

# Automatic hemolysis identification on aligned dual-lighting images of cultured blood agar plates

Mattia Savardi <sup>\*</sup>      Alessandro Ferrari<sup>\*†</sup>      Alberto Signoroni<sup>\*\*</sup>

<sup>\*</sup>*Information Engineering Dept., University of Brescia, Brescia (Italy)*

<sup>†</sup>*Futura Science Park, Copan Group S.p.A., Brescia (Italy)*

---

## Abstract

*Background and Objective.* The recent introduction of full laboratory automation systems in clinical microbiology opens to the availability of streams of high definition images representing bacteria culturing plates. This creates new opportunities to support diagnostic decisions through image analysis and interpretation solutions, with an expected high impact on the efficiency of the laboratory workflow and related quality implications. Starting from images acquired under different illumination settings (top-light and back-light), the objective of this work is to design and evaluate a method for the detection and classification of diagnostically relevant hemolysis effects associated with specific bacteria growing on blood agar plates. The presence of hemolysis is an important factor to assess the virulence of pathogens, and is a fundamental sign of the presence of certain types of bacteria.

*Methods.* We introduce a two-stage approach. Firstly, the implementation of a highly accurate alignment of same-plate image scans, acquired using top-light and back-light illumination, enables the joint spatially coherent exploitation of the available data. Secondly, from each segmented portion of the image containing at least one bacterial colony, specifically designed image features are extracted to feed a SVM classification system, allowing detection and discrimination among different types of hemolysis.

*Results.* The fine alignment solution aligns more than 98.1% images with a residual error of less than 0.13mm. The hemolysis classification block achieves a 88.3% precision with a recall of 98.6%.

*Conclusions.* The results collected from different clinical scenarios (urinary infections and throat swab screening) together with accurate error analysis demonstrate the suitability of our system for robust hemolysis detection and classification, which remains feasible even in challenging conditions (low contrast or illumination changes).

*Keywords:* Digital Microbiology Imaging, Full Laboratory Automation, Hemolysis identification, Machine learning, Image alignment, Image classification

---

## 1. Introduction

Besides its leading role in the development of modern microbiology, bacteria culturing on agar plates remains a gold standard procedure for bacteria identification in the workflow of Clinical Microbiology Laboratories (CML) all around the world. By looking at the bacteria culture, the skilled microbiologist obtains a first (so called presumptive) interpretation of possible ongoing infections affecting the patients' health. This is a fundamental step in the diagnostic process to initiate a prompt and appropriate antibiotic therapy and to guide the selection of specific analytic and diagnostic phases (e.g. antibiogram, MALDI-TOF mass spectrometry, PCR for pathogen DNA or RNA amplification) that allows the confirmation or definition of

15 a more effective treatment. Moreover, there is diagnostically relevant information about pathogenic identification and virulence that is immediately available from bacteria culturing which is impossible or difficult to see from other procedures. One of the most important examples from this is hemolysis associated with bacteria colonies cultured on blood agar plates.

Hemolysis is a peculiar process caused by hemolysin, a group of proteins produced by certain microorganisms, causing the lysis (i.e. the dissolution) of the red blood cell membrane in the growth substrate. The ability to identify the occurrence of this process is very important in order to classify known types of pathogens and, in some type of analysis, such as throat swab screening, is a very effective way of distinguishing positive samples. Moreover, the presence of hemolysis is also a factor of virulence for some bacteria (see for instance E. Coli [1]). Traditionally,

---

<sup>\*</sup>Corresponding author email: [alberto.signoroni@unibs.it](mailto:alberto.signoroni@unibs.it)

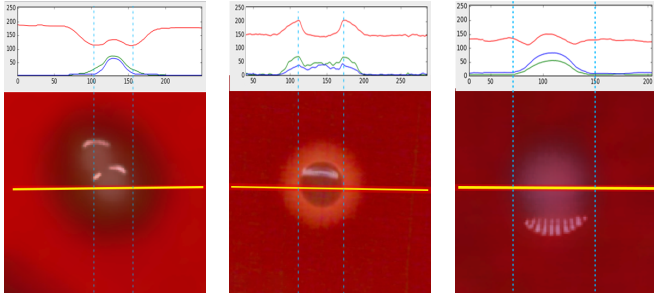


Figure 1: Color distribution (RGB channels) along the colony axis (yellow line) in the cases of Alpha (Left), Beta (Middle) and Gamma (Right) hemolysis. The images are taken from back-lit plates.

microbiologists seek the presence of hemolysis by looking at back-lit plates.

Hemolysis can manifest itself in three different ways [2] (also depicted in Fig.1):

- Alpha ( $\alpha$ ) hemolysis: partial lysis of red blood cell membrane that visually produces a green or brown discoloration in the medium.
- Beta ( $\beta$ ) hemolysis: associated with a complete lysis of red blood cells, it produces a lightened (yellowish) and near transparent halo around the colony.
- Gamma ( $\gamma$ ) hemolysis: despite the apparently contradictory term, this indicates the lack of hemolysis. There should be no reaction in the surrounding medium so that the area near the margin of the colony looks uniform and red (the color of the blood agar substrate).

The assessment of the presence of hemolysis has a significant impact in terms of the speed of presumptive identification of dangerous pathogens but does not involve trivial visual tasks for the microbiologist. The difficulties faced are due to the fact that timely recognition may not be easy because of the mild effects of hemolysis, especially in its early stages. In fact, hemolysis produced by bacteria takes time to form, and so the sooner even the initial stages can be recognized, the faster the diagnostic process will be, because it allows the identification on plates incubated for a shorter time.

The need to find automated solutions for the quantitative assessment of bacteria growth (especially for the tedious but diagnostically relevant counting of the number of Colony Forming Units CFU on a plate) is documented in surprisingly far-dated trials [3][4]. Despite the merit and intuitions of these pioneering works, the complexity and variety of pre-analytical procedures in CMLs have meant that sample preparation and plate reading remained manual and exclusively carried out by skilled technicians and microbiologists until recently. Visual CMLs diagnostics using computer vision/intelligence solutions is now strongly motivated by the ongoing rapid diffusion of Full Laboratory Automation (FLA) systems [5] which are expected to

change radically the working patterns of modern CMLs. FLA systems are able to process automatically hundreds to thousands of samples per day in standardized conditions and produce huge streams of digital images documenting the bacterial growth from the seeded and incubated plates. This new digital revolution is establishing needs and new outlooks of advanced image analysis and understanding tools able to improve the speed of diagnostics and reliability of the widespread clinical procedures involving (digitized) bacteria culture plates. As a reference FLA system for the creation of our experimental database (see Sec.2), we consider the WASPlab<sup>TM</sup> (Copan, Italy), which is able to record high definition plate images at different incubation times and in different lighting conditions, including combinations of front-light and back-light, to produce naturally appearing images for specialist readings on diagnostic workstations. The hemolysis halo can vary greatly in terms of both color and dimension: the width of hemolysis that we want to isolate goes from 0.25 mm to 2.5 mm and over. High spatial resolution images such as those generated by the line scan system mounted on WASPlab enable the observation of phenomena which are barely visible to the naked eye and our objective is to identify hemolysis as early as possible.

There are two main technical achievements of our work: we designed a keypoint based solution to achieve the robust, fast and accurate alignment of images acquired under widely different lighting conditions and displaying critical aspects that make this task particularly challenging. Based on this alignment we developed a joint dual-image classification of bacterial colony segments for the detection and classification of even the subtle effects produced by hemolysis. This is a task that is often problematic even for the skilled specialist and for which we obtained fully satisfactory results in different clinical scenarios (a near perfect recall on unseen images from clinical routine at the cost of an acceptable number of false positives).

