DETECTION OF MALARIA PARASITE IN GIEMSA BLOOD SAMPLE USING IMAGE PROCESSING

Kishor Roy, Shayla Sharmin, Rahma Bintey Mufiz Mukta, Anik Sen

Department of Computer Science & Engineering, CUET, Chittagong, Bangladesh

ABSTRACT

Malaria is one of the deadliest diseases ever exists in this planet. Automated evaluation process can notably decrease the time needed for diagnosis of the disease. This will result in early onset of treatment saving many lives. As it poses a serious global health problem, we approached to develop a model to detect malaria parasite accurately from giemsa blood sample with the hope of reducing death rate because of malaria. In this work, we developed a model by using color based pixel discrimination technique and Segmentation operation to identify malaria parasites from thin smear blood images. Various segmentation techniques like watershed segmentation, HSV segmentation have been used in this method to decrease the false result in the area of malaria detection. We believe that, our malaria parasite detection method will be helpful wherever it is difficult to find the expert in microscopic analysis of blood report and also limits the human error while detecting the presence of parasites in the blood sample.

KEYWORDS

Malaria, HSV segmentation, Watershed segmentation, Giemsa blood sample, RBC.

1. INTRODUCTION

Malaria is one of the severe diseases caused by the protozoan parasites of the genus Plasmodium, transmitted via female Anopheles mosquito. During the process of complex life cycle of parasites in growing and reproducing inside the human body, the red blood cell (RBCs) are used as hosts and destroyed afterwards. World Health Organization estimates that in 2015 mentioning in their website http://www.who.int/en/, 212 million clinical cases of malaria occurred, and 429,000 people died of malaria, most of them children in Africa. Also as malaria causes so much illness and death, the disease hampers on many national economies and WHO also discovers that many countries with malaria are already among the poorer nations it is difficult for them to break the a vicious cycle of disease and poverty. Normally malaria happened because of four types of plasmodium species called *Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, Plasmodium malariae.* Among all of this *Plasmodium falciparum* is responsible for malaria fever in most of the cases.
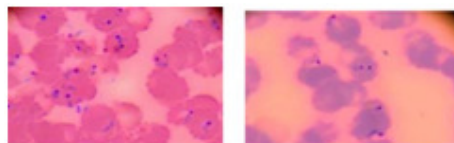

Figure 1: Example of Malaria Parasite in Blood sample

After collecting blood samples, the diagnosis of malaria infection is done by searching for parasites in blood slides through a microscope by experts most of the cases by a pathologist. Figure 1 shows blood sample with the presence of malaria parasites in blood cell. Recognition and detection of parasite in blood sample can be possible by applying a chemical process called (Giemsa) staining. This process slightly colorizes the red blood cell (RBC) and plasmodium parasites. The detection of the Plasmodium parasites requires detection of the stained objects. To prevent the false result, this stained objects need to analysed further to determine whether they are parasite or not. Figure 2 illustrates four species of malaria (*P. falciparum, P. vivax, P. ovale, and P. malariae*) that can cause human infected transmitted via female Anopheles mosquitos.
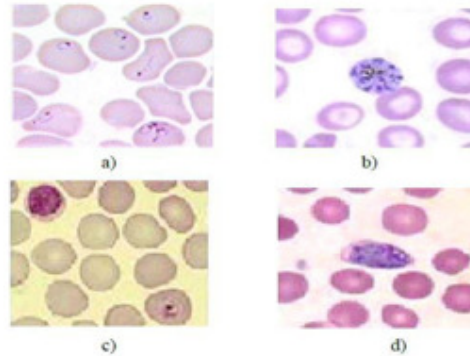

Figure 2: a) *P. falciparum* b) *P. vivax* c) *P. malariae* d) *P. ovale*

It is clear that to detect parasite from giemsa blood sample trained and experienced technicians or pathologists are needed. By digitalizing the approach will reduce the time taken for screening the disease as well as will improve the consistency in diagnosis. This study investigates the use and application of digital image processing for detecting malaria parasites using microscopic colour images from giemsa blood sample also an efficient method is proposed for parasite detection based on colour based pixel discrimination technique and Segmentation (watershed segmentation) operation. The goal of this paper is to propose a fully automated image classification system to positively identify malaria parasites present in thin blood smears. In this paper, we describe an unsupervised approach in which colour and segmentation based algorithms are put together to formulate an algorithm for Plasmodium parasite detection from thin smear slide. This approach has a comparative higher predictability and lower false positive rate. The rest of this paper is organized as follows. Section II presents a discussion of related studies, Section III describes the proposed system methodology, Section IV presents results and related discussions where section V analyse the testing, and finally Section VI concludes the paper.

