An Intelligent Mobile-Enabled Expert System for Tuberculosis Disease Diagnosis in Real Time

Antesar M. Shabut[1], Marzia Hoque Tania[1][*], Khin T. Lwin[1], Benjamin A. Evans[2], Nor Azah Yusof[3, 4], Kamal J. Abu-Hassan[5], M. A. Hossain[1]

[1] Anglia Ruskin IT Research Institute (ARITI), Anglia Ruskin University, Chelmsford, UK

[2] Norwich Medical School, University of East Anglia, Norwich, UK

[3] Institute of Advance Technology, Universiti Putra Malaysia, Serdang, Malaysia

[4] Department of Chemistry, Faculty of Science, Universiti Putra Malaysia, Serdang, Malaysia

[5] Department of Physics, University of Bath, Bath, UK

Abstract

This paper presents an investigation into the development of an intelligent mobile-enabled expert system to perform an automatic detection of tuberculosis (TB) disease in real-time. One third of the global population are infected with the TB bacterium, and the prevailing diagnosis methods are either resource-intensive or time consuming. Thus, a reliable and easy–to-use diagnosis system has become essential to make the world TB free by 2030, as envisioned by the World Health Organisation. In this work, the challenges in implementing an efficient image processing platform is presented to extract the images from plasmonic ELISAs for TB antigen-specific antibodies and analyse their features. The supervised machine learning techniques are utilised to attain binary classification from eighteen lower-order colour moments. The proposed system is trained off-line, followed by testing and validation using a separate set of images in real-time. Using an ensemble classifier, Random Forest, we demonstrated 98.4% accuracy in TB antigen-specific antibody detection on the mobile platform. Unlike the existing systems, the proposed intelligent system with real time processing capabilities and data portability can provide the prediction without any opto-mechanical attachment, which will undergo a clinical test in the next phase.

Keywords: Image processing; machine learning; decision support system; colourimetric tests

1. Introduction

Tuberculosis (TB) is a communicable disease, infecting one third of the world's population. In 2015, 1.8 million TB-related deaths were reported (Centers for Disease Control and Prevention, 2017). On the other hand, every year about 244 million migrants cross international borders (Department of Economic and Social Affairs, 2016). The carriage of TB in a mobile population is a global challenge, which is a particular concern for the border agencies (Posey, Marano, & Cetron, 2017). However, TB is curable with appropriate early diagnosis. The most common diagnosis procedure for TB is a skin test (Mantoux test) or a blood test (Centers for Disease Control and Prevention.; NHS). Despite many commercial test

---

[*] Corresponding Author: Marzia Hoque Tania
E-mail: marzia.hoque@pgr.anglia.ac.uk

schemes, there is still a need for an easy-to-use, effective and feasible point-of-care (POC) TB diagnosis tool, particularly for the remote community where there are very limited or no diagnostic facilities. Such a tool should possess the following features: low cost mobile solution, anytime anywhere access, low energy consumption, ease of use, fast and automatic identification of TB.

The World Health Organization (WHO) prefers diagnostic tools which are inexpensive, disposable and easy-to-use (Khademhosseini, 2011; S. Wang, Xu, & Demirci, 2010). A mobile-enabled expert system can address all these features. Due to the high penetration rate of mobile phones (GSMA Intelligence), such system can reach a wider population, especially those who have limited access to advanced laboratory facilities. Incorporation of the mobile phone can not only facilitate an easy and automatic colour detection but also can enable diagnostic disease decision using machine learning techniques.

In order to establish a widespread application, the mobile-enabled expert systems should possess minimum hardware requirements. To eliminate the necessity of the opto-mechanical attachment, one requires advanced image processing techniques. The difficulty of choosing the right image processing technique for a mobile platform includes the balance between accuracy, robustness and computation cost.

This work aims to develop such a system to provide qualitative TB diagnosis results on the mobile platform in real time. The main contribution of the paper is to 'automatically' detect TB-specific antibodies by analysing digital images (i.e. ELISA images) with colour signals produced by biosensor technology. The plasmonic ELISA tests were conducted in Universiti Putra Malaysia. The proposed system does not require any additional hardware such as an opto-mechanical attachment to enhance the colour detection or guide the illumination source, which makes the system the most conveniently portable. Utilising an intelligent image processing algorithm, the presented system robustly separates the samples from the assay plate and extracts the features, and within a few seconds the system predicts the class label via a machine learning algorithm with high accuracy and ease of use. On a trained model, when a user provide an image to test, the system will require to process the image. Sending this user input directly to the cloud may present certain uncertainty and degradation of the image quality in resource-limited settings. A local analysis can enable TB testing facility for 24/7 even in the remote areas where internet connection is not available or very weak to send the images to the server, conduct the analysis, and send the result back to the smartphone. Although the proposed system is a native application to provide anytime-anywhere access, the presented system can be integrated to a server.

2. Literature Review

2.1 Computational Systems for TB-detection

To the best of authors' knowledge, there is no existing mobile, desktop or server based system for plasmonic ELISA based detection of TB antigen-specific antibodies. In literature, only a few studies employed machine learning techniques to assist in the diagnosis and monitoring of TB to offer a low-cost, simple, rapid and portable platform. Tracey et al. (2011) utilised acoustic signals to track the

recovery of pulmonary tuberculosis patients. The multilayer perceptron (MLP) showed 88.2% accuracy for ambulatory cough analysis.

Osman, Mashor, & Jaafar (2010) proposed a tuberculosis bacteria detection technique from tissue sample by Ziehl-Neelsen staining method. The prepared sample image from an optical microscope was segmented by moving k-mean clustering for tuberculosis bacteria extraction. Both RGB and C-Y colour were utilised to acquire a robust and improved segmentation under various staining condition. The hybrid multilayered perceptron network (HMLP) selected the features among the geometrical features of Zernike moments to detect tuberculosis bacteria. The result showed 98.0%, 100% and 96.19% of accuracy, sensitivity and specificity respectively to find the class of definite and possible TB.

Tsai, Shen, Cheng, & Chen (2013) developed colorimetric sensing using unmodified gold nanoparticles and single- stranded detection oligonucleotides for a TB test. The focus of the work was salt-induced AuNP colourimetric diagnosis for sensing target TB DNA sequences without multiple PCR cycles to amplify specific MTB target DNA sequences from extracted sputum or tissue samples. A smartphone was utilised just to collect the multiple detection results of colour variation from the concentration on cellulose paper and transmit the data to the cloud.

Table 1: TB related mobile applications on the Android platform

User	Region	Aspect	Questionnaire	Intelligent Systems	Ref.
Department of Health	South Africa	Management; TB and HIV diagnostic data	√	X	(Interactive Health Solutions, 2016a)
Specific Users	Bangladesh	Management	√	X	(Interactive Health Solutions, 2017)
Mine community	South Africa	TB screening	√	X	(Interactive Health Solutions, 2016b)
Patients	Pakistan	Control TB and drug-resistance	√	X	(Interactive Health Solutions, 2016)
Clinicians	Global	Decision on rapid diagnosis of TB and resistance	X	X	(Open Medicine Project, 2014)
Clinicians and Patients	Cambodia	Track lab test result	X	X	(Operation Asha, 2017)

There are commercial and endorsed mobile applications for TB in the popular application stores e.g. Google Play (Table 1, searched on 21-09-2017) and Apple app store. When it comes to diagnosis, the applications are for screening purpose only (Interactive Health Solutions, 2016, 2016, 2017). These applications store the screening data via the OpenMRS server. Either the user needs to insert the answers to a series of questions or the lab test results have to be manually inserted by the user or clinician. The available applications can ensure the data portability (Table 1) and in some cases diagnostic decision (Open Medicine Project, 2014), however they lack automation to produce a diagnostic result from the

specimen. Thus, there is a need for a system that does not require any additional hardware e.g. a plate reader and can produce laboratory scale test results.

2.2 Image Processing on mobile platform

An image processing based automatic system to be implemented on mobile platform firstly requires quality assessment and size reduction of the image. The quality assessment of the image will increase the accuracy of the system. The reliability will also increase due to the consistency in the input. The size-reduction and quantisation will make the system faster. For a mobile enabled decision support system, Bourouis et al. (2014) utilised a normalisation function to resize the retinal images to 32x32 pixels before storing 1-dimentional vector of pixel information. Lot of emphasis is provided in literature and commercial mobile applications for adjusting environmental condition and colour and light exposure correction (US9563824 B2, 2017; Wug Oh,Seoung; Kim, 2017).