An overall scheme of the proposed system is presented in Fig.2. This is divided in two consecutive phases. The first issue is related to the fact that the information relevant to the precise detection of hemolytic effects is distributed over images taken with different illumination settings, so these images must first be finely spatially co-registered. This is mainly due to the fact that colony borders are clearly visible in front-light images (while they are blurred in back-light ones), whereas the hemolytic halo is only visible in back-light images. The alignment cannot be taken for granted in practice and, in our case, only a coarse alignment of the two kinds of image is in fact initially provided. We exploit an automatically extracted feature-point based alignment procedure (see Sec.3) that, notwithstanding the high morphological and color heterogeneity of the considered images, actually succeeds, in the large majority of cases, in achieving the desired alignment result. Once the images are co-registered, we adopt a machine learning approach based on specifically designed feature evaluation and SVM classification (see Sec.4) to de-

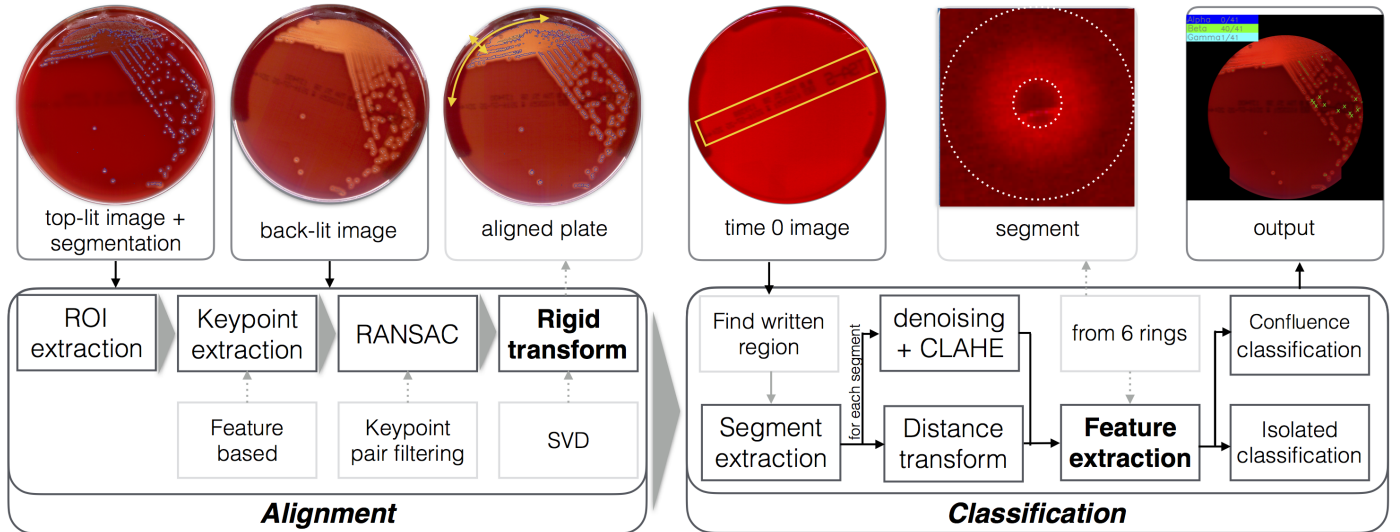


Figure 2: General scheme of the proposed automated hemolysis identification system, comprising plate alignment (on the left) and classification phases (on the right).

130 tect and distinguish between the different types (Alpha, Beta or Gamma) of hemolysis on both single colony and whole plate setups. The experimental tests, documented in Sec.6, were conducted on different clinically relevant scenarios and demonstrate the workability and high performance of our solution. Further considerations and a preliminary evaluation of the clinical impact of the proposed method are given in Sec.7.

## 2. Dataset creation

140 Since the assessment of the effect of hemolysis requires the presence of red blood cells, here we only consider bac-teria cultures on blood agar plates. Blood agar is a generic medium (i.e. it allows the growth of very different types of bacteria) and is one of the most commonly used in microbiology analysis. It is cheaper than other more selective media but, on the other hand, it allows a greater variety of morphology and concentration (load) in the produced growth. To meet the requirements of high reliability and robustness of the diagnostic process, a database of hemolytic and non-hemolytic colonies must be created that actually represents this variety, including many hemolysis examples of different shape, intensity and dimension. Failure to recognize hemolytic segments (i.e. false negatives) must not occur. With this in mind, we collected a dataset of 235 plates, produced by the inoculation on REMEL 5% sheep blood agar media of urine samples collected during routine lab screening tests. Images have been digitized by means of WASPLab automation system which acquires, by linear scanning, 16-mega-pixel color images. These high spatial resolution images (0.0265 mm/pixel) are produced by moving the plate under a trilinear camera equipped with telecentric lenses. Different lighting settings can be used according to laboratory practice. In our case, two different illumination modes were adopted for each plate:

a top-lit image, i.e. where the plate is illuminated from above, and a back-lit image, i.e. where the plate is illuminated from below, but still scanned from the top. With the acquisition of the back-lit images, a moderate top-light component is present in order to allow better visual interpretation of these images (microbiologists are familiar with this type of illumination). As we will see, this also turns out to be useful for the fine plate alignment phase. Moreover, for each plate, we also have a *time-0* image, acquired before the streaking that is used to normalize different acquisitions with respect to variations of the growing substrate. From the digitized plates, we were able to extract (after the segmentation described in [6]):

- 1946 countable colony segments, i.e. containing up to 5 colonies,
- 298 confluent growth segments, i.e. containing more than 5 colonies.

Some examples are shown in Fig.3. Approximately 6% of the dataset is related to  $\alpha$ -hemolysis, 17% to  $\beta$ -hemolysis, and the rest to  $\gamma$ -hemolysis (i.e. non-hemolytic) segments. These proportions reflect both the variability and the distribution encountered in clinical practice. A GUI was created to facilitate the specialists in the labeling process and to reduce interpretation errors as far as possible. In particular, a first microbiologist assigned one of the three labels to each segment, along with an optional free text annotation (mainly used to remark doubtful cases), while a second expert specialist double-checked the label selection and the annotations. In a third step, only the plates with residual uncertainties (about 5% of the total) were discussed jointly and a shared decision was made. Residual disagreement or other difficulties did not arise from the interactions between the domain experts.

We selected segments which contain only one type of

hemolysis, discarding uncertain cases. Metadata was created for each segment comprising a unique segment identifier, its bounding box, the type of hemolysis it exhibits and possible annotations. This information is also linked to the analyzed plate by means of an identifier and saved in *json* format for an easy dataset management.<sup>1</sup>

Since blood agar plates are widely employed in bacteriology labs worldwide and considering that sought pathogens are almost the same in many clinical contexts, our database can be considered a good representation of the clinical interest even beyond urine tract infections screenings. This is discussed in Sec.7, where preliminar validation of the system on plates cultured from throat swab collection specimens is presented.

### 3. Fine plate alignment

As pointed out above, hemolysis is caused by a certain amount of toxins produced by micro-organisms. The action of these agents leads to an effect that is located in a ring virtually extending the colony margin. The dimensions, intensity and shape of the produced halo may vary greatly, but the area of interest is always concentric in relation to the colonies (and, more in general, to colony aggregates of the same strain) and of even thickness along their perimeter. Therefore, an accurate mask of segments is required, in order to analyze only the region where we could expect hemolysis, starting from the margin of the colony (aggregate). Moreover, since the hemolysis classification is independent from the colony appearance, we prefer to exclude the colonies themselves from the learning process, because they could carry information confusing the classifier. Unfortunately, back-lit images are not really suitable for accurate colony boundary identification. This is due to the hemolysis effect itself and to the light filtered and scattered into the blood substrate that blurs and sometimes masks the real border of the colonies. This happens in conjunction with the relatively high noise present in back-lit images, especially if we seek small colonies.

<sup>1</sup>The complete Hemolysis Dataset will be released for research use on the website [www.microbia.org](http://www.microbia.org).

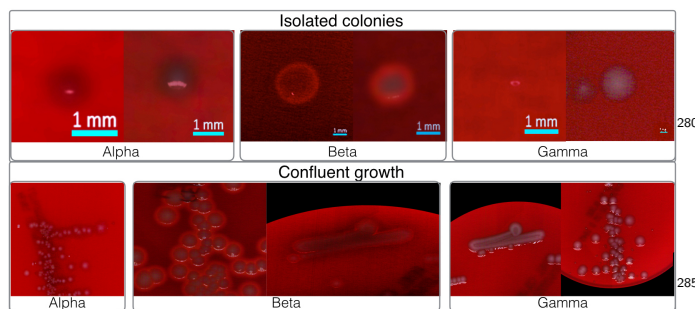


Figure 3: Example of hemolytic segments included in the dataset taken from back-lit images. Here we modified the gamma and contrast of these images to improve the visibility of the hemolytic regions.

A segmentation mask can be much more successfully extracted from the top-lit image, but to do so, we absolutely need to find a robust method of accurately co-registering the two images. This is needed because even just a few pixels of misalignment worsen the performance of the classification in that this would lead to extracting features from incorrect positions; this is relevant since we also want to identify very thin halos (in the order of few pixels). Unfortunately, the desired level of alignment cannot be taken for granted in general. In particular, mechanical vibrations on the conveyor during the transportation of the plates and small stop-and-start imperfections of the scanning intervals cause image misalignments that are present on images of the same plate acquired using different lighting. Although the WASPLab software provides barcode-based coarse image alignment, this is not accurate enough for our purposes.

