsourcing the creation of image segmentation algorithms for connectomics. *Front. Neuroanat.* 9, 142.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (12), 2481–2495.
- Bologna, J., Jorizzo, J., Schaffer 3rd, J., 2012. Dermatology, 3rd edition. Elsevier Health Sciences Publisher ISBN-13: 9780723435716. Cysts," ed. 2012.
- Caicedo, J.C., et al., 2019. Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nat. Methods* 16 (12), 1247–1253.
- Chen, J., Yang, L., Zhang, Y., Alber, M., Chen, D.Z., 2016. Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation. In: Advances in neural information processing systems, pp. 3036–3044.
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4), 834–848.
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: International conference on medical image computing and computer-assisted intervention. Springer, pp. 424–432.
- Ciresan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J., 2012. Deep neural networks segment neuronal membranes in electron microscopy images. In: Advances in neural information processing systems, pp. 2843–2851.
- de Vos, B.D., Wolterink, J.M., de Jong, P.A., Viergever, M.A., Işgum, I., 2016. 2D image classification for 3D anatomy localization: employing deep convolutional neural networks. In: Medical imaging 2016: Image processing. 9784. pp. 97841 Y: International Society for Optics and Photonics.
- Doganay, M., Metan, G., Alp, E., 2010. A review of cutaneous anthrax and its outcome. *J. infect. Public Health* 3 (3), 98–105.
- Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C., 2016. The importance of

- skip connections in biomedical image segmentation. In: Deep Learning and Data Labeling for Medical Applications. Springer, pp. 179–187.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT press.
- Górriz, M., Aparicio, A., Raventós, B., Vilaplana, V., Sayrol, E., López-Codina, D., 2018. Leishmania parasite segmentation and classification using deep learning. In: International Conference on Articulated Motion and Deformable Objects. Springer, pp. 53–62.
- Greenspan, H., Van Ginneken, B., Summers, R.M., 2016. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans. Med. Imaging* 35 (5), 1153–1159.
- Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B., 2009. Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* 2, 147–171.
- Han, Y., Ye, J.C., 2018. Framing U-net via deep convolutional framelets: application to sparse-view CT. *IEEE Trans. Med. Imaging* 37 (6), 1418–1429.
- Hosseini, M.M., Safdari, R., Shahmoradi, L., Javaherzadeh, M., 2017. Better diagnosis of acute appendicitis by using artificial intelligence. *ISMJ* 20 (4), 339–348.
- Hu, M., Liu, Y., Zhang, Y., Guan, T., He, Y., 2019. Automatic detection of *tuberculosis* Bacilli in sputum smear scans based on subgraph classification. In: 2019 International Conference on Medical Imaging Physics and Engineering (ICMIP). IEEE, pp. 1–7.
- Irshad, H., Veillard, A., Roux, L., Racocianu, D., 2013. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE Rev. Biomed. Eng.* 7, 97–114.
- Kant, S., Srivastava, M.M., 2018. Towards automated tuberculosis detection using deep learning. In: 2018 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, pp. 1250–1253.
- Kong, B., Zhan, Y., Shin, M., Denny, T., Zhang, S., 2016. Recognizing end-diastole and end-systole frames via deep temporal regression network. In: International conference on medical image computing and computer-assisted intervention. Springer, pp. 264–272.
- Korez, R., Likar, B., Pernuš, F., Vrtovec, T., 2016. Model-based segmentation of vertebral bodies from MR images with 3D CNNs. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 433–441.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431–3440.
- Maska, M., et al., 2014. A benchmark for comparison of cell tracking algorithms. *Bioinformatics* 30 (11), 1609–1617.
- Miao, S., Wang, Z.J., Liao, R., 2016. A CNN regression approach for real-time 2D/3D registration. *IEEE Trans. Med. Imaging* 35 (5), 1352–1363.
- Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV). IEEE, pp. 565–571.
- Min, S., Lee, B., Yoon, S., 2017. Deep learning in bioinformatics. *Brief. Bioinform.* 18 (5), 851–869.
