Expert Systems With Applications 139 (2020) 112843

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Intelligent image-based colourimetric tests using machine learning framework for lateral flow assays

Marzia Hoque Tania[a,*], Khin T. Lwin[b], Antesar M. Shabut[c], Mohammad Najlah[d], Jeannette Chin[e], M.A. Hossain[b]

[a] Medical Technology Research Centre, Faculty of Science and Engineering & Faculty of Health, Education, Medicine and Social Care, Anglia Ruskin University, Chelmsford, UK
[b] School of Computing and Digital Technologies, Teesside University, Middlesbrough, UK
[c] School of Arts and Communication, Leeds Trinity University, Leeds, UK
[d] School of Allied Health, Anglia Ruskin University, Chelmsford, UK
[e] School of Computing Sciences, University of East Anglia, Norwich, UK

ARTICLE INFO

Article history:
Received 13 November 2018
Revised 5 July 2019
Accepted 25 July 2019
Available online 26 July 2019

Keywords:
Image processing
Histogram thresholding
Feature selection
Machine learning
Deep learning
Colourimetric tests

ABSTRACT

This paper aims to deliberately examine the scope of an intelligent colourimetric test that fulfils ASSURED criteria (Affordable, Sensitive, Specific, User-friendly, Rapid and robust, Equipment-free, and Deliverable) and demonstrate the claim as well. This paper presents an investigation into an intelligent image-based system to perform automatic paper-based colourimetric tests in real-time to provide a proof-of-concept for a dry-chemical based or microfluidic, stable and semi-quantitative assay using a larger dataset with diverse conditions. The universal pH indicator papers were utilised as a case study. Unlike the works done in the literature, this work performs multiclass colourimetric tests using histogram-based image processing and machine learning algorithm without any user intervention. The proposed image processing framework is based on colour channel separation, global thresholding, morphological operation and object detection. We have also deployed aserver-based convolutional neural network framework for image classification using inductive transfer learning on a mobile platform. The results obtained by both traditional machine learning and pre-trained model-based deep learning were critically analysed with the set evaluation criteria (ASSURED criteria). The features were optimised using univariate analysis and exploratory data analysis to improve the performance. The image processing algorithm showed >98% accuracy while the classification accuracy by Least Squares Support Vector Machine (LS-SVM) was 100%. On the other hand, the deep learning technique provided >86% accuracy, which could be further improved with a large amount of data. The k-fold cross-validated LS-SVM based final system, examined on different datasets, confirmed the robustness and reliability of the presented approach, which was further validated using statistical analysis. The understaffed and resource-limited healthcare system can benefit from such an easy-to-use technology to support remote aid workers, assist in elderly care and promote personalised healthcare by eliminating the subjectivity of interpretation.

1. Introduction

There is less than one physician per thousand population for more than 44% of the World Health Organisation (WHO) member states (World Health Organization, 2017). Even in a developed country such as *the* UK, there are only 2.806 doctors for every 1000 people. The longevity of human life has given rise to increasing the understanding of age-related disabilities and diseases, which can create significant burdens on already over-burdened healthcare systems. To support the elderly population, which is expected to increase to 1.91 billion in 2050 (P. D. of the Department of Economic & S. A. of the United Nations Secretariat, 2012), and limit the spread of pandemics, an intelligent, clearer and easier system with the least error-prone diagnosis results is required for both patients and clinicians. In the absence of expert clinical staff, there is a requirement for systems that are easily operable. Such a system could be used by aid workers in remote places to support primary healthcare, in time of epidemic and environmental monitoring for many purposes such as identifying safe drinking water. The easy-to-use system could support the growing need of the elderly population as well. Therefore, the early diagnosis facility, the

* Corresponding author.
E-mail address: marzia.hoque-tania1@anglia.ac.uk (M. Hoque Tania).

https://doi.org/10.1016/j.eswa.2019.112843

disproportional ratio of health professionals (doctor, expert, staff, carer) to patients and the advancement of technology is influencing the field of healthcare prompting the sector of mobile phone-based microscopy, assays, and sensing platforms for Point-Of-Care (POC) diagnostics (Contreras-naranjo, Wei & Ozcan, 2016; Rajan & Glorikian, 2009).

The colourimetric tests are the prominent technologies used in the POC systems. The colourimetric tests for diagnosis purposes are being utilised for decades. The Lateral Flow Assay (LFA), a type of colourimetric test scheme, is more commonly either a qualitative or a semi-quantitative assay. The LFA are mainly popular for POC platforms since they are easy-to-use, fast and low-cost. However, they often suffer from limited specificity and sensitivity due to the limitation of materials including biochemical components (Koczula & Gallotta, 2016).

