Expert Systems With Applications 89 (2017) 41–51

Contents lists available at ScienceDirect

Expert Systems With Applications

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

myStone: A system for automatic kidney stone classification

Joan Serrat[a,*], Felipe Lumbreras[a], Francisco Blanco[b], Manuel Valiente[b], Montserrat López-Mesas[b]

[a] Computer Vision Center and Department of Computer Science, Universitat Autònoma de Barcelona, Edifici O, 08193 Bellaterra, Spain
[b] Centre Grup de Tècniques de Separació en Química, Unitat de Química Analítica, Departament de Química, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain

ARTICLE INFO

Article history:
Received 5 April 2017
Revised 14 July 2017
Accepted 15 July 2017
Available online 17 July 2017

Keywords:
Kidney stone
Optical device
Computer vision
Image classification

ABSTRACT

Kidney stone formation is a common disease and the incidence rate is constantly increasing worldwide. It has been shown that the classification of kidney stones can lead to an important reduction of the recurrence rate. The classification of kidney stones by human experts on the basis of certain visual color and texture features is one of the most employed techniques. However, the knowledge of how to analyze kidney stones is not widespread, and the experts learn only after being trained on a large number of samples of the different classes. In this paper we describe a new device specifically designed for capturing images of expelled kidney stones, and a method to learn and apply the experts knowledge with regard to their classification. We show that with off the shelf components, a carefully selected set of features and a state of the art classifier it is possible to automate this difficult task to a good degree. We report results on a collection of 454 kidney stones, achieving an overall accuracy of 63% for a set of eight classes covering almost all of the kidney stones taxonomy. Moreover, for more than 80% of samples the real class is the first or the second most probable class according to the system, being then the patient recommendations for the two top classes similar. This is the first attempt towards the automatic visual classification of kidney stones, and based on the current results we foresee better accuracies with the increase of the dataset size.

1. Introduction

Medicine and healthcare are among the most important fields where expert systems have found application (Wagner, 2017). For instance, computer aided detection and diagnosis systems can enhance the diagnostic capabilities of physicians or reduce the required time. Many of them are based on diverse imaging modalities like brain CT and MRIs, mammographies, chest X-ray and a long etc. Another type of expert systems are decision support systems, whose purpose is not as much to produce a diagnostic but to analyze data (e.g. images) and present some kind of result so that decisions can be made more easily. At the core of such systems one can often find classifiers, which are typically trained on data samples to generate a discrete prediction (a class label) plus a confidence score or a probability for each class, when presented a new sample. This is the case of the work described in this paper, which deals with the problem of kidney stone classification.

Urinary lithiasis —the formation of kidney stones— shows a steady incidence increase in developed countries. Around 10% of population in developed countries suffer a stone episode at least once in his/her life (Romero, Akpinar, & Assimos, 2010; Scales, Smith, Hanley, & Saigal, 2012). Emphasis should be made on the high prevalence affecting this disease. Some European follow-up studies have quantified the stone recurrence rate (repeated stone episodes for the same patient) at 40% in 5 years (Andreassen, Poulsen, Olsen, Aabeck, & Osther, 2007; Hesse, Brndle, Wilbert, Khrmann, & Alken, 2003). These dramatic numbers reflect not only a disturbing and painful disease but also a considerable burden for the national healthcare systems (Strohmaier, 2012).

Once the stone episode has passed, it is widely agreed that an adequate study of the causes of stone formation is required in order to decrease the high recurrence of this disease (Grases, Costa-Bauzá, Ramis, Montesinos, & Conte, 2002; Kok, 2012; Siener & Hesse, 2012). In fact, it has been pointed out that the correct treatment of stone patients can drop further stone formation as much as 46% (Nolde, Hesse, Scharrel, & Vahlensieck, 1993; Strohmaier, 2011). The urinary stone represents a solid description of the metabolic disturbances suffered by the patient, so it should be regarded as the starting point of an individualized treatment.

* Corresponding author.
E-mail addresses: joans@cvc.uab.es (J. Serrat), felipe@cvc.uab.es (F. Lumbreras), frnblanco@gmail.com (F. Blanco), Manuel.Valiente@uab.cat (M. Valiente), Montserrat.Lopez.Mesas@uab.cat (M. López-Mesas).

http://dx.doi.org/10.1016/j.eswa.2017.07.024
0957-4174/

Aware of the importance of stone characterization, clinical scientists have used a variety of approaches for the description of urinary calculi. Already in the second half of the 19th century, the first classification of stones was devised, based on features such as color, hardness and shape. A number of techniques have been used from that starting point. IR spectroscopy, X-Ray diffraction and stereoscopic microscopy stand as the most extended analysis methodologies (Schubert, 2006). The rest of techniques (electron microscopy, Raman spectroscopy, hyperspectral imaging among them) have been mainly used for research purposes.