To solve our fine alignment problem, we initially considered different approaches. Information-theoretical techniques, like the one in [7][8], besides being quite slow, do not work in our scenario probably because the mutually informative visual portion that could be used for the alignment occupies only a tiny portion of the whole images. We then moved to considering feature-based methods, where the alignment is performed by finding distinctive keypoints in both images [9]. After keypoint location using saliency-based analysis (aka *feature extraction*), an associated description of the nearby region of each keypoint is computed (*feature description*). For image alignment applications, keypoint detection and description should allow candidate correspondence to be found between co-located keypoints in the images to be aligned. The feature extraction phase should return a considerable number of keypoints whose location is robust to image variations such as different illumination and viewpoint. In our case, although top-lit and back-lit images seem to be too morphologically different, the presence of a top-light component in back-lit images, though subsidiary, actually creates the possibility of detecting common feature points on which to estimate reliably an image alignment transformation.

It is clear that, if we cannot extract at least two corresponding keypoint pairs, the problem of rigid transformation estimation does not have a unique solution, and all the following steps are compromised. On the contrary, in presence of multiple correspondences, we should be able to discard outliers and possibly skim less accurate correspondences for the estimation of a solution that minimizes the overall keypoint alignment error. For this reason, the descriptors associated with co-located keypoint must be quite robust to image variations in order to promote easily recognizable correspondences. Another significant problem is the relevant similarity of colonies of the same strain, which is likely to generate fake correspondences. This requires a robust rigid transformation estimation with effective outlier detection and rejection.

Fig.4 shows a diagram of the alignment stage with visual examples of the functional blocks described below.

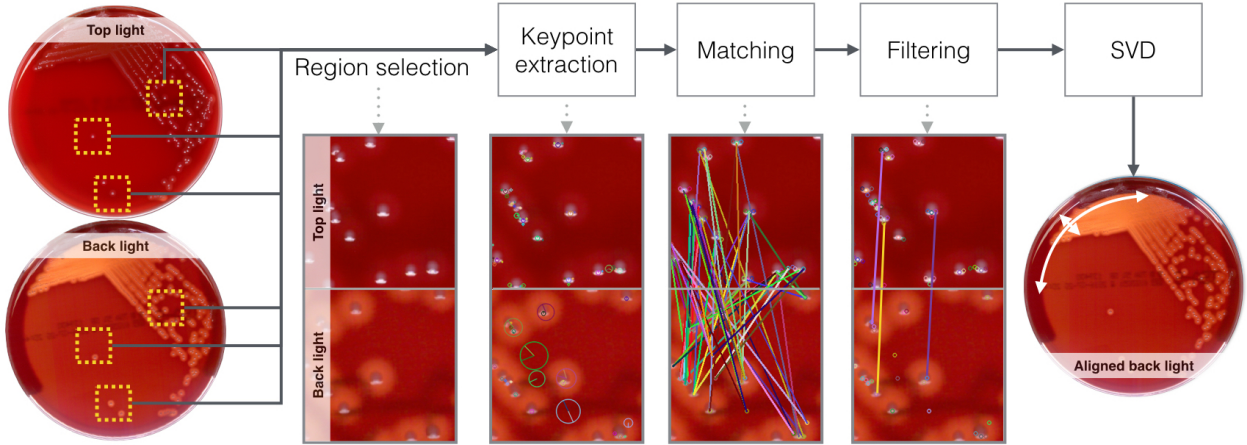


Figure 4: Fine alignment block pipeline. In input, there are the two images from the same plate with different illumination, in which we select the minimum number of regions that leads, at the end of the keypoint extraction, matching and filtering, to obtaining at least 30 keypoint pairs. Finally, we apply SVD to estimate the correct transformation.

### 3.1. Keypoint extraction

We considered many of the most popular feature extraction [10] and description tools [11],[12]. Comparative studies highlight the application-dependency of their performance [13],[14],[9]. We then conducted rapid testing to select the most performing techniques in our context. Given the nature of the images of interest, the most appropriate keypoint detection approaches here are those based on the analysis of a differential multiscale representation of the image, which allows the localized search of visually interesting regions (blob detection) [9],[10]. With regard to keypoint descriptors, binary-valued descriptors [15] did not demonstrate robustness to the changes in image illumination and this is particularly evident for dark images. Moreover, since binary descriptors are usually paired with corner detectors, they often fail to extract enough keypoints in images, as in our case, where rounded shapes prevail.

Real-valued (image gradient based) methods performed better because, in general, they find good matching inside colonies. Indeed, the problem here is mostly that the bottom illumination, as well as causing a higher noise level, highlights effects, such as hemolysis, that visually change the margin of the colonies, introducing discrepancies in the keypoint description.

Therefore, we considered many suitable keypoint extraction and description approaches and selected the two we found to give best performance, which turned out to be SIFT [16] and KAZE [17][18]. Our pre-selection proved to be in line with studies that consider wide benchmark and high resolution images [14].

SIFT features [16] are detected by finding local maxima in a Difference of Gaussian (DoG) pyramid, generated from the image, which works as a blob detector. The SIFT descriptor is based on a Histogram of Oriented Gradients (HOG). Rotation invariance is assessed by assigning the orientation derived from HOG to each keypoint, and then

rotating the descriptor according to the main direction.

In contrast, the KAZE detector builds a nonlinear scale-space, where nonlinear diffusion filters are used instead of Gaussian filters, in which important image details can be highlighted while removing noise [17]. The KAZE descriptor is an adaptation to the nonlinear scale-space of the M-SURF descriptor [19]. The rotation invariance is obtained in a similar way to SIFT and SURF, computing the main orientation of the descriptor.

### 3.2. Keypoint matching

Once the keypoints have been derived from the two images, either with SIFT or KAZE, we need a matching system to test all the possible correspondences. Since the descriptors should be invariant to rotation, translation and illumination, a common method is to adopt an exhaustive approach, trying all the possible matches to find the best one, typically relying on an  $L_2$  distance measure on descriptors. As an alternative, we also tested a solution based on FLANN (Fast Library for Approximate Nearest Neighbors) [20]. This approach is faster because it searches for an approximate nearest neighbor solution, although it may be slightly less accurate. However, many colonies on the plate can be very similar, and this is likely to generate many ambiguous correspondences. Therefore, starting from an ordered list of distances between each keypoint in the first image and all the keypoints in the other image, we introduce a method of retaining only matches that are far enough from the second closest one in each list (as also suggested in [16]). Following appropriate tests, a distance ratio threshold of 0.7 was adopted for this correspondence selection. This turned out to be a good tradeoff between the intrinsic similarity between keypoints of interest and the need to have unambiguous matchings.

### 3.3. Keypoint pair filtering and alignment estimation

The selected matches are used to estimate a mapping function that can be applied to the target image in order to

co-register it to the reference image. Since we are dealing only with possible rotation or shifting along the two axes of the plate due to mechanical vibration induced by the conveyor during acquisition, the mapping function should be a 2D rigid transformation. Therefore, we must guarantee the preservation of distances between every pair of points and, to do that, we exploit the method described in [21] for rigid transformation estimation by means of least-square minimization over a singular value decomposition (SVD). In particular, we want to compute the optimal translation vector  $\mathbf{t}$  and rotation matrix  $R$  that minimize:

$$\sum_{i=0}^n \|(R\mathbf{p}_i + \mathbf{t}) - \mathbf{q}_i\|^2 \quad (1)$$

with  $P = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_n\}$  and  $Q = \{\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_n\}$  the two sets of corresponding keypoint from the target and reference image respectively. This leads to a transformation that, if applied to the back-lit image, produces an aligned version in relation to the top-lit image (note that  $\det(R) = 1$  because we want a rigid transformation).

Unfortunately, a good descriptor matching score is not guaranteed to be spatially correct. In our case, the high dissimilarity of the two acquisitions, in conjunction with the morphological similarity that can exist inside colonies, is likely to lead to fake keypoint correspondences, and is exacerbated by the noise present in the back-lit image. Such outliers must be removed to avoid worsening the rigid transformation estimation. To address this problem, a solution based on the Random Sample Consensus (RANSAC) algorithm [22] can be adopted. This is an iterative resampling technique that generates good candidate solution that, in our case, corresponds to a vector of at least two matches.

A generic RANSAC iteration  $k$  works as follows:

- Randomly select a subset of original keypoint pairs  $(P^{(k)}, Q^{(k)})$ ;
- Compute, using the SVD-based method, a rigid transform estimation  $(R^{(k)}, \mathbf{t}^{(k)})$  using the previously selected keypoint pairs;
- Test all the remaining keypoint pairs  $(P \setminus P^{(k)}, Q \setminus Q^{(k)})$  against the computed model according to the projection error  $e^{(k)}$ . If  $e^{(k)} < T$  (where  $T$  is a predefined threshold) the selected keypoint pairs are considered inliers;

$$e^{(k)} = \sqrt{\sum_{i=0}^{|P^{(k)}|} \|\mathbf{p}_i^{(k)} - (R^{(k)}\mathbf{q}_i^{(k)} + \mathbf{t}^{(k)})\|^2} \quad (2)$$

- $k \leftarrow k + 1$ .

