Custom semantic segmentation neural network architecture in spirochaete detection application

Michał Wieczorek[1,*,†], Natalia Wojtas[2,†]

[1] Faculty of Applied Mathematics, Silesian University of Technology, Kaszubska 23, 44100 Gliwice, Poland
[2] Faculty of Veterinary Medicine, University of Life Sciences in Lublin, Stanisława Leszczyńskiego 7, 20400 Lublin, Poland

Abstract

The kingdom of bacteria is a very diverse group of organisms characterized by high phenotypic variability. This feature is often used in clinical diagnosis. The spirochaetes are a microbes with a characteristic spiral shape of a flagella located within the periplasmic space. Nowadays, there is a high demand for creating a rapid and sensitive method for their detection as many of them performs a high pathogenic risk. Currently used methods lays on combination of clinical examination results, serologic and cultivation methods. There can be also used Polymerase Chain Reaction (PCR) method if needed. Unfortunately this combination can be very time consuming and require a lot of money. This research presents a novel, semantic segmentation neural network architecture designed to quickly create a classification mask, outputting information about the position, shape, and possible affiliation of detected elements. The evaluation method is based on a light microscope imagery and was created to overcome above mentioned problems. Used abstract classes contains erythrocytes, spirochaete and background. The resulted mask can be later mapped to a human-readable form with the inclusion of colors, next to an original image. Such approach allows for semi-automatic recognition of unwanted objects, however still giving the final verdict to the specialist. Developed solution has achieved a high recognition accuracy, while the computer power requirements are kept at a minimum.

The proposed solution can help reduce misclassification rates by providing additional data for the doctor and speed up the entire process with the early diagnosis made by a neural network.

Keywords

Spirochaete, detection, mask, semantic segmentation, neural network

1. Introduction

The spirochetes are a phylum of mostly free living, anaerobic, motile bacteria. Those prokaryotes are large and long spirals. Their shape is slender, helically coiled, spiral, or corkscrew-like [1]. Those gram-negative bacterias contain a distinctive double membrane. Their lengths vary between 3 and 500 m, diameter: 0.09 - 3 m [2]. Beneath the outer membrane, they own a flagella, which number can be highly variable - from 2 in Spirochaeta to more than 300 in Cristospira [3]. The feature that distinguishes them from other phyla is the flagella/axial filament's location - on each pole of the bacteria, within the periplasmic space [4]. During everyday life, veterinary doctors often encounter those bacterias, as many of them produce highly dangerous diseases. Examples can be leptospirosis, lyme boreliosis, treponematoses or brachyspira species, producing swine dysentery. Many of them are zoonotic factors, like Leptospira interrogans, producing flu-like symptoms, renal and hepatic damage and exhibiting serious risk both for wild and domestic animals, especially dogs and their human owners [5]. The most often used methods for detecting those pathogens are the serologic and polymerase chain reaction (PCR) methods. Unfortunately, they are often quite expensive and not available directly in the clinic. In more complicated cases, when there is a time and the owner can afford it the bacterial cultivation can be performed. Often during the routine clinical examination there is performed the blood sampling for haematology and biochemistry evaluation. It gives a chance for a quick and easy accomplishment of executing the blood smears. This may allow the doctor to perform the microscopic method of blood evaluation and classify the pathogen visually, however currently this method is not used as a standard. One of the reasons can be a huge variability among the microbes and often very little optical differences between them. There is also a need for a specific staining for understanding what type of a bacteria the doctor deals with and a cost of specific chemicals. The example can be the spirochaetes, that under the microscope resemble the wiggly hairs and easily may be mistaken with trypanosomes, some protists and other bacteria with a similar shape. The achievement of a direct and quick result may be also influenced by the risk of human mistakes as a result of tiredness, inaccuracy and lack of time and a special interest in this field of medicine. This is why, in our paper we propose a quick and simple method of evaluating the presence of spirochaete bacteria, that may

IVUS 2022: 27th International Conference on Information Technology
*Corresponding author.
† These authors contributed equally.
michal_wieczorek@hotmail.com (M. Wieczorek); natjia@wp.pl (N. Wojtas)

speed up the diagnosis, give the doctors a valuable clue and save money and health of the patient. This can be especially valuable for busy, first contact clinics as a primary method of evaluating the presence of spirochaete microorganisms, whose presence can be predicted after a basic physical examination, before some more detailed and expensive tests.

