

**Intelligent Detection System for diagnosis of Systemic Lupus Erythematosus
(SLE)**

By

SU SING SIONG

FINAL PROJECT REPORT

Submitted to the Electrical & Electronics Engineering Programme
in Partial Fulfillment of the Requirements
for the Degree
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

Universiti Teknologi Petronas
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

© Copyright 2014

by

Su Sing Siong, 2014

CERTIFICATION OF APPROVAL

Intelligent Detection System for diagnosis of Systemic Lupus Erythematosus (SLE)

by

Su Sing Siong

A project dissertation submitted to the
Electrical & Electronics Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirement for the
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

Approved:

Ms Zazilah Bt. May
Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK

2014

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

Su Sing Siong

ABSTRACT

SLE is a disease that is incurable. However, due to the limited technology, some patients has missed the treatment as they aren't been tested to have the positive test having lupus until the disease is uncontrollable and spread through their body and causing death. Therefore, in this paper, we proposed an intelligent detection system for the SLE to have the more accurate diagnostic solution for the patients. In this intelligent detection system, MATLAB is approached and the 4 stages, pre-processing ANA image for the Human Epithelial type 2 (HEP-2) cells, segmentation, feature extraction, and classification are being done to have the accurate detection of the SLE disease. After reading some literature reviews for the research done previously by a lot of researchers, especially Paola Soda, the most appropriate techniques has been applied for this intelligent detection system of SLE. In conclusion, the intelligent detection of SLE by MATLAB approach is going to be a simple and effective ways to help in the medical field for the more accurate results in diagnosis of lupus.

ACKNOWLEDGEMENTS

First and foremost I, Su Sing Siong, matric ID 13925, from Electrical and Electronic Engineering Department would like to thank Universiti Teknologi PETRONAS (UTP) for enabling me to have this precious chance to do a project regarding image processing. My gratitude goes to all the people involved in helping and guiding me in the image processing techniques and studies and leading me to this success.

I would like to extend immense gratitude to my supervisor, Ms Zazilah Bt. May who has been patient enough to guide me. The tips and guidance from her are helpful to help me in this project. Encouragements given from my supervisor are giving me enough courage to move on when I had failed on some path.

Besides that, I would like to thank MIVIA who is willing to share their findings of the HEP-2 cells images to have me proceed with this project.

Because of this project, I was given massive exposures to the MATLAB image processing toolbox and some healthy issues. I learnt how images are achieved and learnt about an image's properties.

Last but not least, I would like to thank Universiti Teknologi PETRONAS (UTP)'s FYP coordinators for enabling all these. I have learned so much from the lecturers and friends in UTP. This allows me to apply the knowledge or information that I have learned to apply on the FYP. Other than that, special thanks go to my sponsor JPA for sponsoring all the cost of my studies.

TABLE OF CONTENTS

| | |
|---|-------|
| CHAPTER 1: INTRODUCTION..... | 1 |
| 1.1 Background study..... | 1-3 |
| 1.2 Problem Statement..... | 3 |
| 1.3 Objective..... | 3 |
| 1.4 Scope of Study..... | 4 |
| CHAPTER 2: LITERATURE REVIEW..... | 5-10 |
| CHAPTER 3: METHODOLOGY..... | 11 |
| 3.1 Research Methodology..... | 11 |
| 3.2 Progress and Activities..... | 12 |
| 3.2.1 Image Acquisition..... | 13 |
| 3.2.2 Image Pre-processing..... | 13-14 |
| 3.2.3 Image Segmentation..... | 14-16 |
| 3.2.4 Feature Extraction..... | 17 |
| 3.2.5 Classification..... | 18 |
| 3.2.6 Key Milestone and Gantt Chart..... | 18-19 |
| 3.3 Software and toolbox..... | 19 |
| CHAPTER 4: RESULTS AND DISCUSSION..... | 20 |
| 4.1 Feature Extraction results..... | 20-21 |
| 4.2 SVM classifier results..... | 22 |
| CHAPTER 5: CONCLUSION AND RECOMMENDATION..... | 23 |
| 5.1 Conclusion..... | 23 |
| 5.2 Recommendation for future work..... | 24 |
| REFERENCES..... | 25-26 |
| APPENDIX A..... | 27-29 |
| APPENDIX B..... | 30 |
| APPENDIX C..... | 31 |
| APPENDIX D..... | 32 |
| APPENDIX E..... | 33-35 |

LIST OF TABLES

| | |
|--|-----|
| Table 1: technique of noise reduction and enhancement of images..... | 6-7 |
| Table 2: technique of segmentation | 9 |

LIST OF FIGURES

| | |
|---|----|
| Figure 1 : 1) Homogeneous, 2) Fine Speckled, 3) Course speckled 4) Nucleolar, 5) cytoplasmatic, and 6) centromere (from left to right)..... | 2 |
| Figure 2: a) original image, b) after convert to gray scale, c) enhanced image. | 14 |
| Figure 3: segmentation of image..... | 14 |
| Figure 4: noise reduction..... | 15 |
| Figure 5: Result of GLCM..... | 17 |
| Figure 6: Graph of contrast from GLCM..... | 20 |
| Figure 7: Graph of correlation from GLCM..... | 20 |
| Figure 8: Graph of energy from GLCM..... | 21 |
| Figure 9: Graph of homogeneity from GLCM..... | 21 |
| Figure 10: Results of Slassifier (fine-speckled) | 22 |
| Figure 11: Results Classifier (centromeric) | 22 |

CHAPTER 1

1. Introduction

1.1. Background Study

One of the autoimmune diseases, Systemic Lupus Erythematosus (SLE), is a well-known disease that leads to chronic inflammation which often harms the heart, lungs, joints, blood vessels, skin, kidneys and neural system. It happens when the immune system of our body is abnormal where it creates inflammation and destruction on our own tissues and organs. [1]

Statistical study of lupus in US is able to show how serious this disease is. From [2], in 2005, it was recorded that there are 1.5 million Americans is suspected to have lupus. 161000 Americans having lupus and women having a higher risk compared to men. Furthermore, there are 77000 annual hospitalization related to SLE from 1988 to 2000. Although, this research does not show the worldwide statistical status about lupus, we can estimate that, the numbers are beyond the numbers that is been collected.

Normally, in our immune system, macrophages (or phagocytes) are the antibodies that attack the foreign invader at the first place. However, sometime they do not want to destroy something that is one part of the body which they cannot recognize. Here, there is the T-helper cells come from lymphocytes and aid to help. If the T-helper cells can determine that the particle is a substance that is recognized to be harmful to our body, the hormones named cytokines will be secreted and will generate the inflammatory reaction. This stimulates the B-cells into action with enzymes that destroy by creating oxidative stress. At the same time, some B-cells will return to the lymph nodes to prepare the antibodies against the foreign invader so as to ready for the next invasion. After the foreign invader has been destroyed, T-suppressor cells will come and calm down the immune response. If the reactive response goes unchecked, the inflammatory response can have possibility gets raged, and hence, can cause the great injury of inflammation. [3]

In order to study about the antibodies in the immune system, immunoassays are needed to be prepared and analyzed. There are several ways in preparing the immunoassays, such as ELISA (Enzyme-linked Immunosorbent assay) and ELONA (Enzyme-linked Oligonucleotide Assay). By establish and maintaining immunoassay capability, we need to validate the assay, prepare sample, obtain matrix effects, get the data processing, maintain the quality assurance, and troubleshooting the method development. Besides, the safety considerations and waste disposal should be addressed along the preparation of the assays.[4] Based on[5] , the ELISA is prepared for the study of the antibody cells.

Anti-nuclear Autoantibody (ANA) is the significant antibody that diagnosis the happening of SLE.[6] Normally, image analysis from the HEP-2 cells allows the identification of ANA.[5] According to[6], there are 97% of the SLE getting a positive test of ANA. To get an ANA test, data mining and image analysis on the HEP-2 cells is required with the technique of immunofluorescence. [5] Regarding the classification of the HEP-2 cells and ANA test to determine the happening of SLE, there are still several techniques and methods needed to be further explored to get the perfect diagnosis and detection of SLE.

By referring to [7], there have been having accuracy of 86.5% by evaluate on an marked dataset of mitotic cells with technique of indirect immunofluorescence (Iif) imaging and aid of computer-aided-diagnosis (CAD). The research and experiment done is a good start of the classification of the HEP-2 cells' pattern in the ANA. However, this early stage of performance cannot be able to handle on larger set of data. From [7], we are able to know that, there are several staining patterns of HEP-2 cells, homogenous, speckled, nucleolar, peripheral nuclear and centromere.