The image segmentation algorithms in the literature can be mainly categorised based on five of the following methods: histogram thresholding, edge detection, clustering, region-based and graph-based methods (X.-Y. Wang, Wu, Chen, Zheng, & Yang, 2016). An alternative to segmentation is often carried out e.g. ELISA Plate Reader (Enzo Life Sciences inc., 2015) and AssayColor (Alidans srl, 2015). Both applications use a guideline e.g. grid or well structure to ensure a better image from a naïve user. Such a guideline can help the user to maintain an adequate distance of the sample from the camera, compromising the flexibility of the assay type. In both cases, the well-to-well distance is restricted because of certain assumptions regarding the plate size. Instead of any intelligent segmentation technique, few works in the literature (Mutlu et al., 2017; Ozkan & Kayhan, 2016) used cropping. It is highly discouraged for two reasons: i) it would require cropping skill from the user, and ii) it reduces the ease of use.

2.3 Colourimetric classification and decision on mobile platform

After processing the images, the features require analysis to generate a diagnostic decision and present it on the mobile platform. The related works done in the literature are mostly for paper based assays (Kim et al., 2017; Solmaz et al., 2018), which are less complex than the wet chemical assays. A cloud based mobile application was demostrated to classify peroxide content from mean RGB, HSV and LAB under diverse lighting environments (Solmaz et al., 2018). The least squares SVM and Random Forest were utilised to provide binary and multi-class classification respectively. The maximum accuracy at training phase (with 10-fold cross-validation) was 95% and on the mobile platform, it was reduced to 90.3%. On the other hand, analysing the colour features e.g. average, mode, median, mean, and centroid from the histogram of four colour spaces, the saliva-alcohol concentration was determined by Linear discriminant analysis (LDA), Support vector machine (SVM) and Artificial neural network (ANN) (Kim et al., 2017). The accuracy varied for different classes. Kim et al. (2017) showed that the stand-alone mobile application is two times faster than the server based application.

There are few mobile applications available for 96 well enzyme-linked immunosorbent assay (ELISA) based colour detection in the commercial and public app stores e.g. Spotxel® Reader (Sicasys Software

GmbH, 2017), Enzo ELISA Plate Reader (Enzo Life Sciences inc., 2015) and AssayColor (Alidans srl, 2015). Enzo ELISA Plate Reader (Enzo Life Sciences inc., 2015) and AssayColor (Alidans srl, 2015) neither provide any automatic complete analysis, nor include any decision support system (DSS) to interpret the colourimetric results. The Spotxel® Reader (Sicasys Software GmbH, 2017) comprising plate annotation and alignment, uses powerful noise processing and signal detection techniques. Instead of intelligent sensing, the application uses a virtual plate which can be laid over the plate image. The application expects the wells to be aligned with the virtual plate. The user is required to match the corner and centre wells with the grid. The virtual plate or grid can be scaled and rotated. However, aligning the wells with the grid requires some image capturing skills, which reduces the ease of use. The developers also acknowledged the limitations in the image processing (Sicasys Software GmbH, 2017). The application is capable of performing statistical analysis to quantify the result. The accuracy of such quantification is yet to be revealed.

Clearly it is evident from the recent literature and commercial mobile application stores (Table 1), that there is no existing low cost mobile solution which can benefit the wider population by anywhere anytime access to perform convenient confirmatory diagnosis of TB. To develop such a system, the critical review of the literature suggests to us the following findings:

- A strong image processing technique is required to eliminate the opto-mechanical attachments.
- Such an image processing technique has to be computationally feasible to be executed in the mobile environment.
- The image processing technique has to be intelligent and robust for wet chemical analysis. It should also consider powerful noise filtering techniques.
- The model needs to be trained off-line before deploying on the mobile platform. A native application would be faster than cloud based solutions, can be availed anytime anywhere and would possess less concern regarding cyber security.

3. Methods

3.1 Data Collection

3.1.1 Sample preparation

The experiments on plasmonic ELISA were mainly conducted in Universiti Putra Malaysia (Abuhassan et al., 2017; Tania, Lwin, Abuhassan, & Bakhori, 2017). However, the TB patient sputum samples were provided by School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Malaysia through their University's hospital. The fresh sputum sample were delivered to their lab and smear microscope analysis were carried out prior to culture method. The ELISA analysis was carried out simultaneously in the same lab.

For the detection of CFP-10, a 10 kDa secreted antigen from *Mycobacterium tuberculosis*, we first coated the ELISA plate with 100 μL of CFP-10 in carbonate buffer and then incubated for 1.5 h. Following the period, the plate was washed three times with PBS pH 7.6 and 0.05 % Tween-20 (PBST) by tapping it

against a clean paper towel. Now the plate was blocked with 370 µL of PBS containing BSA (PBSA) (1 mg/mL) for 1.5 h. All the antibodies and enzyme conjugates were diluted in diluent antibody containing PBST and 1% BSA. The plate was washed with PBST for three times, and kept the plate (invert) at 4°C for 2 h. Now, 100 µL of monoclonal anti CFP-10 antibody as primary antibody was added to the plate at 4°C for 1.5 h. After 1.5 h, the plate was washed with PBST for three times and the plate was added with 100 µL of biotinylated polyclonal secondary antibody and incubated for another 1.5 h at 4°C. The plate then washed three times and 100 µL of catalase-streptavidin conjugate (v/v 1:20) was pipetted into the plates and left for 1.5 h at 4°C. After the period, the wells in the plate were washed three times with PBST, two times with PBS, one time with deionized water and then dried. Now, 100 µL of hydrogen peroxide (in 1 mM MES, pH 6.5) buffer was pipetted into the wells. Immediately, 100 µL of gold ion solution freshly prepared in 1 mM MES buffer was added to the wells prepared in 1 mM MES buffer was added to the wells at room temperature. At this stage, the GNPs formation in the form of coloured solution can be seen and this can be read with microplate reader at an absorbance of 550 nm. For the analysis of real samples, the sputum from positive and negative TB patients were diluted in 4% sodium hydroxide first, and then proceeded to the same coating process as mentioned above.

3.1.2 Image acquisition

The dataset (generated as stated above) contains 252 images and 4 videos (Tania et al., 2017); 106 of them, captured with an iPhone 8-megapixel camera without mobile phone holder, were initially considered for (Abuhassan et al., 2017). Blurry images, images with inadequate camera exposure, observations intended for biosensor optimisation and the initial experiments where the colour widely varied from the final representative colours were removed. Finally, 27 images were selected from 22 independent observations. Among these images, 13, 3 and 2 images were captured by Samsung Galaxy J5 Prime (13-MP), iPhone 7 plus (12-MP) and iPhone 6 (8-MP) cameras respectively. The remaining images were captured with an iPhone (8-MP camera). The dataset contains images of 96 wells, which are partially filled, which means the plates contain both empty wells in addition to wells filled with sample. The final selection of 27 images from 22 independent observations were taken in a laboratory lighting environment. These images contain 266 samples - 81 of them are positive for TB-specific antibody, 181 are negative and three of the samples failed to produce any indicative result, thus 263 samples were finally selected. A mobile phone holder (NJS Telescopic Music Record Mobile Phone iPad iPhone Stand Inc G Clamp Mount 68G) was used while capturing the image. However, the acquired images vary in terms of well size, camera to ELISA plate position, light exposure and mobile phone. Considering a robust application, this variation is expected in the real life incoming images.
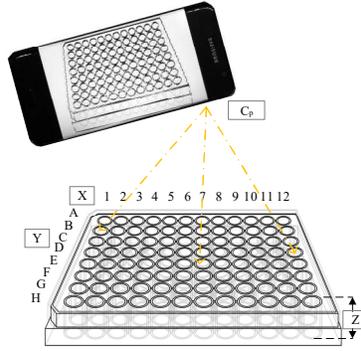
Fig. 1: Impact of sample and camera position with respect to ELISA plate. X and Y are the length and width of the ELISA plate respectively. Z= volume of sample in the well and C_p= camera position.

Let us assume, the assay plate, $A = f(X_x, Y_y, Z_z)$, where $\{X, Y\} \in \mathbb{Z}^+$ and $Z \in \mathbb{R}$ and $Z > 0$. In this work (Fig. 1), $X = \{1,2,...,12\}$ and $Y = \{A, B, ..., H\}$. For the commercially available 96 well plates X and Y will maintain such positions in rows and columns. The space between these wells can vary from plate to plate. Thus, the wells are signified in (x, y, z) coordinates. Each well denoted by $w_{X,Y} \in w_{x,y}$ in the plate and $s_{X,Y} \in w_{X,Y} = sample$, i.e. the well is filled with the sample. Both shape and depth of the well can vary, depending on the specification of the assay plate. Due to the dimension of the well itself, the distance between these wells can differ from plate to plate. Depending on the biochemical protocol, the amount of sample to fill these wells can vary as well. All this information has a direct impact on the imaging. However, the colour of each sample, $s_c(r, g, b) \neq f(x, y, z)$.

We have maintained the camera position (C_p) parallel to the A, giving the wells a uniform exposure to the camera. For a static C_p, the distance between C_p to each $w_{X,Y}$ is not equal. Thus, the sample to camera exposure is not equal. In theory, it would make $s_c(r, g, b)$ appear as $s_{cn}(r, g, b)$. The best exposure would be attained by the median $w_{X,Y}$.

