An Automated Algorithm for the Quantification of hCG Level in Novel Fabric-based Home Pregnancy Test Kits

Manasa K.[1], Manasa Priya K. V. S. N. L.[1], Sadhana Reddy S.[1],
Sumohana S. Channappayya[1*], Sivaramakrishna Vanjari[1], Dhananjaya Dendukuri[2**],
Swathy Sridharan[2], Tripurari Choudhary[2], Paridhi Bhandari[2]
[1]Dept. of Electrical Engineering, IIT Hyderabad, Yeddumailaram, AP - 502205
{EE12P1002, EE12M1020, EE12M1028, sumohana, svanjari}@iith.ac.in
[2]Achira Labs Pvt. Ltd., 57 1st Main Rd, Phase III, J. P. Nagar, Bangalore, Karnataka - 560078
{d.dendukuri, swathy.s, t.choudhary, paridhi.b}@achiralabs.com

***Abstract*— We report a new image processing algorithm that extracts quantitative information about the concentration of human chorionic gonadotropin (hCG), an important early pregnancy marker, from commercially available qualitative home pregnancy kits. The algorithm could potentially be ported onto a simple camera based cell phone making it a low-cost, portable point-of-care device as opposed to costly and time consuming clinical labs for accurate quantitative determination of hCG. The algorithm takes the image of the test result as input, classifies and determines the hCG concentration based on the RGB intensities of the test line. The efficacy of the algorithm is demonstrated using control samples on commercially available strips as well as novel fabric based strips designed for this application.**

I. INTRODUCTION

Home pregnancy test kits have become ubiquitous and are used as the default tool for early pregnancy detection [1]. Typically, the test is performed on urine and is qualitative in nature. The presence of hCG is determined by lateral flow immunoassays that contain antibodies specific to the β subunit of hCG. Though a qualitative test is sufficient to confirm pregnancy, quantification of hCG is necessary to predict different anomalies in pregnancy. The proposed algorithm relies on the acquisition and processing of a digital image of the kit's output to quantify the concentration of hCG. The ultimate goal of this work is to reduce cost by using cheap processing power available on mobile devices as opposed to relatively expensive Enzyme linked immunosorbent assay (ELISA) plate readers or autoanalyzers in identifying the concentration of β-hCG. Further, the algorithm was developed using data collected from kits based on a novel low-cost fabric designed at Achira Labs [3].

II. PROPOSED ALGORITHM

A. Image Acquisition

Sandwich immunoassay for β-hCG test was carried out on pilot batches of fabric chip with positive and negative spiked urine control samples. Analyte (β-hCG) concentrations in the range from 25 mIU/ml to 1000 mIU/ml were spiked into human male urine samples. Test was performed by adding 3 drops (75ul) of reactive sample onto sample pad located upstream of the detection antibody zone of either the fabric strip or a commercial strip (Cipla, an Indian manufacturer). The result of this test (shown in Fig. 1) is captured by the antibodies immobilized in the test and control line producing two distinct red lines. Images of these tested strips were captured under single bright white light lamp source from a distance of 15 cm by Nikon D5100 digital SLR camera fixed at tripod stand.

B. Problem Formulation

The problem can be broken down into two sub-problems: a) identifying control and test lines and b) classification of test line into one of four hCG concentration levels based on color while accounting for intra-class color variability. The four concentrations levels chosen in this work (0, 25, 100, 250 mIU/ml) are representative of the possible concentrations in a real world setting.

C. Solution

The proposed algorithm is summarized in Fig. 2. The algorithm segments the image so as to find the lines of interest by binarization [4] using Otsu's algorithm [5] for the threshold. The resulting noisy binary image is cleaned up using simple majority logic. To locate the control and test lines it is assumed that the control line is always to the right of the image. A set of k contiguous columns are declared a line if k exceeds 10 columns – an empirically determined value. The RGB values of the control line (Fig. 4) are used as feature points to classify it into one of the four chosen hCG concentration levels. Intra-class color variability was handled by training the classifier using random sampling of data within a class. A polynomial kernel SVM [6] classifier gave the best results.