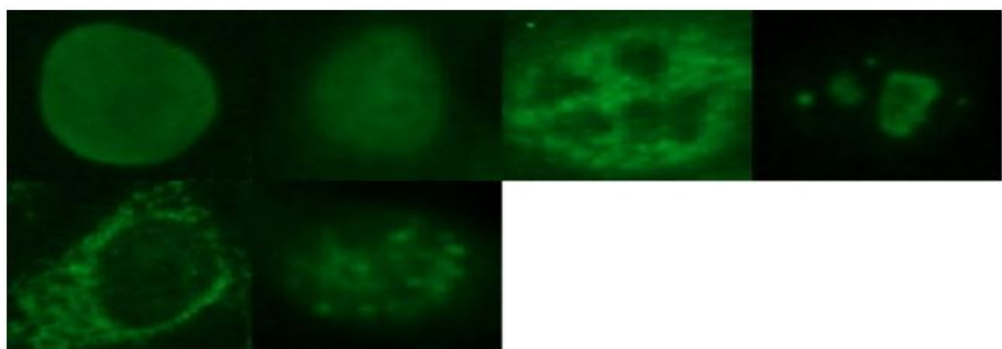


Figure 1: 1) Homogeneous, 2) Fine Speckled, 3) Course speckled 4) Nucleolar, 5) cytoplasmatic, and 6) centromere (from left to right).

Aside from that, according and thanks to Paola Soda, in [8] and [9], the difficulties with the problem of imbalanced dataset have been done with some efforts for the improvement and development on the classification technique. Hence, the imbalanced data can now be handled.

1.2 Problem Statement

SLE, resulting systemic inflammatory disorder that alike as Rheumatoid Arthritis (RA) is an autoimmune disease and incurable.[1] There are currently diagnostic methods by laboratory test. However the diagnostic, detection and classification methods for the SLE are very subjective. Most of them are handled manually, and not standardized. They are very dependable on the high technology and experience of professors dealing with the autoimmune diseases. Without the high technology and experience, that may lead to decreased accuracy and misinterpreting due to the data recording having a loss of information. Furthermore, this may lead to no treatment to a SLE patient.

1.3 Objectives

The main objective of this project is to develop an intelligent system for the diagnosis of SLE from HEP-2 cells' images in the ANA. In order to achieve this there are several sub-objectives needed to be achieved.

1. To classify SLE more accurately from HEP-2 image by image enhancement and edge detection.
2. To perform segmentation on the HEP-2 cells
3. To extract the HEP-2 cells for classification
4. To classify the HEP-2 cells and certify the accuracy of the classification results.

1.4 Scope of study

To achieve the objectives, the scope of study has to be narrow down to get a greater result in a faster time period. The dataset of MIVIA HEP-2 images are studied. The focus point on this project will be at the system part as while the dataset of the image can be calculated, and the images can be identified into several patterns which determining whether the patient is having the SLE disease or not. The MATLAB application, image processing toolbox, in image processing, segmentation, extraction and classification has to be learned and understood. The fine-speckled result from the image is needed to be focused.

CHAPTER 2:

2.1 Literature review

Throughout the study in the research papers and journal, there are multiple ways in helping to determine the SLE in patients, which is analyzing the HEP-2 cells. The methods and testing skills in bio analytical technique are being reviewed and studied throughout the literature review. The performances of the techniques in several medical areas are being deliberated with the results studied.

Critical Analysis

1) Image processing

- Gray scale approach

This approach is widely used as it provides us simplicity. It is simple and useful, as it can reduce the data of an image. We will only process $\frac{1}{4}$ of the RGBA image when we converted it into gray scale. Hence, it allows a large color image's data being stored in MATLAB and shown. However, this approach can have us lost in some information on color.

- Cellular Neural Network (CNN) combined with coupled Non-linear Oscillators

Based on [10], CNN is a technique that emulate the histogram equalization by combining different mapping function for piecewise linear approximation. With the addition of coupled Non-linear oscillators, the processing efficiency can be increased by the coefficient of the non-linear differential equations and hence, the image quality can be enhanced. However, this technique is highly depending on the formulation of the non-linear differential equation.

- Homomorphic Filter model

Reviewing [11] and [12], adaptive filters as high pass and low pass filter is used to enhance the image. The function of the filter is to sharpen the image and prevent noise amplification. This type of technique is simple

and effective. In [12], the ratio rule is an aid to further enhance the color of the image.

- Multi-resolution

From[13], this technique is great in edge enhancement as the data compressed and transmuted has been increased to reduce the loss of the information. The other technique being used at here is using the band-pass filtering.

- Fuzzy logic approach

Modify dynamic range and contrast. This approach can help to remove noise and reduce degradation. It is useful for segmentation. Furthermore, this method still has further approach and improvement can be done for the manipulation of the images.[14]

The table below shows the more common image pre-processing techniques that have been used. The comparisons in between them are done.

| Techniques | Application | Advantages | Disadvantages |
|--|---|---|---|
| Median smoothing | Median filtering is done after the average filtering of the images. It replaces the current point in the image by the median of the brightness in its neighborhood. | Not affected by individual noise spikes Eliminates impulsive noise Does not blur edges much | Damaging thin lines and sharp corners in the image. |
| Edge detection with the zero crossing of the second derivative | The step edge corresponds to an abrupt change in the image function, getting first derivative of the image function to have an extreme | Easier and more precise to find a zero crossing position than an extreme | Smooth the shape too much |

| | | | |
|--------------------------------------|--|---|---------------------------|
| | corresponding to the edge of the image, zero, at 2 nd derivative | | |
| Canny edge detection | Ensure no missed important edges, distance between the actual and located position of the edge is minimal and minimize the multiple responses to a single edge | Optimal for step edges corrupted by white noises. | Very complicated |
| Adaptive neighborhood pre-processing | Detection based on the most homogeneous neighborhood of each pixel. Neighborhood size and shape are depending on the image data on parameters which define measures of homogeneity of a pixel neighborhood. | Self-tune to contextual details in the image. | Lack of unifying approach |

Table 1: technique of noise reduction and enhancement of images

2) Segmentation

The table below shows the several basic techniques that have been used more in advance.

| Segmentation technique | application |
|------------------------|--|
| Threshold based | Histogram thresholding and slicing techniques are used to segment the image. They may be applied directly to an image, but can also be combined with pre-and post-processing techniques. |
| Edge based | With this technique, detected edges in an image are assumed to represent object boundaries, and used to identify the object. |
| Region based | This technique is used where an edge based technique find the object boundaries and then locate the object itself by filling them in. |
| Clustering technique | Clustering methods group the patterns all together which are detected to be similar in some sense. |

Table 2: technique of segmentation

- Watershed Approach

This technique biggest benefit is that it promotes an automatic segmentation system [15]. It is also actually the combination of edge detection technique and region based technique. Its disadvantages are that it will have over segmentation happened. By reviewing [15], the performance and its result accuracy having an average of 92.81%.

- H-minima Transform and Region Merging

This technique is a quite new technique which is proposed to be used in medical field in 2011[16]. This technique used multi-scale alternating sequential filtering to remove noise, h-minima transform and minima imposition technique. Followed by, watershed transform is used for segmentation. Post processing merging technique is the final step to merge the over segmented region. According to[16], the average accuracy of this technique is 96.8%.

With the results of H-minima transform and region merging technique and watershed approach, the H-minima and region merging is tend to be selected and be used as the segmentation technique.

3) Featured extraction

- Extraction of Diagnostic Rules Using Neural Network

Based on [17], Simultaneous Induction of Rules, SIR methodology is being used. This approach is effective but it needs synthetic data and requires a long time periods and patients to get the results.

- Semantic Modeling and Improved SVM-KNN algorithm

Based on [18], this technique has successfully enhanced the quality of the extracted information. However, this technique is applied on the Named Entity Recognition (NER). There is a chance for this technique to be further explored on the feature extraction on the HEP-2 cells.

- Textural feature Extraction

In [19], there are 4 types of feature extraction have been done, Speeded-up Robust Features (SURF), Region of Interest (ROI), texture-based features, and normalized Histogram of Oriented Gradient (NHOG) features. By using a Support Vector Machine (SVM), we can get 91.13% of accuracy.

By comparison these feature extraction techniques, textural feature extraction can have higher accuracy compared to others. Hence, textural feature extraction technique is more likely to have images extracted accurately and faster.