The $s_c(r, g, b)$ can potentially differ due to the ambient conditions such as temperature, weather and geo-location, and certainly for the sample itself. However, this work is conducted in the laboratory environment.

3.2 Image pre-processing and segmentation

3.2.1 Image pre-processing

The goal of this work is to provide TB diagnosis on mobile platforms. Thus, this paper intends to circumvent the limited memory and processing power of the mobile devices, which is why the size of the images need to be reduced. The acquired images were scaled i.e. proportionally resized to reduce the processing time.

After the size reduction, the Gaussian Blur filtering was utilised as the image enhancement technique. This low pass filter (LPF) with parabolic amplitude Bode plot detracted the detail of the image by using a Gaussian function on each pixel.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \ldots\ldots\ldots(1), \text{ where } \sigma = standard\ deviation$$

In an image, x being the distance from the origin in the horizontal axis, y being the distance from the origin in the vertical axis, 2D Gaussian or normal distribution can be written as Eq. (1). Alternatively, a local Laplacian filter with contrast-limited adaptive histogram equalization was also implemented in the desktop application to evaluate the reformation.

Most commonly, the images captured by smartphones are in the RGB format. After smoothing, the image needed to be taken into a more perceptually linear colour space, LAB. This colour space transformation provides the ease to perform Euclidean distance calculation-based clustering at a later stage.

3.2.2 Image segmentation

Initially, a number of segmentation methods were implemented such as Otsu (Otsu, 1979), multi-level Otsu (Liao, Chen, & Chung, 2001), watershed (Meyer, 1994), super pixel (Ren & Malik, 2003) and k-means (Forgy, 1965; Lloyd, 1982; Macqueen, 1967). In the previous study (Abuhassan et al., 2017), k-means clustering showed promising performance, where the number of cluster was 6 and the size reduction was minimum.

Table 2: Major steps of the Algorithm

Input: Images of plasmonic ELISA plates
Output: Result
Steps:
1) Read the images in Red-Green-Blue (RGB) colour space
2) Dynamically scale the image based on the initial size
3) Smooth the image using Gaussian Blur filtering.
4) Convert the image into the CIELAB colour space
5) Use colours in the ab space to measure the Euclidean distance for clustering
6) Select k = 4
7) Dynamically repeat step to avoid local minima
8) For clusters 1 to k, separate the objects using the index clustering. This will produce k images.
9) Convert k images to binary images
10) Use morphological transformation includes Dilation and Erosion
11) Identify the optimum cluster(s) by calculating the difference between the produced images and the white colour (the image with the lowest distance is the optimum cluster)
12) Use Canny edge detector to sharp edges
13) Apply the Find Contours to the optimum cluster. This step will produce images equal to the sample wells in the segmented image
11) Read through the well images and apply size and position noise filtering
12) Extract the histogram features from the image
13) Save the values in the .csv file
14) Repeat steps 11 to 13 if more wells left, If not go to 15
15) Pass the .csv file to the classifier
16) Draw the result of positive or negative
17) Show the result to the user

The qualitative test determines the presence or absence of a substance. Thus, the decision is in the binary form. For the naked-eye tests, these binary classes are supposed to be visually distinguishable. Therefore, in theory there are only 3 relevant colours: background, foreground containing positive samples and alternatively negative samples. Hence, it can be hypothesised that k=3 should provide a perfect segmentation for a qualitative colourimetric test holding both positive and negative samples. However, if the foreground pixels of positive and negative samples are in two different clusters, it would make the sample annotation unnecessarily complex and computationally expensive. Thus, it would be desirable to force the clustering method to keep the positive and negative samples in the same cluster.

During the segmentation process, the random selection of cluster centroid position at initial stage compelled a requirement for the best cluster persuasion. Thus, after the clustering, a series of post-processing techniques were applied. As mentioned in Table 2, these post-processing techniques include morphological operation encompassing Dilation and Erosion followed by object detection. The morphological transformations on a binary image in most cases require two inputs: the image and the kernel which identifies the nature of the operations. The contours were exploited after the segmentation and morphological transformations for size analysis and object detection.
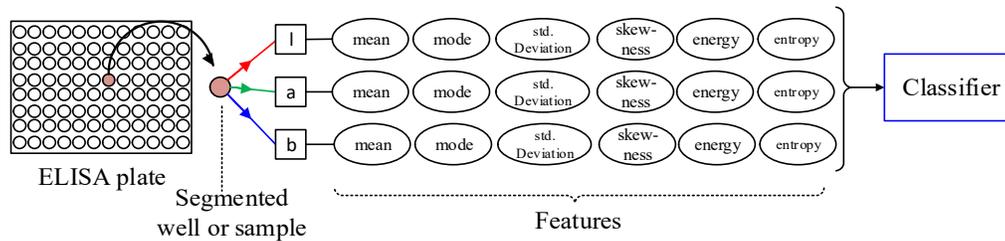
Fig. 2: Feature analysis framework

3.3 Feature analysis and classification

Once the samples (ROI) were separated, the characteristics of these samples were analysed. In this paper, the feature analysis involves measurement of colour moments. This work includes basic features necessary to compute any probability distribution (Sergyan, 2008). The framework is illustrated in Fig. 2. As described in Abuhassan et al. (2017), mean, mode, standard deviation, skewness, energy and entropy in L, a and b channel (18 features in total) were considered to train the model.

3.4 Mobile-enabled expert system

In this work, the plasmonic ELISA-based TB detection was deployed on the Android platform. The mobile application was developed on a Samsung Galaxy S7 edge. The minimum target SDK is 21 (API level 5).
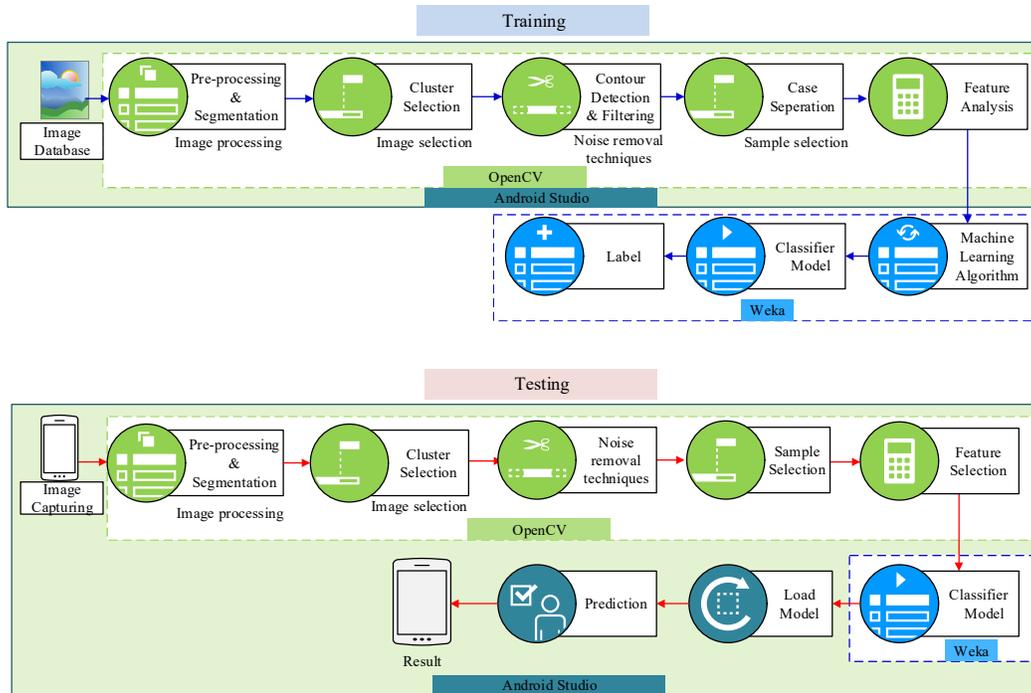
Fig. 3: System framework: Implementation of the Algorithm

The steps outlined in Table 2 were implemented on the Android platform as illustrated in Fig. 3. Due to convenient functionality on the Android platform, OpenCV was utilised to perform the data pre-processing i.e. image processing and feature extraction. The feature values of the segmented, individual sample (well) were stored as text and carried to Weka to train the classifier model offline. The offline training was conducted on a 64 bit Windows system with Intel ® Core ™ i7-4770 CPU at 3.40 GHz processor and 16 GB RAM.

Once the model is trained, it was loaded on the Android platform using Weka library (weka.jar file). At the testing level in Fig. 3, the user can use any new image of the plasmonic ELISA test on an Android device to produce the correct prediction of TB disease in real time.

