Paper

Methods of Classification of the Genera and Species of Bacteria Using Decision Tree

Anna Plichta

Faculty of Computer Science and Telecommunications, Department of Computer Science, Cracow University of Technology, Kraków, Poland

https://doi.org/10.26636/jtit.2019.137419

Abstract—This paper presents a computer-based method for recognizing digital images of bacterial cells. It covers automatic recognition of twenty genera and species of bacteria chosen by the author whose original contribution to the work consisted in the decision to conduct the process of recognizing bacteria using the simultaneous analysis of the following physical features of bacterial cells: color, size, shape, number of clusters, cluster shape, as well as density and distribution of the cells. The proposed method may be also used to recognize the microorganisms other than bacteria. In addition, it does not require the use of any specialized equipment. The lack of demand for high infrastructural standards and complementarity with the hardware and software widens the scope of the method's application in diagnostics, including microbiological diagnostics. The proposed method may be used to identify new genera and species of bacteria, but also other microorganisms that exhibit similar morphological characteristics.

Keywords—bacterial genera and species, decision tree, pattern recognition.

1. Introduction

IT technologies are used in a wide variety of medical and microbiological sciences, with their applications ranging from the solutions used for imaging pathological conditions, systems supporting the processing and analysis of acquired data, dedicated medical instrumentation software, diagnostic support systems, to specialized medical data repositories [1]. In the case of microbiology, computer-based methods are used in the analysis and recognition of laboratory-obtained microscopic data (e.g. bacterial cells). A correct and quick diagnosis is a critical stage in the process of administering the appropriate therapy. At the same time, recognition of the genera and species of bacteria is still typically a task that is performed by experts using specialized manual equipment and diagnostic tests. It is a time-consuming and expensive process. Moreover, due to the lack of a digital data repository, the procedures applied to obtain samples for the analysis and recognition process often must be repeated [1]–[3].

The motivation to undertake research on the use of computational methods to support microbiological diagnostics was drawn from the need to design a computer application using algorithms for automatic recognition of the selected genera and species of bacteria by relying on a digital data repository. Improvement of the process of classifying bacterial material presented in images shortens the time required to successfully perform the recognition. Besides, the proposed application will enable the digital repository to be used without re-acquisition and re-analysis of the samples collected. Computer-based methods employed for classifying bacterial cell images require the use of appropriate algorithms and image processing techniques in order to enable correct identification based on specific extracted features that distinguish the individual bacteria genera and species [4]–[6].

1.1. Selected Issues Related to Microbiological Diagnostics

Microbiology deals with, inter alia, recognition, cultivation and observation of microorganisms – such as bacteria and viruses. It also covers the examination of their role in nature and their impact on living organisms [7].

Bacteria belong to the most common living organisms. They easily adapt to the changing environmental conditions, which may lead to mutations and immunization to the existing protective agents, thus hindering or even preventing effective counteraction of their negative impact on human life and on other living organisms [8]–[11]. However, one should not forget about the beneficial effects that bacteria exert on living organisms, or about their practical uses in biochemical, pharmacological, veterinary, growing and breeding industries, not to mention bio-fuel production [7], [12]–[14].

Microbial diagnostics is a field that is closely related to the recognition of bacterial cells, regardless of the reason for which their identification is required.

1.2. Computer-based Methods Supporting Microbiological Diagnostics

Computer-based methods enabling automatic recognition of selected genera and species of bacteria based on the

analysis of their images facilitate the microbiological diagnostics process, ensuring fast and correct recognition of the tested samples. The preparation of samples having the form of images is preceded by a laboratory stage (culturing bacteria) by the use of biochemical methods and tests required for identification and classification of specific genera and species.

Therefore, diagnostic experience and knowledge are necessary for the proper analysis of bacterial cells. It is a time-consuming process, as a comparative analysis against standard samples from a reference strain bank must be performed. It is also expensive due to the need to perform biochemical tests requiring specialized reagents. The final result of the diagnostic work is the classification of the tested samples, matching them with a specific type and species of bacteria [15]–[18].

Automation of bacterial cell classification significantly reduces test lead time, freeing the specialists from the need to manually analyze and evaluate the samples. Furthermore, if the degree of similarity of the images and the level of diagnostic uncertainty are considerable, computer-based methods are capable of broadening the scope of recognition, indicating the most likely answer without ignoring other potential results.

2. Related Work

Image processing and recognition techniques, combined with different types of classifiers, are used as effective IT tools relied upon to derive medical and microbiological data from images [19]. Computer-based methods used for bacterial classification often employ artificial intelligence, statistical methods or other solutions aiming to automate the process of analyzing and classifying the data obtained. In scientific publications concerning bacterial cell recognition, the most commonly examined systems are dedicated to identifying one species or genera of bacteria (e.g. tuberculosis), or a group of microorganisms, including bacteria sharing similar shapes or other microbiological characteristics.

There are also some computer-based methods that are integrated with an automated microscope or with other types research equipment embedded as components of the entire diagnostic system. Such an approach makes it possible to recognize various microorganisms but, due to hardware and cost limitations, it may limit prevent systems of this type from being used more extensively [20]–[24]. Therefore, interdisciplinary research teams are looking for new solutions offering effective methods that can be relied upon for analyzing microbiological data. Such methods are expected to provide a broad spectrum of applications and to be easily modifiable depending on the diagnostic needs and the samples tested.

One way to identify bacteria is to recognize their geometric features, such as the shape or length-width ratio of the cell. Also, because the shape is not a distinctive feature, since different genera and species of bacteria may share the same morphology, the color of bacterial cells obtained during their staining with the use of biochemical methods is taken into account [23]. Besides, a method based not only on geometrical features but also on the average color of the analyzed images was used to automatically recognize tuberculosis bacilli [25]. In this research, the author tackled the problem of similarities in bacterial morphology and showed that color is a key feature to improve recognition efficiency. However, where polymorphism is involved, bacterial cells are of both purple (Gram-positive) and pink (Gram-negative) color [15], which can hinder or even prevent their correct recognition.

