



Detection of Bacillus spp. Bacteria using Combination of Faster R-CNN and ResNet-50

Ahmed Adnan Badr^{1*}, Thekra Haider Ali Abbas², Mohammed Fadhel AboKsour³

ABSTRACT

Bacteria play an important role in human life since they influence all aspects of life, from vital processes within the human body to the production of medical drugs and vaccines, as well as food production. As this stage is considered one of the basic stages in the diagnosis process, the Bacilli shape is one of the basic forms of bacteria that microbiologists use medical microscopes to diagnose. The purpose of this Article is to develop a bacilliform diagnosis system that employs a pre-trained ResNet-50 algorithm as a feature extraction layer to train the Faster R-CNN detector model. DIBaS (Digital Images of Bacteria Species) dataset, is a public dataset containing 33 different types of bacteria used in training and validating the system. The proposed system achieved 98.99% mini-batch accuracy and 99.10% validation accuracy.

Keywords: Deep learning, ResNet-50, Faster R-CNN, Bacteria, Bacillus spp.

DOI Number: 10.14704/nq.2022.20.8.NQ44908

NeuroQuantology 2022; 20(8): 8858-8867

8858

1. Introduction

Bacteria are single-celled microscopic organisms that are among the oldest known forms of life on Earth. Thousands of distinct varieties of bacteria live in every location and medium on the planet, including soil, seawater, and deep below the earth's crust [1]. They can be classified into three types based on their shape: round (called cocci), linear (called bacilli), and spiral (called spirochetes).

Microbiologists rely on microscopes to identify bacteria shapes, which makes this task time-consuming and effort-intensive and may, in some cases, be inaccurate if the test sample is exposed to contamination. This process is also very important for identifying the type of bacteria [2].

Furthermore, various tools were developed to do automatic bacterial identification based on biochemical testing [3]. These tools are also costly, and they cannot be conducted without the necessary equipment and supplies.

Deep learning has shown promising capabilities in the design of various systems in recent years, including systems for human activity recognition and object detection in images and surveillance cameras [4].

This study used the deep learning methodology and the capabilities of convolutional neural networks to create a system for identifying Bacillus spp. bacteria that intends to decrease the effort and time required by traditional methods while increasing the accuracy of

results for microbiology experts.

Faster R-CNN is a modern deep learning technique that divides an image into various region proposals using a region proposal network (RPN) [5]. This division seeks to guide the training model to regions with a high probability of containing the required object. The feature extraction convolutional layer has been changed by a pre-trained ResNet-50 algorithm that was trained on the ImageNet dataset to extract the features of different objects and shapes [6]. The purpose of using pre-trained algorithms is to decrease the amount of time and effort necessary for system development. Instead of training the algorithm from scratch to solve a given problem, developers use other algorithms that have already been trained on the same or similar problems, which is known as transfer learning [7].

Based on the above, a system for detecting Bacillus spp. bacteria is proposed, which use the DIBaS (Digital Image of Bacterial Species) dataset for training and validation. The proposed system is based on the pre-trained ResNet-50 algorithm for feature extraction and Faster R-CNN algorithm for detecting Bacillus spp. bacteria in images. Data augmentation method is used to cover the lack of available data. The rest of the article is organized as follows:

- The second section is a quick summary of the related works that made the diagnosis of bacteria samples a study topic.

Corresponding author: ^{1*} Ahmed Adnan Badr

Address: ^{1,2} Dept. of Computer Science, Collage of Science, Al-Mustansiriyah University, Baghdad, Iraq.

^{1*} E-mail: ahmed_albndaoui@uomustansirivah.edu.iq

eISSN 1303-5150



- The third section is about feature extraction using the pre-trained ResNet-50 algorithm.
- The fourth section: includes an explanation of the Faster R-CNN layers.
- The fifth Section: includes details about the proposed system implementation.
- The sixth section: explains the experimental results reached by the proposed system.
- Finally, the conclusions of this study.

2. Related works

In recent years, many studies have emerged that have taken the diagnosis of bacteria samples as a topic for research, as there are many techniques and methods in this field, some of which have adopted machine learning techniques, as in B. Zieliński et al. [8], who used the Support Vector Machine (SVM) and Random Forest (RF) algorithms to classify the images of bacterial samples in the DIBaS dataset. Y. He et al. [9] identify three different bacteria species (*E. coli*, *M. semegmatis*, and *P. aeruginosa*) using three different classifiers: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF). J. R. Balbin et al. [10] adopt a different approach as they use an electronic nose with a Support Vector Machine (SVM) to detect (*Enterococcus faecalis*, *Escherichia coli*, and *Staphylococcus aureus*) bacteria in street foods. On the other hand, there is a study that used neural networks to cover the lack of data, as in Y. Du et al. [11], who used a generative adversarial network (GAN) to amplify the data of three foodborne bacteria collected by Raman spectroscopy for classification using a multiclass support vector machine (SVM).