4   Results

4.1   Image acquisition of Plasmonic ELISA

Plasmonic ELISA links the colour of plasmonic nanoparticles to the presence or absence of the analyte (target protein). *Mycobacterium tuberculosis* ESAT-6-like protein esxB (CFP-10) was used as a target protein biomarker for the TB detection Plasmonic is accomplished by linking the growth of gold nanoparticles with the biocatalytic cycle of the enzyme label. The protocol adapts a conventional ELISA procedure with catalase-labelled antibodies. The enzyme consumes hydrogen peroxide (H_2O_2), and then gold (III) ions are added to generate gold nanoparticles. The concentration of hydrogen peroxide dictates the state of aggregation of gold nanoparticles. This allows for the naked-eye detection of analytes by observing the generation of blue- or red-coloured gold nanoparticle solution.
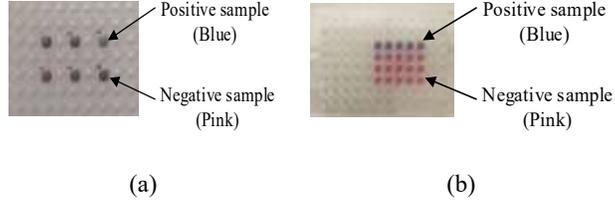
Fig. 4: Samples in a plasmonic ELISA plate. (a) Samples are hard to visually distinguish, (b) Samples are visually distinguishable

In this work, the presence of TB-specific antibodies can be confirmed if the sample turns blue in the ELISA plate. In Fig. 4 (a) gold ions are reduced when H_2O_2 is present. The top 3 samples are free from TB-specific antibodies. In the presence of H_2O_2 non-aggregated nanoparticles are formed turning the solution pink. In the bottom 3 samples, the concentration of H_2O_2 is decreased, turning the samples blue, confirming the presence of TB-specific antibodies.
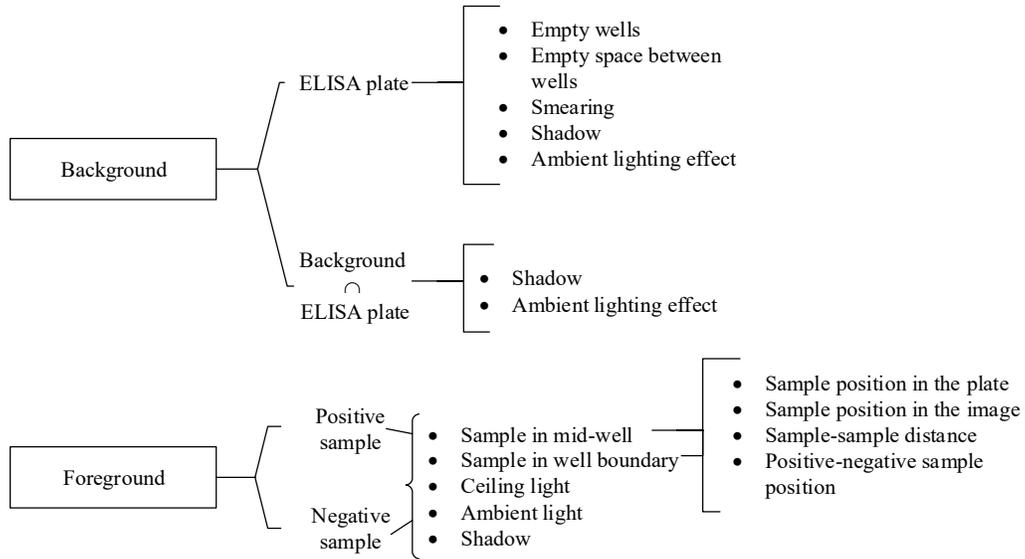
Fig. 5: Observation of the associated colours and key variables in the image

The key observations from the detailed inspection of the dataset are listed below.

Obs. 1: In the presented dataset, the sample-to-sample distance was not constant (Fig. 1). If the wells are filled within a close neighbourhood, there is an unavoidable smearing effect. Thus, the background cluster holds many pixels which are close to the foreground pixels. With varying position(x, y), depending on the class of $s_{X,Y}$, the background cluster is difficult to separate from the foreground clusters.

Obs. 2: In some cases, the positive and negative samples are hardly visually distinguishable. For image e.g. Fig. 4 (b), the samples are adequate for naked eye measurement. For sample image e.g. Fig. 4 (a), the indicated sample pair are hard to differentiate. This issue can worsen if the plate contains only one sample and the colour is as ambiguous as in Fig. 4 (a), which can lead to subjective interpretation. Moreover, there is a conscious variation in the sample colour, $s_c(r, g, b)$ on independent A.

Obs. 3: In the dataset, the value of Z (the volume of sample in a well) had an impact on the size of the sample ($s_{x,y}$). It implies that the 1st colour moment can vary based on how the wells are filled. $s_{cn}(r, g, b) = f(Z)$. A well filled up to the surface would have a better exposure when they are positioned at the far edge of the plate.

Obs. 4: This work is comprised of wet sample, which is not immune to light reflection from its surroundings (Fig. 5). Initially, this ceiling light was not taken into account. Our hypothesis was: the $w_{X,Y}$ with median (x, y) would be the ideal position for the samples. Even a well filled up to the surface (Obs. 3) in the median position can suffer from the ceiling light reflection.

Obs. 5: The impact of 'camera to well position' (Fig. 1) is aligned with our prediction in Sec. 3.1. Such influence can be analysed by the SKEW (Fig. 2).

The observations Obs. 1, Obs. 3 and Obs. 4 have a clear impact on the image processing measures. The Obs. 2 works in our favour. The qualitative colourimetric tests are usually suitable for naked-eye detection, which necessitates (i) adequate biosensors to produce visually distinguishable colours and (ii) a user who has appropriate colour vision. Firstly, the use of intelligent systems can reduce the biochemical complexity without compromising the accuracy, specificity, sensitivity and reliability. Therefore, the positive and negative samples do not require to be visually distinguishable. Secondly, an intelligent system such as the system we presented in this paper can eliminate the subjectivity of interpretation. A robust system should be able to handle the variation of sample colour mentioned in Obs. 2.
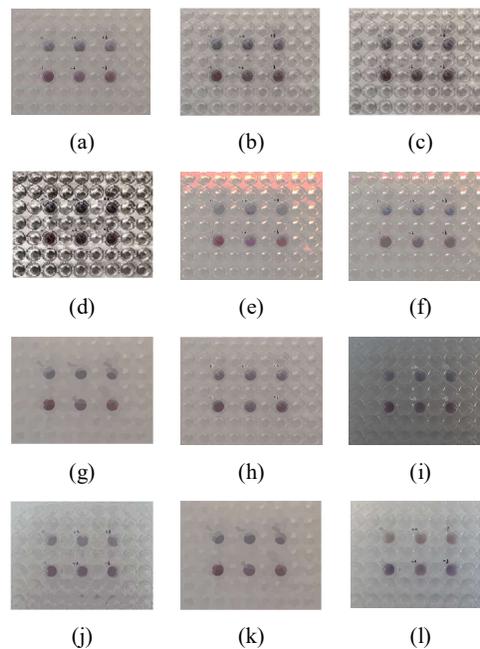
Fig. 6: (a) Samples in a plasmonic ELISA plate. Gradual enhancement of the image: (b) sharpened, (c) smoothened, (d) final enhancement before colour space transformation. Quantisation input: (e) full size quantisation, (f) plane-

by-plane quantisation, (g) Superpixel, (h) JSEG in MATLAB and (i) Gabor filtering, (j) k-means, (k) Superpixel, (l) Gaussian filtering in OpenCV

4.2 Image Analysis

4.2.1 Image pre-processing

At first, the acquired images were scaled and quantised to reduce the size of the image. For a simple method such as Otsu, the impact of scaling on processing time is negligible. On the other hand, to perform numerous iterations, the aid of scaling is obligatory for a heavy segmentation technique such as clustering. In our earlier study (Abuhassan et al., 2017) utilising desktop application, typical images ranging ~3000-4000 pixels were scaled 50%, which requires more reduction to be implemented on mobile platform. Bourouis et al. (2014) utilised 32x32 pixels retinal images, which is not substantial to analyse the colour features of the presented dataset. Moreover, the resizing in Bourouis et al. (2014) was not dynamic. For a known condition, the height and the width of the image will not vary to a great extent. However, it may vary due to factors such as position of the camera, size of the plate, and camera configuration. Thus, the size reduction in this work was performed dynamically (Android Developers, 2018) and proportionally so that the geometry of the ROI was not deformed. The quantisation techniques were carried out to reduce the size of the image by reducing the number of colours in the image (Fig. 6). It was observed that quantisation has insignificant impact on the overall segmentation process. As a result, it was discarded. For a good quality image, such as Fig. 6 (a), the requirement of image enhancement is not high. However, to develop a robust technique applicable to poor quality images, image enhancement is essential. Hence, the images were sharpened, which is a function of resolution and acutance. The radius value of standard deviation of the Gaussian LPF controls the size of the region around the edge pixels that is affected by sharpening. A large value sharpens wider regions around the edges, whereas a small value sharpens narrower regions around edges. The higher value of sharpening will lead toward larger increase in the contrast of the sharpened pixels. A very large value for this parameter may create undesirable effects in the output image, as it may appear as noise. Thus, an edge-aware local contrast alteration was deployed to create more contrast. In the next step, the sharpened image was selectively smoothened to blur the empty wells. The objective of such extensive pre-processing was to ease the segmentation process and minimise the number of iterations. However, these excessively processed images caused separation of cluster 1 and 2 (Eq. 2) in two different clusters at the segmentation stage, as predicted. This led to an elongated object detection step. Hence, the Gaussian Blur filtering (Fig. 6(l)) was utilised as the only image enhancement (negative) technique prior to image segmentation. It assisted to address Obs. 1.

