# BACTERIA IDENTIFICATION FROM MICROSCOPIC MORPHOLOGY: A SURVEY

Noor Amaleena Mohamad[1], Noorain Awang Jusoh[1], Zaw Zaw Htike[1]and Shoon Lei Win[2]

[1,2] Department of Mechatronics Engineering, IIUM, Kuala Lumpur, Malaysia

## ABSTRACT

*Great knowledge and experience on microbiology are required for accurate bacteria identification. Automation of bacteria identification is required because there might be a shortage of skilled microbiologists and clinicians at a time of great need. There have been several attempts to perform automatic background identification. This paper reviews state-of-the-art automatic bacteria identification techniques. This paper also provides discussion on limitations of state-of-the-art automatic bacteria identification systems and recommends future direction of automatic bacteria identification.*

## KEYWORDS

*Bacteria Identification, Cocci,Bacilli, Vibrio, Naïve Bayes, Machine Learning*

## 1. INTRODUCTION

Bacteria, which are prokaryotic microorganisms, are the most abundant and simplest organisms in the world as we know it. Prokaryotes do not possess a nucleus and complex organelles. Because most prokaryotes range in size less than ten micrometers (µm), microscopes are used to study bacteria. Bacteria identification is very important in microbiology and pathology as it serves a basis of understanding diseases. Due to this, various types of methods have been introduced to classify bacteria in microbiology. Clinicians and microbiologists commonly employ the typing schemes which are dependent on the phenotypic typing schemes to develop the bacterial morphology and staining properties of the organism. Vectors, environmental reservoir of organism and pathogen's ways of transmission is important for the clinicians. Therefore, it is extremely essential to perform bacteria classification such that the said information can be obtained. On the other hand, scientists who are interested in microorganisms' evolution are getting more interested in taxonomic techniques which permit the comparison of highly conserved genes among dissimilar species. Therefore, computerized techniques are required for this task [21-37].

The most basic technique used for classifying bacteria is based on the bacterium's shape and cell arrangement. The most ordinary shapes of bacteria include rod, cocci (round), and spiral forms. Cellular arrangements occur singularly, in series, and in groups. Some species have one to numerous projections called flagella which enable the bacteria to swim and move. Cocci or coccus for a single cell are round cells, occasionally flattened when being adjacent to each other. Cocci bacteria can exist individually, in pairs, in groups of four, in chains, in clusters or in cubes consisting of eight cells. Bacilli are rod-shaped bacteria which also can occur individually, in

pairs, or in chains. Figure 1 shows examples of three famousclasses of bacteria namely: cocci, vibrio and bacilli.
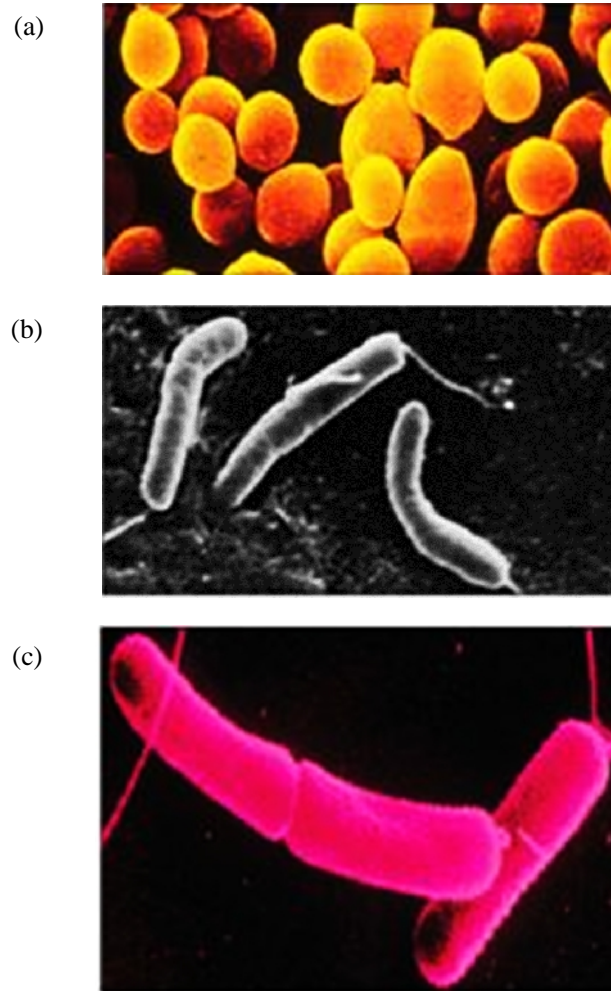
(a)