D. Results

The results of image segmentation and control line identification is shown in Fig. 3. The classification results are summarized in confusion matrices shown in Fig. 5. The classification results are consistently higher than 80% for

*corresponding author.
**DD gratefully acknowledges support through grant no. 0002-02-02-01-01 from Grand Challenges Canada.

both the novel fabric based kits and commercial kits, thereby demonstrating not only its robustness but also an immediate value addition to these kits.

III. Conclusion

A novel and robust algorithm for the classification of the β-hCG concentration is proposed and successfully tested on both a novel fabric kit and a commercial test kit. The proposed algorithm provides a proof-of-concept for converting the qualitative process of early pregnancy detection using self test kits into a semi-quantitative one. Ongoing efforts include the exploration of color features that are more robust to ambient luminance, data normalization, and increasing the number of classes in the classifier to improve its sensitivity. Further, to achieve our bigger goal, we are working on porting the algorithm to a mobile platform.

Fig. 1: Sample image at hCG concentration of 100 mIU/ml.

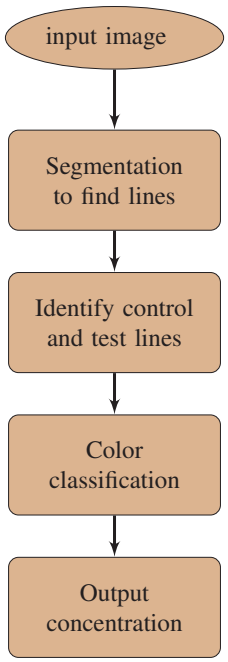
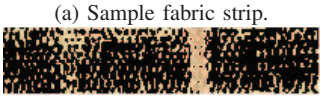


Fig. 2: Flowchart of proposed algorithm.

(a) Sample fabric strip.



(b) Segmented lines before denoising.

(c) Denoised lines.

Fig. 3: Segmentation output for fabric-based kit.

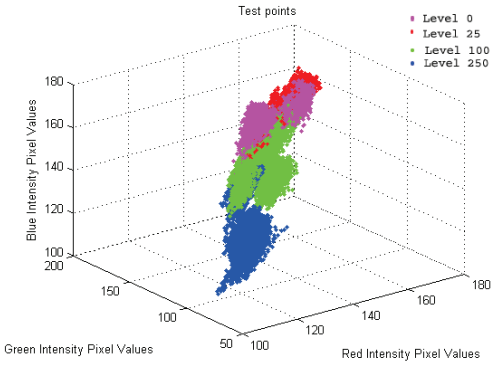


Fig. 4: Test points for the classifier.

References

[1] Womenshealth.gov, "Pregnancy tests." "http://womenshealth.gov/publications/our-publications/fact-sheet/pregnancy-tests.pdf", March 2009.

[2] O. Mudanyali, S. Padmanabhan, S. Dimitrov, I. Navruz, U. Sikora, and A. Ozcan, "Smart rapid diagnostics test reader running on a cell-phone for real-time mapping of epidemics," in *Proceedings of the Second ACM Workshop on Mobile Systems, Applications, and Services for HealthCare*, p. 1, ACM, 2012.

[3] P. Bhandari, T. Narahari, and D. Dendukuri, "fab-chips: a versatile, fabric-based platform for low-cost, rapid and multiplexed diagnostics," *Lab on a Chip*, vol. 11, no. 15, pp. 2493–2499, 2011.

[4] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Prentice Hall, 2008.

[5] N. Otsu, "A threshold selection method from gray-level histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, pp. 62 – 66, jan 1979.

[6] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[7] H. William, A. Teukolsky Saul, T. Vetterling William, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing (3rd ed.)*. New York: Cambridge University Press, 2007.

[8] S. Cheong, S. Oh, and S. Lee, "Support vector machines with binary tree architecture for multi-class classification," *Neural Information Processing-Letters and Reviews*, vol. 2, no. 3, pp. 47–51, 2004.

	Output			
Input	0	25	100	250
0	81.26	18.74	0	0
25	11.29	88.71	0	0
100	0	0	90.60	9.40
250	0	0	1.57	98.43

(a) Achira fabric.

	Output			
Input	0	25	100	250
0	100	0	0	0
25	0	100	0	0
100	0	0	80.92	19.08
250	0	0	4.80	95.20

(b) Commercial strip.

Fig. 5: Confusion matrices of classification results.