



# $\beta$ -Hemolysis Detection on Cultured Blood Agar Plates by Convolutional Neural Networks

Mattia Savardi, Sergio Benini, and Alberto Signoroni<sup>(✉)</sup>

Department of Information Engineering, University of Brescia, Brescia, Italy  
alberto.signoroni@unibs.it

**Abstract.** The recent introduction of Full Laboratory Automation (FLA) systems in Clinical Microbiology opens to the availability of huge streams of high definition diagnostic images representing bacteria colonies on culturing plates. In this context, the presence of  $\beta$ -hemolysis is a key diagnostic sign to assess the presence and virulence of pathogens like streptococci and to characterize major respiratory tract infections. Since it can manifest in a high variety of shapes, dimensions and intensities, obtaining a reliable automated detection of  $\beta$ -hemolysis is a challenging task, never been tackled so far in its real complexity. To this aim, here we follow a deep learning approach operating on a database of 1500 fully annotated dual-light (top-lit and back-lit) blood agar plate images collected from FLA systems operating in ordinary clinical conditions. Patch-based training and test sets are obtained with the help of an ad-hoc, total recall, region proposal technique. A DenseNet Convolutional Neural Network architecture, dimensioned and trained to classify patch candidates, achieves a 98.9% precision with a recall of 98.9%, leading to an overall 90% precision and 99% recall on a plate basis, where false negative occurrence needs to be minimized. Being the first approach able to detect  $\beta$ -hemolysis on a whole plate basis, the obtained results open new opportunities for supporting diagnostic decisions, with an expected high impact on the efficiency and accuracy of the laboratory workflow.

**Keywords:** Digital Microbiology Imaging  
Full Laboratory Automation · Beta-hemolysis detection  
Convolutional Neural Networks · Image classification

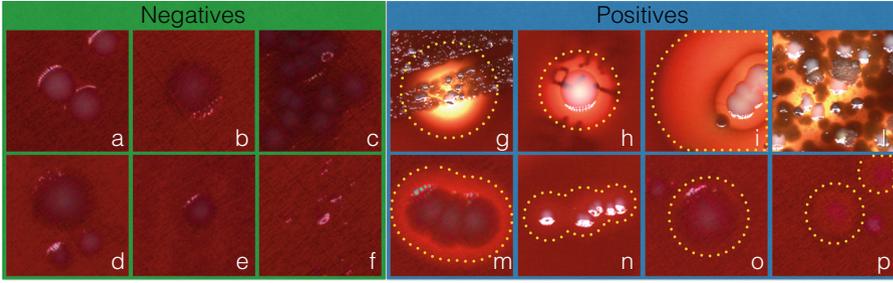
## 1 Introduction

Clinical microbiology is tasked with providing diagnosis and treatment of infectious diseases. The ability to achieve accurate diagnoses in standardized and reproducible conditions is of utmost importance in order to provide appropriate and fast treatment. The gold standard for bacteria identification in the workflow of Clinical Microbiology Laboratories (CML) is bacteria culturing on agar plates. Since traditionally performed almost totally manually, this requires

labor-intensive pre-analytical phases with critical aspects arising with respect to both intra- and inter-laboratory repeatability. Nevertheless, new groundbreaking trends related to the recent diffusion of Full Laboratory Automation (FLA) systems started deeply changing working habits in many clinical microbiology facilities worldwide [1, 2]. A single FLA plant is able to process even thousands of plates per day, generating huge flows of high-resolution digital images (taken during and after plate incubation) to be read on diagnostic workstations. As a consequence, the new field of Digital Microbiology Imaging (DMI) involves high expectations related to the solution of a variety of visual interpretation challenges aimed at supporting and improving the accuracy and speed of the clinical procedures and decisions in CMLs. In this work we focus on the automated detection of one of the main diagnostically relevant features for the assessment of human infections, that is the identification of pathogens'  $\beta$ -hemolytic activity.