Other approaches rely on pre-segmented images obtained from a scanner, as well as on various feature extracting methods, such as the size and shape of cells, using these to assign the bacteria into the appropriate morphotype. A group of researchers gathered around the CMEIAS project (center for microbial ecology image analysis system) [26] used two classifiers for such an analysis. The first classifier analyzes a single feature and is dedicated to processing simple samples containing relatively few morphotypes only – regular sticks or spherical granules, for instance. The other classifier is a hierarchical tree that uses an optimized subset of features to analyze much more complex structures of a greater morphological diversity. They automatically classify each cell into one of the bacterial morphotypes, such as cocci, spiral bacteria, bacilli, u-shaped bacteria, rods, ellipsoids, club-shaped bacteria or comma-shaped bacteria, for instance.

Other identification methods are based on the analysis of bacterial colony patterns (a cluster of bacterial cells created based on single-cell divisions). This analysis uses, inter alia, Fisher vectors, random forest algorithms or the support vector machine (SVM) [19], [27]–[30].

A new approach to recognizing bacteria or other medical images (e.g. x-rays or ultrasound images) relies on methods in which images are analyzed and classified based on a texture model. The texture represents such image properties as pattern direction and porosity. Thanks to this, it is possible to designate areas of a given image that fulfill the specific conditions and, thus, to classify them into a given type of texture based on observations identifying some small patterns and their regular distributions. The mathematical description of the texture is based on parameters describing the properties of the digital image, which are calculated, for example, using statistical methods or signal processing techniques. The numerical representation of texture properties is later used for further analysis and classification.

The use of this approach in analyzing and classifying selected genera and species of bacteria is justified by the fact that different genera and species of bacteria reproduce in a specific manner. They form clusters of a peculiar shape, which may be considered a texture. Such methods rely, inter alia, on Fisher vectors, SVM and deep neural networks. It is noteworthy that deep learning techniques make it possible to implement many subnetworks that specialize in classifying or recognizing features and in processing particular types of input data, such as textures. In

addition, deep convolutional networks allow for gradual filtration of various types of learning data and indicate essential features of the analyzed objects in the pattern recognition or classification process. The connecting layer between successive convolution layers reduces the number of parameters, which provides control over the over-learning of the network. The methods used for recognizing images that are defined as textures, relying on machine learning and neural networks, may therefore be applied in the process of recognizing not only bacteria, but also other microorganisms [30]–[36].

Techniques implementing sensors, i.e. devices used to obtain and process chemical data describing bacterial cells, comprise an innovative group of effective methods used for identifying bacteria. Thanks to the cooperation of experts specializing in biological sciences and statistical analysis, it was possible to design and build devices based on sensors, and thus to acquire and analyze data that has not, until now, been taken into consideration while classifying bacterial cells. Sensors are capable of collecting vast amounts of data in a short period of time. Some data are hardly useful in the classification process. Therefore, it is necessary to use statistical analysis and other mathematical tools to organize the information collected depending on the research direction and needs [37]–[40].

Among the new solutions used in bacteria recognition are the so-called artificial noses or gas-identifying sensors. The effectiveness of an artificial nose commercialized by Cyrano Sciences under the name of Cyranose 320 has been confirmed based on the results of research concerning chemical gas sensors, and on the analysis of data obtained. However, the need to ensure that an appropriate database required for device learning is available and the fact that the device is only able to detect ten different chemical compounds in one sample, limits the use of the artificial nose, even if such devices are successfully used, commercially, for identification of bacteria in diabetic foot infections (Escherichia coli, Pseudomonas aeruginosa and Staphylococcus aureus) that require quick application of effective therapies [41]–[43].

Identification techniques based on Fresnel diffraction spectrum analysis performed by scattering light on bacterial cells, and on statistical methods employed to analyze the information obtained in this manner, are also relied upon to identify bacterial cells [44]–[46].

Other methods that are used ever more often include optical techniques based on the use of light serving as an information carrier. These are divided into spectroscopic methods (direct and indirect fluorescence, infrared spectroscopy, Raman spectroscopy) and methods based on the dispersion of light in laboratory samples (e.g. bacterial cells) containing aqueous solutions, aerosols or bacterial colonies on solid substrates. Fluorescence- or spectroscopy-based techniques, however, are expensive, require the preparation of high-quality samples and rely on appropriate databases of emission spectra of all the bacteria to be identified. The fact that they are time-consuming is an important factor, too [45], [47]–[50].

Most of the methods described above are used to recognize several selected genera and species of bacteria (sometimes only one, e.g. tuberculosis). In many cases, bacteria identification is based on the analysis of the morphological characteristics of their cells, in combination with a particular classification method. This prevents the use of such methods in the recognition of polymorphic bacterial cells, i.e. those that take different shapes within the same type. The need to use specialized equipment computers and diagnostic devices) limits the potential of computer-based methods to support microbiological diagnostics, too.

The method for recognizing selected bacteria genera and species, as proposed and described above, may be used on a standalone basis, but in the case of more complex images containing many different bacterial cells, it is suitable only for the first stage of the classification process. Thanks to this approach, it is possible to reject or narrow down the different types of cells visible in the image, so that other known methods such as, for example, neural networks, may be used in the next step.

The classifiers analyze 7 physical characteristics of bacterial cells. Therefore, they can be used to extract the characteristics of those samples that have not yet been classified by means of the proposed method. Thanks to this, it is possible to apply the method to recognize new genera and species of bacteria or to analyze images of other microorganisms that exhibit characteristics similar to those of bacterial cells.