4) Classification

- Microarray analysis by Machine Learning Procedures

By reviewing [20], this method with the aid of novel statistical technique, can help in analyze the expression of hundreds to thousands genes and it provides a good performances on obtaining genome-wide gene expression profile of SLE. The weakness of this method is that the data quality can be impacted.

- An Adaptive Fuzzy Classification System

Referring to[21], this technique using the IF-THEN rules. This technique is useful as it is able to solve a complex nonlinear problem. It expresses the complexity into a linguistic form.

- K-Nearest Neighbour Classifier

In [22], KNN is used as the classification method due to its good performance and simplicity. It memorizes the training data and matches them. When the test object match with 1 of the training data, it will determine whether that is the normal or abnormal images.

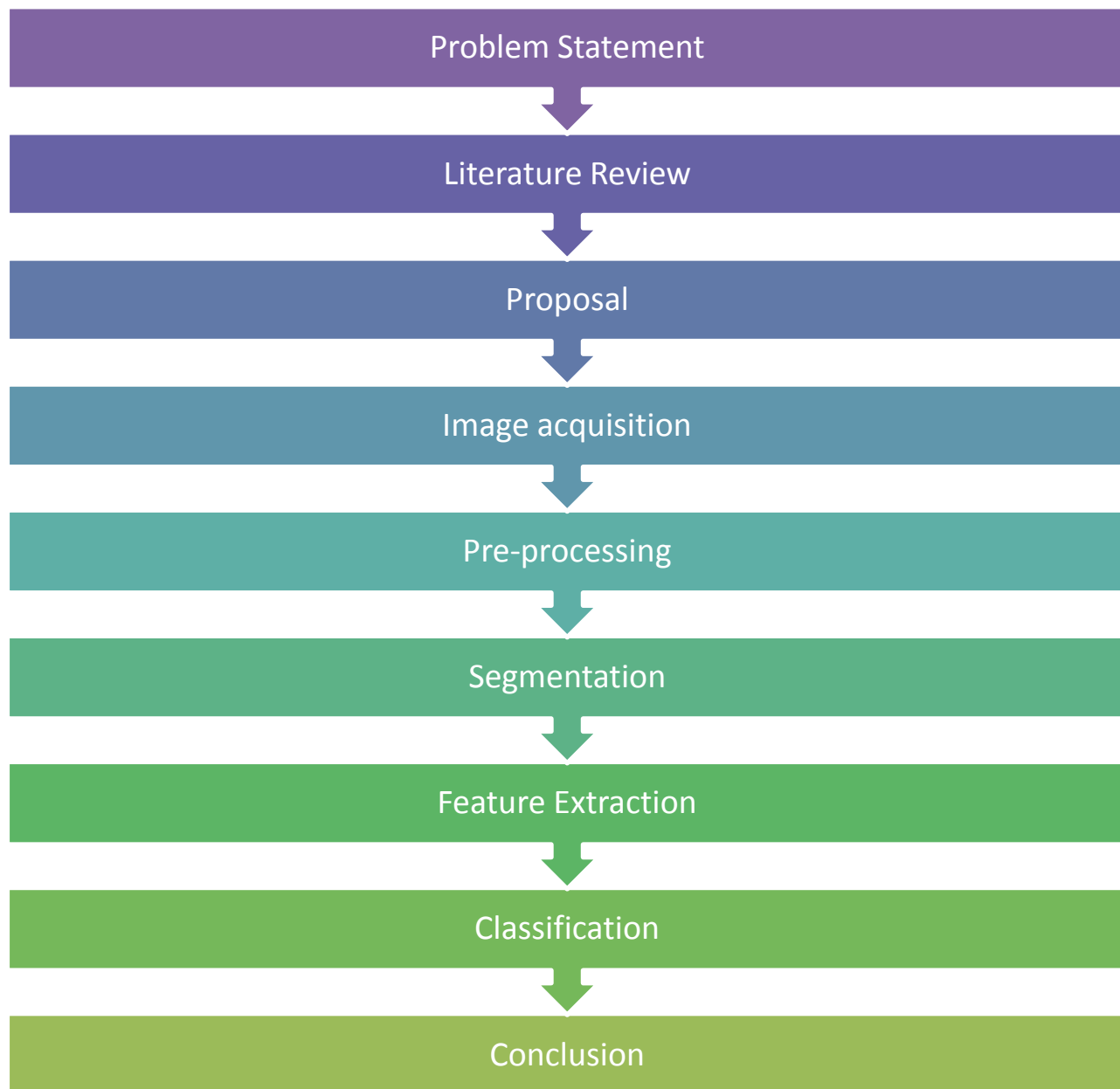
- Support Vector Machine classifier

In[19], the used of SVM is simple and effective. It can have more than 90% accuracy with the feature extracted by texture feature.

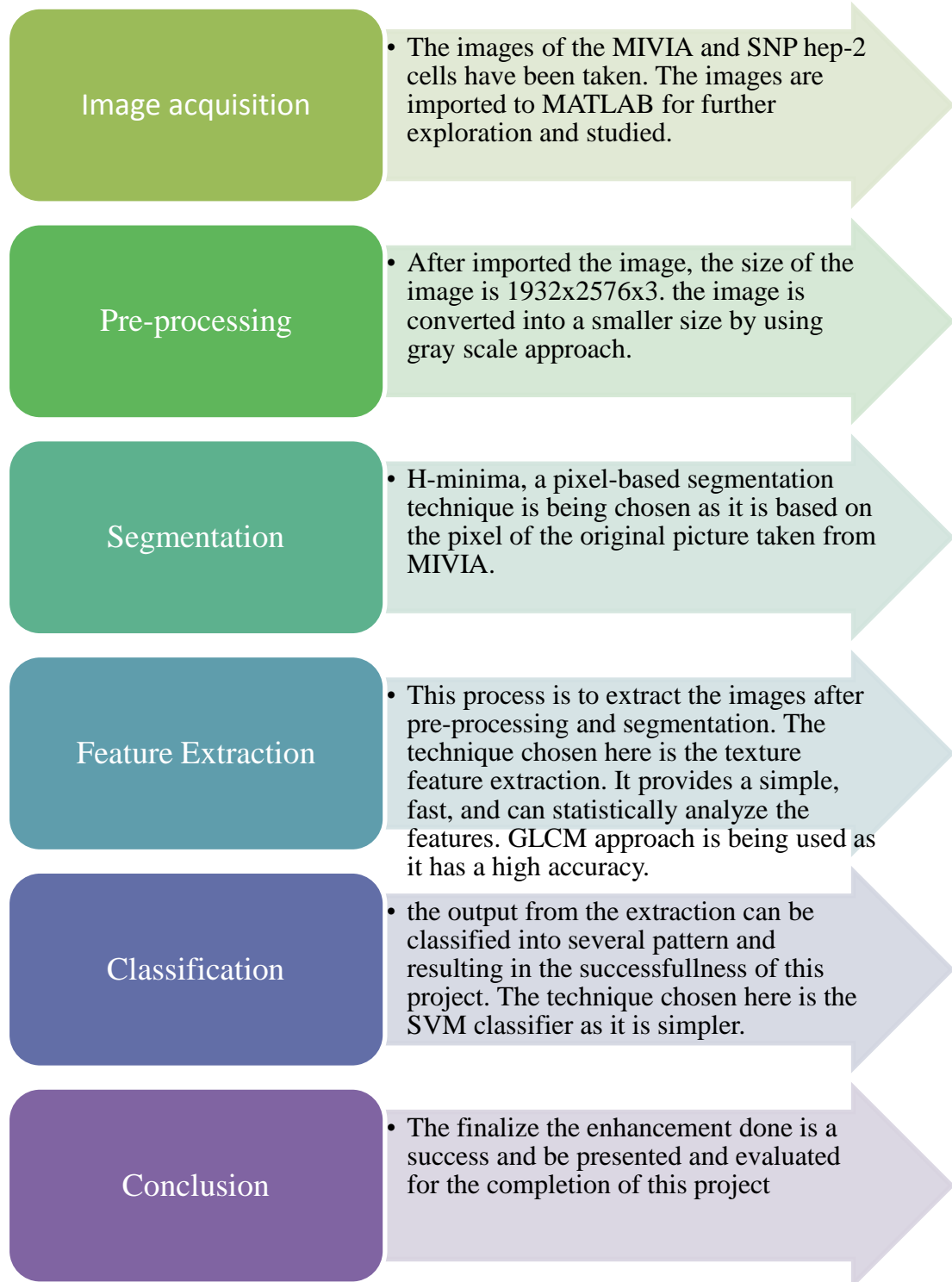
CHAPTER 3

3. METHODOLOGY

3.1 Research Methodology



3.2 Progress and Activities



3.2.1 Image acquisition

The images of the MIVIA and SNP HEP-2 cells have been taken. The images are imported to MATLAB for further exploration and studied. The images formats are .tif and the images are being picked up with the GeneSnap and Genetool (the way to maintain the original data and integrity of the images).