Input	Watershed	Otsu	Multi-level Otsu	K-means
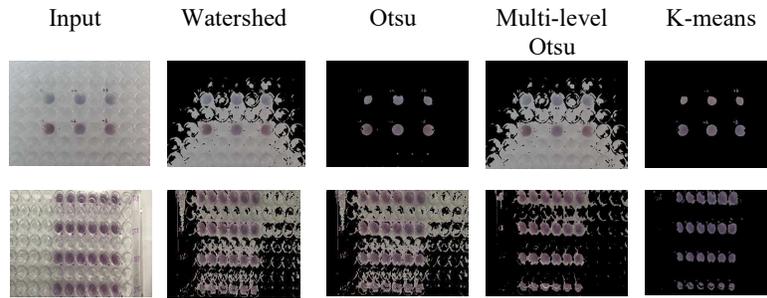
Fig. 7: Image Segmentation using different techniques

4.2.2 Image segmentation

The qualitative performance of different image segmentation techniques can be seen from two different input images as shown in Fig. 7 (OpenCV). The Otsu and watershed transformation were unsuccessful in segmenting adequately. The multi-level Otsu performed well for a good quality image with low smearing effect, where the samples are evenly spaced e.g. Fig. 6 (a). However, it failed for many images e.g. second input of Fig. 7. The JSEG was time consuming, not suitable for implementation in real-time. The k-means showed good segmentation performance resembling our early study (Abuhassan et al., 2017).

(a)   (b)   (c)

(d)   (e)   (f)
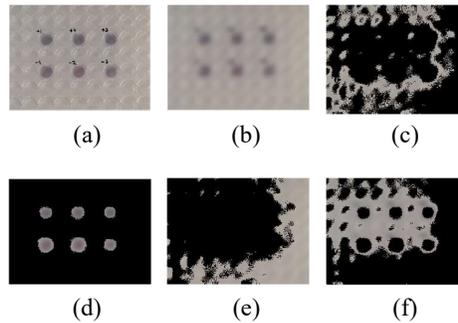
Fig. 8: (a) Input image, (b) Gaussian 2D filtering in MATLAB, (c) first cluster, (d) 2nd cluster with positive and negative samples, (e) 3rd cluster, (f) fourth cluster

As mentioned earlier, we require the positive and the negative samples to belong in the same cluster. The use of Gaussian LPF forced the segmentation process to hold the samples in the same cluster as illustrated in Fig. 8. Such pre-processing and segmentation techniques precisely addressed Obs. 1 and Obs. 4.
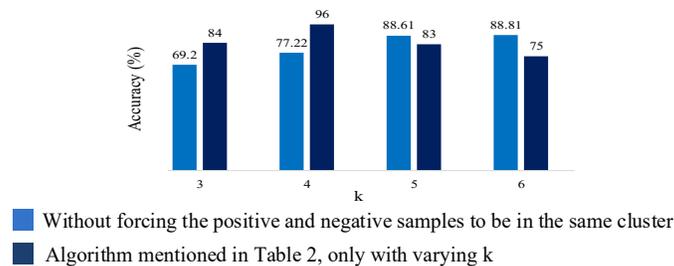
Fig. 9: Performance of the image processing algorithm for different k

In our preliminary study, we performed the k-means with k=6 without any pre-processing and with minimum resizing (Abuhassan et al., 2017). As it is mentioned earlier, theoretically the input image should be segmented into three different clusters, which was later found to be imprecise in this work (Fig. 9). Initially, utilising the desktop application (MATLAB), we analysed the impact of varying k without using Eq. (1). Without forcing the algorithm to keep both positive and negative samples in the same cluster, higher k exhibited better segmentation, which is computationally expensive and less suitable to be performed in the mobile environment. Without utilising Eq. (1), the maximum accuracy achieved was 88.81% with k=6 (Fig. 9).

It was also observed that the required number of k may vary for image to image due to the image quality, filled well-to-well distance, camera position and positive-negative sample position and ratio per image. A range in the required number of clusters was also observed from the silhouette method (Rousseeuw, 1987), which supports the observation. However, it is not feasible to use multiple iterations to choose a different k each time for each image as it would become computationally expensive. If the ELISA plate contains more filled wells, the execution time is likely to be higher, which makes it unsuitable for mobile applications. In future, we will utilise the image histogram to predict the required number of k before starting the iteration. In contrast, many applications in the commercial app stores simplify the image processing portion by utilising a gridline approach, resulting in compromising the freedom of diverse plate size and in some cases the ease of use.

In this paper, we have used k-means with an optimum number of k=4 (as mentioned in Table 2), with complimentary rigorous pre and post processing techniques. With varying k, the overall performance (Fig. 9) of the image processing algorithm (Table 2) varied as well. When k=3, the unsupervised machine learning showed under-segmentation, leading the segmented region to hold image area outside of ROI. When k>4, the algorithm showed over-segmentation, which resulted in more poor performance due to extensive post-processing of the segmented image.
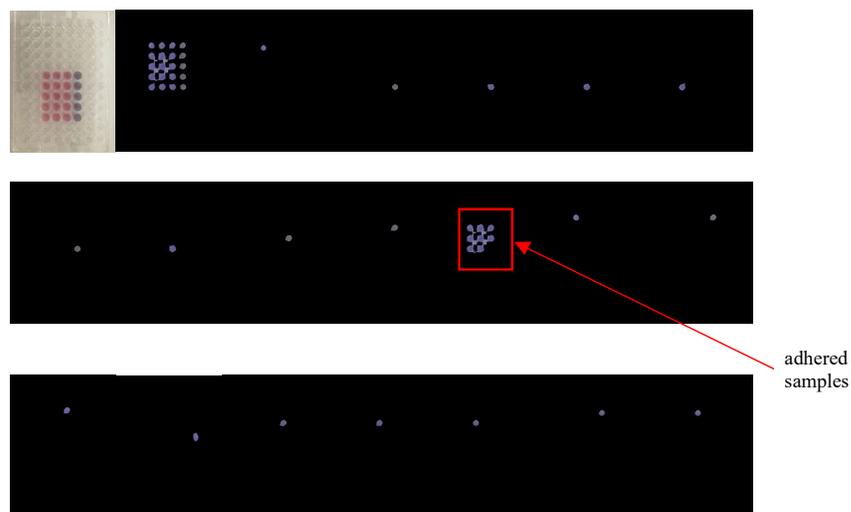
Fig. 10: Constraint in segmentation without post-processing: adhered samples after clustering
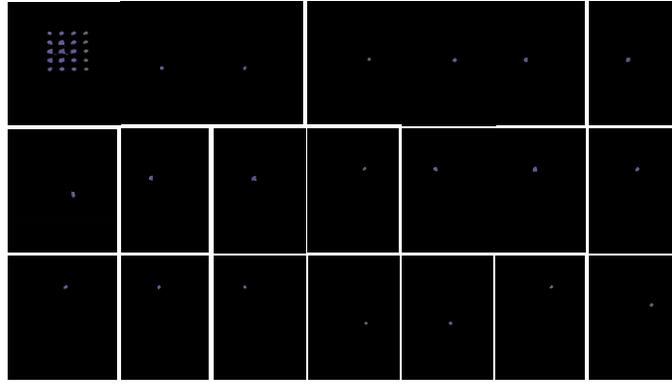
Fig. 11: Image processing utilising morphological processing

