Computers in Biology and Medicine 70 (2016) 12–22

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Automatic image classification for the urinoculture screening

Paolo Andreini, Simone Bonechi*, Monica Bianchini, Andrea Garzelli, Alessandro Mecocci

University of Siena[1] *- Department of Information Engineering and Mathematics, Via Roma 56, I-53100 Siena, Italy*

ARTICLE INFO

Article history:
Received 1 October 2015
Accepted 30 December 2015

Keywords:
Color image processing
Clustering techniques
Artificial neural networks
Support vector machines
Urinoculture screening

ABSTRACT

Urinary tract infections (UTIs) are considered to be the most common bacterial infection and, actually, it is estimated that about 150 million UTIs occur world wide yearly, giving rise to roughly $6 billion in healthcare expenditures and resulting in 100,000 hospitalizations. Nevertheless, it is difficult to carefully assess the incidence of UTIs, since an accurate diagnosis depends both on the presence of symptoms and on a positive urinoculture, whereas in most outpatient settings this diagnosis is made without an ad hoc analysis protocol. On the other hand, in the traditional urinoculture test, a sample of midstream urine is put onto a Petri dish, where a growth medium favors the proliferation of germ colonies. Then, the infection severity is evaluated by a visual inspection of a human expert, an error prone and lengthy process. In this paper, we propose a fully automated system for the urinoculture screening that can provide quick and easily traceable results for UTIs. Based on advanced image processing and machine learning tools, the infection type recognition, together with the estimation of the bacterial load, can be automatically carried out, yielding accurate diagnoses. The proposed AID (Automatic Infection Detector) system provides support during the whole analysis process: first, digital color images of Petri dishes are automatically captured, then specific preprocessing and spatial clustering algorithms are applied to isolate the colonies from the culture ground and, finally, an accurate classification of the infections and their severity evaluation are performed. The AID system speeds up the analysis, contributes to the standardization of the process, allows result repeatability, and reduces the costs. Moreover, the continuous transition between sterile and external environments (typical of the standard analysis procedure) is completely avoided.

1. Introduction

Urinary tract infections (UTIs) are considered to be the most diffuse bacterial diseases, found in common medical practice, and can cause serious health problems. They are mainly due to the presence of Gram-negative microorganisms, with a high prevalence of *Escherichia coli* (E. coli, 70%), usually found in the digestive system, even if complicated infections caused by Gram-positive or multidrug-resistant germs can also occur, especially in hospitalized or elderly patients, on which the common antimicrobial agents are inevitably ineffective, leading to therapeutic failures.

The occurrence of UTIs varies in dependence of age and gender, as well as based on the socioeconomic background. Moreover, specific subpopulations at increased risk of UTIs include infants, pregnant women, the elderly, persons with urological abnormalities, patients with spinal cord injuries and/or catheters, with diabetes, multiple sclerosis, and AIDS. The urinoculture is a screening test[1] in all these cases. In the standard protocol, the urine sample is seeded on a Petri dish equipped with a cell culture substrate, used to artificially recreate the environment required for the bacterial growth, and incubated at 37 °C overnight. Each dish must then be examined by a human expert, adding some more time to the medical report emission. This common situation significantly departs from the requirement to have results in quick time, to set a targeted therapy, avoiding the use of broad-spectrum antibiotics and improving the patient management.[2] Moreover, the traditional analysis procedure suffers from possible errors arising in the visual, qualitative, inspection of the dishes — due to the skills and the expertise of each operator, whereas difficulties also arise in the traceability of samples and results [1].

* Corresponding author. Tel.: +39 0577 233609.
E-mail address: simo_bone@alice.it (S. Bonechi).
[1] http://www.diism.unisi.it

[1] The term *screening* is used here in its common meaning of a routine test performed on a large population, to identify those who are likely to develop a specified disease. Instead, in in vitro diagnostics it stands for preventive analyses aimed at establishing if a sample is positive or not.

[2] Rapid reporting is crucial especially when pediatric patients are involved since in this case, the infection symptoms are not always specific, while it is urgent to decide if an antibiotic therapy is necessary or not and when to start it.

http://dx.doi.org/10.1016/j.compbiomed.2015.12.025
0010-4825/

In recent years, significant improvements in biology and medicine applications and decision support systems [2] have been obtained by using hybrid methods, based on a combination of advanced image processing techniques [3,4], artificial intelligence tools [5–7], machine learning [8], fuzzy logic [9], genetic algorithms [10], and Bayesian modeling [11]. In particular, the development of automated tools for results assessment (screening systems) has attracted increasing research interest during the last decade, due to their higher repeatability, accuracy, reduced staff time (that are the main limiting factors of manual screening), and lower costs [12]. Automated urinalysis devices improve the capacity of the laboratory to screen more samples, producing results in less time than by manual screening. Moreover, the redeployment and lower grading of staff with the increased turnover and speed of urine screening, gave economic advantages of automated screening over manual screening [13].