The method proposed for recognizing bacterial images using a decision tree enables automatic identification of the analyzed samples, based on photographs of twenty different genera and species of bacteria, and their classification into appropriate genera and species. The results obtained during tests indicate that the method is effective, although it requires further implementation-related work to eliminate the cases of incorrect classifications or situations where no classification is possible at all.

3. Research Methodology

Computer-based methods require the use of appropriate algorithms and image processing techniques to enable correct classification while relying on such extracted features as color, size, shape, number of clusters, cluster shape, density and distribution of cells, which differentiate the specific genera and species of bacteria. Sample images of the analyzed bacteria genera and species are presented in Fig. 1.

Some features are used, at the beginning of the classification process, to optically differentiate the tested samples. These include, in particular, the color of the bacterial cell – purple for Gram-positive (G+) and pink for Gram-negative (G-) cells – and the shape or size of the cell. However, the remaining features require a more sophisticated analysis. These are, primarily, the typical shapes of cell clusters characteristic of the particular genera and species, as well as the manner in which the the number of bacterial cells, density or their distribution are assessed.

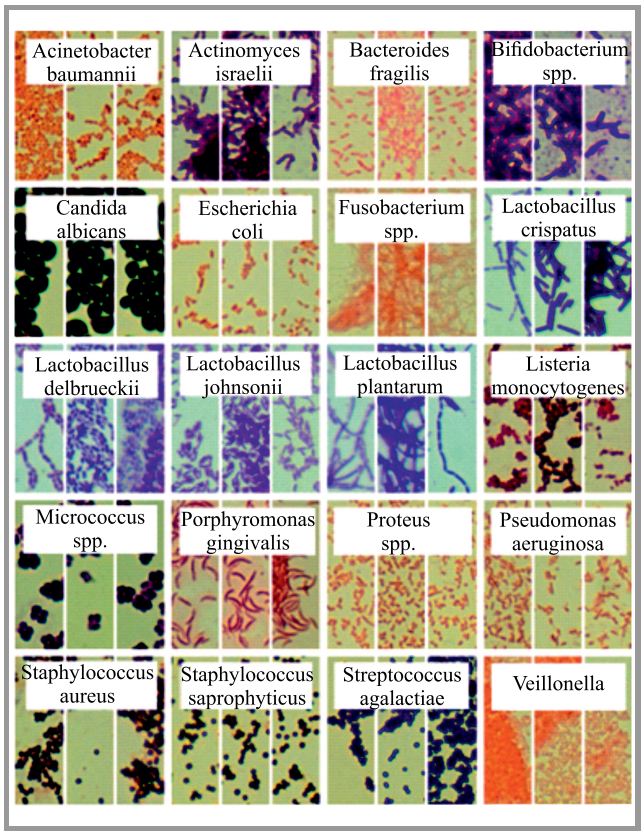

Fig. 1. Sample images of the analyzed genera and species of bacteria. (For color pictures visit www.nit.eu/publications/journal-jtit)

Therefore, after consultation with experienced microbiologists, seven features were selected: color, shape of a single bacterial cell, cell size, number of clusters formed, cluster shape, density and distribution of cells in the image. These features were the basis for the classifiers dedicated to distinguishing between the selected genera and species of bacteria.

4. Decision Tree Supporting the Classification Process

400 images from the DIBAS DB (Digital Image Of Bacterial Species) database were used to analyze and recognize bacteria images – 20 images for each of the 20 different bacteria genera and species. The database created for the purpose of the research was related to bacterial identification. The images were based on the microbiological preparations used for microscopic and biochemical analysis.

A decision tree was used in the classification studies based on the physical characteristics of bacterial cells. The method uses 7 classifiers based on the physical characteristics of bacterial cells. Within each classifier, the analyzed bacterial cells have been divided into such categories as:

- color classifier: purple for Gram-positive (G+) and pink for Gram-negative (G-),
- classifier related to the shape of a single bacterial cell: round, rod-shaped, stick, club, donut and boat,
- size classifier: large and small,
- classifier related to the number of clusters formed: single cells, diplococci, tetrads, larger,
- cluster shape classifier: parquet, snake and others,
- density classifier: rare, dense, very dense,
- classifier related to the distribution of cells in the image: even, uneven, very uneven.

The decision tree, used as a classification method, allows for swift identification of the analyzed bacterial image based on the verification of the aforementioned, selected features of bacterial cells present in a given sample. The decision is made unconditionally, and no confidence levels attached to the classifiers are taken into account. As a result, one obtains information only about the final classification, whereas the confidence level is omitted. Figure 2 presents an example of a decision tree, applied to a selected type and species of bacteria, that minimizes the number of decisions necessary to classify a given sample based on the analyzed features.

It is worth noting that an the answer concerning the first classifier only is sufficient to identify several genera and species of bacteria. At the same time, it can be seen how an error related to this particular classifier affects all samples tested. An error made by any classifier in the decision tree results in an incorrect final classification.

5. Performed Tests

The primary factor determining the correctness of this classification method is the number of correctly classified images showing selected genera and species of bacteria, in which the classifiers have detected all the characteristics of bacterial cells necessary for correct classification. As a result, the confusion matrix shown in Table 1 is obtained. The correctness of this tree classification for all analyzed samples is 76.85%, and the sensitivity result is 95.94%.

The obtained results show that this is not the best classification outcome that may be achieved by means of a decision tree. The decision tree described above was designed to minimize the number of decisions needed to uniquely classify a given sample. Therefore, it optimizes the number of decisions, but not the quality of the classification process.