2. LITERATURE REVIEW

In 2010, Frean [1] build a morphologic based analysis system to count parasites from individual microscopic images. But there was some inconsistency of detecting the parasite which is in primary state. In 2011, Somasekar [2] proposed a linear programming based Image segmentation and morphological operations for detection of malarial parasites. But there was no complete system for detecting different parasites. In 2011, Prescott WR, Jordan [3] proposed Shape and colour pattern of parasite and RBC (Red Blood Cells) for Detection of malaria parasite from thick and thin smear image. But there some difference of colour pattern between the resultant value and WHT value. In 2011, Edison M, Jeeva J, Singh M [4] proposed Image filtering and edge detection to analyze the changes of *Plasmodium vivax* in thin smear erythrocytes images. In 2013, Pallavi T. Suradkar [5] proposed flood fill algorithm and the colour range of RBC and parasite to

detect parasite from the blood sample. But there was no system for detecting the *P. falciparum* parasite.

3. METHODOLOGY

The proposed system model is implemented by using two segmentation that are HSV segmentation and watershed segmentation. At first we applied two method separately and we experienced that HSV segmentation failed to detect some parasite on the other hand watershed can detect those but in some cases watershed failed to do so but HSV segmentation detect those parasite easily. To avoid this problem we map the result of both segmentations so that we can count almost all malaria parasites in a blood cell which makes our desired output more accurate. Figure 3 depicts the whole system architecture of the model.
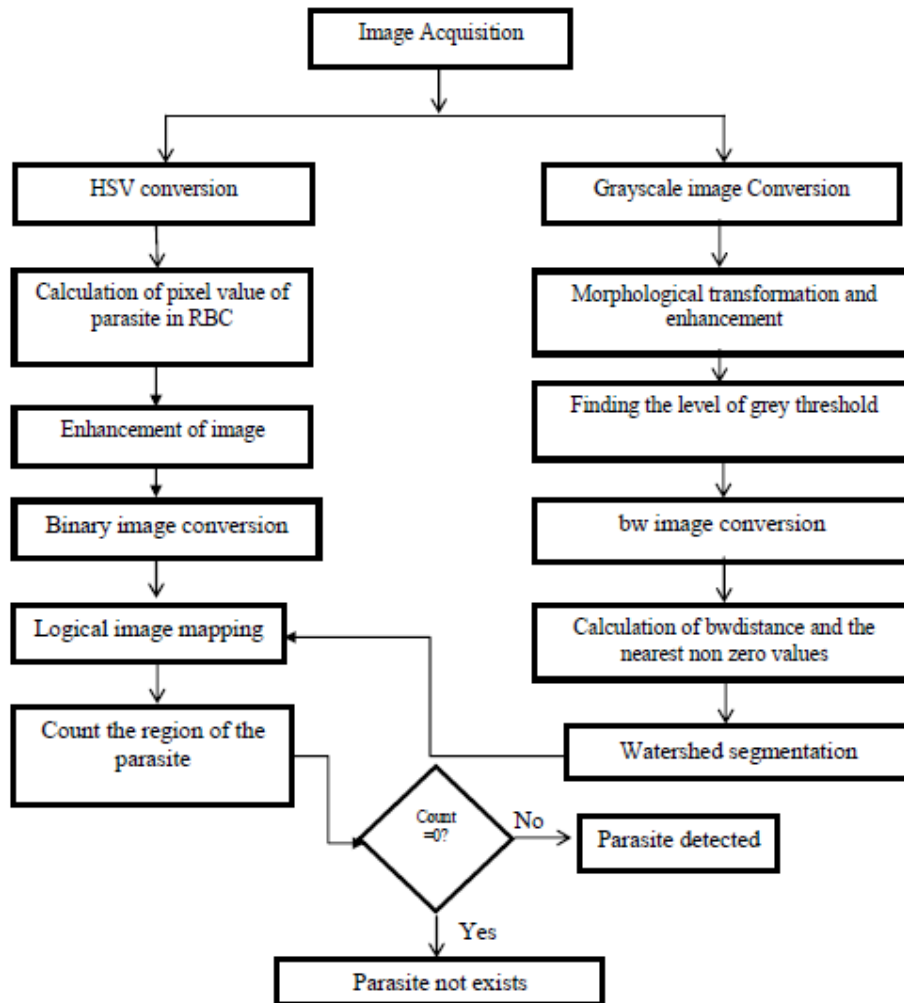