3.2.2 Image pre-processing

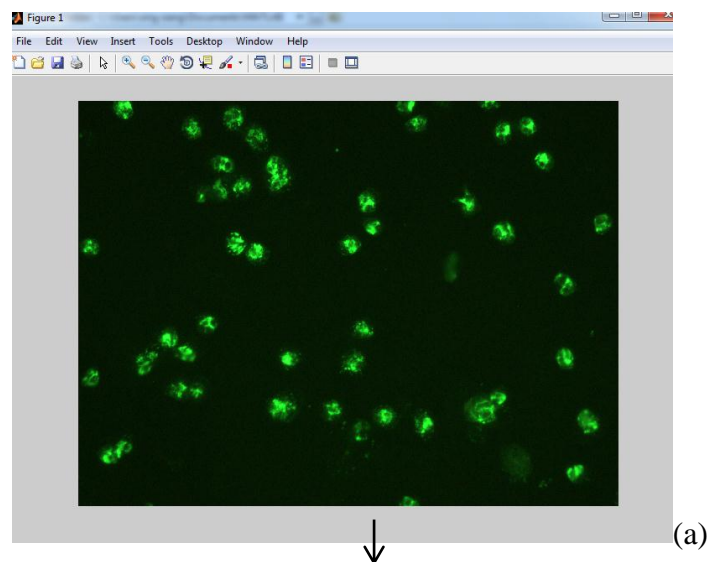
After imported the image, the size of the image is 1932x2576x3. It is found that, the images' dataset can't be shown as MATLAB has its limit on organising the data of a large colour image. Hence, the gray scale approach is being used to process the image into a smaller datasets. The function used is `rgb2gray()`, and the coding used to convert the image into gray scale is

```
~ Image2 = rgb2gray(Image);
```

After converting the image into gray scale, in order to have a clearer view, image enhancement is done with the function, `imadjust()`. This function is to adjust the contrast of the previous image. The coding used is

```
~Image3= imadjust(Image2);
```

The images will look like as below.



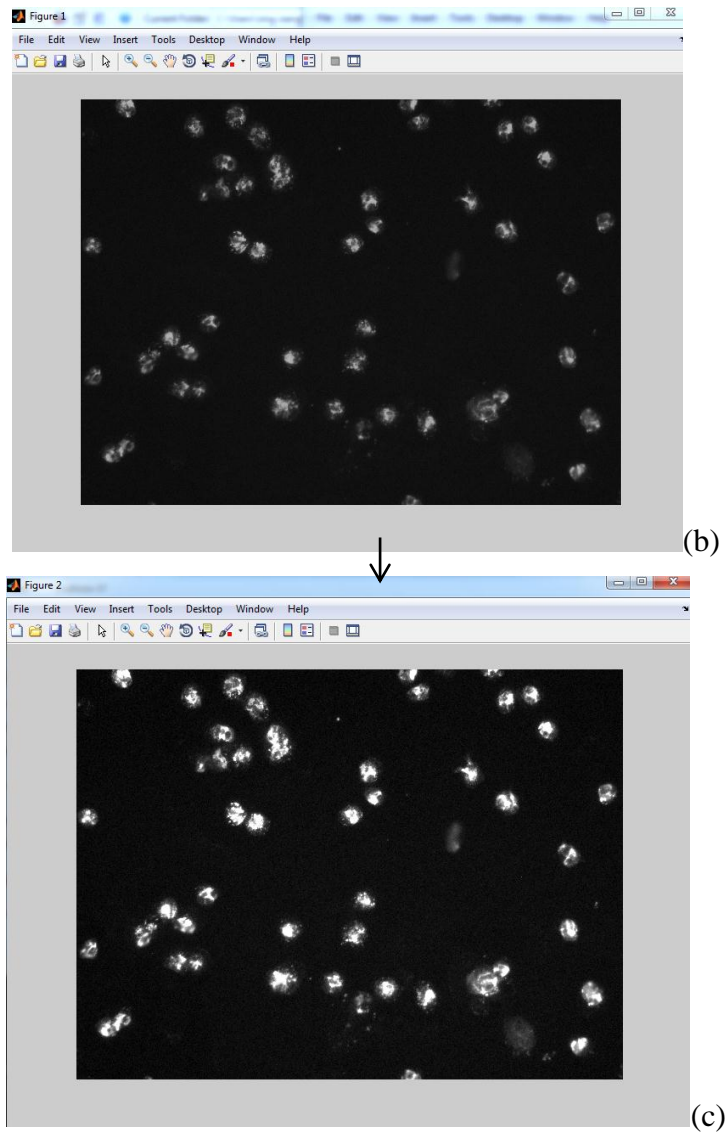


Figure 2: a) original image, b) after convert to gray scale, c) enhanced image.

3.2.3 Image Segmentation

In this segmentation part, the Histogram minima approach is used and it is actually a pixel-based method in segmentation.

~ `histeq()` function is being used to get the histogram data from the images.

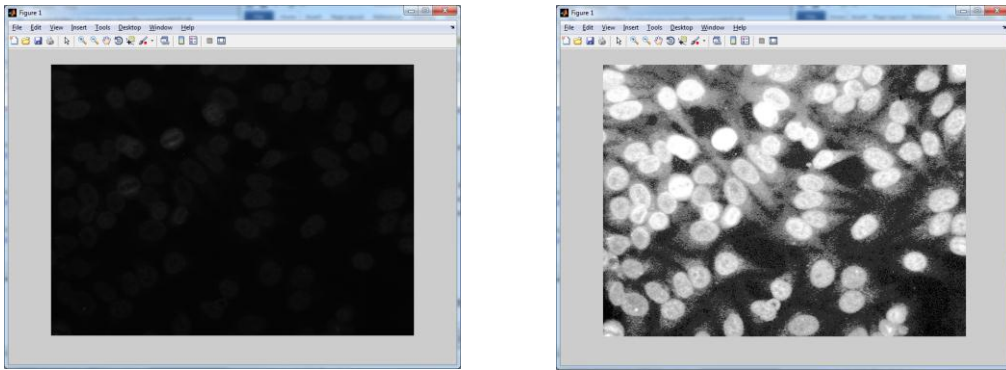
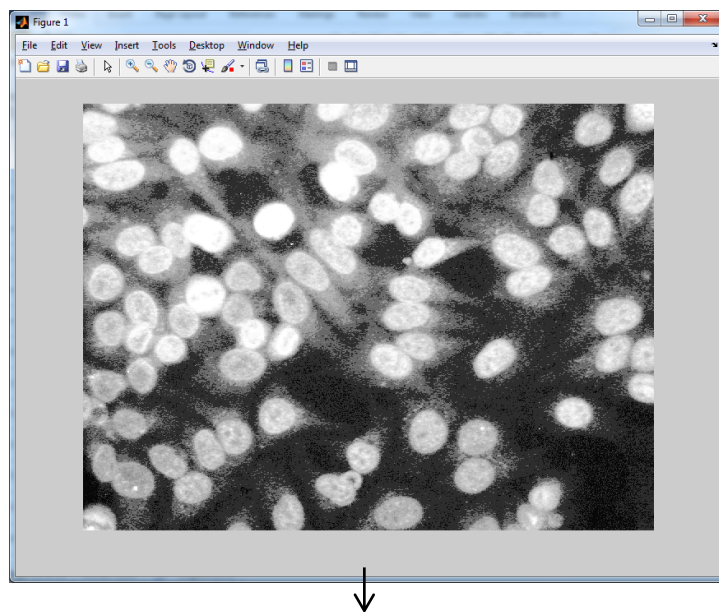


Figure 3: segmentation of image

However, in this project, it was found that the function of `histeq()` approach is not really suitable, as the noises are added and will affect the results of the properties of the original images.

In order to perform noise reduction, a simple filtering is needed. The simplest filter is by using the median filter. Median filter is an effective smoothing technique and compared to mean filter technique or convolution filtering technique (Gaussian filter). It does not build any new unrealistic pixel values and preserving sharp edges.