This is repeated until an iteration limit  $K$  is reached on  $k$ . Thanks to the low complexity of RANSAC, high values of  $K$  (order of hundreds) can be adopted without significantly increase the computation time. Eventually, the SVD is recomputed with all the keypoint pairs found as inliers to find the final  $(R, \mathbf{t})$  transformation.

### 3.4. Random region selection speed-up

Since there are many steps in the analysis of each plate image and laboratories (and thus FLA systems) are required to handle as many as one thousand plates per day, the processing time becomes a critical parameter. Thus we are interested in reducing the time needed to estimate the  $(R, \mathbf{t})$  transformation as much as possible. The process can be accelerated by making use of the relevant *a-priori* information we have: the colony segmentation mask. The segmentation mask components can be considered as regions of interest to localize the keypoint extraction. Therefore, we can start to extract features only from a small and randomly chosen number of regions around colonies instead of working on the whole image.

Using keypoint-pair extraction as a central element, we implement a co-registration method that automatically selects the minimum number of keypoints in order to accelerate this phase, without any unwanted compromise in terms of robustness and precision of the results. This works by recursively adding regions until we reach at least 30 keypoint pairs before estimating the rigid transformation. The choice of 30 correspondences represents a cross-validated value: we empirically verified that the error decreases exponentially in relation to valid keypoint pairs and, beyond that threshold, we have almost no improvement.

## 4. Hemolysis classification

Once the front- and back-lit images of the same plate are finely co-registered, we are ready to jointly exploit the information deriving from both of them. Observing a line passing through the center of the colony, it is immediately clear that there are quite different trends in the trichromatic pixel levels for the three types of hemolysis (Fig.1). This is, with good approximation, invariant with respect to the chosen angle, if we do not consider the light reflections on the colonies. In particular, observing the Alpha hemolysis case and focusing on the red component, we observe a decreasing trend, from a background value, along the hemolysis halo, until the margin of the colony is reached. In the Beta case, we observe dual behavior, since the pixels become brighter approaching the colony. Finally, in presence of Gamma hemolysis, the red color distribution around the colony remains, as expected, in the range of the background value. These observations formed the starting point for the design of the automatic hemolysis classification system described in the following.

### 4.1. Preprocessing

Back-lit images are quite noisy and this can cause problems, especially for the interpretation of very small colonies. For this reason, all the segments to be classified are pre-processed with a bi-lateral filter [23] for noise reduction. This guarantees edge-preserving filtering that is computationally fast and that works without any particular assumption. This way, the border of the colonies, especially

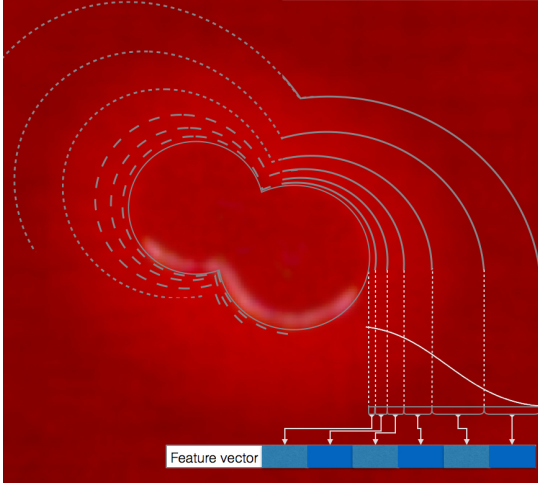


Figure 5: Feature extraction from six concentric rings.

the light sparkle, will not be faded, leaving the region  
 around the colonies, from which we will extract relevant  
 features, without unwanted components from the margins.  
 Moreover, this choice almost avoids altering the computa-  
 tional load of the classification pipeline. Since the images  
 have low contrast, and in order to highlight the hemolysis  
 effect, we also apply Contrast Limited Equalization  
 (CLAHE) [24] to the red channel only. CLAHE works  
 only on small regions in the image, enhancing the con-  
 trast of each of them in an adaptive way. Adjacent re-  
 gions are thus combined using bi-linear interpolation to  
 get rid of the artificial boundaries created by equalizing  
 each region in a different way. Since we apply only soft  
 equalization, this does not introduce any arbitrary varia-  
 tion that could create difficulties for the learning phase.  
 CLAHE limits amplification by clipping the histogram at  
 a predefined value, thus it does not discard a peak in the  
 histogram, but redistributes it among all the bins. The  
 parameters should be chosen cautiously because, as a side  
 effect, CLAHE also amplifies noise and we clearly do not  
 want to trade equalization for a problematic increase of  
 noise. However, thanks to the application a soft equaliza-  
 tion only, this problem does not occur.

#### 4.2. Feature selection

The previously emphasized behavior of the three types  
 of hemolysis prompts us to extract features from an area  
 surrounding the colony segment margin, excluding the seg-  
 ment itself (that does not carry information about hemolysis),  
 in a way that is invariant to the considered direction.

In order to implement this idea, we first divide the  
 area surrounding the colony (colony segment) into six con-  
 centric rings (curves), as shown in Fig.5. These curves  
 are created by setting thresholds for a distance transform  
 computed from the margin of the colony segment. The  
 thresholds are chosen to be initially close to the colonies,  
 and increasingly distant from each other while we become  
 more distant from the colony segment boundary. From

each region we extract the mean value, the variance and  
 also a color histogram with eight bins for each color chan-  
 nel. This guarantees shape independence, making features  
 robust to different types of colony segments. A Feature  
 Vector  $\mathbf{FV}$  is created by taking the difference of the mean  
 and variance of consecutive regions  $R_i$ , and by concate-  
 nating them with the histograms:

$$\begin{aligned}
 \mathbf{M} &= [M_i^c]_{i \in [1, \dots, 6], c \in \{R, G, B\}} \\
 M_1^c &= \text{mean}(R_1)^c \\
 M_i^c &= \text{mean}(R_i)^c - \text{mean}(R_{i+1})^c, i \neq 1 \\
 \mathbf{V} &= [V_i^c]_{i \in [1, \dots, 6], c \in \{R, G, B\}} \\
 V_1^c &= \text{var}(R_1)^c \\
 V_i^c &= \text{var}(R_i)^c - \text{var}(R_{i+1})^c, i \neq 1 \\
 \mathbf{H} &= [H_i^{(b,c)}]_{i \in [1, \dots, 6], b \in [1, \dots, 8], c \in \{R, G, B\}} \\
 H_i^{(b,c)} &= \text{ColHist}(R_i) \\
 \mathbf{FV} &= [\mathbf{M}, \mathbf{V}, \mathbf{H}]
 \end{aligned} \tag{3}$$

for a total feature dimension of 3108.

In order to make the classification as independent as  
 possible from factors affecting the background color of the  
 agar (different agar manufacturer, plate conservation con-  
 ditions, illumination variations, color calibration settings),  
 we take advantage of *time-0* images available from WAS-  
 PLab (acquisitions of the plate just before streaking) by  
 normalizing values extracted from each ring in relation to  
 the background mean and variance values derived from  
*time-0* images. As we will verify, subtracting the mean  
 from each value coming from the concentric rings and di-  
 viding by the variance leads to a considerable increase in  
 possibility of adapting our method to different plates and  
 acquisition parameters.

#### 4.3. Classification

Given the high number of  $\mathbf{FV}$  components, dimen-  
 sionality reduction is advisable prior to machine learning.  
 Principal Component Analysis (PCA) [25] allows only the  
 most significant vectors to be kept by projecting the in-  
 put data using Singular Value Decomposition, so that the  
 maximal data variance (energy) is retained. Then, we clas-  
 sified the reduced feature vectors using a Support Vector  
 Machine (SVM) [26] approach with Radial Basis Function  
 (RBF) kernel. In particular, to address our multi-class  
 classification problem, we used a *one-against-one* commit-  
 tee, where three binary classifiers are trained to discrimi-  
 nate between every possible couple of different classes, and  
 a majority decision is taken.

Since there are appreciable visual differences between  
 isolated (countable) and confluent colony segments in the  
 dataset, we keep these categories separated to achieve bet-  
 ter results. Therefore, two independent classifier trainings  
 are implemented accordingly. For each classifier, the re-  
 lated dataset portion is split into a 70% training set and a  
 30% validation set. Cross-validation is performed on the  
 30% of the training set, then included for training with

the final model. Finally, a Recursive Feature Elimination (RFE) algorithm [27] is applied to select features by recursively considering smaller and smaller subsets: at first, the training is performed with the initial set of features (retained PCA components); subsequently, RFE considers smaller and smaller feature sets by pruning features with the lowest weight. The weights correspond to the coefficients used by SVM to make its prediction. This procedure is repeated recursively on the pruned set, keeping only those coefficients that do not cause the (cross-validation) classification score to decrease.