IR spectroscopy is easy to use and yields quantitative results, but the sample needs to be grinded and the distribution of components in the stone, related to its etiology, is inevitably lost. X-Ray diffraction presents the same drawback, as well as less available instrumentation. Hyperspectral imaging has proven to be a high-performance alternative. Like in IR spectroscopy, hyperspectral images are suitable to detect the presence of chemical components based on the spectral signature of the sample but, at the same time, to know their spatial distribution on the imaged surface of the stone. A few works have attempted to perform a pixel-wise classification according to some kidney stone classes with infrared and near infrared spectral imaging, like (Blanco et al., 2012; Piqueras, Duponchel, Tauler, & DeJuan, 2011) and notably (Blanco et al., 2015). However, they do not perform a large scale study with hundreds of stones, like us. Additionally, the problem of this approach is that the equipment needed for its implementation is costly and not always available even in clinical laboratories.

In stereoscopic microscopy (Cloutier, Villa, Traxer, & Daudon, 2015; Daudon, Bader, & Jungers, 1993; Grases et al., 2002) the external surface and a section of fragments are observed at a magnification of 10X–40X. The color, 3D shape (e.g. lobulations, surface roughness and smoothness), size, shape and orientation of crystals, deposition layers, the existence of a core etc. provide cues to the expert with regard the stone class. Thus, stereoscopic microscopy relies on the expertise of the technician who performs the analysis. While very precise, it is time consuming, cannot be offered at a competitive price and generally the analysis is carried out in laboratories external to the hospital, so the waiting time for the results may become an issue.

This paper presents a system composed of a device and automatic classification method, specifically conceived for the fast and on-site analysis of urinary stones. The system is based on the principles of stereoscopic microscopy, so the sample is classified attending to the amount and distribution of mineralogical components, as they appear on images captured by a standard camera. To the best of our knowledge, this is the first attempt to automate the visual classification of kidney stones. Hence, we consider as the main contributions of this work :

1. The construction of a fully functional device, including hardware, user interface and classification software, for the visual recognition of renal calculi.
2. The first extensive dataset for this kind of samples, available upon request at Lumbreras, Serrat, and Rotger (2017). It consists of 14,500 images of 908 stone fragments from 454 producers, recording separately both the external and internal side of each fragment, under visible and near infrared light sources.
3. The identification of discriminant color and texture features to train a state-of-the art classifier that attains a baseline accuracy of 63% for the top class and 83% for the top-2 classes, in spite of the large intraclass variability combined in some cases with considerable inter-class similarity.
4. We show that a boost in performance is possible with the use of the urinary pH level, obtaining 70% and 89% top-1 and top-2 accuracy, respectively. Moreover, confusions align with classes judged more similar by the annotator expert.

Table 1
Percentage of kidney stone classes and subclasses in our dataset and naturally appearing according to Grases et al. (2002). Even though figures vary depending on the world location of the study, COM, COD and UA calculi stand for the vast majority and their proportions are similar to those in the dataset.

Main	Class		Dataset		Natural
components	Scheme 1	Scheme 2	Samples	Percent	Frequency
		2	31	6.8	
COM	2	2b	29	6.4	29.3
		2codt	30	6.6	
		3b	27	5.9	
COD	3	3t	53	11.7	33.8
		3bt	59	13.0	
CO + HAP	4	4	63	13.9	11.2
HAP	5	5	19	4.2	7.1
STR	6	6	38	8.4	4.1
BRU	7	7	14	3.1	0.6
UA	8	8	61	13.4	8.2
UA + CO	9	9	30	6.6	2.6
CYS and others					3.1
Total			454		

The organization of this paper is as follows. Section 2 presents a widespread taxonomy of kidney stones. They are characterized by the presence of certain chemical components that show up as color and textural features. Section 3 describes the device we have built to acquire images of kidney stone fragments. We follow a certain procedure to record a sample, which is a collection of images of a pair of fragments from one same patient. Accordingly, we have built a large dataset with samples of all the classes but the one less frequently found (Section 4). On the images of the dataset we have computed a set of visual features related to color and texture which then are fed into a random forest classifier (Section 5). In Section 6 we combine the class probabilities given by this classifier with those obtained from the urinary pH level, a non-visual feature that helps to distinguish some classes. Section 7 reports the results for four variants of the classifier, depending on the scheme of classes used (8 main classes or 12 fine-grained classes) and the use or not of the urinary pH level as an additional feature. Finally, Section 8 draws the main conclusions and avenues of future work.

2. Kidney stone taxonomy

There is a well known taxonomy proposed by Daudon et al. (1993). It is a hierarchical classification whose first level considers the main or two main chemical species in the stone (calcium oxalate, uric acid etc.). A second level is provided to account for different etiologies or pathologies that such broad classes do not discern well. Thus, urologists consider also what the minor components are and their spatial distribution. The later means, for instance, whether they are on the surface or the inside of the stone, forming a core, layers, radial structures, lobules or uniformly spread. Table 1 relates the two taxonomies.

This second level has a total of 21 classes, making it difficult to adapt to the clinical practice. Moreover, since our goal is to train a classifier, we will need as much samples as possible per class. Unfortunately, many second level classes (and even some of the first level) have a low natural frequency of occurrence, so they may not be well represented in a dataset. Grases et al. (2002) simplified this classification scheme and we draw from them our first scheme of classes which are:

- calcium oxalate monohydrate (COM)
- calcium oxalate dihydrate (COD)
- mixed calcium oxalate and hydroxiapatite (CO - HAP)
- hydroxiapatite (HAP)
- struvite (STR)
- brushite (BRU)