2. Proposed deep learning solution

The concept of microbe detection can be approached using different various techniques, however some of the most accurate methods include visual classification. Normally such process is performed by a human specialist and contains manual checking of hundreds of objects within previously selected frames per patient. This process is extremely slow and requires full focus of the doctor for the whole time in order to avoid misclassification and oversight.
As in the nature of such examination are high repeatability and small amount of additional stimulus, it is very difficult for any human being to maintain the peak detection performance during the whole process, especially considering some external factors like tiredness, small amount of time, very little visual differences between microbes or lack of special interest in the field.
Above mentioned problems lead to high average detection error rate.

More common deep learning techniques contain rectangle masking, such as in [6] and [7], however for this task the output needs to be more precise. Because of that as a partial solution to this issue in this research a custom semantic segmentation neural network architecture has been created. It provides additional data, in the form of a mask with initial elements classification, to the doctor next to the original image for easy and fast verification. Such approach can highly reduce error rate by providing additional diagnosis and pointing out suspicious elements, as well as speed-up the entire diagnosis process by reducing the time needed to analyse the image.
Additionally, by choosing segmentation architecture over the classical rectangle masking one, the classification is made per-pixel and thus there is clarity and accuracy improvement on images with higher amount of overlapping objects.

Final architecture is based on the U-Net shape and the final parameters were selected empirically. Final shape is presented in Fig. 3. The input layer has a shape of 256x256 and is reduced by max-pooling by a factor of 2 after each block. The internal signal addition within the convolutional blocks contains several branches with different layers count and combinations of batch normalization layers. The final block output is combined using concatenate and add layers for better signal fidelity. Sample scheme is presented in Fig. 4. The training has been optimized using NAdam algorithm with learning rate of 0.0078 and the selected loss function was Categorical Cross-entropy with custom class weights computed before training to balance the training.

2.1. NAdam algorithm

To improve model's performance in terms of final accuracy performance and the training times the NAdam training algorithm has been used. The formula can be described as follows:

$$z_s = \gamma_1 z_{s-1} + (1-\gamma_1) g_s, \tag{1}$$

$$k_s = \gamma_2 k_{s-1} + (1-\gamma_2) g_s^2, \tag{2}$$

where γ parameters are constant hyper-parameters and g is the current gradient value of an error function. Values z_s and k_s are used later for computing the correlations marked as $\hat{z}_s$ and $\hat{k}_s$ according to below equations:

$$\hat{z}_s = (1-\gamma_1) g_s + \gamma_{1_{s+1}} z_s \tag{3}$$

$$\hat{k}_s = \frac{k_s}{1-\gamma_2^s}. \tag{4}$$

Finally, using previously calculated variables, the final formula can be defined as:

$$w_s = w_{s-1} - LR \frac{\hat{z}_s}{\sqrt{\gamma_{2_s}} + \epsilon} \tag{5}$$

where ϵ is a small, constant value and LR is a learning rate.

3. Training dataset

There were several factors needed to be taken into consideration while searching for the dataset:

- The data have to contain microscopy imagery of both microbes and healthy cells,
- The dataset has to be free for academical use,
- The images need to have appropriate masks.

During the research phase the most suitable one, containing masked images of the spirochaete microorganisms mixed with the red blood cells was the "Bacteria detection with darkfield microscopy" dataset gathered and annotated as part of a bachelor thesis of university Heilbronn, Germany. The dataset contains

Algorithm 1 NAdam training process

```
Generate random weights,
while global error value ε < error_value do
    Shuffle the training dataset,
    for each batch inside training dataset do
        Compute gradient vector g on the batch,
        Update vector m eq. (1),
        Update vector v eq. (2),
        Rescale vector m̂ eq. (3),
        Rescale vector v̂ eq. (4),
        Update variable ŵ_t eq. (5).
        Step = Step + 1,
    end for
    Calculate global error ε,
end while
```

366 images from the darkfield microscopy with manually created masks labelling 3 abstract classes: background, spirochaete and erythrocytes.