Testing relies on the colony counting described in [6] to establish to which of two types a given segment belongs and to select the proper classifier.

#### 4.4. Written portion issues

Ad-hoc solutions may be required to deal with specific issues that can arise depending on the plate manufacturer. This problem concerns the presence of written portions printed by the manufacturer on the bottom of the plate (Fig.6). These writings are almost invisible in top-lit images but their shadow, blurred by the light scattering within the agar, is clearly visible with back-lit images. Any bacteria growing in this area, can prevent correct hemolysis interpretation (especially in case of type  $\alpha$ ). Therefore, where present, these regions, should be excluded from the detection and interpretation of hemolysis on back-lit images. The information printed on the plates is extremely important (type of agar, best-by date and plate manufacturer), but it would be advisable to have this data printed on the side, so as not to interfere with computer vision techniques. In our case, in order to estimate the portion of the plate where the classifier should refrain from taking decisions, it is sufficient to adopt specifically tuned adaptive thresholds for the *time-0* back-lit images in relation to the background reference level. We have estimated that 10% of the segments, on average, lie above the written area (with a worst-case of 67% in our dataset). Because the problem is structural and also affects human interpretation, this is beyond what can be reasonably addressed here and calls for further discussion about plate manufacturing processes in the digital microbiology imaging era.

## 5. Implementation

We provide<sup>2</sup> a Python and Jupyter Notebook [28] implementation of the proposed software, divided into three easy-to-use notebooks that follow the presented pipeline.

First, the fine alignment software (Sec.3), which automatically takes care of the complete alignment process. In input, it needs the back- and top-lit images of the plate, and the segmentation mask, and returns an aligned version

<sup>2</sup>All the notebooks, python code and the complete dataset (both train and test) is attached as additional material and will be released for research use on the website [www.microbia.org](http://www.microbia.org) upon acceptance.

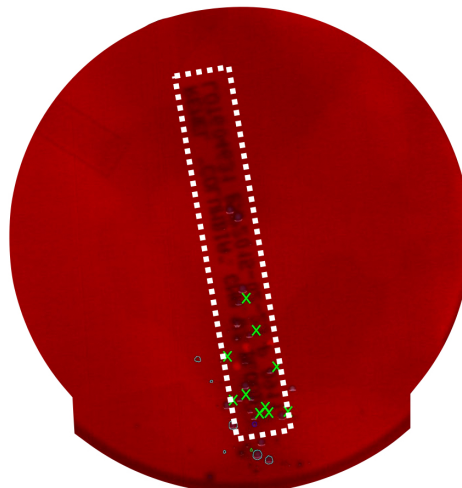


Figure 6: Output of the system with writing on the back (white dotted line) that masks bacterial growth and discarded colonies (green crosses).

and debug images to visualize how the method is working. Then, the feature extraction and dimensionality reduction, and an example of parameter selection in order to tune the SVM classification model. This script loads the dataset and metadata in order to train an SVM classifier (Sec.4). Finally, the hemolysis detection pipeline, starting from the full plate dataset provided, comprehensive of annotations, counting and segmentation masks, gives the classification results as an overlay of the plate images.

## 6. Results

We provide experimental results on the proposed methods of hemolysis identification and interpretation: the front- and back-lit image alignment (Sec.6.1), and the isolated and confluent segment classifiers (Sec.6.2). To evaluate how well the proposed method generalizes to new images, we will also validate the whole pipeline on new clinical plates (coming from the same WASPLab equipment but taken at another time), including a discussion based on an exhaustive error analysis (Sec.6.3). Finally, a computational load breakdown and related observations are given in Sec.6.4.

### 6.1. Fine plate alignment

All the tests to evaluate the robustness and accuracy of our image alignment solution have been carried out using the images of the dataset described in Sec.2. Careful manual alignments of all the image pairs were produced by an expert technician and used as the benchmark for measuring the automatic alignment error. This is defined as the modulus of the maximum displacement error (with reference to the segment centers) on the plate compared to the ground-truth position. In fact, we are not interested in breaking down the rotation and shift error, but



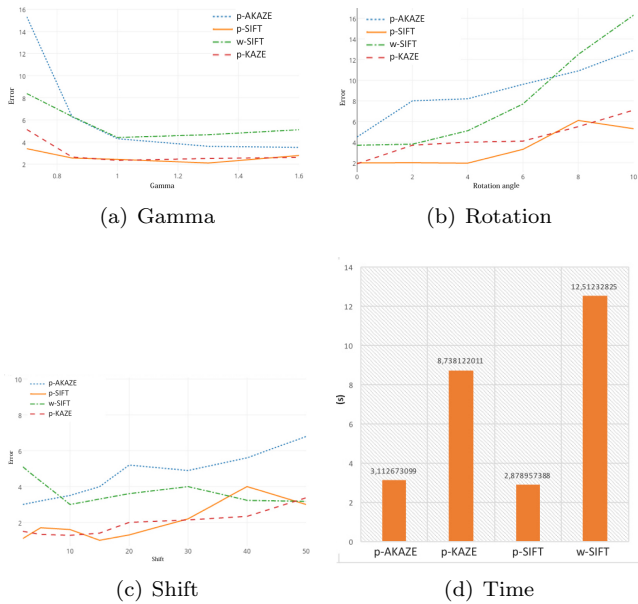


Figure 7: Comparison of the error vector length versus the tested parameters. For comparisons we consider the proposed methods composed by keypoint extraction from a subset of selected regions with the three tested algorithms: SIFT (p-SIFT), KAZE (p-KAZE) and AKAZE (p-AKAZE). To make a comparison we also plot the results obtained with SIFT-based keypoint extracted from the whole image (w-SIFT)

in evaluating the maximum displacement that they jointly produce on the plate.

Firstly, we observed that the use of FLANN, as an alternative to brute force feature matching, leads to only minor variations in both accuracy ( $<0.5\%$  average error increase) and computational time (about 3% gain). This can be justified by the fact that we had already pre-selected of good regions in which to search for keypoints. However, since even small timing savings are relevant in lab-automation, and virtually no accuracy degradation was observed, we kept the FLANN option operating in our system.

*Robustness and computational comparison.* At first, we are interested in comparing the feature extraction methods selected in Sec.3.1, i.e. SIFT [16] and KAZE [17] (and its speed-up version AKAZE [18]), especially from a robustness point of view. To this end, we deliberately introduced variations to our data by modulating the image contrast and by using a range of known rotations and translations of the reference alignments. In particular, gamma correction from 1/0.7 (darker) to 1/1.6 (lighter), artificial rotations up to 10 degrees and pixel shifts up to 50 pixels were used on pre-aligned images of our dataset. Another important comparison we are interested in is execution time.

Fig.7 shows the results of the tested variations, in particular p-SIFT, p-KAZE and p-AKAZE refer to the proposed pipeline, with region selection before the keypoint extraction (see Sec.3.4). Instead, w-SIFT refers to SIFT descriptor extracted from the whole images, and it is used

as a reference. We play with gamma correction in order to evaluate robustness against images with different illumination and contrast. The results (see Fig.7a) highlights the fact that both p-KAZE and p-SIFT respond almost in the same way, and that, with very dark images we get poor results: this leads us to apply a contrast enhancement to the images as a pre-processing phase, and this is why the pre-processing described in 4.1 is actually executed before the alignment.

If we consider the problem of rotation (Fig.7b), we can see an expected error increase with the rotation angle, which is steeper for w-SIFT solution. In fact, the controlled extraction of keypoint from interest regions significantly improves the performance, especially for large rotations lets focus the extraction on informative regions. On the other hand, p-AKAZE displays an error of about 6 pixels more than the other proposed methods almost anywhere.

We get good shift results (Fig.7c) with both SIFT and KAZE features with our pipeline. As was the case before, p-AKAZE performs worse when considering large shift values. Finally, our method, given that it is focused on keypoint extraction only in meaningful regions, is shown to outperform significantly w-SIFT in terms of computational speed (Fig.7d). p-KAZE remains rather time-consuming and leads to approximately the same degree of error performance as p-SIFT. p-AKAZE is almost as fast as p-SIFT but with a definitely worse error performance. Thus, here we choose to adopt SIFT feature extraction because it is the best compromise reached in terms of precision and speed. However, p-KAZE is another valid option if compatible with application timing constraints.

*Accuracy assessment.* Here, we validate the selected alignment method against the real misalignments on the database images. To better evaluate the improvements, in Fig.8 we plot the histograms, with their probability density functions, in relation to the displacement error before (in light blue) and after (in green) the proposed fine alignment. The residual misalignment is divided into three different classes of error, according to the length of the maximum displacement error vector (dashed line in Fig.8):

- Class I: perfect match; deviation is less than 0.13 mm (5 pixels), which is not significant in terms of the subsequent classification.
- Class II: acceptable deviation; deviation is up to 0.40 mm (15 px), which can produce erroneous classification on smaller colonies.
- Class III: poor match; deviation up to 1.33 mm (50 px), the colonies are almost out of alignment and this is likely to generate hemolysis detection errors even on bigger colonies.