Figure 3: System overview

3.1 Image Database

For making the image database we have collected two categories of sample images. Samples are collected both from the internet and diagnostic centre. There were 200 images from the internet of

affected and clean blood cell and 100 images from the diagnostic centre of plasmodium affected blood cell which we used for testing in our experiment.

3.2 Image Acquisition

To collect the image blood samples we collected images from diagnostic centre and internet. The collected images are acquired from a USB microscope which is connected to a computer via a USB port. It is a low-powered digital microscope and is widely available at low cost for use at diagnostic centre and hospital. The range of the cost is varies from hundred to thousand. Normally, USB microscopes are a webcam having a high-powered macro lens. It generally does not use transmitted light, but it can rely on incident light from in-built LEDs light where these LEDs light situated next to the lens and these lights reflected from the sample following reflection it enters to the camera lens. As these cameras are highly sensitive they do not need additional lighting. The camera attaches directly to the USB port of a computer for this reason the images are shown directly on the computer monitor. Figure 4 shows an USB microscope collected from the Wikipedia.

Figure 4: USB microscope

3.3 Image Segmentation

For detecting the parasite more accurately we have mapped two segmentation procedures which consist of HSV Segmentation and Watershed segmentation. The procedure is described below:

3.3.1 HSV Segmentation

HSV stands for Hue, Saturation and Value. By this segmentation calculate the parasite hue, saturation and value component and by this value, the parasite position in RBC can be identified. This procedure is divided into several steps which are described below.

(i) Converting the image into HSV format and calculate the indices of the parasite:

First we convert the image into HSV format. After converting it into HSV format the hue, saturation and the value are calculated. For the parasite hue components lies greater than 0.3 to less than 0.9 and the value component lies under 0.8.We can calculate this value by the RGB value of the parasite and use the blue pixel from that. We analyse 150parasite images and found the hue component value 0.3<hue<0.9. As the max value will be the value component so we take the value component less than 0.8. The hue, saturation and value images are presented in figure 5.
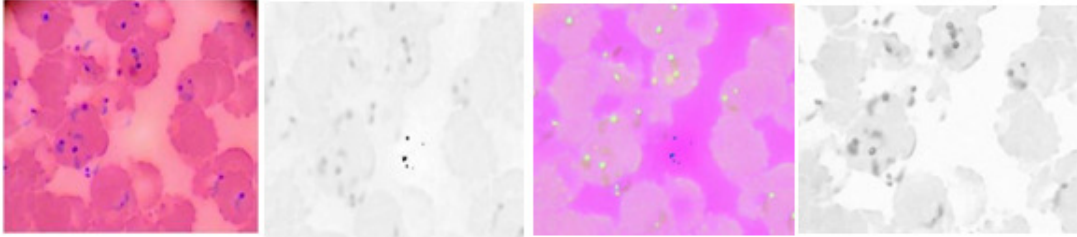
Figure 5: a) RGB image b) Hued image c) Saturation Image d) Value image

(ii) Enhancing the Image

Following the previous step, the system multiply the values of the pixel indices of the parasite by 9.5 and that gives an enhanced image which is used for the next step of the procedure. We checked for various values but in 9.5 maximum number of parasite were visible in the enhanced image. The sample of enhanced image is shown in figure 6.
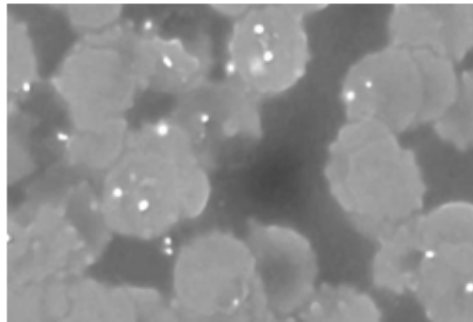

Figure 6: Enhanced Image

(iii) Binary image conversion

Now this enhanced image is converted into binary image. As for the result the parasite position in the RBC are identified. The binary image is shown in figure 7. Here, we can see that our region of interest that means the parasite are labelled as white which will be more helpful for calculation of the next step.