3.1. Data augmentation

Although the data are high quality and each image consists of many microbes and red blood cells, the number of training examples is relatively small to train a highly accurate model without the use of data augmentation. After several trials including variety of simple image transforms, as well as state of the art methods based on Generative Adversarial Networks (GAN), such as Principal Component Resampling presented in [8], the best combination in this case includes horizontal and vertical flip, random image rotation and random zooming.

4. Model's performance

4.1. Used hardware

During this research all computations were made on a PC with specification below:

- CPU: Ryzen Threadripper 2950X 16c/32t,
- RAM: 128GB,
- GPU: NVidia RTX 3090 24GB.

4.2. Performance

During this research one of the main goals was to create not only an accurate model but also to reduce its memory and power requirements to the bare minimum. Such approach is crucial, as it helps to spread the use of similar models in real-world applications. Although computer hardware is becoming more powerful each year, very little people, especially in smaller clinics are able to update it fast enough and even when updating, the costs needs to be small so most of the time there is only a mediocre CPU with small amounts of RAM and integrated GPU. Very often those computers are also laptops.

With that in mind some compromises has been made, mainly on the training length side, however after the final reduction the model consists of 7,921,534 parameters and weights around 94MB. The evaluation times are below 0.1 second on the GPU and around 0.87 second on the CPU per image.
The training plots are presented in Fig. 1.

5. Results visualization

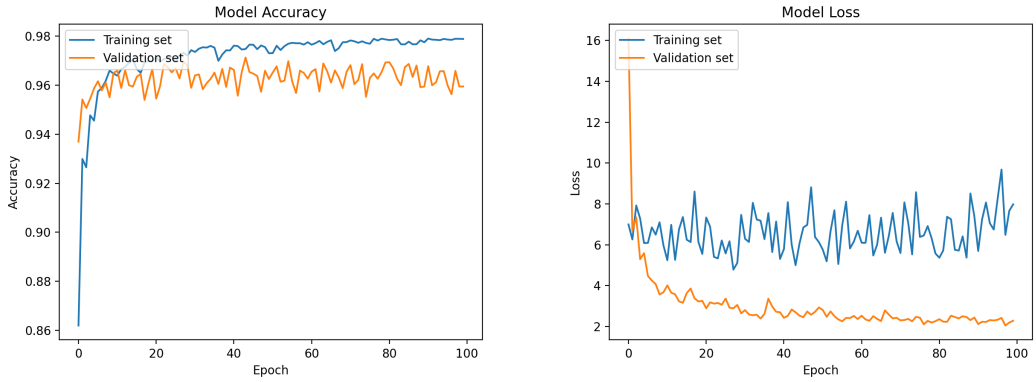
The network originally outputs the data in a form of a two dimensional matrix with sparse representation of classes using integer values. Such data are optimal for being stored and analyzed by the computer, however presents no useful value for the non-technical user and requires further processing to create an informative image. That's why, in order to make it readable, the matrix has been expanded by 3 additional color channels and integer values from the $[0, 2]$ range has been mapped to red, green and blue channels. Based on basic human psychology blue has been chosen as a background, green as harmless blood cells and red as dangerous microbes. Such prepared mask is presented next to the original image for fast and easy validation by the user.

Other methods of visualization has been considered, such as merging both original and mask into one image, however the level of clarity has been highly reduced and the validation became much more difficult as some of the original data has been compromised.

Sample final results produced by the network can be seen in Fig. 2.
In the above examples there can be seen that, although the network in some cases struggles to find the exact shape of the microbe or the red blood cell, it is still able to perform really well in most cases, even the more extreme ones where the image is not the highest quality, there is high amount of overlapping elements, the contrast is very low or the microbe is small relative to the whole image.