The function of the median filter, `medfilt2()`.



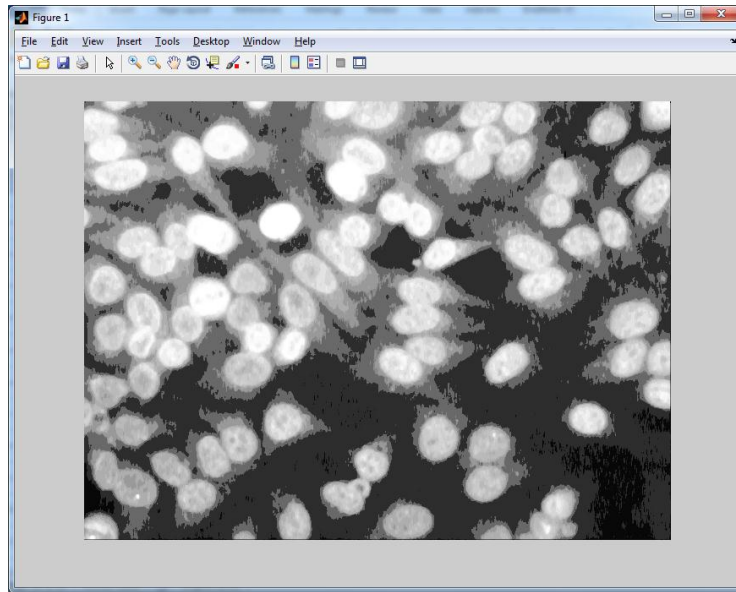


Figure 4: noise reduction

After the pixel-based segmentation has been done, the next step is to crop the cells out for the feature extraction.

The testing to auto-crop the image has been done. The coding is shown in APPENDIX A. However, this process is not recommended in this project as we need the properties of this image to identify their patterns' name. With this auto-crop function, the threshold segmentation is used, and we need to change the dataset into a binary set, which will make most of the data of the images lost. We are able to diagnosis the shape of the cells only with the auto-crop technique. Hence, the using of auto-crop technique is not been applied in this project.

The overall image pre-processing, segmentation and noise reduction coding are shown in APPENDIX B.

3.2.4 Feature Extraction

The technique selected to extract the feature is by using textural feature extraction. The texture information of the images is needed to be collected and analyzed. A statistical method is applied to distinguish difference between the patterns of the HEP-2 cells. The contrast, correlation, energy, and homogeneity, will be achieved after the Matlab calculation. The coding will be shown in the APPENDIX C.

After those formulas, a GLCMs is needed to be generated

```
~ glcms = graycomatrix (Image, 'Numlevels',8,'G',[ ] )
```

and the 'graycoprops' function is used to extract the gray-level occurrence matrices to determine the contrast, homogeneity and correlation of the HEP-2 cells image.

```
~ stats = graycoprops(glcm);
```

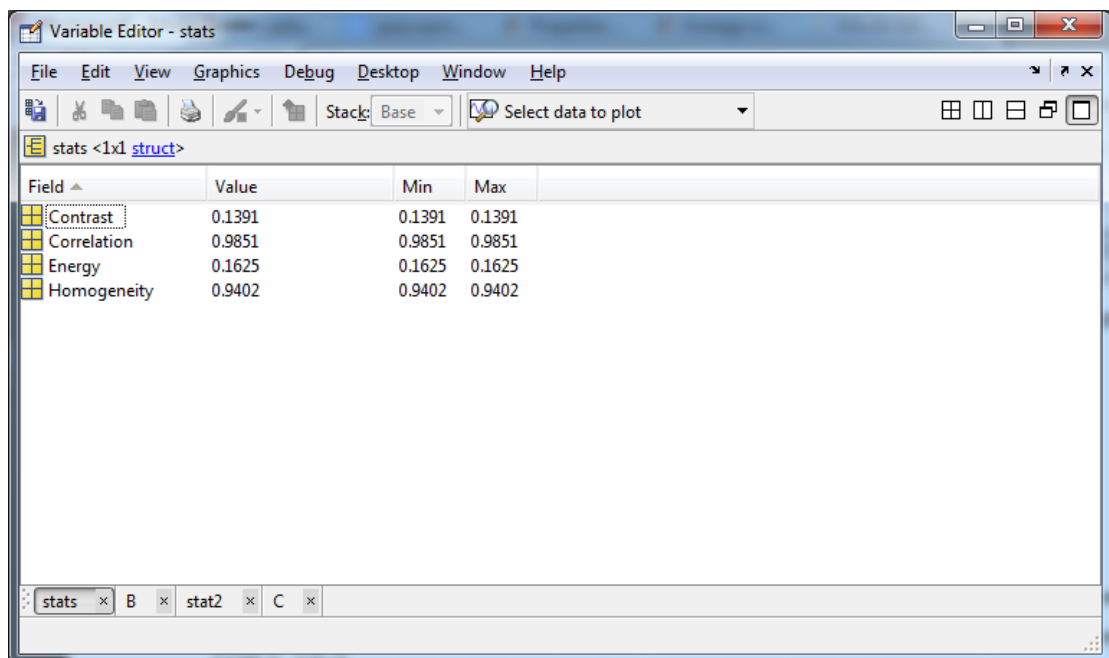


Figure 5: Result of GLCM

3.2.5 Classification

In classification, the method used is by using SVM, Support Vector Machine, classifier, and with If- Then rules. With this classifier, the images are able to be classified into categories and determine whether the images collected from the patient is having the autoimmune disease or not. This classifier is a simple but effective technique.

The MATLAB coding of the classification by using SVM is shown in APPENDIX E. From the MATLAB coding, the energy value of each cell has been used from the result of GLCM calculation. The value is defined and the pattern groups are defined either 0 or 1. For speckled pattern, the pattern is 1, and for centromeric, the pattern is 0.

3.2.6 Key Milestones and Gantt chart

| Activities | No of week (FYP 1) | | | | | | | | | | | | | | |
|---|--------------------|---|---|---|---|---|---|---|---|----|----|----|----|----|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | |
| Topic Selection | ■ | ■ | | | | | | | | | | | | | |
| Background study and literature review | | ■ | ■ | ■ | ■ | ■ | | | | | | | | | |
| Submission of Extended Proposal | | | | | | ● | | | | | | | | | |
| Extended study on the pre-processing technique and segmentation technique | | | | | | ■ | ■ | | | | | | | | |
| Proposal Defense | | | | | | | | ■ | ■ | | | | | | |
| Further research on the latest advanced technique used. | | | | | | | | | | ■ | ■ | ● | | | |
| Submission of Interim Draft Report | | | | | | | | | | | | | ■ | | |
| Submission of Interim Report | | | | | | | | | | | | | | ● | |

| Activities | No of week (FYP 2) | | | | | | | | | | | | | |
|--|--------------------|---|---|---|---|---|---|---|---|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| Apply techniques of image enhancement, segmentation, and feature extraction on the images. | | | | | | | | | | | | | | |
| Classification of the images. | | | | | | | | | | | | | | |
| Results and discussion | | | | | | | | | | | | | | |
| Progress Report | | | | | | | | | | | | | | |
| Draft report | | | | | | | | | | | | | | |
| Dissertation (soft bound) | | | | | | | | | | | | | | |
| Technical paper | | | | | | | | | | | | | | |
| Oral Presentation | | | | | | | | | | | | | | |
| Dissertation (hard bound) | | | | | | | | | | | | | | |

3.3 Tools and software

Throughout the completion of the project, the tools and software needed are:

- 1) MATLAB
- 2) Image processing toolbox

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Feature Extraction.

After the feature extraction is completed with several images from MIVIA, the properties, contrast, correlation, energy and homogeneity is being recorded in APPENDIX D. Graphs are drawn for analysis for the properties.