Figure 7: Binary Image

3.3.2 Watershed Segmentation

By HSV all the parasites cannot be detected because of some colour distortion of the input images. So we apply another segmentation procedure called watershed segmentation for the

detection of parasite which is difficult to detect from previous segmentation. Before applying this segmentation some pre-processing steps are applied which are described below with the explanation of watershed segmentation.

(i) Grayscale conversion

After converting the RGB image into grayscale image the contrast of the grayscale image is enhanced using local histogram equalization to enhance the visibility of the parasite in the image. The grayscale image is given in figure 8.
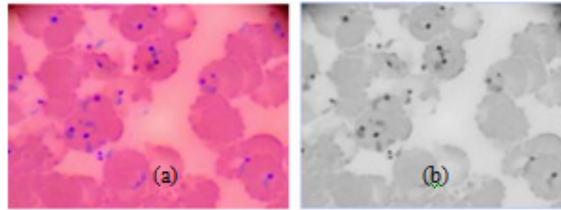

Figure 8: a) RGB image b) Grayscale image

(ii) Morphological Transformation

After converting into grayscale image we apply a morphological transformation on the grayscale image and for that we use the structuring element of disk of size 50. The purpose of morphological transformation is to remove the noise from the image which is lesser than the structuring element. The morphological transformation is given in the figure 9.
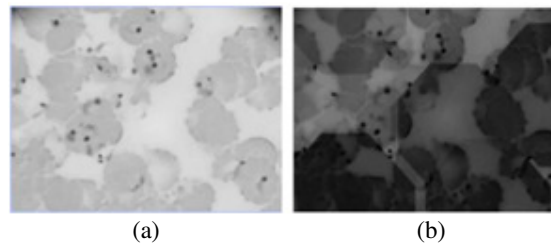

(a)   (b)

Figure 9: a) Grayscale image b) Morphological operation

(iii) Adjustment of the Image

After morphological transformation adjustment of the morphological image is done in this step. As a result the structuring elements were enhanced. The Adjustment of the image is given in figure 10.
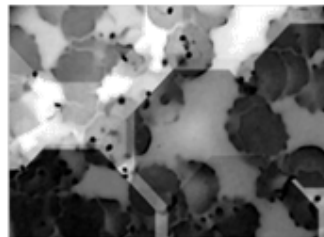

Figure 10: Adjustment of the image

(iv) Converting into black & white image:

After finding the level of gray threshold and by using this gray threshold value the system convert it into black and white image. The black & white image is given in figure 11.

Figure 11: Black & White image

(v) Finding the bwdistance and nearest nonzero values

In this step the system computes the bwdistance of the image which stands for zero to non-zero distance so that it can be helpful to find out the nearest non-zero values where the distance transform of a binary image is the distance from every pixel to the nearest nonzero-valued pixel [9].

Table 1: Binary Image and its distance transformation

1	1	0	0	0		0.00	0.00	1.00	2.00	3.00
1	1	0	0	0		0.00	0.00	1.00	2.00	3.00
0	0	0	0	0		1.00	1.00	1.41	2.00	2.24
0	0	0	0	0		1.41	1.00	1.00	1.00	1.41
0	1	1	1	0		1.00	0.00	0.00	0.00	1.00

The extended maxima pixels are only local minima in the image. These local minima in the image were the parasites in the RBC.

(vi) Watershed Segmentation

The term watershed refers to a ridge that divides areas drained by different river systems [10]. A catchment basin is the geographical area draining into a river or reservoir [10]. By considering a topographic surface, water would collect in one of the two catchment basins [10]. Water falling on the watershed ridge line separating the two basins would be equally likely to collect into either of the two catchment basins [10]. Watershed algorithm finds the catchment basins and the ridge lines in an image.