For post-processing, subsequent to clustering, the images were converted into binary images, followed by morphological operation. Without morphological processing, many images would suffer from incorrect object reckoning. After segmentation, in a few cases the samples joined together in association with the noise. If this phenomenon occurs in the right cluster, then the ROI separation process will fail. This problem can be better visualised with Fig. 10, where the samples could not be adequately separated. Due to Obs. 1, few samples are attached together as illustrated in the highlighted image (marked with a red box). All these samples were categorised as a single sample, whereas they are 8 adhered samples. Therefore, the binary dilation and erosion was used to ease the object detection. Erosion operation was iterated four times more than the dilation in order to isolate individual wells and overcome Obs. 1. The size of each (processed) sample in Fig. 11 is much smaller than Fig. 10, which helped to reduce noise, and the samples were no longer adhered (Fig. 11). However, it presented a small possibility of over segmentation due to camera–to-plate position and the type of the plate. To achieve a higher degree of freedom regarding the plate size, a conversation table is required to transform the physical dimension of the assay plate into image pixels.
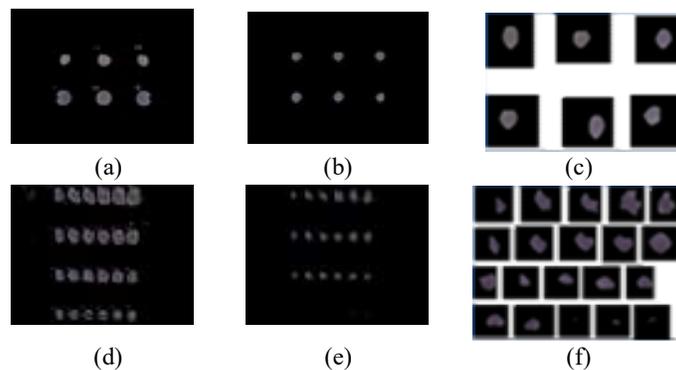
Fig. 12: (a) Segmented image using Fig. 4(a) as the input. (b) Post-processing after segmentation. (c) Final output after contour detection. (d) Segmented image using Fig. 5(a) as the input. (e) Post-processing after segmentation. (f) Final output after contour detection

The next challenge was to automatically recognise the best cluster among the four clusters, which was accomplished by exploring the well-to-well background. According to our research, the best cluster is the one that has less white background. As explained earlier, the adversity of Obs. 1 is worsened if this

phenomenon happens after the segmentation in the chosen best cluster (Fig. 10). This background acting as noise needed to be filtered from the best cluster (Fig. 11). Finally, the ROIs i.e. samples are separated using contour detection technique.

To address all the challenges listed as the observations in Sec. 4.1, selection of a precise post-processing technique was crucial. The $s_{X,Y}$ was diverse for the entire dataset, which is expected for a robust use of the application. We demonstrated an intelligent inspection after segmentation to correctly extract the sample from the noise (Fig. 12). The object detection technique functioned accurately even in the case of blurry images in which a larger number of wells were detected and used.

4.3 Feature Analysis and Classification Result

The colour moments of the extracted ROI were analysed to train the system offline. The reported articles (Mutlu et al., 2017; Solmaz et al., 2018) mostly feed the mean colour values to the classifiers. We have considered 18 histogram features listed in Table 2 to ensure all the variables are being considered for a robust operation.
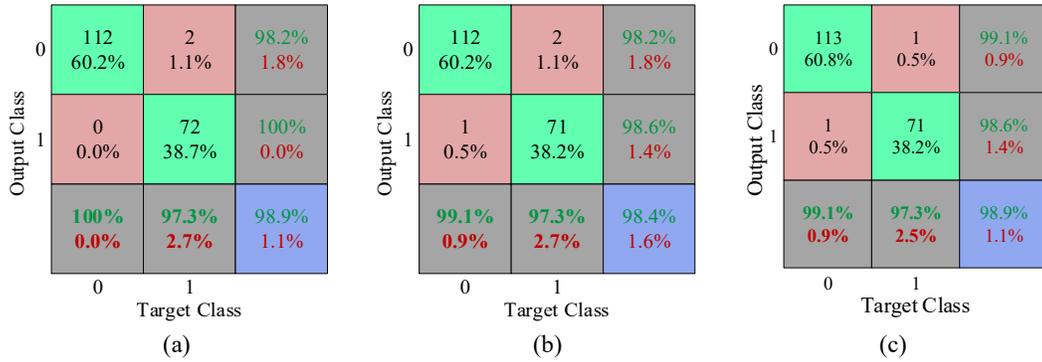
Fig. 13: Confusion matrix of (a) Random Forest, (b) Random Tree, (c) Random Committee

The non-parametric classifiers such as random forest (RF), decision tree, k-nearest neighbours algorithm (kNN), and cubic support vector machine (CSVM) performed better than the parametric classification method e.g. linear discriminant and logistic discrimination. Without cross-validation, all these non-parametric methods produced 100% accuracy.

The Multilayer Perceptron (MLP) with backpropagation was comparatively slow and the classification performance was poor as well. It provided 95.2% accuracy. The learning rate was 0.3. In order to circumvent the backpropagation algorithm to be trapped in the local minima, the momentum rate was chosen to be 0.2. There were 500 epochs to train through without decaying the learning rate. The network was allowed to be reset. Both attributes and classes were normalised before training the model. The required number of hidden layers were calculated from the number of attributes and classes $\left(\frac{attributes+classes}{2}\right)$. The nodes of these 10 hidden layers were sigmoid. No validation set was used to terminate the training. The model was built in 0.35 seconds.

With 30 weak learners, the RF provided a high accuracy (97.2%) in our preliminary study (Abuhassan et al., 2017). The RF showed consistent performance in this study as well. In this work, the bag size in RF was chosen to be 100 without storing out-of-bag predictions in internal evaluation object. The bagging was conducted with 100 iterations and base learner. Only one seed was taken for random number generator. The maximum depth of the trees were kept unlimited and minimum one instances per leaf was allowed to occur. The desired batch size for prediction was chosen to 100. It took 0.09 seconds to build the model in Weka.

Table 3. Result of different classifiers in Weka platform

Classifier	κ	TP rate	FP Rate	Precision	F-Measure	ROC Area	Class
Random Forest	0.9775	0.982	0.00	1.00	0.991	1.00	Negative
		1.00	0.018	0.973	0.986	1.00	Positive
Random Tree	0.9661	0.982	0.014	0.991	0.987	0.984	Negative
		0.986	0.018	0.973	0.979	0.984	Positive
Random Committee	0.9773	0.991	0.014	0.991	0.991	0.999	Negative
		0.986	0.009	0.986	0.986	1.00	Positive
Bagged Tree	0.9098	0.956	0.042	0.973	0.965	0.996	Negative
		0.958	0.044	0.932	0.945	0.997	Positive
Multilayer Perceptron	0.8988	0.947	0.042	0.973	0.960	0.977	Negative
		0.958	0.053	0.920	0.939	0.981	Positive

In this paper, the RF and Random Committee (RC) both showed 98.9% accuracy with stratified cross validation (10-fold) in the Weka platform. Keeping the batch size, number of seeds, minimum number of instances per leaf as same as RF, the RC was built in 0.01 seconds using default number of iterations (10). The size of the tree varied in each iteration. The Random Tree (RT), a decision tree built on a random subset of columns achieved 98.4% accuracy. Keeping the parameters as same as RC, the Bagged Trees consisting unpruned binary trees provided 95.7% accuracy.

The Cohen's kappa coefficient (Cohen, 1960) (κ) is a statistic which compares an observed accuracy with an expected accuracy that can be seen as a random chance. It can be calculated as, $\kappa = \frac{p_o - p_c}{1 - p_c}$, where p_o is the prequential accuracy of the classifier and p_c is the probability that a chance-classifier makes a correct prediction. The result of κ being 1 would signify that the classifier is always correct and 0 would mean that the predictions coincide with the correct ones as often as those of the chance classifier. The κ can provide more precise evaluation than the traditional accuracy metric. Moreover, it can aid in evaluating the classifiers among themselves. From the κ measurement (Table 3), the RF is the best classifier. The κ of RF is in agreement with the accuracy of our previous study as well (Abuhassan et al., 2017).

The true positive (TP) rate provides the instances where the samples are correctly classified as the given class. TP rate $= \frac{TP}{TP+\text{False negative (FN)}}$. It can also be expressed as the sensitivity or recall. The highest TP rate was attained by the RF. The false positive (FP) rate or Fall-out provides the instances when the samples are falsely classified as the given class. The precision is the fraction of relevant instances among the retrieved instances i.e. Precision $= \frac{TP}{TP+FP}$. The F-Measure provides a combined measure for precision and recall. It can be expressed as, $F - \text{measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision}+\text{Recall}}$. The detection ability of the classifier can be better perceived by the receiver operating characteristic (ROC) area (Table 3). Considering the ROC area, the RF is the best classifier for our dataset. The accuracy, specificity and sensitivity can be better visualised from the confusion matrix. The confusion matrix of the top three classifiers are illustrated in Fig. 13.