(b)

(c)

Figure 1. Three classes of bacteria (a) Cocci (b) Vibrio (c) Bacilli [21]

## 2. BACTERIA IDENTIFICATION

Bacteria classification plays important role in yielding information for disease control. Bacterial species are usually sub-grouped to different types and is used for many crucial pathogenic bacteria such as Salmonellae, E Coli, and Vibriones [1]. H.C. Gram in 1884 discovered the Gram stain classification remains an important and useful technique until today. This technique classifies bacteria as either Gram positive or negative based on their morphology and differential staining properties [2]. Table I shows the general phenotypic classification of bacteria for the Gram positive [3].

Other types of bacteria classifications that are commonly used are based on the prokaryotes of the bacteria which include their function and structures such as the slime, capsule, peptidoglycan,

cytoplasmic membrane, flagella, pili, and the secreted products. The main objectives of phenotypic classification are to generate clusters of strains and establish the hierarchy of bacteria species as members of different species may share high levels of resemblance [3].

Increasing bacterial adaptation level to human environments proves that the identification of pathogens in a bacteria species level is unsatisfactory [1].Therefore bacterial type diagnosis is required, which involves the classification of pathogens below species level. To date, biological and microbiological data analysis entails an extensive amount of human intervention [1]. The manual procedures are susceptible to inconsistency and are a tedious and complicated work which needs abundant correlative data. [3].These procedures also consume a lot of time and energy and are of great cost [1]. It is therefore important to reduce the amount of human intervention in order to handle the rising data volume besides trying to achieve adequate level data accuracy. Due to inefficient manual procedures in classifying bacteria, a considerable amount of research has been done on automatic bacteria classification using various approaches which include pattern recognition technique.

Table 1: General Phenotypic Classification of Bacteria

| Name | Morphology | Type of infection |
|---|---|---|
| Staphylococci | Cocci in grapelike clusters | Soft tissue, bone, joint, food poisoning |
| Streptococci | Cocci in pairs, chains | Skin pharyngitis, endocarditis, toxic, shock |
| Enterococci | Cocci in pairs, chains | UTI, GI, catheterrelated, infections |
| Bacilli | Rods, sporeforming | Anthrax, food poisoning, catheter-related infections |
| Clostridia | Rods, spore formers | Tetannus, diarrhea, gas gangrene |

Pattern recognition which is a field of machine learning is an act of taking in raw data and taking action based on pattern category [4][5]. The algorithm of pattern recognition normally consists of image segmentation, feature extraction, and classification [5]. Desired object is first segmented and the selected features are extracted from the image. Then, the classifier can be trained using these features as input data. Example of algorithm which can be used in bacteria classification is shown in figure 1[5].
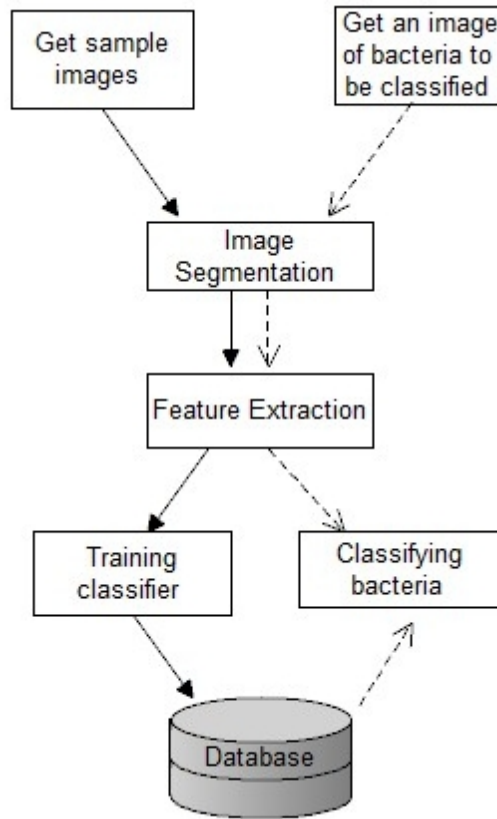
Figure 2.  Algorithm used in bacteria classification technique.