The deep learning approach has a significant role in the design of bacterial diagnostic systems, such as T. Beznik et al. [12] employ U-Net to classify bacteria samples into two groups based on blood agar fermentation for vaccine production: viral strains (bvg+) and avirulent strains (bvg-). T. Treebupachatsakul et al. [13] classify two types of bacteria (*Staphylococcus aureus* and *Lactobacillus delbrueckii*) using LeNet. Ö.F. Nasip et al. [14] use VggNet and AlexNet to classify 33 bacterial species in the DIBaS dataset.

Finally, there are studies that used data mining methods or hybrid methods that combined deep

learning and machine learning, as in X. Gu et al. [15], where a fully convolutional network (FCN) was used in segmentation process and a Deep Convolutional Neural Network (DCNN) to extract features and then fed into a binary Support Vector Machine (SVM) classifier. E. Bonah et al. [16] use a support vector machine (SVM) that optimized using "genetic algorithm (GA)", "particle swarm optimization (PSO)", and "grid searching (GS)" algorithms to classify five types of pathogenic bacteria in food. A. RAJAEI et al. [17] use data mining to diagnose *E. coli* bacteria in urine samples of kidney patients. The Fuzzy C-Mean clustering (FCM) and Differential Evolution (DE) algorithms are used to classify the activity of bacteria into three categories: normal, active, and semi-active. Table 1 shows the summary of the related works.

8859

3. ResNet-50 for features extraction

ResNet-50 is one of the Residual Networks versions, and its name refers to the number of layers that comprise it, as seen in Fig. 1 [18]. The motive behind the design of residual networks is to overcome the "vanishing gradient" problem that afflicts shallow networks [19], by using "Residual Connections", thus enabling the construction of deeper networks.

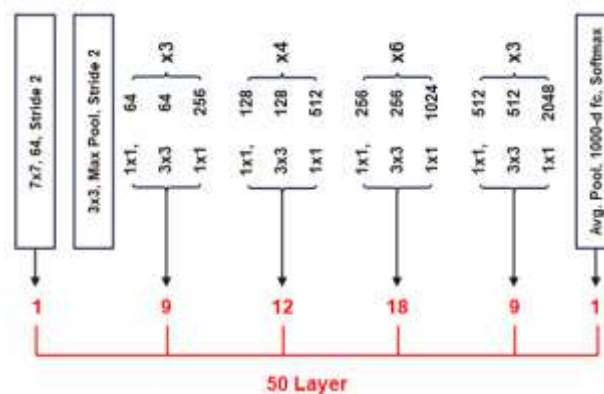


Figure. 1 ResNet-50 layers

In traditional Forward Propagation neural networks, data travels sequentially through each layer. By skipping some layers, residual connections

Table 1. Related works summary

| Authors | Methods | Dataset | Accuracy |
|----------------------------------|----------------|--------------------|----------|
| B. Zieliński et al. [8] | SVM, RF | DIBaS | 97% |
| Y. He et al. [9] | SVM, KNN, RF | N-SIM | 98% |
| J. R. Balbin et al. [10] | SVM | ----- | ----- |
| Y. Du et al. [11] | GAN, SVM | Raman spectroscopy | 90% |
| T. Beznik et al. [12] | U-Net | GSK Labs. | 96.1% |
| T. Treebupachatsakul et al. [13] | LetNet | Collected Dataset | 75% |
| Ö.F. Nasip et al. [14] | VggNet | DIBaS | 98.25%, |
| | AlexNet | | 97.53% |
| X. Gu et al. [15] | FCN, DCNN, SVM | Collected Dataset | 80% |
| E. Bonah et al. [16] | PSO-SVM | E-nose | 98.95% |
| | GA-SVM | | 96.87% |
| | GS-SVM | | 94.79% |
| A. RAJAEI et al. [17] | FCM, DE | Imam Ali Hospital | ----- |

8860

give alternative way for data to reach later parts of the neural network [20] as shown in Fig. 2.

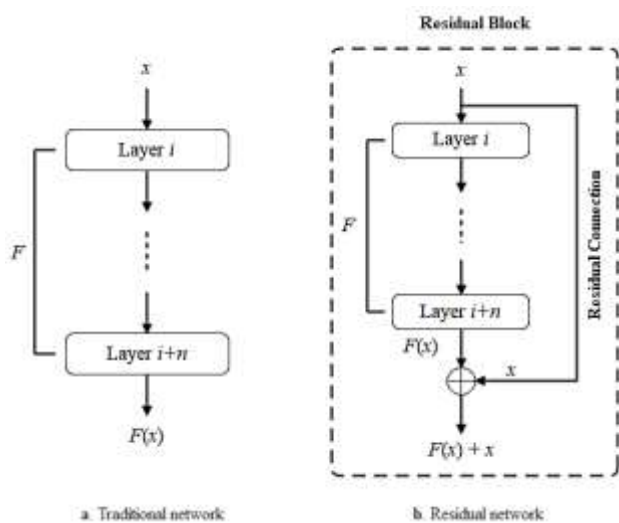


Figure. 2 Residual connection

Residual networks apply the principle of residual connections by using two types of blocks (identity blocks and convolution blocks) [20]. The use of each of them is due to the size of the outputs and inputs for each layer. If $F(x)$ and x have the same dimensions, then the Identity Block is used, which simply combines x and $F(x)$, but in the case of a difference, the convolution block is used by passing x through the (1×1) convolution layer to reduce its dimensions and then combine it with $F(x)$ [20]. The whole architecture that takes an input x and delivers an output $F(x) + x$ is commonly referred to as a Residual Block or a Building Block [21].