6. Conclusion

This paper presents a novel solution for fast spirochaete detection using a custom Semantic Segmentation Neural Network. The output is presented in a clear and easy to understand way, also allowing for quick validation if



(a) Accuracy Plot

(b) Loss Plot

Figure 1: Training plots

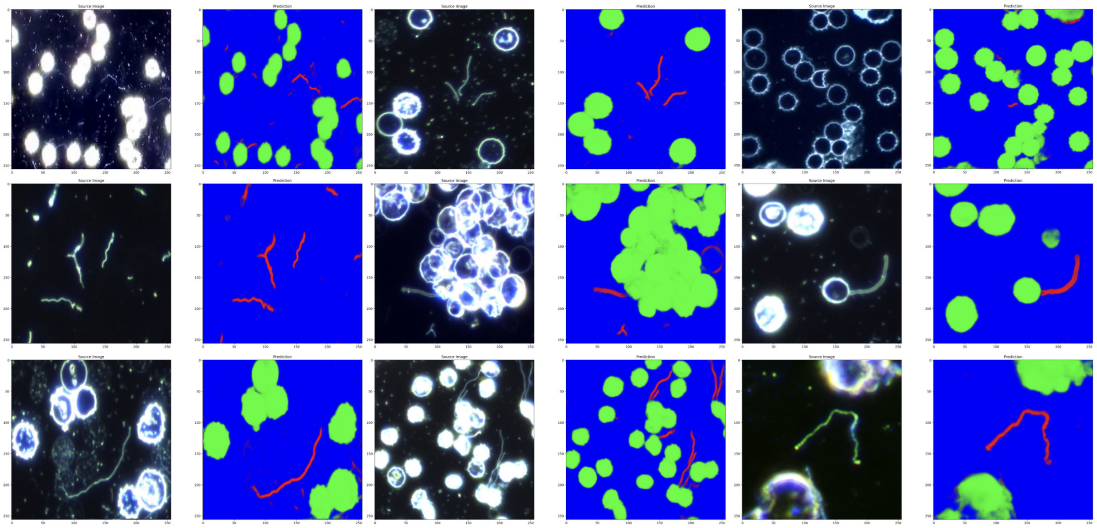


Figure 2: Results visualization

needed. Although the training has been performed on a powerful GPU, the evaluation could be also done on a CPU from the budget level computer, making it accessible for almost everyone. Presented solution is able to speed up the detection time by providing additional data to the original image, helping the human performing the evaluation spot potential bacteria. This could not only allow for testing more animals at the same time but also drastically reduce the costs of such operation.

7. Future possibilities

In the future there are many paths of improvements both in terms of functionality and accuracy performance. One of them is expanding the current dataset with new images captured on more diverse conditions, such as more noisy backgrounds, different bacteria shapes, lower contrast ratio between elements, etc. This approach would lead to much higher accuracy on validation data as the network will understand the wider context, thus the feature extraction should work correctly on more cases than the current model. Another way of improving the network would be by not only adding more images to the training