Contrast

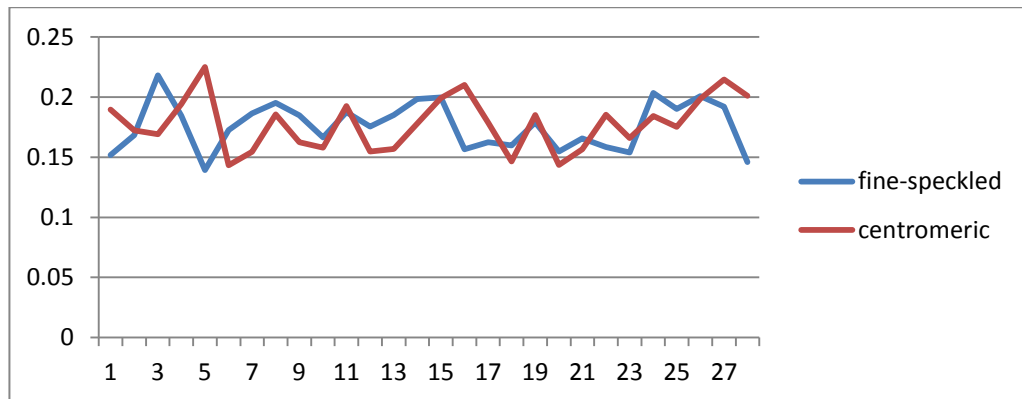


Figure 6 : Graph of contrast from GLCM

Correlation

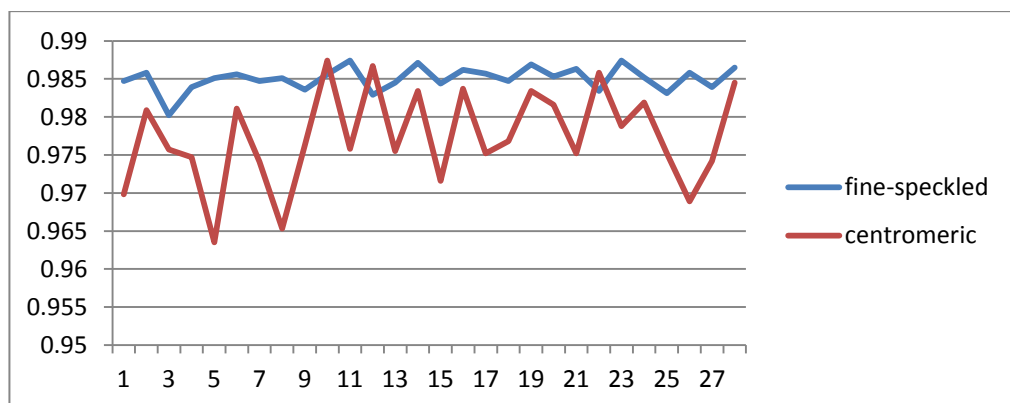


Figure 7 : Graph of correlation from GLCM

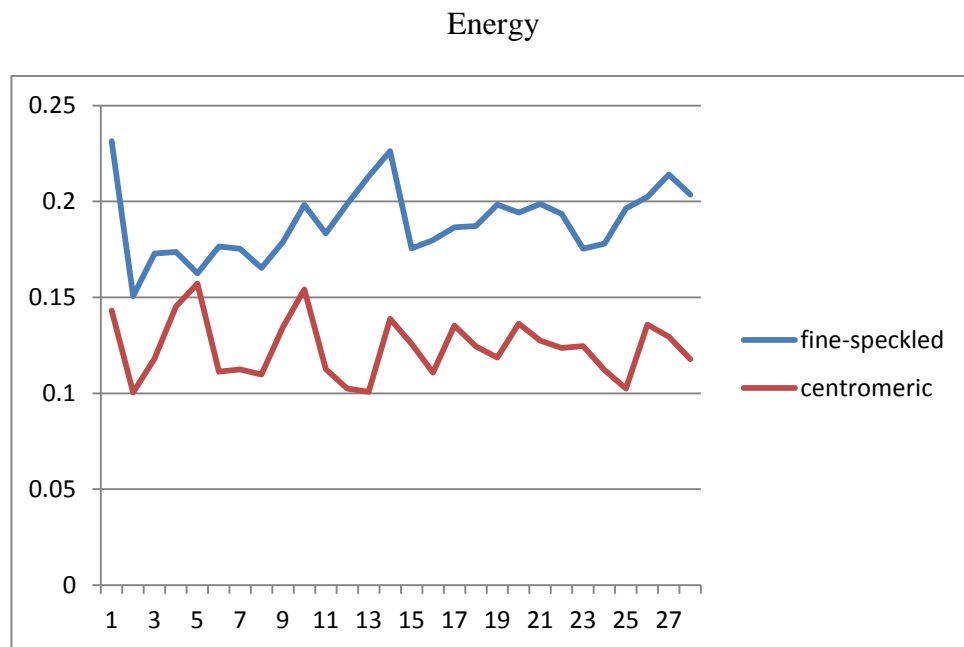


Figure 8 : Graph of energy from GLCM

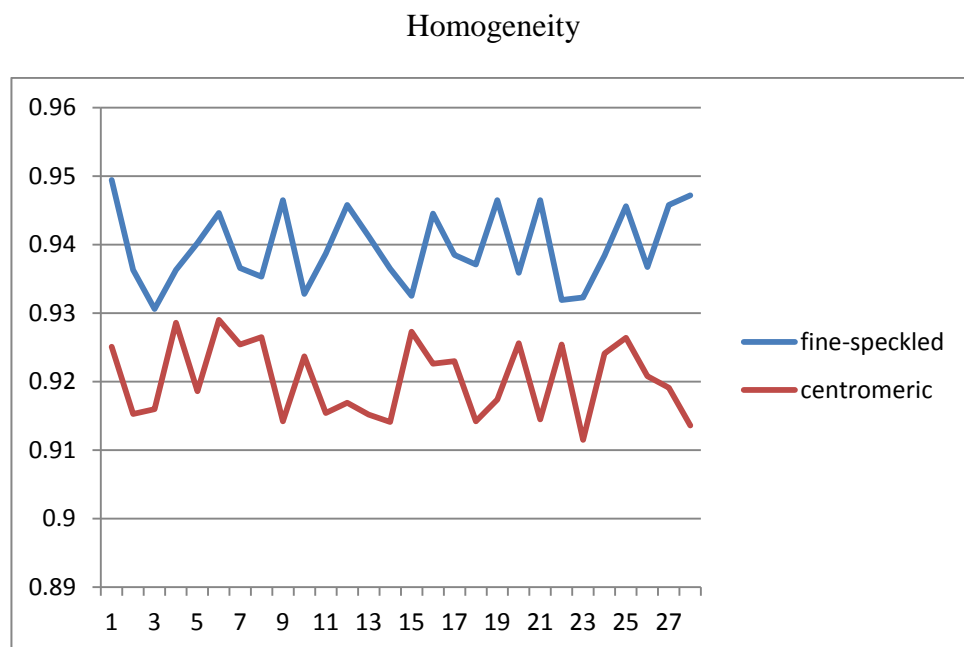


Figure 9 : Graph of homogeneity from GLCM

From the results shown in graph, the energy and homogeneity of the properties of the images have shown a more significant difference between the pattern of fine-speckled and centromeric. Therefore, the SVM approach programming can be done to classify

the pattern of the HEP-2 cells by using the data of energy or homogeneity from the GLCM.

4.2 SVM Classifier

When homogeneity is a random image from MIVIA, with value of 0.9425,

```
>> Pattern = svmclassify(ImageTrain,0.9425);
>> if Pattern == 1
display('the HEP-2 cell is fine-speckled pattern, this patient has SLE')
else
disp('the HEP-2 cell is centromeric pattern, this patient has no SLE but may have suffer from other disease')
end
the HEP-2 cell is fine-speckled pattern, this patient has SLE
|
```

Figure 10: Results of SVM Classifier (fine-speckled)

When homogeneity is a random image from MIVIA, with value of 0.9112,

```
>> Pattern = svmclassify(ImageTrain,0.9112);
>> if Pattern == 1
display('the HEP-2 cell is fine-speckled pattern, this patient has SLE')
else
disp('the HEP-2 cell is centromeric pattern, this patient has no SLE but may have suffer from other disease')
end
the HEP-2 cell is centromeric pattern, this patient has no SLE but may have suffer from other disease
|
```

Figure 11: Results of SVM Classifier (centromeric)

When the other images are being tested and the homogeneity is being tested, the results having a high accuracy compared with the actual classification from the MIVIA.

CHAPTER 5

5.1 Conclusion

Based on the studies, the intelligent detection system for SLE can be done with MATLAB approach. According to the studies, the intelligent detection system is going to prove to have a great influence in the medical field. By achieving the proper steps in pre-processing images, segmentation, feature extraction, and classification, the intelligent detection system is able to get the MIVIA dataset classified into their patterns, and the homogeneous pattern is the positive test of ANA from the patients having the disease of SLE. Hence, the patients diagnosed to have SLE can have an early treatment. However, this intelligent detection system can also be further improved in the future as there will have more accurate result with the higher technology to indicate the ANA positive test.