The naked-eye and colour chart based colourimetric tests including LFA expect the user to possess a perfect colour vision, whereas colour blindness is a common genetic deficiency. Globally, one in twelve men is colour blinded (NHS Choices, 2016). There are more than 2.7 million colour blind people in the United Kingdom (Colour Blind Awareness, 2018). Moreover, perception of colour can vary from person to person and reading from colour charts can be complicated for non-clinicians. As a POC system, the integration of computational system to LFA can enhance the overall diagnosis experience such as the research conducted by Ozkan (2017). In this work, we explored the computational solutions to provide an automatic colourimetric decision that fulfils our evaluation criteria.

This paper aims to provide a proof-of-concept for a dry-chemical based or microfluidic, stable and semi-quantitative assay using a larger dataset with diverse conditions. At first, this paper provides a context of the evaluation criteria and presents the rationale for the evaluation criteria (Section 2), followed by defining the assay types from the perspective of computer vision (Section 3). The current point of view in the field of POC systems is from a number of disciplines; dominated by biochemistry, nanotechnology and optoelectronics. Findings concerning such systems have presented the prospect of isolated individual colourimetric components but often lack the rigorous detailing of how such a system can be and should be designed. Therefore, there is a need for an extensive study to deal with the inadequacy to perceive colourimetric tests from the frame of computer vision. The challenges include the quest for searching a suitable image processing technique for robust operation of colourimetric tests. There is a requirement of knowledge exploration for such techniques to develop a better understanding of colourimetric test data, which can facilitate better management of the computational complexity of such data. Hence, experiments are designed accordingly; the detail of sample preparation is provided in Section 4, followed by the proposed image processing framework to separate the region of interest (ROI) from images of colourimetric tests.

Analysing the extracted features from ROIs could help to create domain-specific knowledge. Identifying the key features and how the features are being analysed can play a crucial role in the core model of a colourimetric Decision Support System (DSS). Thus, feature optimisation and feature analysis techniques would be a promising contribution, which is investigated in Section 6, and extended to exploring classification and regression algorithms and further expanded to consider the potential of utilising more advanced machine learning techniques in Section 7.

The purpose of this paper is to deliberately examine the scope of an intelligent colourimetric test that can reason about and interpret the colourimetric data, fulfils our defined evaluation criteria and demonstrate the claim as well. The paper also asserts the contribution regarding the pseudo-control colour.

2. Evaluation criteria

WHO prefers the diagnostic system to be inexpensive, disposable and easy-to-use (Khademhosseini, 2011; Wang, Xu & Demirci, 2010). Such a diagnostic system should follow the criteria called ASSURED (Affordable, Sensitive, Specific, User-friendly, Rapid and robust, Equipment-free, and Deliverable) (Kettler, White & Hawkes, 2004). This paper studies, how a computational system can act as an expert to perform colourimetric test complying with the ASSURED criteria.

2.1. Affordable technology

The mobile phones have a high penetration rate (GSMA Intelligence, 2017), making it widely accessible and affordable technology to the resource-limited setting. By 2019, the number of mobile phone users is expected to reach 5.07 billion (Statista, 2015). From the computational context, the use of a mobile phone can act as an affordable-ASSURED technology. The mobile phone can effectively eliminate the operating cost by minimising the requirement of plate readers and analysers.

In general, paper-based assays such as Sicard et al. (2015) are more affordable and suitable for less trained personnel. A mobile-enabled paper-based assay can enhance the processing of result (Lopez-Ruiz et al., 2014; Roda et al., 2016), ease the effort to interpret the result and make the result conveniently communicable (Sicard et al., 2015). The objective of such a system is not aimed to replace the biochemical systems but instead to assist (Kim, Awofeso, Choi, Jung & Bae, 2017), simplify (Lopez-Ruiz et al., 2014) or accelerate (Tania, Lwin, Abuhassan & Bakhori, 2017) the process. For example, when it is difficult to provide visually distinguishable colours, such systems can aid to provide a better decision (Abuhassan et al., 2017; Shabut et al., 2018). Therefore, this work aims to develop a system which is computationally efficient to be deployed on the mobile platform, making it an affordable system.

2.2. Specific and sensitive performance

The next ASSURED criteria are specificity and sensitivity, which require the system to have low false negative and false positives. Although, it is a common practice for the computation systems to present the result in terms of accuracy, evaluating the performance of the system only with accuracy can be misleading.

2.3. User-friendly system

The ASSURED criteria put emphasis on the minimum requirement of training from the users. The ratio between health professionals and patients are imbalanced worldwide. The global understaffed health systems can benefit from technologies that provide ease of use. These easy to use systems can support the associated need of growing elderly population, provide more autonomy to users for personalised healthcare at home settings, and more importantly in remote locations where there is a scarcity of trained medical personnel.

Exploring the existing literature, this research suggests, the user-friendly system should not only require less medical training but also should not demand high technical skills from the user. Therefore, the system should require minimum user interactions or interventions with the system in order to provide a decision. In the reported articles on colourimetric tests, there are systems that require users to aid with the data pre-processing techniques e.g. cropping and seed points by the user (Mutlu et al., 2017; Rahmat et al., 2018; Solmaz et al., 2018). No detail description was provided in the article (Mutlu et al., 2017) regarding the cropping