- Mithra, K., Emmanuel, W.S., 2019. Automated identification of *Mycobacterium tuberculosis* from sputum images for tuberculosis diagnosis. *SIVIP* 13 (8), 1585–1592.
- Moeskops, P., et al., 2016. Deep learning for multi-task medical image segmentation in multiple modalities. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 478–486.
- Mohamed, E.H., El-Behaidy, W.H., Khoriba, G., Li, J., 2020. Improved white blood cells classification based on pre-trained deep learning models. *J. Commun. Sof. Syst.* 16 (1), 37–45.
- Nie, D., Zhang, H., Adeli, E., Liu, L., Shen, D., 2016. 3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients. In: International conference on medical image computing and computer-assisted intervention.
- Springer, pp. 212–220.
- Ning, F., Delhomme, D., LeCun, Y., Piano, F., Bottou, L., Barbano, P.E., 2005. Toward automatic phenotyping of developing embryos from videos. *IEEE Trans. Image Process.* 14 (9), 1360–1371.
- Panicker, R.O., Kalmady, K.S., Rajan, J., Sabu, M., 2018. Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. *Biocybern. Biomed. Eng.* 38 (3), 691–699.
- Payer, C., Štern, D., Bischof, H., Urschler, M., 2016. Regressing heatmaps for multiple landmark localization using CNNs. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 230–238.
- Poudel, R.P., Lamata, P., Montana, G., 2016. Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. In: Reconstruction, segmentation, and analysis of medical images. Springer, pp. 83–94.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer, pp. 234–241.
- Rosli, F.A., Mashor, M.Y., Noor, S.S.M., 2019. An automated intelligent identification and counting system procedure for tuberculosis. *J. Phys. Conf. Ser.* 1372 (1), 012079 IOP Publishing.
- Saedi, S.I., Khosravi, H., 2020. A deep neural network approach towards real-time on-branch fruit recognition for precision horticulture. *Expert Syst. Appl.* 113594.
- Setio, A.A.A., et al., 2016. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans. Med. Imaging* 35 (5), 1160–1169.
- Simonovsky, M., Gutiérrez-Becker, B., Mateus, D., Navab, N., Komodakis, N., 2016. A deep metric for multimodal registration. In: International conference on medical image computing and computer-assisted intervention. Springer, pp. 10–18.
- Song, Y., et al., 2016. Segmentation, splitting, and classification of overlapping bacteria in microscope images for automatic bacterial vaginosis diagnosis. *IEEE J. Biomed. Health Inform.* 21 (4), 1095–1104.
- Song, Z., et al., 2018. A sparsity-based stochastic pooling mechanism for deep convolutional neural networks. *Neural Netw.* 105, 340–345.
- Stollenga, M.F., Byeon, W., Liwicki, M., Schmidhuber, J., 2015. Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation. In: Advances in neural information processing systems, pp. 2998–3006.
- Swetha, K., Sankaragomathi, B., Thangamalar, J.B., 2020. Convolutional neural network based automated detection of *Mycobacterium Bacillus* from sputum images. In: 2020 International Conference on Inventive Computation Technologies (ICICT). IEEE, pp. 293–300.
- Veta, M., Pluim, J.P., Van Diest, P.J., Viergever, M.A., 2014. Breast cancer histopathology image analysis: a review. *IEEE Trans. Biomed. Eng.* 61 (5), 1400–1411.
- Xie, Y., Zhang, Z., Sapkota, M., Yang, L., 2016. Spatial clockwork recurrent neural network for muscle perimysium segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 185–193.
- Xing, F., Yang, L., 2016. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. *IEEE Rev. Biomed. Eng.* 9, 234–263.
- Yang, A., Mullins, J.C., Van Ert, M., Bowen, R.A., Hadfield, T.L., Blackburn, J.K., 2020. Predicting the geographic distribution of the *Bacillus anthracis* A1. A/Western North American sub-lineage for the continental United States: new outbreaks, new genotypes, and new climate data. *Am. J. Trop. Med. Hyg.* 102 (2), 392–402.
- Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J., 2018. Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 3–11.