### 1.1 Problem Definition

$\beta$ -hemolysis is an effect caused by certain bacteria species growing on blood agar plates that leads to the dissolution of the blood substrate surrounding the colony. The produced visual effect is a yellowish halo visible by holding the plate against the light [3] or on back-lit images acquired by FLA systems under proper plate illumination settings. In many clinical microbiology protocols,  $\beta$ -hemolysis has high impact and it is the (almost) first step of a chain that needs a high sensitivity. This is true especially in throat swabs culture, and when it is important to address streptococci. Moreover, there is a diagnostically relevant information about virulence (see for instance *E. Coli* [4]) that is promptly available from hemolytic activity assessment which is not possible or difficult to achieve by other diagnostic procedures. However, to accurately distinguish  $\beta$ -hemolysis with the naked eye in all possible manifestations is difficult even for a skilled microbiologist. This requires caution and experience and, especially under high labs load, it is an error prone procedure. In Fig. 1 some examples of negative (a-f) and positive (g-p) cases are shown. In the first line of positive examples (g-l)  $\beta$ -hemolysis is easily recognizable, even if appearing in a variety of different morphological forms and textures: in the middle of a confluence growth (g), over a written portion of the plate (h), or forming multiple rings (i) or in heavy mixed situations inside big confluences (l). These situations and their variability configure a first main challenge (we refer to it as *Multiform Challenge, MC*) for a machine which is asked to reliably classify images, usually well interpretable by a microbiologist, by containing false positives (FP) while maintaining a high recall. A second challenge (we refer to it as *Detectability Challenge, DC*) occurs in cases with soft hemolysis like in (m, n) and particularly in diagnostically relevant cases when early detection by humans start being difficult due to the presence of very thin halos (check for example (b) with respect to (o)), or when  $\beta$ -hemolysis is barely visible because hidden under the colony (see (p)). In this case both humans and machine-based techniques are particularly committed to prevent false negatives (FN) by maintaining a suitable degree of precision.



**Fig. 1.** Negative and positive  $\beta$ -hemolysis examples on blood agar plates. The dotted yellow line highlights the  $\beta$ -hemolytic regions.

## 1.2 Related Work and Contribution

So far, there has been only one and very recent work dealing with the problem of automated detection of hemolysis on agar plates [5], where a machine-learning method based on hand-crafted features is able to accurately classify hemolysis on image patches representing single colonies. On the one hand, this previous work does not handle detection on a whole plate basis and is not even able to handle most of the frequently occurring range of cases exemplified in Fig. 1 (under-the-colony, within confluent areas, very thin halo, over the written part cases) which characterize the clinical problem in its real complexity. On the other hand, classification in [5] comprises  $\alpha$ -hemolysis (which generates a brownish halo), which is however virtually absent and of no diagnostic interest for the throat-swab clinical context considered in our work.

Deep learning (DL) approaches, especially those based on Convolutional Neural Networks (CNN), have recently been shown to outperform feature-based machine learning solutions whenever difficult visual tasks and large datasets are involved. Applications of deep learning to medical image analysis started to appear consistently only very recently and nowadays are rapidly spreading [6]. Concerning DL methods in the field of DMI, Ferrari et al. [7] already proposed a system for bacterial colony counting, while Turra et al. [8] started investigating bacterial species identification by using hyper-spectral images. More in general, DL detection methods in Computer Assisted Diagnosis (CAD) contexts have been recently proposed for classification of skin cancer [9], cells and mitosis detection [10, 11], and mammographic lesions [12], to name a few.

In this work, by exploiting a dataset created for the purpose (as described in Sect. 2), we present a  $\beta$ -hemolysis detection technique, based on a region proposal stage (Sect. 3.1) followed by a CNN (Sect. 3.2) which classifies image patches as  $\beta$ -hemolytic or not. Our system is able to effectively cope with the highly diversified behaviour that  $\beta$ -hemolysis displays in the considered CML procedures involving throat swab cultures finalized to respiratory tract infections identification. Our approach overcomes all the limitations of [5] thus resulting the first one capable to work in real complexity conditions (i.e., facing both the above defined challenges *MC* and *DC*). We eventually validate the effectiveness

of the method according to both patch-based and whole-plate tests to evaluate the quality of the classification stage and of the overall system, respectively (Sect. 4).

## 2 Throat Swab Culture Dataset

We collected a dataset from 1,500 culture plates coming from routine lab screening tests and produced by the inoculation of throat swab samples on REMEL 5% sheep blood agar media. Images came from a WASPLab FLA system (by Copan Diagnostics Inc.) which acquires, by linear scanning, 16-mega-pixel RGB color images. For each plate we retrieved both back-light and top-light acquisitions. The ground truth data for the training process consists in throat swab (1,200 plates), randomly selected from a one week of work in a medium size lab, and comprises the segmentation maps produced for the purpose by expert microbiologists that delineated  $\beta$ -hemolytic regions. This dataset is composed of 160 positive plates and 1,040 negative ones. In order to create a blind test-set for the overall evaluation of the system, we labelled another batch composed of 300 new plates acquired two weeks after with respect to the training one. In this case we only required specialists to give information accounting for the presence or not of  $\beta$ -hemolysis. In this case we had a proportion of 51 positive plates and 249 negative ones.