## 3. IMAGE SEGMENTATION

Image segmentation is the process of dividing an image into multiple regions [6]. In this step, image is separated into different regions in which each region is almost homogenous making the union of the two region to be impossible.Image segmentation is very important in pattern recognition and image processing area as it is widely used in recognizing object and tracking, face detection, and other computer-vision-related applications [7]. Generally, image segmentation techniques can be classified into few categories [6].

### 3.1. Clustering Method

Clustering is a process wherein pixels of same colour, texture, etc is replaced by a cluster. The K-means algorithm is a type of clustering method used to divide an image into K clusters. The quality of the algorithm depends on the initial set of clusters and the value of K [6].

### 3.2. Thresholding Method

Thresholding converts multilevel image into binary image. It compares each pixel value with some threshold value, T and assign the value of either 0 (background) or 1(foreground) to the pixel [6].

## 3.3. Edge detection Method

An edge is a boundary in an image in which significant change occurs in the physical aspect of the image [2]. Features of edges are commonly chosen as it is fairly easy to be extracted and is non-susceptible to light and other noise. Edge detector helps to reduce the complexity of algorithms by cutting down the quantity of input data. Its main purpose is to identify and locateabtrupt discontinuities .There are many ways to perform edge detection which can be summarised in the Table 2 [8].

Table 2. Edge Detection Techniques[8]

| Operator | Advantages | Disadvantages |
|---|---|---|
| Classical (Sobel, Prewitt, Kirsch,…) | • Simplicity<br>• Detection of edges and their orientations | • Sensitivity to noise<br>• Inaccurate |
| Zero Crossing (Laplacian, Second directional derivative) | • Detection of edges and their orientations •  Having fixed characteristics in all directions | • Responding to some of the existing edges<br>• Sensitivity to noise |
| Laplacian of Gaussian (LoG) (Marr-Hildreth) | • Finding the correct places of edges<br>• Testing wider area around the pixel | • Malfunctioning at corners, curves and where the gray level intensity function varies<br>• Not finding the orientation of edge because of using the Laplacian filter |
| Gaussian (Canny, Shen-Castan) | • Using probability for finding error rate<br>• Localization and response<br>• Improving signal to noise ratio<br>• Better detection especially in noise conditions | • Complex computations<br>• false zero crossing<br>• time consuming |
| Colored Edge Detectors | • Accurate<br>• more efficient in object recognition | • Complicated<br>• complex computations |

Bacteria are segmented such that the objects of interest, *i.e.* the cells, are separated from the background. Research made by Rusuuvari et al, segmented bacteria by first converting it to grey scale [9]. Then, adaptive thresholding is applied in which a smoothed version is subtracted from original image, leaving only details [9].

## 4. FEATURE EXTRACTION

Since classification is highly dependent on features, it is essential that the features extracted carry significant information about the studied objects [9]. In bacteria classification, feature extraction is done to find the morphology of mycobacterium by their shape. Geometry features are used to

measure the perimeter, area, radii, circularity, compactness, eccentricity, and tortuosity of the samples [10].