5.2 Recommendation and Future Work

There should have a way on the part of the automated system in cropping the image of the HEP-2 cells without reducing the properties of it. With the help of this automated function, it can ease the user to crop the images into a single cell without doing it manually. Besides, it can help to improve the standard of the cell recognition.

Aside from this, there should have another way about how the ANA cells are different from one another despite their properties only. This is still need to be developed further with the help of MIVIA, as the dataset are all provided from them. Therefore, the images from them are also very important for the further development of the diagnosis of diseases lead by the ANA cells.

In the future, other classifier should be used to classify the patterns of ANA cells as the current used, SVM approach, has its limitation which is only for 2 patterns.

REFERENCES

- [1] A. Demin and T. Sentjakova, "New synchronized therapy with plasmapheresis and high-dose cyclophosphamide in severe systemic lupus erythematosus," in *Science and Technology, 2004. KORUS 2004. Proceedings. The 8th Russian-Korean International Symposium on*, 2004, pp. 331-335 vol. 3.
- [2] Quintiles. *statistic of the lupus 2005*. Available: <http://www.clinicalresearch.com/conditions/lupus/by-the-numbers/>
- [3] D. K. W. Ray D. Strand, *What Your Doctor Doesn't Know about Nutritional Medicine May Be Killing You*: Thomas Nelson, 2002.
- [4] M. Weller, "Jeanette M. Van Emon (Ed.): Immunoassay and other bioanalytical techniques," *Analytical and Bioanalytical Chemistry*, vol. 389, pp. 1653-1654, 2007/11/01 2007.
- [5] P. Perner, H. Perner, and B. Müller, "Mining knowledge for HEP-2 cell image classification," *Artificial Intelligence in Medicine*, vol. 26, pp. 161-173, 9// 2002.
- [6] T. L. F. o. America. (2013). *What are the laboratory tests for lupus?* Available: <http://www.lupus.org/answers/entry/lupus-tests>
- [7] P. Foggia, G. Percannella, P. Soda, and M. Vento, "Early experiences in mitotic cells recognition on HEP-2 slides," in *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*, 2010, pp. 38-43.
- [8] P. Soda and G. Iannello, "Decomposition Methods and Learning Approaches for Imbalanced Dataset: An Experimental Integration," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 3117-3120.
- [9] P. Soda, "An experimental comparison of MES aggregation rules in case of imbalanced datasets," in *Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on*, 2009, pp. 1-6.
- [10] K. Kyamakya, J. C. Chedjou, M. A. Latif, and U. A. Khan, "A novel image processing approach combining a 'coupled nonlinear oscillators'-based paradigm with cellular neural networks for dynamic robust contrast enhancement," in *Cellular Nanoscale Networks and Their Applications (CNNA), 2010 12th International Workshop on*, 2010, pp. 1-7.
- [11] M. M. Hadhoud, "Image contrast enhancement using homomorphic processing and adaptive filters," in *Radio Science Conference, 1999. NRSC '99. Proceedings of the Sixteenth National*, 1999, pp. C5/1, C5/7.
- [12] S. Ming-Jung and V. K. Asari, "Homomorphic processing system and ratio rule for color image enhancement," in *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, 2004, pp. 2507-2511 vol.4.
- [13] I. Chakravarty and B. Nichol, "A multi-resolution representation of 3D imagery for efficient image enhancement and processing," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86.*, 1986, pp. 2083-2086.
- [14] T. Ching-Yu and P. P. Wang, "Image processing-enhancement, filtering and edge detection using the fuzzy logic approach," in *Fuzzy Systems, 1993., Second IEEE International Conference on*, 1993, pp. 600-605 vol.1.
- [15] C. Chung-Chuan, T. Jin-Shiuh, H. Tsu-Yi, and T. Chin-Wang, "Segmentation of anti-nuclear antibody images based on the watershed approach," in *Industrial Electronics and Applications (ICIEA), 2010 the 5th IEEE Conference on*, 2010, pp. 1695-1700.
- [16] K. Ali, A. Jalil, M. U. Gull, and M. Fiaz, "Medical Image Segmentation Using H-minima Transform and Region Merging Technique," in *Frontiers of Information Technology (FIT), 2011, 2011*, pp. 127-132.

- [17] I. K. Sethi, J. H. Yoo, and C. M. Brickman, "Extraction of diagnostic rules using neural networks," in *Computer-Based Medical Systems, 1993. Proceedings of Sixth Annual IEEE Symposium on*, 1993, pp. 217-222.
- [18] H. Xia and R. Rao, "The Method of Medical Named Entity Recognition Based on Semantic Model and Improved SVM-KNN Algorithm," in *Semantics Knowledge and Grid (SKG), 2011 Seventh International Conference on*, 2011, pp. 21-27.
- [19] S. Ghosh and V. Chaudhary, "Feature analysis for automatic classification of HEP-2 florescence patterns : Computer-Aided Diagnosis of Auto-immune diseases," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, 2012, pp. 174-177.
- [20] R. Armananzas, B. Calvo, I. Inza, M. Lopez-Hoyos, V. Martinez-Taboada, E. Ucar, *et al.*, "Microarray Analysis of Autoimmune Diseases by Machine Learning Procedures," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, pp. 341-350, 2009.
- [21] G. Nai Ren, K. Chao-Lin, T. Tzong-Jiy, and C. Shi-Jaw, "An adaptive fuzzy classification system," in *Soft Computing in Industrial Applications, 2008. SMCia '08. IEEE Conference on*, 2008, pp. 377-381.
- [22] R. V. Rajesh, J. Veerappan, S. K. Sujitha, and E. A. Kumar, "Classification and retrieval of images using texture features," in *Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on*, 2012, pp. 1-5.

APPENDIX

APPENDIX A:

AUTO-CROPING CODING

```
function preprocessing
%preprocessing image

%read the image
I = imread('image.bmp');
%%figure(1)
%%imshow(I)
%convert to grayscale
Igray = rgb2gray(I);
%%figure(2)
%%imshow(Igray)

%convert to binary image
Ibw = im2bw(Igray,graythresh(Igray));
%%figure(3)
%%imshow(Ibw)

%detect edges
Iedge = edge(uint8(Ibw));
%%figure(4)
%%imshow(Iedge)

%perform image dilation
se = strel('square',2);
Idil=imdilate(Iedge,se);
%%figure(5)
%%imshow(Idil)

%perform image filing
Ifill = imfill(Idil,'holes');
%%figure(6)
%%imshow(Ifill)

%blobs analysis - display the number of objects
[Ilabel num] = bwlabel(Ifill);
disp(num)
Iprops = regionprops(Ilabel);
Ibox = [Iprops.BoundingBox];
Ibox = reshape(Ibox,[4 num]);
%%figure(7)
%%imshow(I)

%plot the object
%%hold on;
%%for cnt = 1:num
%%rectangle('position',Ibox(:,cnt),'edgecolor','r');
%%end

%crop each object into figures
for cnt = 1:num
img_crop{cnt} = imcrop(I,Ibox(:,cnt));
%%figure(cnt)
%%imshow(img_crop{cnt})
end
```

```

%change each cropped image to binary
for cnt = 1:num
Igray2{cnt} = rgb2gray(img_crop{cnt});
Ibw2{cnt} = im2bw(Igray2{cnt},graythresh(Igray2{cnt}));
%%figure(cnt)
%%imshow(Ibw2{cnt})
end

% Find the boundary of the image
for cnt = 1:num
[y2temp x2temp] = size(Ibw2{cnt});
x1=1;
y1=1;
x2=x2temp;
y2=y2temp;

% Finding left side blank spaces
cntB=1;
while (sum(Ibw2{cnt}(:,cntB))==y2temp)
    x1=x1+1;
    cntB=cntB+1;
end

% Finding right side blank spaces
cntB=1;
while (sum(Ibw2{cnt}(cntB,:))==x2temp)
    y1=y1+1;
    cntB=cntB+1;
end

% Finding upper side blank spaces
cntB=x2temp;
while (sum(Ibw2{cnt}(:,cntB))==y2temp)
    x2=x2-1;
    cntB=cntB-1;
end

% Finding lower side blank spaces
cntB=y2temp;
while (sum(Ibw2{cnt}(cntB,:))==x2temp)
    y2=y2-1;
    cntB=cntB-1;
end

% Crop the image to the edge
Ibw3{cnt}=imcrop(Ibw2{cnt},[x1,y1,(x2-x1),(y2-y1)]);
%%figure(cnt)
%%imshow(Ibw3{cnt})
end

for cnt = 1:num
im_resize{cnt} = imresize(Ibw3{cnt},[70,50]);
%%figure(cnt)
%%imshow(im_resize{cnt})

for cnt1=1:7
    for cnt2=1:5
        Atemp{cnt}=sum(im_resize{cnt}((cnt1*10-9:cnt1*10),(cnt2*10-
9:cnt2*10)));