From [10], let $f \in C(D)$ have minima $\{m_k\}_{k \in I}$, for some index set I. The catchment basin $CB(m_i)$ of a minimum m_i is defined as the set of points $x \in D$ which are topographically closer to m_i than to any other regional minimum m_j :

$$CB(m_i) = \{x \in D \mid \forall j \in I\backslash\{i\} : f(m_i) + T_f(x, m_i) < f(m_j) + T_f(x, m_j)\}\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Let W be some label, W $6\in$ I. The watershed transform of is a mapping

$$\lambda : D \rightarrow I \cup \{W\}, \text{ such that } \lambda(p) = i \text{ if } p \in CB(m_i), \text{ and } \lambda(p) = W \text{ if } p \in Wshed(f)\ldots\ldots\ldots(2)$$

61

So the watershed transforms of assigns labels to the points of D, such that different catchment basins are uniquely labelled, and a special label W is assigned to all points of the watershed [10].

Then watershed segmentation is applied so that maximum number of parasites can be identified. The watershed segmentation image is given in figure 12.
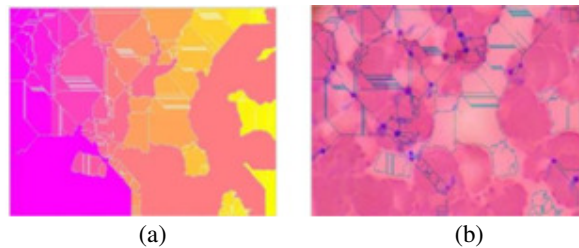

| (a) | (b) |

Figure 12: a) watershed segmentation b) Parasite detection

3.4 Parasite Detection

To detect the parasite the system converts HSV and watershed segmented image into logical image. Then measure the white property of the logical image. These white properties mean that in the provided image parasites do exist. If the property is not zero then there will be parasite detected and if the property is zero that means the cell is clean and is not affected. These images is given in figure 13

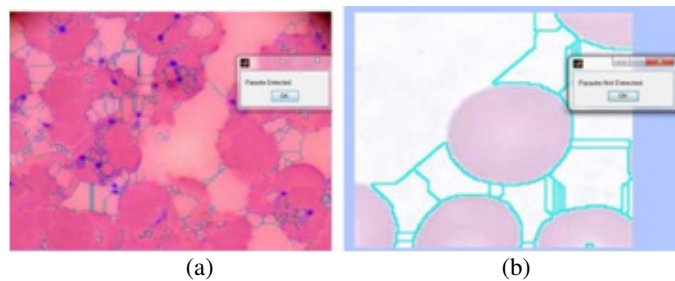| (a) | (b) |

Figure 13: a) Parasite Detected b) Parasite not exists

4. RESULT ANALYSIS

For our system we have used both the real sample collected from the diagnostic centre and sample that are obtained from the internet. For HSV segmentation and watershed segmentation we used 100 real sample and 200 internet based sample were used. Therefore, by watershed segmentation process from 85 real samples among of 100, parasites are detected and from 164 photos of 200 from internet parasite are detected successfully. By HSV segmentation process from 70 real samples among of 100, parasite are detected and from 160 photos of 200 from internet parasite are detected successfully. But when we mapped both segmented images before counting the parasite the number of correctly detected parasite increased. The system can detect all parasites when the image contains intensity of light at a standard and equal value. But as the input images are different in intensity of light among the parasites, which is the reason why our system failed to detect parasites from some affected image.

Table 2: Result Analysis

		Samples from	
		Diagnostic Centre	Internet
	No. of Test images	100	200
Successful Recognition	Watershed segmentation	85	164
	HSV segmentation	70	160
	Mapping logical image	90	164
Accuracy	Watershed segmentation	85%	82%
	HSV segmentation	70%	80%
	Mapping logical Image	90%	82%

5. TESTING

The technique of testing without having any knowledge of the interior workings of the application is Black Box testing where the tester is oblivious to the system architecture and does not have access to the source code. Typically, when performing a black box test, a tester will interact with the system's user interface by providing inputs and examining outputs without knowing how and where the inputs are worked upon.