Even if some interesting research has been carried out in recent years for the urinoculture screening, obtaining an overview of the state-of-the-art in image processing/AI solutions to the automatic analysis of Petri dishes is difficult, since results are published in various domains — from food and beverage safety to environmental control and specific clinical analyses [14–19], based on different data sets, and they are also often related to subtle variations of the core problem (e.g., in [20], where the colony classification problem is addressed with promising results, but with respect to a very small number of images and based only on the determination of isolated colonies).

In this paper, we propose the AID (Automatic Infection Detector) software, an automated tool that provides a decision support system for the urinoculture analysis. Dish images are acquired from a color camera and, thereafter, through a suitable preprocessing phase, involving spatial clustering, colonies are isolated from the culture substrate, even in the presence of ground disuniformities. Detected bacterial infections are then classified based on both artificial neural networks (ANNs) and support vector machines (SVMs). Finally, besides the infection identification and classification, the AID system also performs the bacterial count, giving an estimate of the number of microorganisms per milliliter of urine.

AID actually allows a substantial speedup of the whole analysis procedure, also avoiding the continuous transition between sterile and external environments which is typical in the standard protocol. The final outcomes are directly stored, along with the related analysis records (the image, the infection type and the colony count). Experimental data have been provided by DIESSE Ricerche Srl, Siena, a world class manufacturer of innovative biomedical devices and reagents for infectious disease testing. Preliminary experiments show very promising results, in terms of both classification accuracy of the infections and estimation of the bacterial count.

The paper is organized as follows. Section 2 introduces the problem and describes the preprocessing procedure, from the image acquisition phase to the background removal. Section 3 presents the classification methods and shows experimental results, whereas Section 4 defines the procedure used to estimate the infection severity and Section 5 introduces a specific approach to recognize the Candida infection. Finally, conclusions are drawn in Section 6.

2. Automatic image analysis of Petri dishes

The system model for the AID application considered in this paper is shown in Fig. 1. All the processing steps, from image acquisition to classification and bacterial load count, will be detailed throughout the following sections.

First of all, in order to correctly recognize the infection type and to precisely estimate the bacterial load, it is fundamental to acquire a good quality image of the Petri dish. To avoid imperfections due to manual plate handling, images are captured by an automatic camera setup (Fig. 2). After the acquisition, a suitable preprocessing step is applied to locate the region of interest (the Petri dish), and to grant that it is in an appropriate position inside the field of view. At this point the image is saved along with auxiliary information.

The automatic acquisition is performed as follows: a simple and fast algorithm, based on change detection [21] and morphological filtering [22], is applied and the image is acquired only when the dish is correctly positioned, the scene is well illuminated, and no movements are observed. Before saving the image, the Petri dish is isolated from the rest of the scene, using a Random Hough circle transform [23], to detect the circular Region of Interest (RoI) (see Fig. 3).

The acquisition setup has been used in a real application scenario at DIESSE Ricerche premises, to collect a dataset of 253 images, subsequently divided into a training, a validation, and a test set, containing 154, 64, and 35 images, respectively (see Table 1). As a requirement, eight different classes of infection were detected, namely: E. coli, KES (Klebsiella, Enterobacter, Serratia), Enterococcus faecalis, Streptococcus agalactiae, Pseudomonas, Proteus, Staphylococcus aureus, and Candida.

2.1. Selection of the chromatic space

Since a chromogenic medium (UriSelect 4) is used as ground seed, the color of the pixels is the most important feature for classifying different colonies. Therefore, the background color distribution has been analyzed in four different color spaces (i.e., RGB, HSV, CIE-Lab, and YCrCb). A supervised training process has been adopted, during which a human expert selected about 40 different regions belonging to the background and to the foreground. The chromatic components of the pixels belonging to such regions are accumulated to represent the typical background and foreground chromatic values. Then, the Dunns Index has been used to give a quantitative ranking (based on the Centroid Linkage distance and the Centroid Diameter dispersion [24]):

$$DI(X) = \frac{\min_{1 \le i \le j \le k} d(C_i, C_j)}{\max_{1 \le s \le k} \{\Delta(C_s)\}}$$

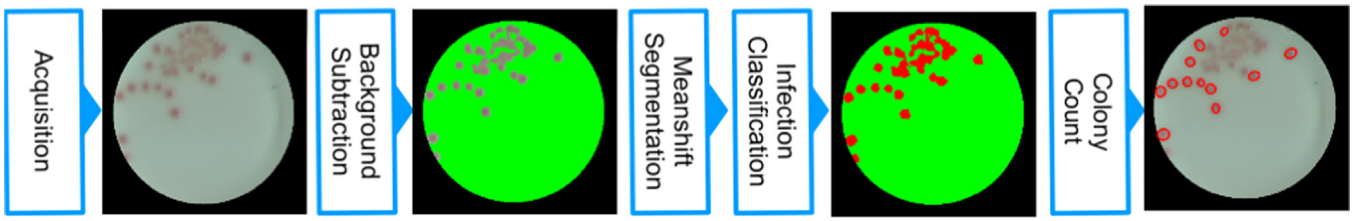

Fig. 1. The system pipeline.