```

```
        lett{cnt}((cnt1-1)*5+cnt2)=sum(Atemp{cnt});
    end
end

lett{cnt}=(100-lett{cnt})/100;
lett{cnt}=lett{cnt}';

out(:,cnt) = lett{cnt};
end

P = out(:,1:cnt);

for cnt = 1:num
figure(cnt)
plotchar(P(:,cnt))
end
```

APPENDIX B:

IMAGE SEGMENTATION AND NOISE REDUCTION

```
% image processing
ImA = ('file location');
%Change to gray scale
ImAG= rgb2gray(ImA);
%enhance image contrast;
ImAGE = imadjust(ImAG);

%equalize histogram (segmentation)
ImAGEH = histeq(ImAGE);
% display image segmented
% imshow(ImAGEH);

% noise reduction, median filter
ImAGEHF = medfilt2(ImAGEH);

%display image-filtered
%imshow(ImAGEHF);
```

APPENDIX C:

MANUAL CROP AND FEATURE EXTRACTION

```
% manual crop the Image
Icrop= imcrop(ImAGEHF);

%display image-cropped
%imshow(Icrop);

% feature extraction
GLCM = graycomatrix(Image, 'NumLevels', 8, 'G', []);
Stat = graycoprops (GLCM)
```


**APPENDIX D:
RESULTS FROM FEATURE EXTRACTION OF GLCM**

| contrast | | correlation | | energy | | homogeneity | |
|-------------------|-------------|-------------------|-------------|-------------------|-------------|-------------------|-------------|
| fine- speckled | centromeric | fine- speckled | centromeric | fine- speckled | centromeric | fine- speckled | centromeric |
| 0.1516 | 0.1895 | 0.9847 | 0.9698 | 0.2314 | 0.1431 | 0.9494 | 0.9251 |
| 0.1682 | 0.1722 | 0.9858 | 0.9809 | 0.1506 | 0.1004 | 0.9363 | 0.9153 |
| 0.218 | 0.169 | 0.9802 | 0.9757 | 0.1728 | 0.1181 | 0.9306 | 0.916 |
| 0.1845 | 0.1939 | 0.9839 | 0.9747 | 0.1737 | 0.1454 | 0.9363 | 0.9286 |
| 0.1391 | 0.2249 | 0.9851 | 0.9635 | 0.1625 | 0.1572 | 0.9402 | 0.9186 |
| 0.1725 | 0.1433 | 0.9856 | 0.9811 | 0.1766 | 0.1113 | 0.9446 | 0.929 |
| 0.1864 | 0.1545 | 0.9847 | 0.9741 | 0.1754 | 0.1125 | 0.9366 | 0.9254 |
| 0.1952 | 0.1855 | 0.9851 | 0.9653 | 0.1654 | 0.1098 | 0.9353 | 0.9265 |
| 0.1847 | 0.1625 | 0.9836 | 0.9763 | 0.1789 | 0.1345 | 0.9465 | 0.9142 |
| 0.1664 | 0.1578 | 0.9856 | 0.9874 | 0.1982 | 0.1542 | 0.9328 | 0.9237 |
| 0.1875 | 0.1925 | 0.9874 | 0.9758 | 0.1833 | 0.1126 | 0.9387 | 0.9154 |
| 0.1753 | 0.1547 | 0.9829 | 0.9867 | 0.1987 | 0.1025 | 0.9458 | 0.9169 |
| 0.1851 | 0.1569 | 0.9845 | 0.9755 | 0.213 | 0.1008 | 0.9412 | 0.9152 |
| 0.1984 | 0.1778 | 0.9871 | 0.9834 | 0.2263 | 0.1388 | 0.9365 | 0.9141 |
| 0.1996 | 0.1992 | 0.9844 | 0.9716 | 0.1756 | 0.1258 | 0.9325 | 0.9273 |
| 0.1564 | 0.21 | 0.9862 | 0.9837 | 0.1798 | 0.1108 | 0.9445 | 0.9226 |
| 0.1623 | 0.1785 | 0.9857 | 0.9752 | 0.1865 | 0.1354 | 0.9385 | 0.923 |
| 0.1598 | 0.1465 | 0.9847 | 0.9768 | 0.1872 | 0.1245 | 0.9371 | 0.9142 |
| 0.1789 | 0.1851 | 0.9869 | 0.9834 | 0.1984 | 0.1187 | 0.9465 | 0.9174 |
| 0.1547 | 0.1435 | 0.9853 | 0.9816 | 0.1942 | 0.1363 | 0.9359 | 0.9256 |
| 0.1656 | 0.1564 | 0.9863 | 0.9752 | 0.1987 | 0.1274 | 0.9465 | 0.9145 |
| 0.1585 | 0.1852 | 0.9834 | 0.9858 | 0.1935 | 0.1236 | 0.9319 | 0.9254 |
| 0.1539 | 0.1657 | 0.9874 | 0.9788 | 0.1754 | 0.1247 | 0.9323 | 0.9115 |
| 0.2034 | 0.1842 | 0.9852 | 0.9819 | 0.1781 | 0.1122 | 0.9384 | 0.9241 |
| 0.1899 | 0.1752 | 0.9831 | 0.9752 | 0.1963 | 0.1024 | 0.9456 | 0.9264 |
| 0.2007 | 0.1985 | 0.9858 | 0.9689 | 0.2023 | 0.1359 | 0.9367 | 0.9208 |
| 0.1919 | 0.2144 | 0.9839 | 0.9742 | 0.2141 | 0.1295 | 0.9458 | 0.9191 |
| 0.1458 | 0.2009 | 0.9865 | 0.9845 | 0.2036 | 0.1178 | 0.9472 | 0.9136 |

**APPENDIX E:
SVM CLASSIFIER CODING**

```
% training image  
H1=0.9494  
H2=0.9363  
H3=0.9306  
H4=0.9363  
H5=0.9402  
H6=0.9446  
H7=0.9366  
H8=0.9353  
H9=0.9465  
H10=0.9328  
H11=0.9387  
H12=0.9458  
H13=0.9412  
H14=0.9365  
H15=0.9325  
H16=0.9445  
H17=0.9385  
H18=0.9371  
H19=0.9465  
H20=0.9359  
H21=0.9465  
H22=0.9319  
H23=0.9323  
H24=0.9384  
H25=0.9456  
H26=0.9367  
H27=0.9458  
H28=0.9472  
H29=0.9251;  
H30=0.9153;  
H31=0.916;
```

```

H32=0.9286;
H33=0.9186;
H34=0.929;
H35=0.9254;
H36=0.9265;
H37=0.9142;
H38=0.9237;
H39=0.9154;
H40=0.9169;
H41=0.9152;
H42=0.9141;
H43=0.9273;
H44=0.9226;
H45=0.923;
H46=0.9142;
H47=0.9174;
H48=0.9256;
H49=0.9145;
H50=0.9254;
H51=0.9115;
H52=0.9241;
H53=0.9264;
H54=0.9208;
H55=0.9191;
H56=0.9136;

```

```

Traning = cat(1, H1, H2, H3, H4, H5, H6, H7, H8, H9 ,H10, H11, H12,
H13, H14, H15, H16, H17, H18, H19, H20, H21, H22, H23, H24, H25,
H26, H27, H28, H29, H30, H31, H32, H33, H34, H35, H36, H37, H38,
H39, H40, H41, H42, H43, H44, H45, H46, H47, H48, H49, H50, H51,
H52, H53, H54, H55, H56 );

% define pattern for each homogeneity. Fine-speckled = 1,
centromeric=0

Group= [1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1;
1; 1; 1; 1; 1; 1; 1; 1; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0;
0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0;];

```

```
ImageTrain= svmtrain(Training, Group);
%classifying test image's properties of homogeneity
Pattern = svmclassify(ImageTrain, 'value');
If pattern ==0
    disp('the HEP-2 cell is fine-speckled pattern, this patient
has SLE')
Else
    disp('The HEP-2 cell is centromeric pattern, this patient may
have suffer from other disease')
end
```