While analyzing the differences between the genera and species of bacteria in the incorrectly classified samples, one may notice

- 26% of Actinomyces israelii cell samples are mistakenly identified as belonging to Lactobacillus johnsonii. The decision tree does not take into account the fact that those species differ in terms of the shape of the clusters of bacterial cells they form;
- 48% of Bacteroides fragilis cell samples are incorrectly classified as Acinetobacter baumannii bacterial

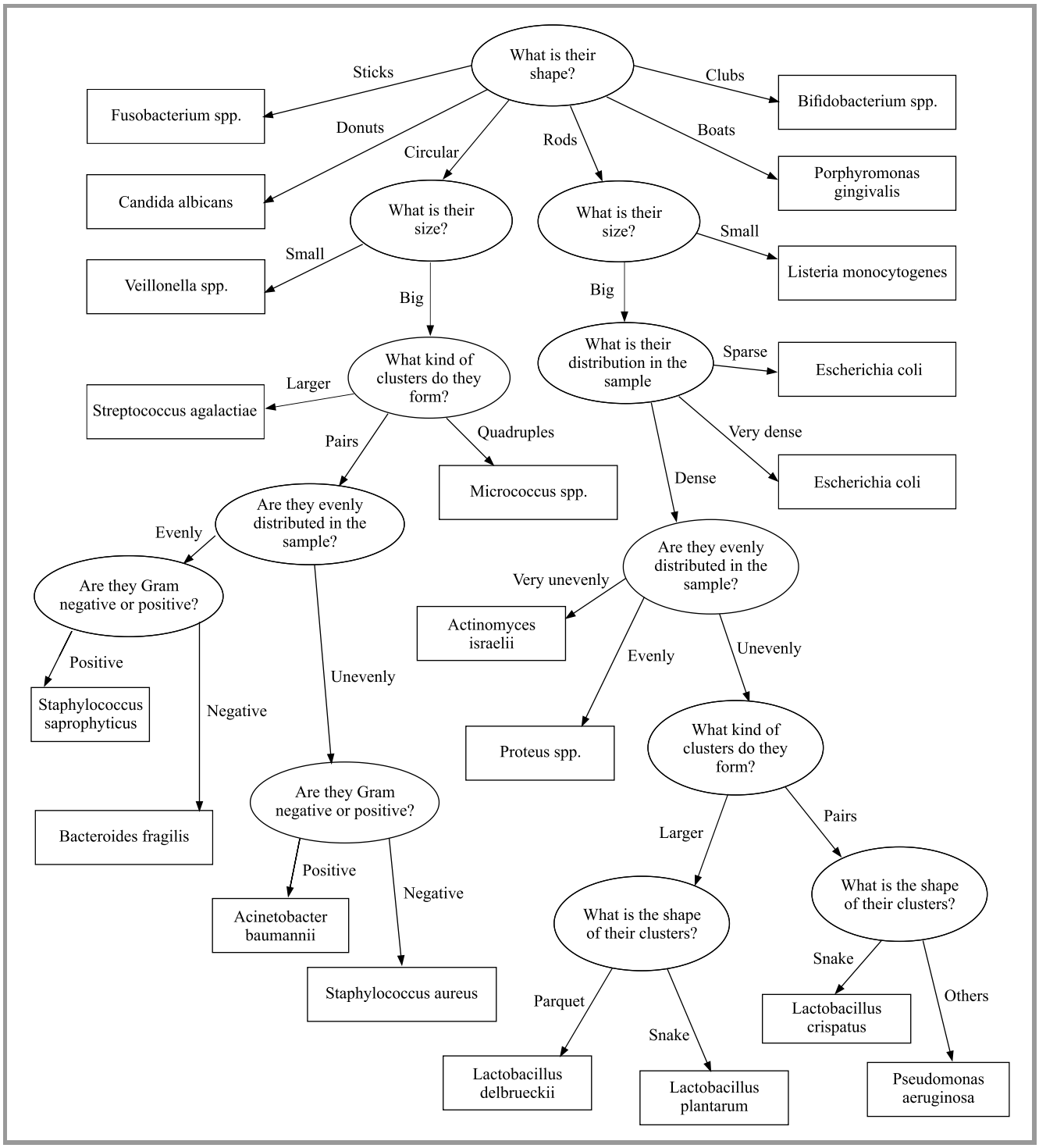

Fig. 2. Example of a decision tree for selected genera and species of bacteria.

samples. The decision tree does not take into account the fact that they differ in the density of bacterial cells in the sample;

- 64% of cell samples of Veillonella spp. bacteria are incorrectly recognized as Streptococcus agalactiae samples, which can be improved by taking into account differences in the color of these two genera and species of bacteria.

Hence, it is necessary to increase the correctness of classification by introducing changes to the decision tree. The solution is to add additional tree branches. An example of a modified decision tree is presented in Fig. 3. The corrections introduced to the decision tree from Fig. 2 are marked using bolded lines.

The confusion matrix of the modified decision tree is shown in Table 2. The classification of this tree for all