The ResNet-50 method was used to extract the features in the proposed system, as it was previously trained on the ImageNet dataset, which contains millions of images of various objects and shapes. This algorithm can extract features for more than 1000 objects and shapes. The use of pre-trained algorithms adheres to the principle of transfer learning, which allows pre-trained algorithms to be used on a problem similar to the current situation rather than training a new algorithm from scratch. [22].



4. Faster R-CNN

Faster Region-based Convolutional Neural Network (Faster R-CNN) is a modern algorithm that works by dividing the feature map into several region proposals (Anchor Boxes) of varying scale and aspect ratio [5]. This division seeks to direct the training model to regions with a high probability of containing the object to be diagnosed. This algorithm's architecture consists of four levels, from bottom to top, as shown in Fig. 3:

- Level for feature extraction and feature maps generation.
- Level for generating region proposals using Region Proposal Network (RPN).
- Level that uses both the proposed regions and the feature map to generate the feature vector by Region of Interest (RoI) Pooling Layer.
- And finally, Fast R-CNN classifier.

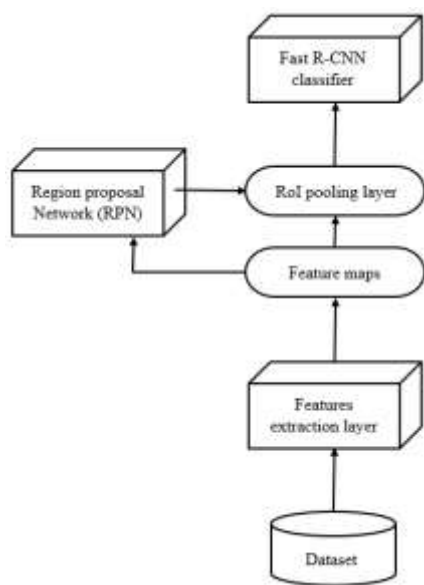


Figure. 3 Faster R-CNN architecture

Region Proposal Network (RPN) operates on the feature-map, an $(n \times n)$ sliding window passed on this feature map to generate Several multi-scale and aspect ratio Anchor boxes from each Anchor point [23], which is the center of the sliding window as seen in Fig. 4.

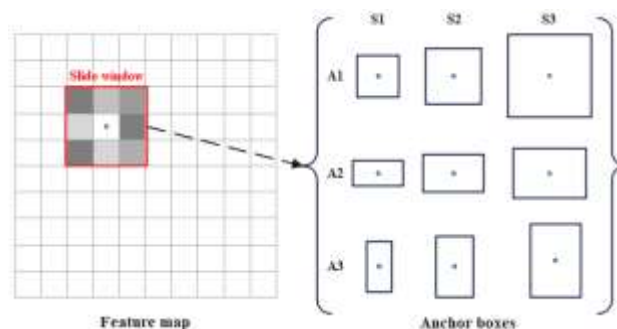
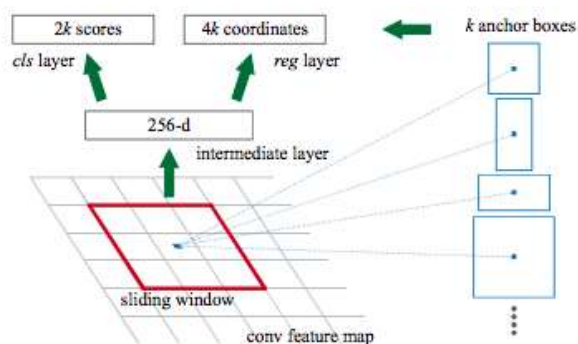


Figure. 4 Anchor boxes generation

The Region Proposal Network (RPN) includes both a classifier (*cls*) and a regressor (*reg*) [24]. The classifier defines the probability that a proposed region (anchor boxes) contains the target object. The regressor regresses the proposed region coordinates, as shown in Fig. 5.



8861

Figure. 5 Region proposal network (RPN)

The classifier (*cls*) determines whether the generated anchor box contains the target object by using Intersection Over Union (IOU) between the generated anchor box and ground truth bounding boxes (BBboxes) [23,24], as illustrated in Eq. (1). The anchor box that owns an $IOU \geq 0.7$ considered a positive overlap region or a target object region [25].

$$IOU = (A \cap B) / (A \cup B) \quad (1)$$

Where: (A) is Anchor box, (B) is Bounding boxes

The regressor (*reg*) extracts the coordinates of the positive overlap regions. The RoI Pooling Layer then uses both the feature maps and Region Proposal Network (RPN) outputs to generate the feature vector for all positive overlap regions and sends them to the Fast R-CNN classifier [25]. Generating multiple scales and aspect ratio anchor boxes for each region makes the



detection model able to detect the target object in different sizes.