The random feature selection in RF, makes the trees more independent from one another than Bagging, which led to higher accuracy and better bias-variance trade-off. Each tree is able to learn only from a certain subset of features, making it a faster ensemble method as well. Moreover, the RF showed a consistent better performance in various metrics. Similar performance to RF was attained in the MATLAB platform as well. Therefore, we trained our model with RF on the mobile platform.
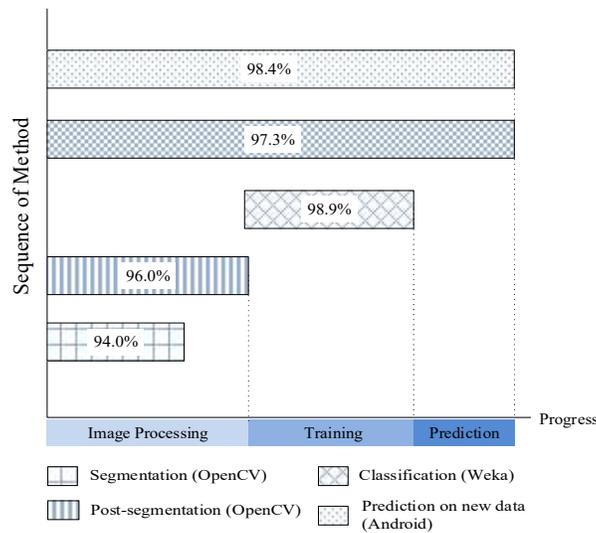
Fig. 14: Accuracy at different stage of the system

4.4 Testing and Validation on the Mobile platform

In this work, we demonstrated automatic, real-time TB disease decision making on a mobile platform.

The trained model was deployed on the Android platform as illustrated in Fig. 3. To test the efficiency of this mobile-based intelligent algorithm for detecting TB, a separate dataset was used than Sec. 3.1. This new dataset is unknown to the system and contained 61 samples. Among these samples, 20 were positive, 41 were negative and one failed to produce a colour. This held-out validation on the mobile platform ensures the reliability of the system.

Fig. 15: Confusion matrix of testing performance at mobile platform

On the mobile platform, for this unseen data, the system provided correct prediction for 60 samples. Thus, a final accuracy i.e. from image processing up to TB detection on the mobile platform, of 98.4% was achieved (Fig. 14). The necessity of balanced data can be perceived from the confusion matrix (Fig. 15). The performance of the classifier on a balanced dataset in Weka platform is shown in the Supplementary document. In the absence of a larger dataset, over-sampling or multiple resampling (Estabrooks, Jo, & Japkowicz, 2004) can shed some more light on the performance of the mobile platform.
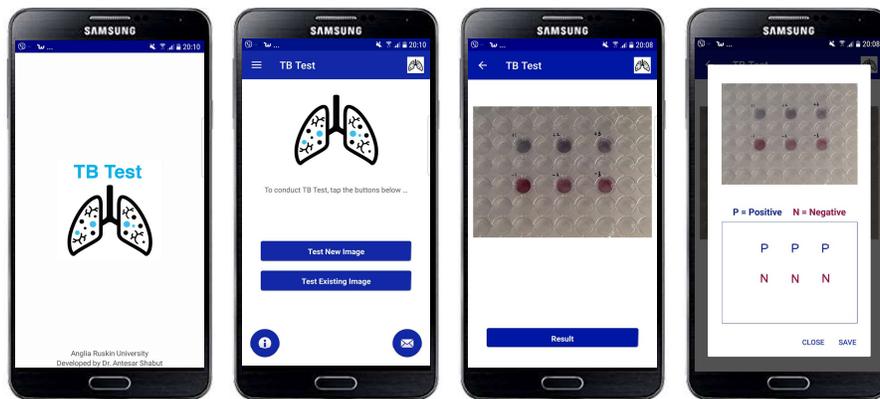
Fig. 16: TB disease detection application

One of the biggest challenges of this work was to provide this diagnostic decision on the mobile platform in real-time. The prediction on the mobile platform requires the performing of image processing of the incoming image on the mobile device itself. The processing time is a subject to the number of iterations during image segmentation and object detection and is heavily influenced by Obs. 1, Obs. 2 and Obs. 4. We embraced careful pre-processing techniques to minimise the number of iterations. The k-means uses a random initial value and it is sensitive to the size of the image, thus scaling has a direct impact on this clustering technique regarding the number of k and how the image is being segmented, which justifies the pre-processing used in here. Moreover, we managed to confine the best cluster containing multi-sample of different classes within the same cluster with aid from the Gaussian filtering. This algorithm (Table 2) possess the robustness to deploy the image processing scheme on the mobile platform for other assay plates, which can be further authenticated for the quantitative colourimetric tests.

The execution time was recorded for all the images in the dataset. Due to fewer iterations, the image processing occurs in real time liberating the implementation on the mobile platform (Fig. 16). The input image of Fig. 10, containing 20 samples, took 23 seconds to produce the result. The image with only 6 samples e.g. Fig. 6(a) provided the result within 9 seconds. Therefore, it can be concluded that our system is capable of delivering TB disease decision from the plasmonic ELISA image on the mobile platform within ~1-2 seconds/sample (Fig. 16).

An image annotation technique was used in the embedded system to identify each sample individually. The Android memory management was utilised to enhance the heap performance by following the correct life cycle of activities in the Android platform. The memory management includes actions with garbage collector (GC), memory optimisation, and tree dominator (Android Developers, 2018). In order to reduce memory leak, the elements used by the system were scaled. The application persistently searched for the objects that were no longer required (garbage) after the life cycle, or reachable which were needed for references. The heap dumps were accumulated over the period of time to determine if there was any growing memory leak. The allocation tracker facilitated a better understanding of the memory usage.

In this work, we utilised 8-bit channel ARGB_8888 configuration for the bitmap of image. Although it occupies considerable amount of memory and immediately allocated in the heap, quickly exhausting the memory, this is an optimised choice to maintain the quality of the scaled images. Moreover, this work involves series of image conversations. Therefore, redundant images had to be simultaneously deleted. For garbage collection, a mechanism to remove unnecessary objects to the java application using a virtual machine, Dalvik GC was utilised (Ehringer, 2010). The type of garbage collectors were also examined closely.

The application was tested on Samsung Galaxy S6, Samsung Galaxy Note 3 and Samsung Galaxy J3 Prime. On most of these devices, the application performs similarly in terms of the classification accuracy and processing time.

5   Limitations and scope of improvement

For any diagnostic system, it is important to note its limitations as well as its capabilities. In the image processing section, in spite of multi-step filtering, the noise due to Obs. 1 needs to be further adjusted. In future, to train the model we will conduct feature optimisation and bias-variance trade-off. The future work should also focus on rectifying the variation between κ and accuracy (Sec. 4.3). We will also explore non-parametric NN with backpropagation, deep NN, hybrid decision tree and naïve Bayes classifiers (Farid, Zhang, Rahman, Hossain, & Strachan, 2014) to investigate a potential enhancement of the accuracy.

6   Conclusion

This paper has presented a mobile enabled plasmonic ELISA based TB antigen-specific antibodies detection scheme using smartphones with the integration of machine learning techniques. Using a robust image processing technique comprised of clustering and object detection, our system can detect samples

(wells) without any guide or virtual plate. The decision components facilitated selection of the right cluster among the multiple number of clusters, detection of wells and transcending the samples from noise. Therefore, unlike the reported articles, the system does not require the user to provide seed points or perform cropping. Moreover, the system is capable of reading multiple samples and classifying them as positive or negative in real time.

The plasmonic ELISA based technique produce colours for positive and negative samples. However, making of a final decision based on the colour appearance is not accurate in all the cases. Therefore, in this work, we demonstrated a smartphone-based POC platform that takes the final decision based on colour analysis. This work incorporated supervised machine learning to free the TB test result from the colour perception of individuals and its subjectivity of interpretation. Utilising 18 histogram features, we achieved 98.9% accuracy with the Random Forest classifier. This fully automated and self-contained system with image capturing, analysing and classification service is then embedded into the Android system. Using a completely new dataset, we demonstrated 98.4% accuracy to diagnose TB-positive samples on the mobile platform. In the absence of any existing automatic platform without an opto-mechanical attachment, to the best of our knowledge, it is the best performance for TB diagnosis on the mobile platform.

The portability, technical and financial feasibility in automatic TB diagnosis in the presented system can benefit millions of people, especially in remote locations where few experts are available. This technique can be applied to other colourimetric qualitative tests, especially for ELISA and paper-based assays. Moreover, this system can be a guide in providing properly distinguishable colour by minimising the complexity of a chemical method utilising a powerful algorithm, which will reduce the dependency on perfect colour vision for naked-eye evaluation. The scheme shows great potential in evolving healthcare applications to benefit wider communities. A polythetic approach and subsequently a clinical trial will be executed in future to enhance the expert system with better precision and reliability.