The fine alignment results can be considered very satisfactory, as shown in Fig.8 (green), since 98.1% of alignments showed a residual error of under 0.13 mm (5 pixels),

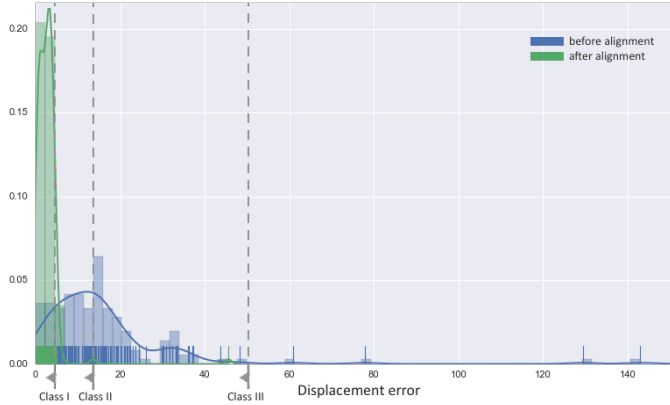


Figure 8: Displacement error before and after the proposed method. Histogram height shows a density such that the area under the curve is unitary.

which is satisfactory considering the size of the smaller colonies where to detect hemolysis. Only 0.5% of the tested plates show a Class III deviation and 1.4% Class II. Furthermore, no plates exceeded 1.33 mm deviation.

This method has proved to be surprisingly robust to images that display very marked differences due to complex unconventional effects, caused by the change in illumination. As far as we could observe, this achievement is not documented in other works.

### 6.2. Dataset segment classification

Here we present the results of the two different classifiers for the countable and confluent classes in our initial database (Sec.2). First, we consider the 2000 countable colony segments extracted from more than 500 clinical plates. 20 PCA components are enough to retain 99% of the variance of the original 3108-dimension feature vectors. During the learning phase, model selection aimed at high recall, in order to minimize failure to detect hemolysis (false negatives). This comes at the price of some additional, but less critical, false positives. The results are highly satisfactory with a precision of 98.1%, a recall of 98.5% and an accuracy of 98.1%. Looking at Fig.9, where we selected representative visual results of both correct and incorrect classifications, we can observe some significant aspects. Alpha hemolysis displays high variety, correctly classified but, if the region showing alpha hemolysis is too wide, with little contrast or very small, it is classified as Gamma. Almost the same happens in the case of Beta hemolysis. In the case of Gamma hemolysis, on the other hand, if we are looking at a gamma colony close to a Beta hemolysis colony, we could wrongly classify the segment as Beta, due to the high sensitivity of the proposed solution.

For confluent case, the dataset consists of 300 segments from the same 500 plates. 22 PCA components were retained. The precision here is 88.6% with a recall of 87.5% and an accuracy of 87.5%. This, although significantly lower than the previous case, is due to the complexity of the task in crowded areas.

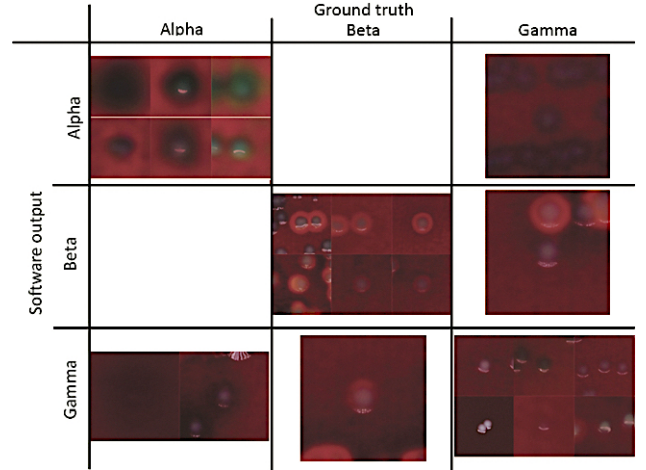


Figure 9: Some visual results of the isolated model in a confusion matrix-like grid.

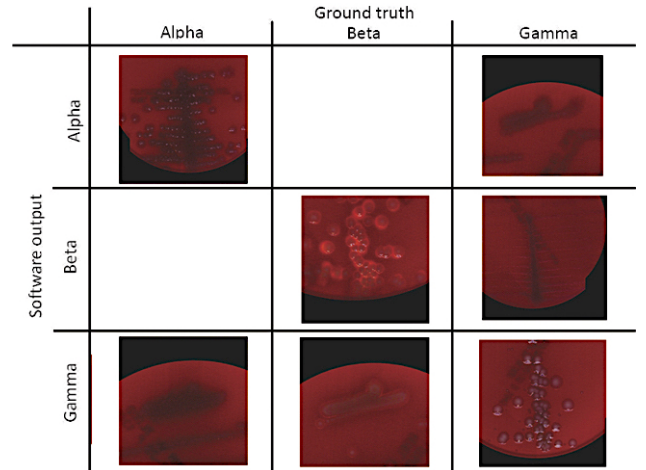


Figure 10: Some visual results of the confluent model in a confusion matrix-like grid.

Fig.10 shows some examples of good and tricky classifications. In particular, the difficulty of classifying hemolysis in confluent areas if the effect produced is too soft is a considerable problem. Other classification ambiguities can be due to color and texture mix-ups caused by too many neighboring colonies surrounding the target colony or aggregate.

In conclusion, the proposed method can assess the problem of hemolysis classification on all segment types, with more criticism for the confluent case. This allows us to produce a classification of the whole plate, which is very important from a clinical point of view, as we discuss in Sec. 7.2.

### 6.3. System validation and error analysis

Aiming at validating the system for the main target clinical scenario, we run our complete analysis and classification pipeline over another set of 350 unseen plates, for a total of 1743 segments (including both countable and confluent). As was the case for the dataset discussed in Sec.2,

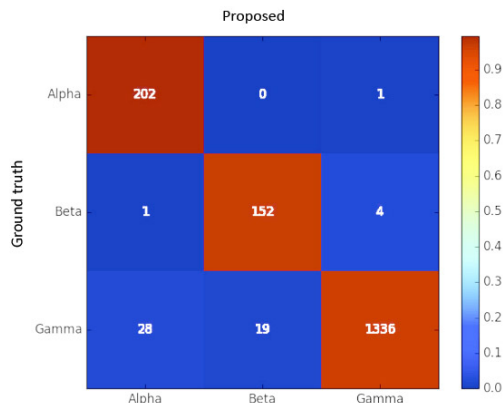


Figure 11: Confusion matrix of the hemolysis classification pipeline, using the validation dataset.

Table 1: Erroneous classification distribution.

Alignment	Confluence	Near border	Other
48%	25%	18%	10%

they were digitized with WASPlab, randomly selected and collected in anonymized form, from the microbiology lab screening routine, but at different time. Looking at the confusion matrix of Fig.11, we can see that only 5 segments with hemolysis are not recognized (0.3% false negatives). Moreover, if we look at the overall plate results, there is no hemolytic plate (plates in which there is at least one alpha or beta hemolysis colony) where our algorithm failed to identify at least one segment as hemolytic. On the other hand, there are nearly 2.75% false positives. Precision however, considering the strict recall requirements, reaches 88.3% (See Fig.12) with a recall of 98.6%.

Now we perform more detailed error analysis considering the whole image processing pipeline, where a classification of the error causes has been conducted and shown in Table 1.

Even though, as shown, the proposed co-registration method is quite robust, about half of the errors are due to residual misalignments. In fact, especially for smaller colonies, even a very small error could potentially produce a incorrect classification. There are two main effects of residual misalignment that can generate classification errors.

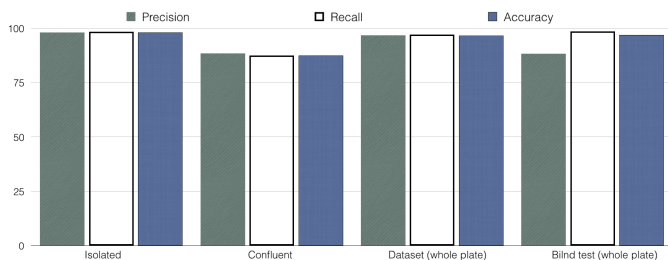


Figure 12: Result comparison on the analyzed setups: isolated, confluent and both segments on the test-set; blind test (on both kinds of segments) from a real clinical scenario.