5. Proposed system

To diagnose Bacillus spp. In the DIBaS (Digital Images of Bacteria Species) dataset, a diagnostic system for three different types of Bacillus bacteria (Clostridium perfringens, Lactobacillus salivarius and Lactobacillus crispatus) is presented. As shown in Figure 6, this system is divided into three phases. The first step consists of data pre-processing procedures. Before the data is split and transmitted to the Faster R-CNN, the second step proposes to perform several transformations on the preprocessed data in order to prepare it for training and validation.

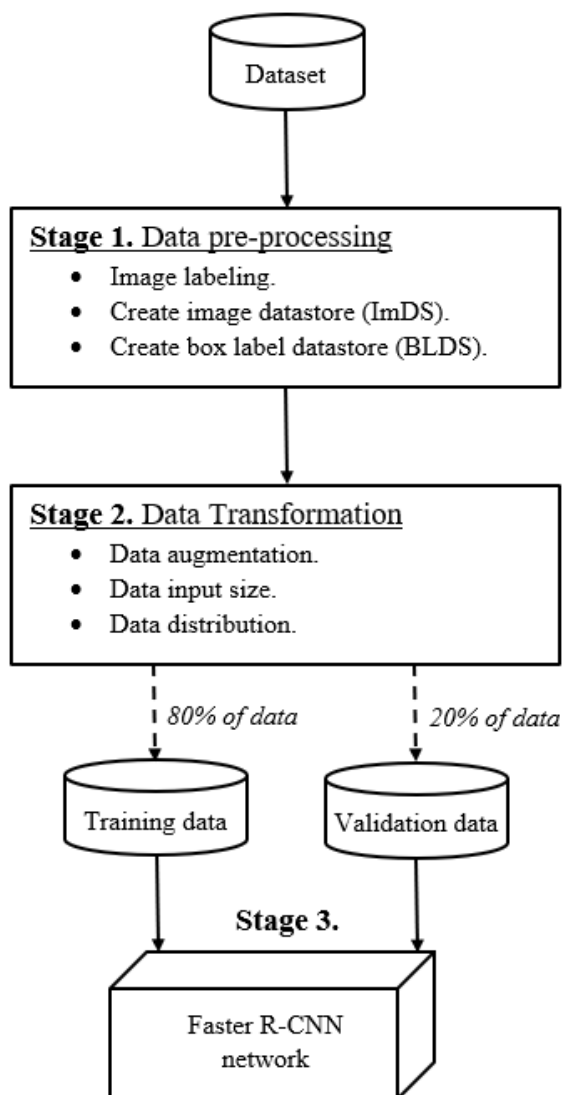
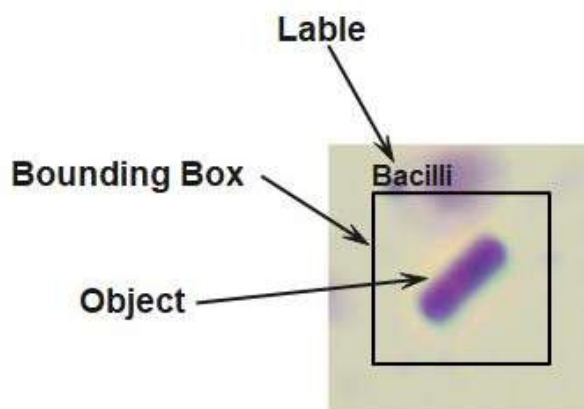


Figure. 6 Proposed system architecture

5.1 Data Pre-processing

The DIBaS dataset contains a set of images with dimensions of (2048×1532). This set of images must be preprocessed to suit the requirements of the training network as it first requires labeling the targeted objects in these images.

Image labeling is one of the modern methods for locating the target objects in an image [26]. There are several ways to do this, including the Bounding Boxes (BBoxes) method, where the target objects from which the features were extracted are enclosed in a square or rectangular shape, with the label referring to the type of the target object [26], as shown in Fig. 7.



8862

Figure. 7 Image labeling with BBox.

After identifying all the locations of the target objects to extract the features from, these locations are exported to the Ground Truth Label Table that includes the path of the images and the locations of the Bounding Boxes (BBoxes) in each image. This Ground Truth Label Table is converted into two data structures called Images Datastore (ImDS) that contains the images, and Box Label Datastore (BLDS) that contains the locations of the bounding boxes (BBoxes) in each image.

5.2 Data Transformation

Data augmentation Fig. 8, is a set of methods for increasing the amount of data by producing additional data from existing data [27]. This involves some changes on data or using deep-learning to generate new data.