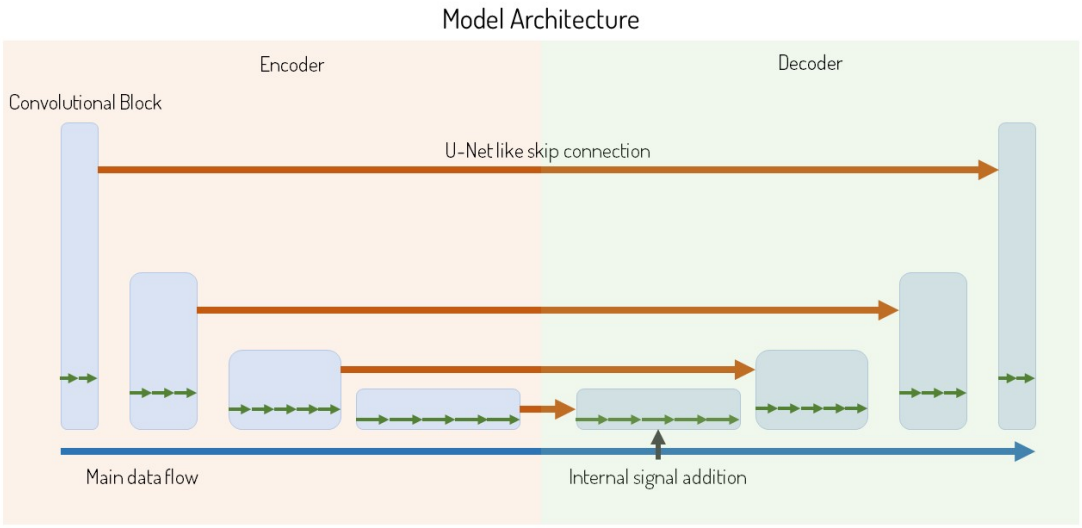


Figure 3: Deep learning model scheme

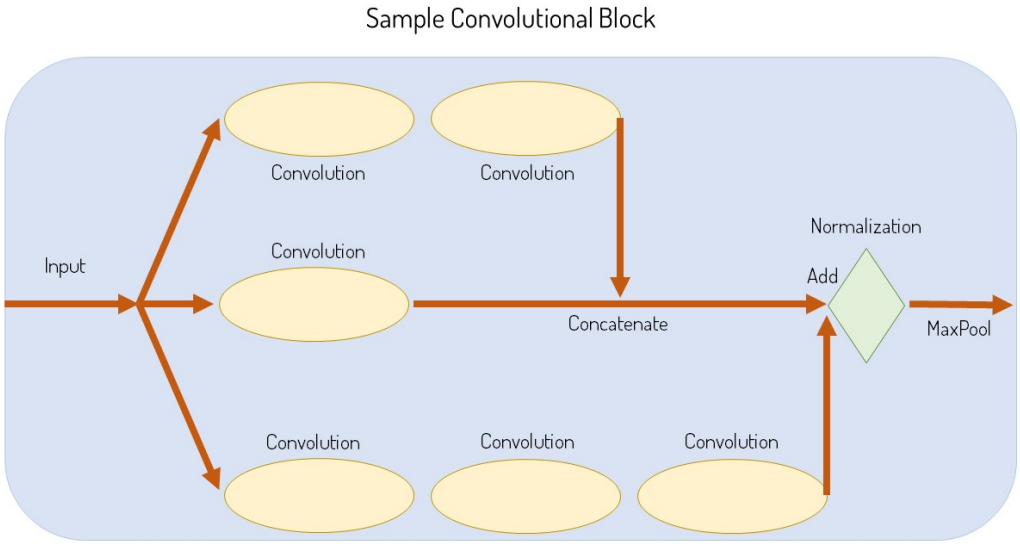


Figure 4: Sample block model scheme

set but also by expanding the number of abstract classes providing examples of other microbes. This would lead to better understanding of the world by the model but could also reduce the misclassification rate. To better fit the potential new dataset some changes in the Deep Learning architecture could be necessary to further improve the accuracy.

References

[1] E. Jawetz, J. Melnick, E. Adelberg, G. Brooks, J. Butel, L. Ornston, Spirochetes and other spiral microorganisms, Medical microbiology, 18th ed. Appletion and Lange, Norwalk, Conn (1989) 267–271.

[2] L. Margulis, J. B. Ashen, M. Sole, R. Guerrero, Composite, large spirochetes from microbial mats: spirochete structure review, Proceedings of the National Academy of Sciences 90 (1993) 6966–6970.

[3] K. H. Hougen, A. Birch-Andersen, Electron microscopy of endoflagella and microtubules in treponema reiter, Acta Pathologica Microbiologica Scandinavica Section B Microbiology and Immunology 79 (1971) 37–50.

[4] M. Madigan, J. Martinko, P. Dunlap, D. Clark, Brock biology of microorganisms 12th edn. microbiol. 2008; 11: 65-73, 2019.

[5] D. C. Alexander, P. N. Levett, C. Y. Turenne, Molecular taxonomy, in: Molecular Medical Microbiology, Elsevier, 2015, pp. 369–379.

[6] M. Wieczorek, J. Sika, M. Wozniak, S. Garg, M. Hassan, Lightweight cnn model for human face detection in risk situations, IEEE Transactions on Industrial Informatics (2021).

[7] M. Woźniak, J. Siłka, M. Wieczorek, Deep learning based crowd counting model for drone assisted systems, in: Proceedings of the 4th ACM MobiCom Workshop on Drone Assisted Wireless Communications for 5G and Beyond, 2021, pp. 31–36.

[8] O. O. Abayomi-Alli, R. Damaševičius, M. Wieczorek, M. Woźniak, Data augmentation using principal component resampling for image recognition by deep learning, in: International Conference on Artificial Intelligence and Soft Computing, Springer, 2020, pp. 39–48.