From the image database an image patch dataset can be created by considering  $150 \times 150$  pixel patches extracted from the 1200 fully annotated plate images, labeling them as positives if at least one pixel from the delineated  $\beta$ -hemolytic regions falls inside a  $100 \times 100$  pixel region centered with respect to the patch. Every patch is taken with a 33% of overlap so that every portion of the image falls inside the 100-pixel region. We choose 150 pixels as patch dimension as a good trade-off between the colony dimensions and the required computational effort. To the above patches extracted on a regular grid basis, we added all the patches generated by the differential region proposal approach that we explain in Sect. 3.1, resized to  $150 \times 150$  pixels if needed, where again each patch is labeled negative or positive according to the same rule above. This is done to add more examples similar to those that will be encountered during test-time, when only patches coming from the region proposal are considered. At the same time, the use of a sliding window guarantees to collect a suitable amount of training material, being the region proposal tailored to the reduction of the analyzed patches. Moreover, since following natural CML proportions, negative patches would be about 50 times more than the positive ones, we randomly sample just a portion of them, from each plate, until the patch dataset results balanced. Finally, the full set of patches (about 160k in total) is further divided on a plate basis in two additional sets for the CNN training (70%) and validation (30%) processes.

## 3 $\beta$ -Hemolysis Detection Method

In this section, we describe our approach to  $\beta$ -hemolysis detection, consisting of a patch extraction (region proposal) phase followed by a classification stage

based on a specific CNN architecture. An overall scheme of the proposed solution is depicted in Fig. 2.

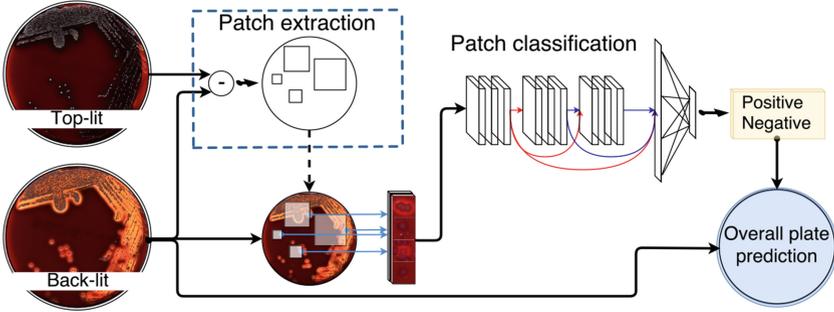


Fig. 2. The overall system for  $\beta$ -hemolysis detection.

### 3.1 Patch Extraction

In a common scenario the plate is covered by colony growth only in a minority portion with respect to the whole substrate. Moreover hemolysis usually involve (with few exceptions) only a small portion of the growth. This is why a sliding window patch extraction mechanisms for hemolysis detection and classification would be highly inefficient. To significantly increase the computational efficiency of our method we exploit the physical effect that hemolysis produces i.e., an erosion of the blood film, which results in a region in which more light is transmitted from below when acquired back-lit. Thus we adopt a region proposal solution which works on a differential image obtained by subtracting the back-lit image from the top-lit one. We process this image by bilateral filter denoising and morphological filtering in order to produce a map composed of high probability  $\beta$ -hemolytic blobs. Specifically, this map is obtained as  $\max(|\text{Img}_{\text{top}} - \text{Img}_{\text{back}}|, t) \bullet K$ , where  $\bullet$  is a morphological closure operating on the denoised differential image with a circular  $5 \times 5$  structuring kernel  $K$ , and where  $t$  is a parameter impacting on the recall of the patch-proposal that mainly depends on the FLA illumination settings and plate manufacturer. All the parameters, including  $t$ , are tuned by using the patch database so as to produce a 100% recall region proposal (no FNs). As a last step we use this map to create a list of possible hemolytic regions to be extracted from the back-light plate for the subsequent classification phase. In particular  $150 \times 150$  patches are created with smaller regions in their centre or by subdivision of larger regions.