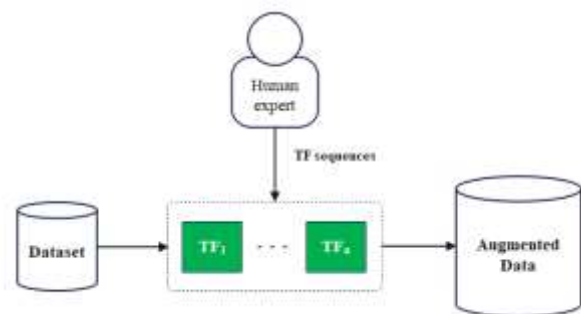


Figure. 8 Data augmentation (TF – Transformation Functions)

In this case, the X-reflection transformation function is used on both ImDS and BLDS, this operation will be applied on both the images and their bounding boxes (BBoxes) simultaneously to protect the data from corruption, as shown in Fig. 9.

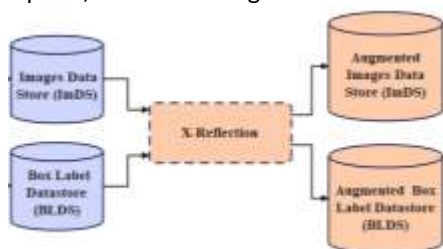


Figure. 9 Data Augmentation

The images in the dataset have a size of (2048x1532) for each image, which is a relatively large size when used causes a slowdown in the training process. To accelerate the training process while preserving image details, the images were reduced to a size of (1536x1149) as shown in Fig. 10.

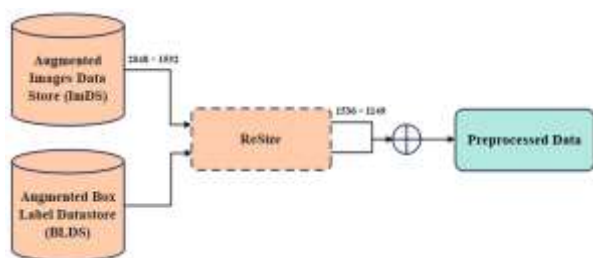


Figure. 10 Data Re-Size

When changing the size of the image in ImDS, as with the data augmentation, the size of the bounding boxes (BBoxes) in BLDS must be changed to match the new size. The percentage difference between the original image and the reduced image is used to calculate the amount of change in the size of the bounding boxes (BBoxes). At

the end of this stage, the pre-processed data will be divided into two parts: 80% of data for training and, 20% of data for validation.

5.3 Features extraction

Faster R-CNN network can extract features automatically due to the presence of a convolution layer; this layer can be modified by employing another feature extraction algorithm. The ResNet-50 network is a 50-layer deep, pre-trained on the ImageNet database that contains more than one million images [18] used as a feature extraction layer in the proposed system. Using a pre-trained algorithm in system design is known as "Transfer Learning". Developers use pre-trained algorithms to reduce development time while modeling the current task [28].

5.4 Setting Faster R-CNN layers

This stage intends to construct Faster R-CNN layers. This network has two models: the Region Proposal Network (RPN) for producing region proposals and the fast R-CNN classifier. The layers of this network have been developed utilizing the following steps:

- Step 1.** Define Region Proposal Network (RPN) using the following:
 - a. Generate nine anchor boxes with three scales= {8x8, 16x16, 32x32}, and three aspect ratios= {0.5, 1, 2}.
 - b. Negative overlap range= {0, 0.6}.
 - c. Positive overlap range= {0.7, 1}.
- Step 2.** Set number of classes= 1.
- Step 3.** Set input size = [1536x1149].
- Step 4.** Use ResNet-50 network for feature extraction.
- Step 5.** Use training data to train Fast R-CNN classifier.
- Step 6.** Use validation data to validating Fast R-CNN classifier.

6. Experimental results

The proposed system was developed using MATLAB R2021a as it was implemented on a personal computer with the following specifications:

- CPU: Core™ i7 H, 12 gen.
- RAM: 32 GB.
- VGA: Asus GeForce RTX 3060, 16 GB.
- Storage: NVMe™ SSD 1024 GB.
- OS: Microsoft Windows 10 Pro, x64, Ver. 21H2



Table 2. Mini-Batch and Validation Accuracies

| Epoch | Mini-Batch | Validation |
|-------|------------|------------|
| 1 | 52.11% | 96.44% |
| 2 | 97.78% | 98.94% |
| 4 | 88.14% | 86.54% |
| 5 | 96.21% | 95.41% |
| 7 | 98.98% | 97.75% |
| 8 | 97.47% | 98.09% |
| 10 | 96.28% | 97.10% |
| 10 | 98.99% | 99.10% |

Images of three types of Bacillus bacteria (Clostridium perfringens, Lactobacillus salivarius, and Lactobacillus crispatus) in DIBaS dataset are labeled using “Image Labeler” which is software package provided by the MATLAB to locating the target objects in images using bounding boxes, the information’s of labels and bounding boxes coordinates is exported in the form of a table called the Ground Truth Label Table that includes the path of the images and the coordinates of the bounding boxes in each image. This table is divided into two data structures, the image datastore (ImDS) contains images information’s such as (path, size, type, etc.) and box label datastore (BLDS) contains the number of bounding boxes and their locations in each image.