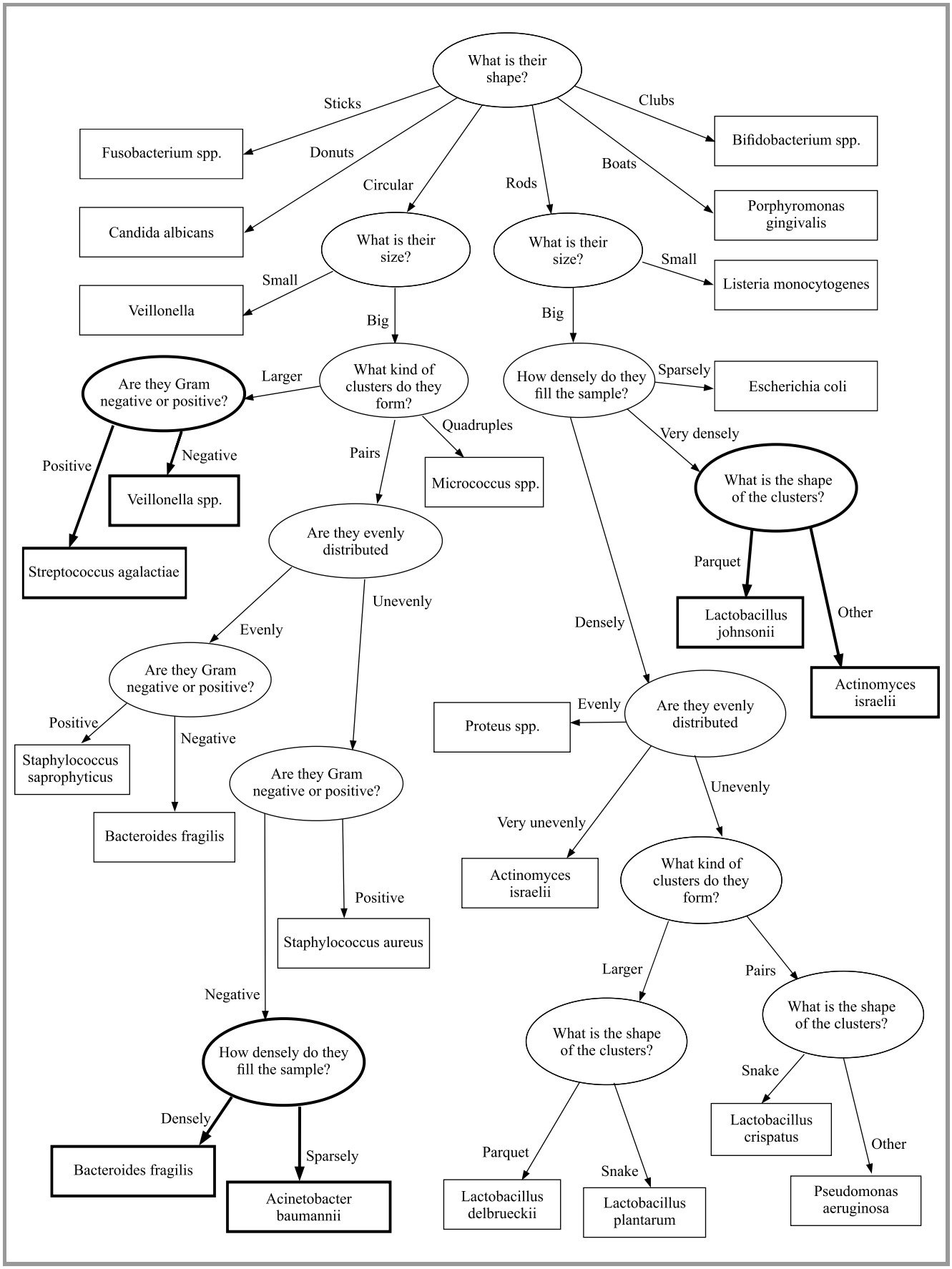

Fig. 3. Example of a modified decision tree.

analyzed genera and species of bacteria is 83.77%. The sensitivity has not changed and remains at 95.94%.

The obtained results show that every introduced correction has brought about positive results. The new decision tree uses, more frequently, highly correct classifiers, such as the bacterial cell color, which has the highest correctness of all. This decision tree may be optimized further to provide the best possible results using the boosted decision trees method. By adopting this approach, one may obtain a decision tree that optimizes correctness of the classific-

Table 1
Confusion matrix of the decision tree

No.	Genera and species of bacteria	Acinetobacter baumannii	Actinomyces israelii	Bacteroides fragilis	Bifidobacterium spp.	Candida albicans	Escherichia coli	Fusobacterium spp.	Lactobacillus crispatus	Lactobacillus delbrueckii	Lactobacillus johnsonii	Lactobacillus plantarum	Listeria monocytogenes	Micrococcus spp.	Porphyromonas gingivalis	Proteus spp.	Pseudomonas aeruginosa	Staphylococcus aureus	Staphylococcus saprophyticus	Streptococcus agalactiae	Veillonella spp.	No classification
1	Acinetobacter baumannii	90%		5%																		5%
2	Actinomyces israelii		48%		4%			4%			26%											18%
3	Bacteroides fragilis	48%		48%																	4%	
4	Bifidobacterium spp.				96%			4%														
5	Candida albicans					100%																
6	Escherichia coli				20%		75%															5%
7	Fusobacterium spp.							96%														4%
8	Lactobacillus crispatus				15%			5%	45%		10%		5%									20%
9	Lactobacillus delbrueckii				5%		5%	5%		60%	15%	10%										
10	Lactobacillus johnsonii				10%			5%		20%	65%											
11	Lactobacillus plantarum		5%		10%			5%		5%	5%	70%										
12	Listeria monocytogenes				5%			5%					90%									
13	Micrococcus spp.													95%				5%				
14	Porphyromonas gingivalis														100%							
15	Proteus spp.							5%								95%						
16	Pseudomonas aeruginosa															5%	80%					15%
17	Staphylococcus aureus																	90%	5%			5%
18	Staphylococcus saprophyticus																	15%	75%			10%
19	Streptococcus agalactiae																			95%	5%	
20	Veillonella spp.	9%																		64%	27%	