### 3.2 Patch Classification

For the patch-classification phase we need a state-of-the-art CNN architecture particularly suited to be used on datasets with similar dimension and complexity

to ours. This is why we selected DenseNet [13] which is composed of a fully-interconnected series of layers that ensure maximum information flow and force an efficient use of the learned representation (Fig. 2 top-right). DenseNet exposes two parameters: the number of layers  $L$  which controls the vertical scale, while the growth rate  $k$  accounts for the horizontal scale (i.e., the number of filters). Moreover, to increase the computational efficiency we add a bottleneck layer before each convolutional layer (solution referred to as DenseNet-BC). We train the network from scratch following Xavier weight initialization. In this case in fact, due to the new type of images, fine-tuning approaches would lead to no performance improvement. We adopt Adam as optimizer, Keras framework with TensorFlow, and a Nvidia GPU. We perform 120 training epochs with an initial learning rate set to 0.01 and factor-two reduction on plateaus.

## 4 Results and Discussion

After a quantitative assessment of the complexity reduction factor produced by the region proposal method, we evaluate the obtained detection performance according to two different criteria: (1) *Patch-based*: we consider the ability to correctly identify and classify patches that present  $\beta$ -hemolysis from negative ones. This metric is useful to evaluate and guide CNN hyper-parameter tuning and training, and accounts for the high performance of the implemented solution in response to both *MC* and *DC* challenges. (2) *Plate-based*: we investigate the ability to correctly classify the whole plate, which is the ultimate clinically relevant target.

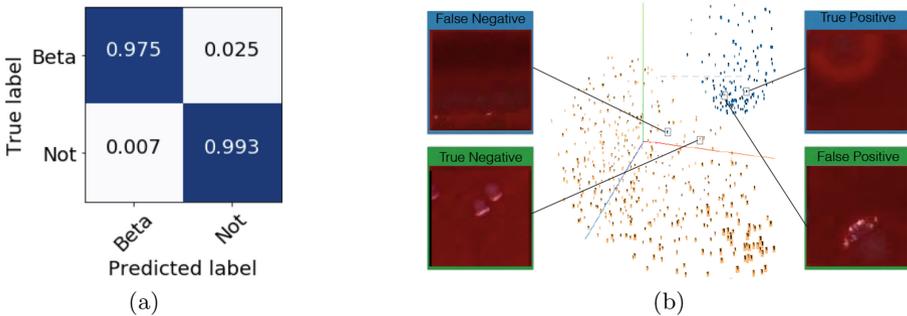
*Patch Extraction.* The adopted region proposal allows to extract image patches containing all regions with a high probability of  $\beta$ -hemolysis occurrence. Following the parameter selection described in Sect. 3.1 we indeed obtained no FN, with a concurrent 20 $\times$  reduction in the number of patches to classify with respect to the sliding window generation used for dataset creation (Sect. 2).

*Patch Classification.* In Table 1 we report some results obtained with different configurations of DenseNet. We achieve best result with a medium capacity model, either using or not the bottleneck layer BC. This can be explained observing that medium-sized models have a number of trainable parameters which is more compatible to the dimension of our dataset. Bigger models tend to overfit and prevent to reach a good generalization. The adoption of conventional radiometric and geometric data augmentation techniques accounts for an improvement of about 0.2% already included in the final score. In Fig. 3(a) we show the confusion matrix of the best classifier (BC-Medium). FN errors are mainly due to borderline cases, which are also very difficult to discriminate to the naked eye, while FP patches are typically caused by light reflections creating misleading color effects on the plate. In the additional material we included both correctly classified patches as well as FP and FN cases. Results are very promising with both recall and precision approaching 99%. This demonstrates a highly

satisfactory response to both the *MC* and *DC* challenges defined in Sect. 1.1. In Fig. 3(b), we show the CNN internal representation of the last hidden layer by using a reduced dimensionality visualization based on t-SNE, where a random portion of the validation patches is taken in input. This allows to appreciate the good level of separability of the two classes (with isolated rare exceptions).

**Table 1.** DenseNet models comparison on patch classification task ( $L$  is the depth and  $k$  the growth-rate as in [13]).