The first, mainly responsible for false positives, is when the residual displacement is smaller than the segment dimension. In this case, slices of the interior segment are misplaced onto the supposed halo region where discrimination features are computed. For certain types of bright colored colonies, such as *Staphylococcus aureus*, this can cause  $\beta$ -hemolysis misclassifications, whilst for dull-colored ones, such as *Streptococcus agalactiae*, this could produce  $\alpha$ -hemolysis errors. If, on the other hand, the residual displacement is comparable or superior to the segment dimension, the discrimination region erroneously falls over the background area and thus it is identified as  $\gamma$ -hemolytic, possibly causing false negatives.

Colony growth could lead to project shadows onto other colonies, especially near a confluent area. In this case, shadows produce an effect similar to that caused by the  $\alpha$ -hemolysis. Another type of error leading to similar effects is due to a particular kind of diffuse covering growth, called swarming, which can occur with certain kinds of bacteria (e.g. some *E. coli* serovars). Both effects cause darkened areas possibly leading to misclassifying  $\gamma$ - as  $\alpha$ -hemolysis (false positive), while usually the result is not affected in presence of  $\beta$ -hemolysis. Together, these effects cause 30-35% of errors. There is also a portion of errors determined by glare near the border of the plate, or by the high noise present when evaluating very small colonies, accounting for another 20-25% of errors (again likely to be false positives). Finally, we have errors due to very noisy regions and micro-colonies or doubtful cases that account for the last 10%.

#### 6.4. Computational Aspects

Finally, in Fig.13, we present the timings of the proposed solution for all the described blocks. The alignment is definitely the more time consuming part, while the classification phase is very fast. In particular, while many algorithmic parts were taken from optimized C++ libraries or were easy to optimize, the region selection is implemented in Python and it is the only part working on the whole high-resolution image. For alignment, the majority (> 90%) of the computation time is taken up by the SIFT OpenCV implementation.

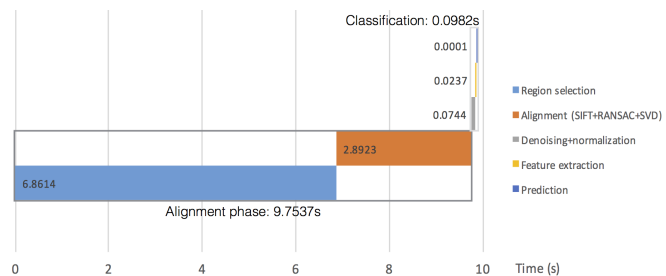


Figure 13: Computational speed of all the proposed stages. Based on a Python implementation running on an Intel i7 6700k CPU.

## 7. Discussion

The relevance of the proposed method and of the obtained results is twofold. On the one hand, this study is one among the first in the context of Digital Microbiology Imaging (DMI) and related to the emerging field of Full Laboratory Automation systems. On the other hand, it can be considered with a view to the clinical applications on which it may have most impact.

### 7.1. Hemolysis identification in a FLA context

Few recent studies have already demonstrated the opportunities of modern image understanding for CML tasks, especially in the DMI context of modern FLA systems [6, 29, 30]. In [6] the problem of bacteria colony counting is addressed using a Deep Learning approach. The complex task of bacteria species identification has been first approached in [29] with a modular *divide et impera* solution, where a chain of supervised classification stages solves different nested tasks, leading progressively to the target identification problem. Another processing and classification chain has been proposed in [30] for the classification of major urinary infections. Other works approach bacteria identification using hyperspectral datasets [31, 32, 33] or forward-scattering techniques [34, 35]. All the above works concentrate on bacterial colonies without considering their possible interactions with the growing medium, while here we address for the first time the problem of automated hemolysis analysis, which involves the interaction between growing pathogens and blood agar which is of high diagnostic significance. Moreover, differently from other works, we also exploit a combination of plate illuminations targeting both detection and classification issues. The complexity of the task and the necessary sensitivity crucially require a timely and accurate support for the clinical diagnosis. With the proposed system, despite some errors, the accuracy results are already highly satisfactory as well as being produced at a speed compatible to the analysis and FLA workflow. The nature and causes of errors have been analyzed and specific solutions should be studied for them in future works.

### 7.2. Clinical impact and throat swab screening scenario

It is important to point out that hemolysis is associated with bacterial strain (serovar) that must be carefully brought to the attention of the clinical microbiologist because of its intrinsic virulence and the fact that it is produced by pathogens of utmost importance like Streptococci Groups A, C and G. In general, the more virulence factors a strain expresses, the more severe the infection that it causes. This is why the proposed solution is eligible to be used in combination with other rapid diagnostic tools, resulting in a valuable complementary tool. For instance, mass spectrometry Maldi-TOF [36], which is nowadays the gold standard for the rapid identification of bacteria, cannot distinguish between serovars of bacterial

species or hemolytic production. Thus, with our complementary analysis, we can automatically produce a relevant diagnostic factor for the timely assessment of the virulence of certain bacterial strains and ultimately save valuable time in the identification of the correct antibiotic therapy for the patient. A significant impact of the proposed solution can be in the screening of throat swabs. The importance of investigating throat swab specimens cultured on blood agar plates is related to the fact that seeking the presence of  $\beta$ -hemolysis is the gold standard for the identification of Streptococci, which are the most common causes of bacterial pharyngitis [37]. Throat swab culture is a very widespread test for the diagnosis of bacterial throat infections [38]. These infections can be very aggressive and must be identified without errors, because they can result in a number of conditions, from strep throat to pneumonia or tonsillitis, or even meningitis. If, for instance, Group A streptococcal bacteria appear (in most of the cases *Streptococcus pyogenes*), the individual is probably affected by strep throat. Meningitis could be caused by *Neisseria meningitidis* and *Candida albicans*, a fungus responsible for mouth and tongue infections. Streptococcal bacteria are very contagious: they can be spread through airborne droplets and, an infected person can easily pass the bacteria on, through coughing, sneezing, sharing food or drinks. With screening, we refer to a medical protocol used to identify infections in a supposedly healthy population, that is before any signs or symptoms appear. This is a very important method used in sensitive environments, like hospitals, to guarantee prompt intervention and management. In order to avoid over-diagnosis, a test used in a screening program must have good sensitivity in addition to acceptable specificity, especially for a disease with low incidence.

Here we want to assess primarily the capabilities of the proposed system to be properly applied in this scenario, where main interest is focused on  $\beta$ -hemolysis detection. The substrate and the method of analysis are the same as the ones described for UTIs, so we use the same proposed and trained classification pipeline that we exploited for the core part of this work. Therefore, this different clinical setting has the dual advantage of being a relevant generalization test bed and of representing a highly relevant case study.

Since, in this context, the outcome is binary (positive, negative), with a positive result for even just one  $\beta$ -hemolytic colony, we analyzed the whole set of segments of a plate with an ensemble block that takes as input the identification performed respectively by the two different classifiers (on countable and confluent segments), divided by the number of colonies present in a segment. Therefore, we stack the prediction made by the countable classifier in 5 bins, according to the number of colonies in the segment (from 1 to 5), and in 5 bins also for the confluent classifier, again according to the estimated counting (from 6 to 10) and where the 10th bin also includes segments with more than 10 colonies. This block is composed of a

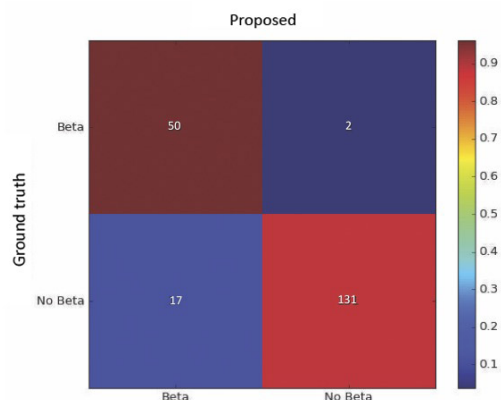


Figure 14: Confusion matrix of hemolysis identification in a screening scenario.

decision tree classifier trained on 100 plates which is tuned so as to maintain the recall as high as possible with the aim of reliably increase the precision. The performance analysis is based on an experimental dataset obtained in standardized laboratory conditions, for the task of throat swab screening, and is carried out by selecting 200 plates featuring usual throat flora and pathogens. The obtained results are already promising as shown see in Fig.14: false negative plates are only 1%, with a precision of 75% and a recall of 96%. In a screening scenario like this, the achieved precision is valuable for creating priority lists or for an automatic discarding policy that leads to a load reduction for the laboratory technician. Moreover, there is still room for improving these results, especially by adopting more powerful ensemble learning approaches and by targeting false positive reduction. Many erroneous cases are plates with soft  $\beta$ -hemolysis and a light background, presenting a confluent grown region with a lot of noise, while plates that feature well-formed hemolysis on many segments are never missed.