At times, bacilli shape is not an adequate indication to be used as a discriminator feature as other bacteria species and particles share the same morphology. Therefore besides their shape, it is also essential to consider the color as to improve the discrimination precision [20]. Figure shows the procedure flow chart of bacilli discrimination [2].

This technique is also based on the new segmentation method followed by an identification process. The segmentation permits the elimination of a large amount of debris objects, and only those having a similar bacilli color are preserved. One of the key factors of the existing system is the analysis and screening of the bacilli shapes [20]. Even if the descriptors used to compute the cluster centroids are needed for a recognition system, some cases show failures [20]. Because of that, we used the heuristic information about the shape of bacilli to construct a classification tree as well as to improve the overall performance of the classification process [11,20].
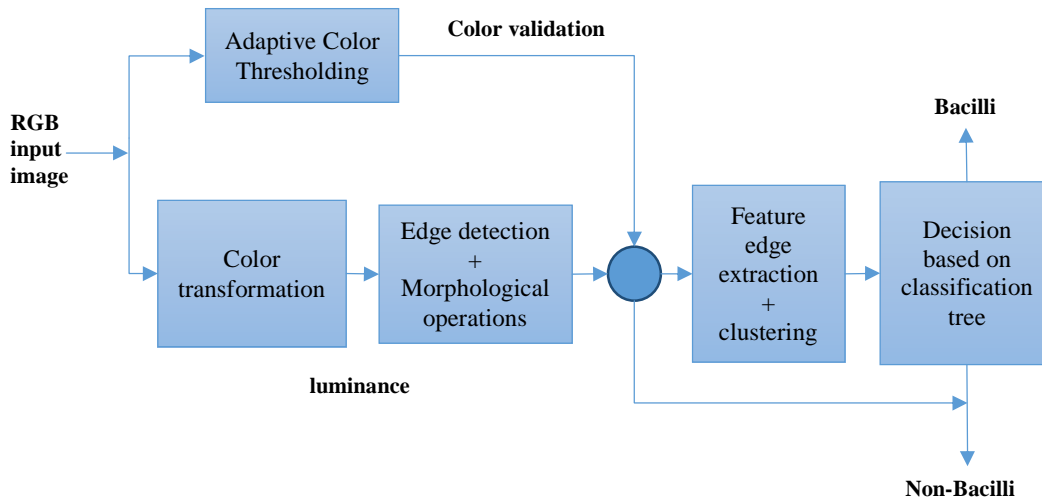
Figure 3. Bacilli discrimination procedures

## 5. FEATURE DESCRIPTORS

In image matching, extraction of features is important in providing reliable matching of an image with different viewpoints. Feature detection in an image aims to depict part of image with significant or unique information (feature descriptors) [12]. Recent trends of feature detection and matching which revolves around local features provide an invariant description of image. Local features can be points, edges or small patches, with intensity, colour and texture being the common properties of it .Features detected should remained unchanged under different conditions so that proper image matching can be obtained. Such features are said to be invariant features [13]. Feature descriptors are important in both training and classification phase. They are extracted from images, and stored during training phase, while in classification phase, image query will be matched with all the trained image features, and those which has the maximum correspondence is considered the best match. Feature descriptor matching can be calculated using Euclidean, or Mahalanobisor distance ratios [12-13].

Various feature detection algorithms have been proposed in previous research such that a reliable and robust descriptors can be obtained for image matching purpose. Scale Invariant Feature Transform (SIFT), and Speeded Up Robust Features(SURF) are the most common descriptors used in machine vision literature.

## 5.1. SIFT detector

The SIFT detector consists of four main stages namely, scalespace extrema detection, key point localization, orientation computation and key point descriptor extraction. During the first stage, Difference of Gaussians (DoG) is used to identify the potential key points. From the input images, several Gaussian blurred images are produced and DoG are computed from their neighbours [13]. During the second stage, candidate key points are located by finding extrema in the DoG images. Spatially unstable key points and low contrast key points are removed while the remaining key points are localised by interpolating across the DoG images. In the third stage, a principal orientation is assigned to each key point. Finally, in the fourth phase, a highly distinctive descriptor are computed for each key point. SIFT descriptors are invariant to rotation, scale, contrast and partially invariant to other transformations[12-13].