Table 2
Confusion matrix for the modified decision tree

No.	Genera and species of bacteria	Acinetobacter baumannii	Actinomyces israelii	Bacteroides fragilis	Bifidobacterium spp.	Candida albicans	Escherichia coli	Fusobacterium spp.	Lactobacillus crispatus	Lactobacillus delbrueckii	Lactobacillus johnsonii	Lactobacillus plantarum	Listeria monocytogenes	Micrococcus spp.	Porphyromonas gingivalis	Proteus spp.	Pseudomonas aeruginosa	Staphylococcus aureus	Staphylococcus saprophyticus	Streptococcus agalactiae	Veillonella spp.	No classification
1	Acinetobacter baumannii	90%		5%																		5%
2	Actinomyces israelii		74%		4%			4%														18%
3	Bacteroides fragilis	9%		87%																	4%	
4	Bifidobacterium spp.				96%			4%														
5	Candida albicans					100%																
6	Escherichia coli				20%		75%															5%
7	Fusobacterium spp.							96%														4%
8	Lactobacillus crispatus				15%			5%	45%		10%		5%									20%
9	Lactobacillus delbrueckii				5%		5%	5%		60%	15%	10%										
10	Lactobacillus johnsonii				10%			5%		20%	65%											
11	Lactobacillus plantarum		5%		10%			5%		5%	5%	70%										
12	Listeria monocytogenes				5%			5%					90%									
13	Micrococcus spp.													95%				5%				
14	Porphyromonas gingivalis														100%							
15	Proteus spp.							5%								95%						
16	Pseudomonas aeruginosa															5%	80%					15%
17	Staphylococcus aureus																	90%	5%			5%
18	Staphylococcus saprophyticus																	15%	75%			10%
19	Streptococcus agalactiae																			95%	5%	
20	Veillonella spp.	9%																			**9**1%	

ation process. However, even the most optimized decision tree cannot offer completely satisfactory results due to its high sensitivity to identification errors. Even one wrong recognition of one classifier automatically prevents the correct classification of a given sample.

It is also worth considering the problem of Escherichia coli bacterial samples. Some of them were incorrectly classified as Bifidobacterium spp. based on the bacterial shape classifier. This classifier for Escherichia coli bacterial cells has a confidence level of 75%. This is a very good example showing how important high-quality classifiers are in this method. In this case, the only way to improve the quality of identification is to improve the effectiveness of the bacterial shape classifier.

The results of the performed tests show that the final classification of the analyzed samples is significantly better in a few cases only. The number of correctly classified bacterial samples of Veillonella spp. equals 91% of all samples analyzed, while the previous result is equal to 27%. The number of correctly classified bacterial samples of Actinomyces israelii is 74%, while the previous result is 48%. The number of correctly classified bacterial samples of Bacteroides fragilis is 87%, while the previous result is 48%. The use of such classifiers as density or cell distribution in the sample enabled more images to be correctly classified as belonging to appropriate bacteria genera and species. The correctness of the proposed classification method increased from 76.85% to 83.77%. The obtained results, however, are still not satisfactory.

6. Summary

The limitations of the method presented serve as a source of motivation for further research in the field computer-based methods supporting microbiological diagnostics. The next task will be to increase the degree of correctness and sensitivity of the classifiers implemented, or to develop and implement new ones, so that the method may be verified for these genera and species of bacteria that could not be correctly recognized so far.

The results of the tests regarding the classification of images of selected genera and species of bacteria with the use of a decision tree confirm the effectiveness of the proposed method. However, it is still possible to obtain better identification results and to expand the number of analyzed images. Hence, there is a need to improve the method and to conduct further related research.

References

[1] P. Lutomski *et al.*, "Wykorzystanie informatyki medycznej w Polsce i na świecie" (An Application of the medical computer sciences in Poland and worldwide), Przedsiębiorczość i Zarządzanie, vol. 3, no. 15, pp. 83–92, 2014 [Online]. Available: http://piz.san.edu.pl/docs/e-XV-12-3.pdf [in Polish].

[2] R. Rudowski, *Informatyka Medyczna*. Wydawnictwo Naukowe PWN, 2003 (ISBN: 9788301140564) [in Polish].

[3] R. Tadeusiewicz and W. Wajs, *Informatyka Medyczna*. Uczelniane Wyd. Naukowo-Dydaktyczne Akademii Górniczo-Hutniczej, 1999 [in Polish].

[4] F. Sahba and H. R. Tizhoosh, "Filter fusion for image enhancement using reinforcement learning", in *Proc. Canadian Conf. on Elec. and Comp. Engin. CCECE 2003*, Montreal, Quebec, Kanada, 2003, vol. 2, pp. 847–850 (doi: 10.1109/CCECE.2003.1226027).

[5] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*. Cengage Learning, 2014 (ISBN: 9781133593690).

[6] Ch. Hau, *Handbook of Pattern Recognition and Computer Vision*. World Scientific, 2015 (ISBN: 9789814656535).

[7] P. R. Murray, K. S. Rosenthal and M. A. Pfaller, *Medical Microbiology*. *Elsevier Health Sciences*, pp. 16–348, 2015 (ISBN: 9780323299565).

[8] H. J. Busse, E. B. M. Denner, and W. Lubitz, "Classification and identification of bacteria: current approaches to an old problem. Overview of methods used in bacterial systematics", *J. of Biotechnol.*, vol. 47, no. 1, pp. 3–38, 1996 (doi: 10.1016/0168-1656(96)01379-X).

[9] A. R. Hall, D. C. Angst, K. T. Schiessl, and M. Ackermann, "Costs of antibiotic resistance separating trait effects and selective effects", *Evol. Appl.*, vol. 8, no. 3, pp. 261–272, 2015 (doi: 10.1111/eva.12187).

[10] J. Penterman *et al.*, "Rapid evolution of culture-impaired bacteria during adaptation to biofilm growth", *Cell Reports*, vol. 6, no. 2, pp. 293–300, 2014 (doi: 10.1016/j.celrep.2013.12.019).

[11] G. G. Perron, R. F. Inglis, P. S. Pennings, and S. Cobey, "Fighting microbial drug resistance: a primer on the role of evolutionary biology in public health", *Evol. Appl.*, vol. 8, no. 3, pp. 211–222, 2015 (doi: 10.1111/eva.12254).