The image datastore (ImDS) and box label datastore (BLDS) the passed to data augmentation step to increase their size by using X-reflection transformation function, and then to data re-size step to generate pre-processed data.

The pre-processed data was separated into two parts: 80% for training and 20% for validation. The pre-trained ResNet-50 network was used as feature extraction layer to extract features from training data and generate the feature maps.

The Region Proposal Network (RPN) will use (nxn) slide window to generate nine different scales and aspect ratios region proposals (anchor boxes) for each point in feature maps then this window will shift by 16 point (stride = 16), the classifier (*cls*) determines wither the generated region contains the target object by using the Intersection Over Union (IOU) between generated anchor box and bounding boxes in box label datastore (BLDS), the anchor box that have an $IOU \geq 0.7$ will be considered asa positive overlap region or a target object region, else it will be considered as anegative overlap region or background region. The regressor (*reg*) extracts the coordinates of the positive overlap regions.

The RoI Pooling Layer then uses both the feature maps and Region Proposal Network (RPN) outputs to generate the feature vector for all positive overlap regions and sends them to the Fast R-CNN classifier.

6.1 Training and validation results

The faster R-CNN network does training and validation at the same time, as well as calculating training and validation accuracy and loss for each epoch. The proposed system was trained for 10 epochs, the total time for the training phase was 9 hours and 53 minutes. This system achieved a Mini-Batch Accuracy of 98.99% and a Validation Accuracy of 99.10%, as shown in Table 2.

Fig. 11 shows the accuracy results reached by the proposed system during the training and validation phase.

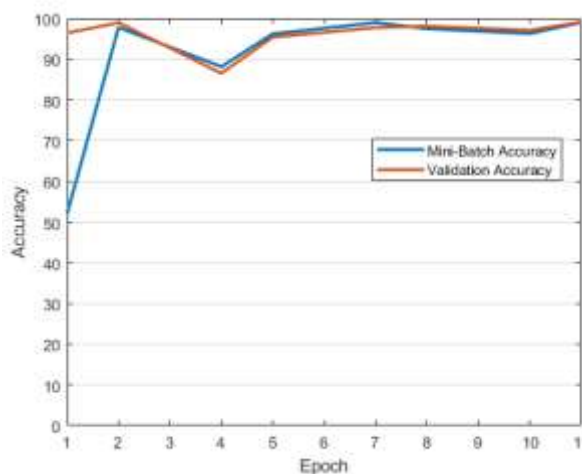


Figure. 11 Mini-Batch and Validation Accuracies

On other hand the Faster R-CNN network calculate the loss of training and validation for each epoch as shown in Table 3.

Table 3. Mini-Batch and Validation loss

| Epoch | Mini-Batch | Validation |
|-------|------------|------------|
| 1 | 2.7556 | 1.4803 |
| 2 | 1.7673 | 0.8866 |
| 4 | 0.6.63 | 0.5248 |
| 5 | 0.2054 | 0.2722 |
| 7 | 0.2024 | 0.2217 |
| 8 | 0.1270 | 0.2007 |
| 10 | 0.1364 | 0.2145 |
| 10 | 0.1092 | 0.3210 |

Fig. 12 shows the loss results reached by the proposed system during the training and validation phase.



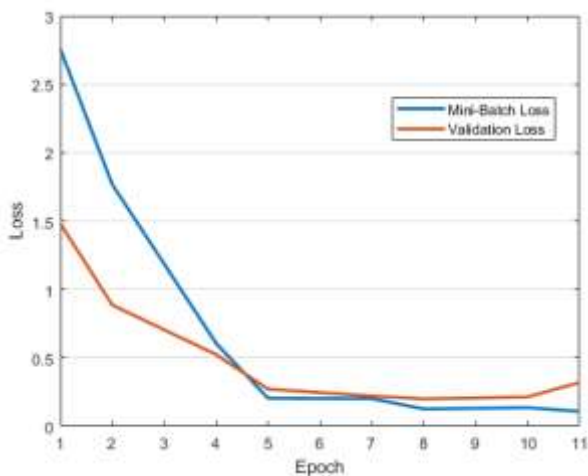


Figure. 12 Mini-Batch and Validation loss

6.2 Testing results

The proposed system was tested on images that were not part of the training and validation sets; Table 4 illustrates the testing results reached by the proposed system for three types of Bacillus spp. bacteria (Clostridium perfringens, Lactobacillus salivarius, and Lactobacillus crispatus) in DIBaS dataset.