Table 3: Test cases of Black box testing

		Samples from	
		Diagnostic Centre	Internet
	No. of Test images	100	200
Successful Recognition	Watershed segmentation	85	164
	HSV segmentation	70	160
	Mapping logical image	90	164
Accuracy	Watershed segmentation	85%	82%
	HSV segmentation	70%	80%
	Mapping logical Image	90%	82%

6. CONCLUSION

Since the detection of malaria parasites is done by pathologists manually using microscopes, the chances of false detection due to human error are high, which in turn can result into fatal condition. Considering this fact in mind we proposed a method which curbs the human error while detecting the presence of malaria parasites in the blood sample by using image processing. We achieved this goal using image segmentation smoothing processing techniques to detect malaria parasites in images acquired from giemsa stained peripheral blood samples. The real samples collected from diagnostic centre provide better result than collected images from internet because the images from diagnostic centre contain similar color intensity and other image quality where internet images vary in quality. For future work, classification process can be introduced for more accurate and automated detection of malaria parasite. And we also hope that if we want

to apply image processing in any other detection in blood sample this study will help a lot as a pioneer.

REFERENCES

[1] Frean J,(2010) "Microscopic determination of malaria parasite load: role of image analysis". Micrsocopy: Science, Technology, Applications, and Education 862-866.

[2] Somasekar J, Reddy B, Reddy E, Lai C, (2011) "Computer vision for malaria parasite classification in erythrocytes", International Journal on Computer Science and Engineering 3: 2251-2256.

[3] Prescott WR, Jordan RG, Grobusch MP, Chinchilli VM, Kleinschmidt I, et al. (2012) Performance of a malaria microscopy image analysis slide reading device. Malar J 11: 155.

[4] Edison M, Jeeva J, Singh M, (2011) "Digital analysis of changes by Plasmodium vivax malaria in erythrocytes", Indian Journal of Experimental Biology 49: 11-15.

[5] Pallavi T. Suradkar "Detection of Malarial Parasite in Blood Using Image Processing", International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 10, April 2013.

[6] Deepali A. Ghate, Prof. Chaya Jadhav "Automatic Detection of Malaria Parasite from Blood Images", May, 2012.

[7] F. B. Tek, A. G. Dempster, and I. Kale, "Malaria parasite detection in peripheral blood images," in Proc. British Machine Vision Conference, Edinburgh, September 2006.

[8] Varsha Waghmare, Syed Akhter ,"Image analysis based system for automatic detection of malarial parasite in blood images", International Journal of Science & Research(IJSR),ISSN(Online):2319-7064, July, 2015.

[9] P. Pratim Acharjya and M.Santiniketan, ," Watershed Segmentation based on Distance Transform and Edge Detection Techniques", International Journal of Computer Applications (0975 – 8887) Volume 52– No.13, August 2012

[10] Jos B.T.M. Roerdink , Arnold Meijster, "The Watershed Transform: Definitions, Algorithms and Parallelization Strategies", Institute for Mathematics and Computing Science, University of Groningen, The Netherlands, Fundamenta Informaticae 41 (2001) 187–228 1 IOS Press.

AUTHORS

Kishor Roy received his B.Sc degree in Computer Science & Engineering in the year 2017, from Chittagong University of Engineering & Technology. His interested areas of working are machine learning, image processing, data mining, artificial intelligence, computer vision and IOT

Shayla Sharmin completed her B.Sc. Engineering in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Bangladesh in 2014 and currently pursuing her M.Sc. Engineering from the same department. She is also a Lecturer in the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh. Her research interest includes image Processing and human robot/ computer interaction

Rahma Bintey Mufiz Mukta received her B.Sc. Engineering and M.Sc. Engineering in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Bangladesh in 2013 and 2017 respectively. She is currently working as an Assistant Professor in the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh. Her research interest includes privacy preserving data mining, multilingual data management and machine learning.

Anik Sen received his B.Sc degree in Computer Science & Engineering in the year 2014 from Chittagong University of Engineering & Technology. He started his professional career as a web developer. He currently owns his own software company and pursuing M.Sc degree in Computer Science & Engineering from Chittagong University of Engineering & Technology. His interested areas are machine-learning, computer vision, data mining and advanced database management system management and machine learning.