[12] S. H. Gillespie, *Medical Microbiology Illustrated*. Butterworth-Heinemann Ltd., 2014, pp. 1–12 and 146–159 (ISBN: 978-1483177823).

[13] E. Goldman and L. H. Green, *Practical Handbook of Microbiology*. CRC Press, 2015, pp. 19–77 and 135–153 (ISBN: 9780429168932).

[14] F. J. Baker, R. E. Silverton, and E. D. Luckcock *An Introduction to Medical Laboratory Technology*. Butterworth-Heinemann, 2014, pp. 441–451 (ISBN 978-1483179605).

[15] J. Wójkowska-Mach, A. Różańska, T. Gosiewski, M. Brzychczy-Włoch, and A. Chmielarczyk, *Mikrobiologia z parazytologią: skrypt dla studentów II roku Wydziału Lekarskiego Uniwersytetu Jagiellońskiego Collegium Medicum*. Krakowska Oficyna Naukowa Tekst, 2015 (ISBN: 9788394273019).

[16] M. R. Arabestani, H. Fazzeli, and B. N. Esfahani, "Identification of the most common pathogenic bacteria in patients with suspected sepsis by multiplex PCR", *The J. of Infection in Develop. Countries*, vol. 8, no. 4, pp. 461–468, 2014 (doi: 10.3855/jidc.3856).

[17] M. Ford, *Medical Microbiology*, 3rd ed. Oxford University Press, 2014, pp. 1–32 (ISBN: 9780198818144).

[18] J. R. Jamison, *Man Meets Microbes: An Introduction to Medical Microbiology*. Butterworth-Heinemann, 2014, pp. 1–65 (ISBN: 9781483141626).

[19] W. M. Ahmed *et al.*, "Classification of bacterial contamination using image processing and distributed computing", *IEEE J. of Biomed. and Health Inform.*, vol. 17, no. 1, pp. 232–239, 2013 (doi: 10.1109/TITB.2012.2222654).

[20] S. Trattner, H. Greenspan, G. Tepper, and S. Abboud, "Automatic identification of bacterial types using statistical imaging methods", *IEEE Trans. on Medi. Imag.*, vol. 23, pp. 807–820, 2004 (doi: 10.1109/TMI.2004.827481).

[21] N. Blackburn *et al.*, "Rapid determination of bacterial abundance, biovolume, morphology, and growth by neural network based image analysis", *Appl. and Environmen. Microbiol.*, vol. 64, no. 9, pp. 3246–3255, 1998 [Online]. Available: https://aem.asm.org/content/aem/64/9/3246.full.pdf

[22] P. Perner, "Classification of he–2 cells using fluorescent image analysis and data mining", in *Medical Data Analysis Second International Symposium, ISMDA 2001, Madrid, Spain, October 8-9, 2001. Proceedings*, J. Crespo, V. Maojo, and F. Martin, Eds. *LNCS*, vol. 2199, pp. 219–224 Springer, 2001 (doi: 10.1007/3-540-45497-7_33).

[23] P. S. Hiremath and P. Bannigidad, "Automated Gram-staining characterization of digital bacterial cell images", in *Proc. 6th Int. Conf. on Sig. and Image Process. ICSIP 2009*, Amsterdam, The Netherlands, 2009, pp. 209–211.

[24] B. K. De Bruyne *et al.*, "Bacterial species identification from maldi-tof mass spectra through data analysis and machine learning", *System. and Appl. Microbiol.*, vol. 34, no. 1, pp. 20–29, 2011 (doi: 10.1016/j.syapm.2010.11.003).

[25] M. G. Forero, G. Cristłbal, and M. Desco, "Automatic identification of mycobacterium tuberculosis by Gaussian mixture models", *J. of Microscopy*, vol. 223, pp. 120–132, 2006 (doi: 10.1111/j.1365-2818.2006l.01610.x).

[26] J. Liu, F. B. Dazzo, O. Glagoleva, B. Yu, and A. K. Jain, "Cmeias: a computer-aided system for the image analysis of bacterial morphotypes in microbial communities", *Microbial Ecol.*, vol. 41, no. 3, pp. 173–194, 2001 (doi:10.1007/s002480000004).

[27] C. Cortes and V. Vapnik, "Support-vector networks"', *J. Machine Learn.*, vol. 20, pp. 273–297, 1995 (doi: 10.1023/A:1022627411411).

[28] M. Holmberg *et al.*, "Bacteria classification based on feature extraction from sensor data", *Biotechnol. Tech.*, vol. 12, no. 4, pp. 319–324, 1998 (doi: 10.1023/A:1008862617082).

[29] H. Ates and O. N. Gerek, "An image-processing based automated bacteria colony counter", in *Proc. Int. Symp. on Comp. and Inform. Sci. ISCIS 2009*, Guzelyurt, Cyprus, 2009, pp. 18–23 (doi: 10.1109/ISCIS.2009.5291926).

[30] Ch. Sommer and D. W. Gerlich, "Machine learning in cell biology-teaching computers to recognize phenotypes", *J. of Cell Sci.*, vol. 126, pp. 1–11, 2011 (doi: 10.1242/jcs.123604).

[31] M. Cimpoi, S. Maji, I. Kokkinos, and A. Vedaldi, "Deep filter banks for texture recognition, description, and segmentation", *Int. J. of Comp. Vision*, vol. 118, no. 1, pp. 65–94, 2016 (doi: 10.1007/s11263-015-0872-3).

[32] A. Signorini *et al.*, "Combining the use of CNN classification and strength-driven compression for the robust identification of bacterial species on hyperspectral culture plate images", *IET Comp. Vision*, vol. 12, no. 7, pp. 941–949, 2018 (doi: 10.1049/iet-cvi.2018.5237).