Table 4. Testing results

| Bacteria | No. detected objects | Shape:Accuracy |
|--------------------------|----------------------|-------------------|
| Clostridium Perfringens | 4 | Bacilli: 92.2777% |
| | | Bacilli: 99.1909% |
| | | Bacilli: 98.8651% |
| | | Bacilli: 93.3127% |
| Lactobacillus Crispatus | 6 | Bacilli: 98.6856% |
| | | Bacilli: 99.2697% |
| | | Bacilli: 98.0971% |
| | | Bacilli: 94.3266% |
| | | Bacilli: 93.9349% |
| | | Bacilli: 91.502% |
| Lactobacillus Salivarius | 17 | Bacilli: 99.8776% |
| | | Bacilli: 99.1029% |
| | | Bacilli: 98.8198% |
| | | Bacilli: 99.0928% |
| | | Bacilli: 99.533% |
| | | Bacilli: 97.9031% |

| |
|-------------------|
| Bacilli: 99.1378% |
| Bacilli: 89.1122% |
| Bacilli: 99.7906% |
| Bacilli: 99.6303% |
| Bacilli: 99.411% |
| Bacilli: 89.7531% |
| Bacilli: 98.9833% |
| Bacilli: 98.956% |
| Bacilli: 99.7332% |
| Bacilli: 99.2863% |
| Bacilli: 82.1363% |

7. Conclusions

In this study, a real time bacillus spp. bacterium diagnosing system is proposed using DIBaS dataset, all bacteria object in images are labeled in data pre-processing stage, and employ data augmentation method to increase data size in data transformation stage. For features extraction this system used pre-trained ResNet-50 network as feature extraction layer in faster R-CNN network, the combination of ResNet-50 and Faster R-CNN networks has the best overall accuracies it reached 98.99% mini-batch accuracy and 99.10% validation accuracy.

Finally, diagnosing bacteria shape is one of the stages in the process of diagnosing bacteria samples. This stage was chosen because it is based on determining the external shape of bacteria, which is applicable because it requires some image processing using artificial intelligence.

Several studies have been conducted on the issue of diagnosing bacteria samples, and the results of these studies will be compared to the results of the proposed system. Some of these studies used the same dataset that we used to train the proposed system, while others used a different dataset, as shown in Table 5.

Table 5. A comparison between previous studies and the proposed system.

| Authors | Dataset | Method | Accuracy |
|-------------------------|---------|--------------|----------|
| Ö. F. Nasip et al. [14] | DIBaS | VggNet | 98.25% |
| | | AlexNet | 97.53% |
| B. Zieliński et al. [8] | DIBaS | SVM, RF | 97% |
| Y. He et al. [9] | N-SIM | SVM, RF, KNN | 98% |
| Proposed System | DIBaS | Faster R-CNN | 99.10% |



Based on the foregoing conclusions, the following points are suggested for future work:

- Increasing the number of bacteria shapes to include Spirilla and Vibrio and converting it into an application.
- Increasing the amount of training data using more advanced data augmentation techniques such as the Generative Adversarial Networks (GANs) algorithm.
- Using other feature extraction algorithms like Xception, ResNet152, InceptionV3, VGG19, MobileNet model etc.

Acknowledgment

The authors would like to thank Mustansiriya University (www.uomustansiriya.edu.iq Baghdad-Iraq) for supporting this work.