Acknowledgement

References

Abuhassan, K. J., Bakhori, N. M., Kusnin, N., Azmi, U. Z. M., Tania, M. H., Evans, B. A., … Hossain, M. A. (2017). Automatic Diagnosis of Tuberculosis Disease Based on Plasmonic ELISA and Color-based Image Classification. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4512–4515). Jeju Island, South Korea. https://doi.org/10.1109/EMBC.2017.8037859

Alidans srl. (2015). AssayColor. Retrieved January 10, 2017, from https://play.google.com/store/apps/details?id=com.alidans.assaycolor

Android Developers. (2018). Overview of memory management. Retrieved May 29, 2018, from https://developer.android.com/topic/performance/memory-overview

Bourouis, A., Feham, M., Hossain, M. A., & Zhang, L. (2014). An intelligent mobile based decision support system for retinal disease diagnosis. *Decision Support Systems*, *59*(November 2015), 341–350. https://doi.org/10.1016/j.dss.2014.01.005

Centers for Disease Control and Prevention. (n.d.). Tuberculosis (TB) | CDC. Retrieved September 18, 2017, from https://www.cdc.gov/tb/

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Department of Economic and Social Affairs. (2016). *International Migration Report 2015*. https://doi.org/ST/ESA/SER.A/384

Ehringer, D. (2010). *The Dalvik virtual machine architecture*. Retrieved from http://show.docjava.com/posterous/file/2012/12/10222640-The_Dalvik_Virtual_Machine.pdf

Enzo Life Sciences inc. (2015). Enzo ELISA Plate Reader. Retrieved September 21, 2017, from https://play.google.com/store/apps/details?id=com.enzo.elisaplatereader

Estabrooks, A., Jo, T., & Japkowicz, N. (2004). A Multiple Resampling Method for Learning from Imbalanced Data Sets. *Computational Intelligence*, *20*(1), 18–36. https://doi.org/10.1111/j.0824-7935.2004.t01-1-00228.x

Farid, D. M., Zhang, L., Rahman, C. M., Hossain, M. A., & Strachan, R. (2014). Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks. *Expert Systems with Applications*, *41*(4 PART 2), 1937–1946. https://doi.org/10.1016/j.eswa.2013.08.089

Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, *21*, 768–769.

GSMA Intelligence. (n.d.). Definitive data and analysis for the mobile industry. Retrieved August 7, 2017, from https://www.gsmaintelligence.com/

Interactive Health Solutions. (2016). Global Fund TB. Retrieved September 21, 2017, from https://play.google.com/store/apps/details?id=com.ihsinformatics.tbr3mobile_sa&hl=en

Interactive Health Solutions. (2016). MINE TB. Retrieved September 21, 2017, from https://play.google.com/store/apps/details?id=com.ihsinformatics.tbr4mobile

Interactive Health Solutions. (2016). TB REACH 4 - Kotri. Retrieved September 21, 2017, from https://play.google.com/store/apps/details?id=com.ihsinformatics.tbr4mobile_pk&hl=en

Interactive Health Solutions. (2017). Childhood TB-Bangladesh. Retrieved September 21, 2017, from https://play.google.com/store/apps/details?id=com.ihsinformatics.childhoodtb_mobile&hl=en

Khademhosseini, A. (2011). Nano/microfluidics for diagnosis of infectious diseases in developing countries. *Adv Drug Delivery Rev*, *62*(4–5), 449–457. https://doi.org/10.1016/j.addr.2009.11.016.Nano/microfluidics

Kim, H., Awofeso, O., Choi, S., Jung, Y., & Bae, E. (2017). Colorimetric analysis of saliva--alcohol test strips by smartphone-based instruments using machine-learning algorithms. *Appl. Opt.*, *56*(1), 84–92. https://doi.org/10.1364/AO.56.000084

Liao, P.-S., Chen, T.-S., & Chung, P.-C. (2001). A Fast Algorithm for Multilevel Thresholding. *Journal of Information Science and Engineering*, *17*, 713–727.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. https://doi.org/10.1109/TIT.1982.1056489

Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 281–297). Berkeley, California: University of California Press. Retrieved from https://projecteuclid.org/euclid.bsmsp/1200512992%0A%0A

Maggio, Emilio;Pan, Qi; Reitmayr, G. (2017). *US9563824 B2*. Retrieved from https://www.google.com/patents/US9563824

Meyer, F. (1994). Topographic distance and watershed lines. *Signal Processing*, *38*(1), 113–125. https://doi.org/10.1016/0165-1684(94)90060-4

Mutlu, A. Y., Kılıç, V., Özdemir, G. K., Bayram, A., Horzum, N., & Solmaz, M. E. (2017). Smartphone-based colorimetric detection via machine learning. *The Analyst*, *142*(13), 2434–2441. https://doi.org/10.1039/C7AN00741H

NHS. (n.d.). Tuberculosis (TB) - NHS Choices. Retrieved September 18, 2017, from http://www.nhs.uk/Conditions/Tuberculosis/Pages/Introduction.aspx

Open Medicine Project. (2014). FIND TB. Retrieved September 21, 2017, from https://play.google.com/store/apps/details?id=tompsa.findtb&hl=en

Operation Asha. (2017). eAlert Cambodia. Retrieved September 21, 2017, from https://play.google.com/store/apps/details?id=org.opasha.eCompliance.ecomplianceLabCambodia

Osman, M. K., Mashor, M. Y., & Jaafar, H. (2010). Detection of mycobacterium tuberculosis in Ziehl-Neelsen stained tissue images using Zernike moments and hybrid multilayered perceptron network. *Conference Proceedings - IEEE*

International Conference on Systems, Man and Cybernetics, 4049–4055. https://doi.org/10.1109/ICSMC.2010.5642191

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Ozkan, H., & Kayhan, O. S. (2016). A Novel Automatic Rapid Diagnostic Test Reader Platform. *Computational and Mathematical Methods in Medicine*, *2016*. Retrieved from http://dx.doi.org/10.1155/2016/7498217

Posey, D. L., Marano, N., & Cetron, M. S. (2017). Cross-border solutions needed to address tuberculosis in migrating populations. *The International Journal of Tuberculosis and Lung Disease*, *21*(5), 485–486. https://doi.org/10.5588/ijtld.17.0187

Ren, & Malik. (2003). Learning a classification model for segmentation. In *Proceedings Ninth IEEE International Conference on Computer Vision* (pp. 10–17 vol.1). IEEE. https://doi.org/10.1109/ICCV.2003.1238308

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sergyan, S. (2008). Color Histogram Features Based Image Classification in Content-Based Image Retrieval Systems. In *2008 6th International Symposium on Applied Machine Intelligence and Informatics* (pp. 221–224). Herlany. https://doi.org/10.1109/SAMI.2008.4469170

Sicasys Software GmbH. (2017). Spotxel® Reader. Retrieved January 12, 2018, from https://play.google.com/store/apps/details?id=com.sicasys.spotxel&hl=en

Solmaz, M. E., Mutlu, A. Y., Alankus, G., Kılıç, V., Bayram, A., & Horzum, N. (2018). Quantifying colorimetric tests using a smartphone app based on machine learning classifiers. *Sensors and Actuators B: Chemical*, *255*, 1967–1973. https://doi.org/10.1016/J.SNB.2017.08.220

Tania, M. H., Lwin, K. T., Abuhassan, K., & Bakhori, N. M. (2017). An Automated Colourimetric Test by Computational Chromaticity Analysis: A Case Study of Tuberculosis Test. In *Advances in Intelligent Systems and Computing* (Vol. 616, pp. 313–320). Springer, Cham. https://doi.org/10.1007/978-3-319-60816-7

Tracey, B. H., Comina, G., Larson, S., Bravard, M., López, J. W., & Gilman, R. H. (2011). Cough detection algorithm for monitoring patient recovery from pulmonary tuberculosis. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, (day 0), 6017–6020. https://doi.org/10.1109/IEMBS.2011.6091487

Tsai, T.-T., Shen, S.-W., Cheng, C.-M., & Chen, C.-F. (2013). Paper-based tuberculosis diagnostic devices with colorimetric gold nanoparticles. *Science and Technology of Advanced Materials*, *14*(4), 44404. https://doi.org/10.1088/1468-6996/14/4/044404

UKVI. (n.d.). Tuberculosis tests for visa applicants. Retrieved September 24, 2017, from https://www.gov.uk/tb-test-visa

Wang, S., Xu, F., & Demirci, U. (2010). Advances in developing HIV-1 viral load assays for resource-limited settings. *Biotechnology Advances*, *28*(6), 770–781. https://doi.org/10.1016/j.biotechadv.2010.06.004

Wang, X.-Y., Wu, Z.-F., Chen, L., Zheng, H.-L., & Yang, H.-Y. (2016). Pixel classification based color image segmentation using quaternion exponent moments. *Neural Networks : The Official Journal of the International Neural Network Society*, *74*, 1–13. https://doi.org/10.1016/j.neunet.2015.10.012

Wug Oh,Seoung; Kim, S. J. (2017). Approaching the computational color constancy as a classification problem through deep learning. *Pattern Recognition*, *61*, 405–416. https://doi.org/10.1016/J.PATCOG.2016.08.013

Yetisen, A. K., Martinez-Hurtado, J. L., Garcia-Melendrez, A., da Cruz Vasconcellos, F., & Lowe, C. R. (2014). A smartphone algorithm with inter-phone repeatability for the analysis of colorimetric tests. *Sensors and Actuators B: Chemical*, *196*, 156–160. https://doi.org/10.1016/j.snb.2014.01.077