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", 2014 [Online]. Available: https://arxiv.org/abs/1409.1556.

[34] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge", *Int. J. of Comp. Vision*, vol. 115, no. 3, pp. 211–252, 2015 (doi: 10.1007/s11263-015-0816-y).

[35] A. Buetti-Dinh *et al.*, "Deep neural networks outperform human expert's capacity in characterizing bioleaching bacterial biofilm composition", *Biotechnol. Rep.*, vol. 22, 2019 (doi: 10.1016/j.btre.2019.e00321).

[36] B. Liu, S. Wang, R. Long, and K.-Ch. Chou, "iRSpot-EL: identify recombination spots with an ensemble learning approach", *Bioinformatics*, vol. 33, no. 1, pp. 35–41, 2016 (doi: 10.1093/bioinformatics/btw539).

[37] O. Garner *et al.*, "Multi-centre evaluation of mass spectrometric identification of anaerobic bacteria using the VITEK MS system", *Clinical Microbiol. and Infection*, vol. 20, no. 4, pp. 335–339, 2014 (doi: 10.1111/1469-0691.12317).

[38] J. A. Branda *et al.*, "Multicenter validation of the VITEK MS v2.0 MALDI-TOF mass spectrometry system for the identification of fastidious Gram-negative bacteria", *Diag. Microbiol. and Infect. Disease*, vol. 78, no. 2, pp. 129–131, 2014 (doi: 10.1016/j.diagmicrobio.2013.08.013).

[39] G. C. Green, A. D. C. Chan, and M. Lin, "Robust identification of bacteria based on repeated odor measurements from individual bacteria colonies", *Sensors and Actuators B.: Chemical*, vol. 190, pp. 16–24, 2014 (doi: 10.1016/j.snb.2013.08.001).

[40] A. Alvarez-Ordonez, D. J. M. Mouwen, M. Lopez, and M. Prieto, "Fourier transform infrared spectroscopy as a tool to characterize molecular composition and stress response in foodborne pathogenic bacteria", *J. of Microbiol. Methods*, vol. 84, no. 3, pp. 369–378, 2011 (doi: 10.1016/j.mimet.2011.01.009).

[41] X. D. Wang and O. S. Wolfbeis, "Fiber-optic chemical sensors and biosensors (2008–2012)", *Anal. Chemistry*, vol. 85, no. 2, pp. 487–508, 2012 (doi: 10.1021/ac303159b).

[42] A. A. Abdullah *et al.*, "Rapid identification method of aerobic bacteria in diabetic foot ulcers using electronic nose", *Adv. Sci. Lett.*, vol. 20, no. 1, pp. 37–41, 2014 (doi: 10.1166/asl.2014.5306).

[43] N. Yusuf *et al.*, "Comparison of various pattern recognition techniques based on e-nose for identifying bacterial species in diabetic wound infections", *Trans. on Inform. and Commun. Technol.*, vol. 53, pp. 43–59, 2014 (doi: 10.2495/Intelsys130061).

[44] A. Suchwałko, I. Buzalewicz, and H. Podbielska, "Computer-based classification of bacteria species by analysis of their colonies Fresnel diffraction patterns", in *Proc. of SPIE – The Int. Soc. of Opt. Engin.*, San Francisco, CA, USA, 2012, vol. 8212, pp. 82120R–82120R13 (doi: 10.1117/12.907420).

[45] H. Podbielska, I. Buzalewicz, A. Suchwałko, and A. Wieliczko, "Bacteria classification by means of the statistical analysis of fresnel diffraction patterns of bacteria colonies", *in Proc. of Conf. Biomed. Optics*, Miami, FL, USA, 2012, pp. 1–3 (doi: 10.1364/BIOMED.2012.BSu5A.5).

[46] H. Kim, I. J. Doh, A. K. Bhunia, G. B. King, and E. Bae, "Scalar diffraction modeling of multispectral forward scatter patterns from bacterial colonies", *Opt. Express*, vol. 23, no. 7, pp. 8545–8554, 2015 (doi: 10.1364/OE.23.008545).

[47] J. R. Carey *et al.*, "Rapid identification of bacteria with a disposable colorimetric sensing array", *J. of the American Chemical Soc.*, vol. 133, no. 19, pp. 7571–7576, 2011 (doi: 10.1021/ja201634d).

[48] A. Suchwałko, I. Buzalewicz, and H. Podbielska, "Bacteria identification in an optical system with optimized diffraction pattern registration condition supported by enhanced statistical analysis", *Opt. Express*, vol. 22, no. 21, pp. 26312–26327, 2014 (doi: 10.1364/O.E.22.026312).

[49] E. Baee *et al.*, "Portable bacterial identification system based on elastic light scatter patterns", *J. of Biological Engin.*, vol. 6, no. 1, pp. 1–11, 2012 (doi: 10.1186/1754-1611-6-12).

[50] A. Suchwałko, I. Buzalewicz, A. Wieliczko, and H. Podbielska, "Bacteria species identification by the statistical analysis of bacterial colonies fresnel patterns", *Opt. Express*, vol. 21, no. 9, pp. 11322–11337, 2013 (doi: 10.1364/OE.21.011322).

Anna Plichta, Ph.D., graduated comparative literature at the Jagiellonian University in 2007. She also graduated computer science at Cracow University of Technology in 2010. Currently, she is an Assistant Professor at Tadeusz Kościuszko Cracow University of Technology. The main topics of her research are pattern recognition, artificial intelligent systems and e-learning technologies.

https://orcid.org/0000-0001-6503-308X
E-mail: aplichta@pk.edu.pl
Faculty of Computer Science and Telecommunications
Department of Computer Science
Cracow University of Technology
Warszawska 24
31-155 Kraków, Poland