References

- [1] S. Kotwal, P. Rani, T. Arif, J. Manhas, and S. Sharma, "Automated Bacterial Classifications Using Machine Learning Based Computational Techniques: Architectures, Challenges and Open Research Issues." Archives of Computational Methods in Engineering (2021), Oct. 12, 2021. doi: <https://doi.org/10.1007/s11831-021-09660-0>.
- [2] P. S. Hiremath, P. Bannigidad, and P. Hiremath, "Automatic Classification of Bacterial Cells in Digital Microscopic Images" vol. 7546. Second International Conference on Digital Image Processing, Singapore, Feb. 26, 2010. doi: <https://doi.org/10.1117/12.853303>.
- [3] "API®," bioMérieux. <https://www.biomerieux-usa.com/clinical/api>.
- [4] L. Alzubaidi et al., "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions" vol. 8, no. 53. Journal of Big Data, Mar. 31, 2021. doi: <https://doi.org/10.1186/s40537-021-00444-8>.
- [5] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," vol. 28. Curran Associates, Inc., Jun. 04, 2015. doi: <https://doi.org/10.48550/arXiv.1506.01497>.
- [6] B. Li and D. Lima, "Facial expression recognition via ResNet-50," vol. 2. International Journal of Cognitive Computing in Engineering, pp. 57–64, 00, 2021. doi: <https://doi.org/10.48550/arXiv.1512.03385>.
- [7] F. Yang, W. Zhang, L. Tao, and J. Ma, "Transfer Learning Strategies for Deep Learning-based PHM Algorithms," vol. 7, no. 2361. Appl. Sci., Mar. 30, 2020. doi: <https://doi.org/10.3390/app10072361>.
- [8] B. Zieliński, A. Plichta, K. Misztal, P. Spurek, M. Brzychczy-Włoch, and D. Ochońska, "Deep Learning Approach to Bacterial Colony Classification," vol. 12, no. 9. PLoS ONE, Sep. 14, 2017. doi: <https://doi.org/10.1371/journal.pone.0184554>.
- [9] Y. He, W. Xu, Y. Zhi, R. Tyagi, Z. Hu, and G. Cao, "Rapid Bacteria Identification Using Structured Illumination Microscopy and Machine Learning," vol. 11, no. 01. | Journal of Innovative Optical Health Sciences, Sep. 22, 2017. doi: <https://doi.org/10.1142/S1793545818500074>.
- [10] J. R. Balbin, J. T. Sese, C. V. R. Babaan, D. M. M. Poblete, R. P. Panganiban, and J. G. Poblete, "Detection and classification of bacteria in common street foods using electronic nose and support vector machine." 2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, pp. 247–252, Nov. 2017. doi: <https://doi.org/10.1109/ICCSCE.2017.8284413>.
- [11] Y. Du et al., "Raman spectroscopy-based adversarial network combined with SVM for detection of foodborne pathogenic bacteria," vol. 237. Talanta, Jan. 15, 2022. doi: <https://doi.org/10.1016/j.talanta.2021.122901>.
- [12] T. Beznik, P. Smyth, G. de Lannoy, and J. A. Lee, "Deep learning to detect bacterial colonies for the production of vaccines," vol. 470. Neurocomputing, pp. 427–431, Jul. 22, 2021. doi: <https://doi.org/10.1016/j.neucom.2021.04.130>.
- [13] T. Treebupachatsakul and S. Poomrittigul, "Bacteria Classification using Image Processing and Deep learning," 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 2019, pp. 1-3, doi: <https://doi.org/10.1109/ITC-CSCC.2019.8793320>.
- [14] Ö. F. Nasip and K. Zengin, "Deep Learning Based Bacteria Classification," 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2018, pp. 1-5, doi: <https://doi.org/10.1109/ISMSIT.2018.8566685>.
- [15] X. Gu, L. Pan, H. Liang, and R. Yang, "Classification of Bacterial and Viral Childhood Pneumonia Using Deep Learning in Chest Radiography." In Proceedings of the 3rd International Conference on Multimedia and Image Processing (ICMIP 2018), New York, USA, pp. 88–93, Mar. 16, 2018. doi: <https://doi.org/10.1145/3195588.3195597>.
- [16] E. Bonah, X. Huang, R. Yi, J. H. Aheto, R. Osae, and M. Golly, "Electronic nose classification and differentiation of bacterial foodborne pathogens based on support vector machine optimized with particle swarm optimization algorithm," vol. 42, no.



6. Journal of Food Process Engineering, Aug. 23, 2019. doi: <https://doi.org/10.1111/jfpe.13236>
- [17] A. RAJAEI and V. HEYDARI, "Diagnosis of escherichia coli bacteria patient by data mining," vol. 97, no. 21. Journal of Theoretical and Applied Information Technology, Oct. 15, 2019.
- [18] B. Li and D. Lima, "Facial expression recognition via ResNet-50," vol. 2. International Journal of Cognitive Computing in Engineering, pp. 57–64, 0 0, 2021. doi: <https://doi.org/10.48550/arXiv.1512.03385>.
- [19] H. H. Tan and K. H. Lim, "Vanishing Gradient Mitigation with Deep Learning Neural Network Optimization," 2019 7th International Conference on Smart Computing & Communications (ICSCC), 2019, pp. 1-4, doi: <https://doi.org/10.1109/ICSCC.2019.8843652>.
- [20] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: <https://doi.org/10.1109/CVPR.2016.90>.
- [21] Wen, Long, Xinyu Li, and Liang Gao. "A transfer convolutional neural network for fault diagnosis based on ResNet-50." Neural Computing and Applications 32.10 pp. 6111-6124, 2020.
- [22] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," in IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, Oct. 2010, doi: <https://doi.org/10.1109/TKDE.2009.191>.
- [23] D. König, M. Adam, C. Jarvers, G. Layher, H. Neumann and M. Teutsch, "Fully Convolutional Region Proposal Networks for Multispectral Person Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 243-250, doi: <https://doi.org/10.1109/CVPRW.2017.36>.
- [24] Bo Li, Junjie Yan, Wei Wu, Zheng Zhu, Xiaolin Hu, "High Performance Visual Tracking with Siamese Region Proposal Network", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8971-8980.
- [25] BRAZIL, Garrick; LIU, Xiaoming. "M3d-rpn: Monocular 3d region proposal network for object detection". In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9287-9296, 2019.
- [26] Åström, F., Petra, S., Schmitzer, B., & Schnörr, C., "Image labeling by assignment. Journal of Mathematical Imaging and Vision", Vol.58, No. 2, pp. 211-238, 2017.
- [27] B. Zoph, E. D. Cubuk, G. Ghiasi, T.-Y. Lin, J. Shlens, and Q. V. Le, "Learning Data Augmentation Strategies for Object Detection", ECCV 2020: Computer Vision – ECCV 2020, vol. 12372. Springer, Cham, pp. 566–583, Nov. 19, 2020. doi: https://doi.org/10.1007/978-3-030-58583-9_34.
- [28] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C., "A survey on deep transfer learning", In International conference on artificial neural networks, Springer, Cham, pp. 270-279, 2018.