Model	Parameters ( $10^6$ )	Accuracy
Small ( $L = 13, k=12$ )	0.085	0.964
Medium ( $L = 22, k = 24$ )	0.961	0.989
Big ( $L = 40, k = 24$ )	4.182	0.968
BC-Small ( $L = 13, k = 12$ )	0.071	0.988
BC-Medium ( $L = 22, k = 24$ )	0.688	<b>0.990</b>
BC-Big ( $L = 40, k = 24$ )	1.906	0.954



**Fig. 3.** (a) Normalized confusion matrix of the patch-based classifier. (b) T-SNE visualization of the last CNN hidden layer for  $\beta$ -hemolysis discrimination.

*Plate Classification.* We now apply the proposed pipeline to the 300 unseen plates (blind test-set). Without any post-processing we reach 83% precision and 99% recall with only 3 FN plates with a single and very light  $\beta$ -hemolytic colony in challenging conditions (in our cases near the plate border or below a colony). All the images of FN plates are given in the additional material, with TP, TN and FP meaningful examples as well. By using instead one third of the blind test-set to tune the classification threshold, and test again on the remaining, we reach a 90% precision with the same recall.

Finally, we compare our solution against the one in [5], using their publicly available plate based test-set. In a fair comparison, which requires the exclusion

of all the colonies grown over the written portion, we reach almost the same recall of 96% with a significantly increased precision by 12% up to 87%. Beyond this improvement, our method handles  $\beta$ -hemolysis detection inside confluences (not considered in [5]) and over the usually large written plate portions as well, thus standing as a system able to better cope with the real problem complexity.

## 5 Conclusion

We presented a fully automatic method for  $\beta$ -hemolysis detection on blood agar plate images. We operated with a complexity reduction region proposal and with a representation learning approach based on DenseNet CNN for the classification of both single patches and full plates. Our solution evidenced highly satisfactory performance on a blind test-set and overcomes performance and functional limitations of a previous work. As a next step, we would like to integrate our method in a diagnostic workflow with the microbiologist-in-the-loop. Our feeling is that thanks to the achievements reached on both the multiform and detectability challenges, further impact can be expected in terms of consistency and efficiency as suggested in [14], where the combination of deep leaning predictions with the human diagnostic activity led to significantly improve the total error rate.

## References

1. Bourbeau, P.P., Ledebor, N.A.: Automation in clinical microbiology. *J. Clin. Microbiol.* **51**(6), 1658–1665 (2013)
2. Doern, C., Holfelder, M.: Automation and design of the clinical microbiology laboratory. *Manual of Clinical Microbiology* (2015)
3. Hogg, S.: *Essential Microbiology*. Wiley, Chichester (2005)
4. Jorgensen, J., Pfaller, M., Carroll, K.: *Manual of Clinical Microbiology*, 11th edn. ASM Press, Washington (2015)
5. Savardi, M., Ferrari, A., Signoroni, A.: Automatic hemolysis identification on aligned dual-lighting images of cultured blood agar plates. *Comput. Methods Programs Biomed.* **156**, 13–24 (2017)
6. Litjens, G., et al.: A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017)
7. Ferrari, A., Lombardi, S., Signoroni, A.: Bacterial colony counting by convolutional neural networks. In: *Proceedings of the IEEE Engineering in Medicine and Biology Society Conference*, pp. 7458–7461 (2015)
8. Turra, G., Arrigoni, S., Signoroni, A.: CNN-based identification of hyperspectral bacterial signatures for digital microbiology. In: Battiato, S., Gallo, G., Schettini, R., Stanco, F. (eds.) *ICIAP 2017*. LNCS, vol. 10485, pp. 500–510. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-68548-9\\_46](https://doi.org/10.1007/978-3-319-68548-9_46)
9. Esteva, A., et al.: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**(7639), 115 (2017)
10. Xie, Y., Xing, F., Shi, X., Kong, X., Su, H., Yang, L.: Efficient and robust cell detection: A structured regression approach. *Med. Image Anal.* **44**, 245–254 (2018)
11. Li, C., Wang, X., Liu, W., Latecki, L.J.: Deepmitosis: mitosis detection via deep detection, verification and segmentation networks. *Med. Image Anal.* **45**, 121–133 (2018)

12. Kooi, T., et al.: Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* **35**, 303–312 (2017)
13. Huang, G., Liu, Z., Weinberger, K.Q., van der Maaten, L.: Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, p. 3 (2017)
14. Wang, D., Khosla, A., Gargeya, R., Irshad, H., Beck, A.H.: Deep learning for identifying metastatic breast cancer. arXiv preprint [arXiv:1606.05718](https://arxiv.org/abs/1606.05718) (2016)