## 8. Conclusion

In this study, we presented technical solutions for automatic hemolysis detection and classification on blood agar plates. This had never been addressed before, nor had it been associated with freely taken images nor with respect to the huge stream of images generated by the newest FLA systems. The main achievements are the robust dual-light image alignment and the hemolysis classification modules. The former addresses a very challenging problem, due to the difference in terms of illumination and appearance between the images to be aligned, with very good results both in terms of timing and, mostly, reliability. In particular, we achieved perfect matching in more than 98% of the tested images, with the worst case featuring an error of less than 1.33 mm. The hemolysis was classified for every colony segment, achieving good recall results, though at the cost of a certain number of false positives. We reached

a precision of 88% and a recall of 98% using unseen images. Possible causes of error were investigated and, in particular, we found that alignment errors are indeed the most frequent cause of misclassification. A problem with written portions under the plate can also lead to the loss of important hemolytic colonies. We only partially tackle this aspect by detecting the printed area and avoiding making any decision on it. Being related to plate manufacturing processes, we recommend manufacturers solve the problem by producing plates with written parts located in less crucial regions (like the plate border) in order prevent obstruction of image analysis and understanding techniques. We also tested our hemolysis classification tool to a relevant clinical microbiology screening scenario, where the ability to identify  $\beta$ -hemolysis is fundamental for the diagnosis. To this end, we used the segment-based information of hemolysis to suggest  $\beta$ -hemolytic plates. The ultimate goal here is to accelerate and simplify throat swab screening and to reduce the turnaround time for the identification of the best antibiotic therapy, allowing no stone to be left unturned in the use of all the information present on the culture plate. Promising results have been obtained in this generalization scenario strictly relating to our core technology.

## Acknowledgments

This work was partially supported by the Italian Ministry of Education, Universities and Research (MIUR) under the Smart Factory Cluster initiative, Adaptive Manufacturing Project: CTN01 00163 216730. The authors would also like to express their sincere thanks to the scientific and technical staff of Copan SpA (Brescia, Italy) for their essential support in providing anonymized material for the creation of the image database. A special thanks to dr. Pietro Casella for his valuable clinical support.

## References

- [1] K. C. James Jorgensen, Michael Pfaller, Manual of Clinical Microbiology, Eleventh Edition, ASM Press, 2015.
- [2] S. Hogg, Essential microbiology, John Wiley & Sons, 2013.
- [3] H. Mansberg, et al., Automatic particle and bacterial colony counter., American Association for the Advancement of Science. Science 126 (1957) 823–7.
- [4] H. E. Kubitschek, Electronic counting and sizing of bacteria, Nature 182 (4630) (1958) 234–235.
- [5] C. Doern, M. Holfelder, Automation and design of the clinical microbiology laboratory, Manual of Clinical Microbiology, 2015.
- [6] A. Ferrari, S. Lombardi, A. Signoroni, Bacterial colony counting with convolutional neural networks in digital microbiology imaging, Pattern Recognition 61 (2017) 629–640.
- [7] P. Viola, W. M. Wells III, Alignment by maximization of mutual information, International Journal of Computer Vision 24 (2) (1997) 137–154.
- [8] J. P. Pluim, J. A. Maintz, M. A. Viergever, Mutual-information-based registration of medical images: a survey, IEEE Transactions on Medical Imaging 22 (8) (2003) 986–1004.
- [9] M. Hassaballah, A. A. Abdelmegeid, H. A. Alshazly, Image Features Detection, Description and Matching, Springer International Publishing, 2016, pp. 11–45.

- [10] Y. Li, S. Wang, Q. Tian, X. Ding, A survey of recent advances in visual feature detection, *Neurocomputing* 149, Part B (2015) 736 – 751.
- [11] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (10).
- [12] M. H. Lee, M. Cho, I. K. Park, Feature description using local neighborhoods, *Pattern Recognition Letters* 68, Part 1 (2015) 76 – 82.
- [13] D. Mukherjee, Q. M. Jonathan Wu, G. Wang, A comparative experimental study of image feature detectors and descriptors, *Machine Vision and Applications* 26 (4) (2015) 443–466.
- [14] K. Cordes, L. Grundmann, J. Ostermann, Feature evaluation with high-resolution images, in: G. Azzopardi, N. Petkov (Eds.), *Computer Analysis of Images and Patterns, CAIP 2015*, Springer International Publishing, 2015, pp. 374–386.
- [15] J. Heinly, E. Dunn, J.-M. Frahm, Comparative evaluation of binary features, in: *Computer Vision – ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part II*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 759–773.
- [16] D. G. Lowe, Distinctive image features from scale-invariant keypoints, *International journal of computer vision* 60 (2) (2004) 91–110.
- [17] P. F. Alcantarilla, A. Bartoli, A. J. Davison, Kaze features, in: *European Conference on Computer Vision, Springer, 2012*, pp. 214–227.
- [18] P. F. Alcantarilla, T. Solutions, Fast explicit diffusion for accelerated features in nonlinear scale spaces, *IEEE Trans. Pattern Anal. Mach. Intell* 34 (7) (2011) 1281–1298.
- [19] M. Agrawal, K. Konolige, M. R. Blas, Censure: Center surround extremas for realtime feature detection and matching, in: D. Forsyth, P. Torr, A. Zisserman (Eds.), *Computer Vision – ECCV 2008: 10th European Conference on Computer Vision, Springer Berlin Heidelberg, 2008*, pp. 102–115.
- [20] M. Muja, D. G. Lowe, Fast approximate nearest neighbors with automatic algorithm configuration., *VISAPP* (1) 2 (331-340) (2009) 2.
- [21] M. R. Sorkine Olga, Least-squares rigid motion using SVD, Technical notes, Department of Computer Science, ETH Zurich (2016).
- [22] M. A. Fischler, R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM* 24 (6) (1981) 381–395.
- [23] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: *Computer Vision, 1998. Sixth International Conference on, IEEE, 1998*, pp. 839–846.
- [24] A. M. Reza, Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement, *Journal of VLSI signal processing systems for signal, image and video technology* 38 (1) (2004) 35–44.
- [25] I. Jolliffe, *Principal component analysis*, Wiley Online Library, 2002.
- [26] C. Cortes, V. Vapnik, Support-vector networks, *Machine learning* 20 (3) (1995) 273–297.
- [27] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, *Machine learning* 46 (1-3) (2002) 389–422.
- [28] H. Shen, Interactive notebooks: Sharing the code, *Nature* 515 (7525) (2014) 151.
- [29] A. Ferrari, A. Signoroni, Multistage classification for bacterial colonies recognition on solid agar images, in: *2014 IEEE International Conference on Imaging Systems and Techniques (IST) Proceedings, IEEE, 2014*, pp. 101–106.
- [30] P. Andreini, S. Bonechi, M. Bianchini, A. Garzelli, A. Mecocci, Automatic image classification for the urinoculture, *Computers in Biology and Medicine* 70 (2015) 12–22.
- [31] G. Turra, N. Conti, A. Signoroni, Hyperspectral image acquisition and analysis of cultured bacteria for the discrimination of urinary tract infections, in: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015*, pp. 759–762.
- [32] S. Arrigoni, G. Turra, A. Signoroni, Hyperspectral image analysis for rapid and accurate discrimination of bacterial infections: A benchmark study, *Computers in Biology and Medicine* 88 (Supplement C) (2017) 60 – 71.
- [33] D. F. Leroux, R. Midahuen, G. Perrin, J. Pescatore, P. Imbaud, Hyperspectral imaging applied to microbial categorization in an automated microbiology workflow, in: *Proc. SPIE (vol.9537)*, 2015, pp. 953726(1–9).
- [34] U. Minoni, A. Signoroni, G. Nassini, On the application of optical forward-scattering to bacterial identification in an automated clinical analysis perspective, *Biosensors and Bioelectronics* 68 (2015) 536–543.
- [35] W. M. Ahmed, B. Bayraktar, A. K. Bhunia, E. D. Hirleman, J. P. Robinson, B. Rajwa, Classification of bacterial contamination using image processing and distributed computing, *IEEE Journal of Biomedical and Health Informatics* 17 (1) (2013) 232–239.
- [36] E. Carbonnelle, C. Mesquita, E. Bille, N. Day, B. Dauphin, J.-L. Beretti, A. Ferroni, L. Gutmann, X. Nassif, Maldi-tof mass spectrometry tools for bacterial identification in clinical microbiology laboratory, *Clinical biochemistry* 44 (1) (2011) 104–109.
- [37] UK standards for microbiology investigations, investigation of throat related specimen, *Standards Unit* (9).
- [38] D. Mertz, R. Frei, B. Jaussi, A. Tietz, C. Stebler, U. Flückiger, A. F. Widmer, Throat swabs are necessary to reliably detect carriers of staphylococcus aureus, *Clinical Infectious Diseases* 45 (4) (2007) 475–477.