## 5.2. SURF descriptor

SURF which is commonly known as approximate SIFT generate key points and descriptors very efficiently by employing integral images and efficient scale space construction. It consists of two stages namely key point detection and description [13]. Unlike SIFT, SURF does not use Difference of Gaussian(DoG) to identify keypoint detection. Instead, it uses integral images to allow fast computation of Gaussian images using box filter. The first stage brings about scale and location invariance. During the final stage, each detected key point is assigned to a reproducible orientation and oriented along the orientation as obtained before. The resulting SURF descriptor is invariant to rotation, scale, and contrast [12-13].

## 6. CLASSIFICATION

Classification is one of important machine learning areas which collect unprocessed data and categorizes it into certain classes according to the set of parameters[11]. Classification structure commonly uses either statistical or syntactic approach depending on the pattern's statistical characteristics, assuming a probabilistic system [11]. In recognizing pattern, an extensive range of algorithms can be applied from the simplest Bayesian classifiers to the much more complex neural networks [11].

The main task in a classification problem is to decide in which finite dataset category an object belongs to. A classification task normally involves training and testing datasets which contains several traits or features such that pattern can be recognized [14]. Supervised and unsupervised classifiers are two techniques commonly used in image classification. Supervised classifier, also known as parametric methods require an intensive training phase of the classifier parameters [15]. It involves the act of obtaining the information class of data sets, and classifying the test data by identifying training data's best possible match [14]. On the contrary, unsupervised classifier does not need training and relies directly on data to perform classification[15].It analyses a huge amount of datasets and split it into a number of classes based on the natural groupings which is available within datasets [11].

## 6.1. Bayesian theory

Bayesian learning algorithm which is the most practical learning approach involves the evaluation of explicit probabilities for hypotheses[11]. Since it does not explicitly manipulate probabilities, it helps to provide unique perspective in understanding many learning algorithms [11].

## 6.2. Naive-Bayes Classifier

The Naive Bayes classifier simplifies the assumption that the attribute values are conditionally independent given target value [11]. It does not support a continuous data, so the independent variable or variables whose values are continuous values first be divided [16]. The formula to calculate probability is shown below:

$$P(H|E) = \frac{[P(E|H) \text{X} P(H)]}{P(E)}$$

Where

> H is all expected event or hypothesis
> E is all events possible
> P(H) – probability for event H to happen
> P(H|E) – probability for an event H to happen when event E happens

## 6.3. Decision tree classifier

Decision tree classifier is a classifier which implement an effective hierarchical classifiers. Its tree structure is easily understandable, besides being able to execute automatic feature selection, and can help reduce complexity [17].

## 6.4. KNearest Neighbor classifier

K-nearest neighbor (k-NN) which is said to be the simplest algorithm in machine learning, is a classifier that classifies objects based on the closest training examples in the feature pace [14]. It computes the distance from an unknown test pattern to every training pattern and selects the K nearest training samples to base the classification on [18,26]. It has been implemented on pattern recognition systems because of its good performance and simple algorithm [27].

> *Finding k value*:
>
> One of the most important things in applying k-nearest neighbour is selecting a suitable value for k parameter. The apposite value of k generally depends on the dataset. Value of k which is too small may contribute to over fitting, while overly large k value will increase computational complexity thus affecting decision making process [14]. To break ties, it is common to select small and odd value of k, typically 1, 3 or 5 [11].
>
> *Finding distance metric*:
>
> The performance of *k*-NN essentially depends on the distance metric used to identify the nearest neighbours. Study shows that classification accuracy of *k*-NN can be significantly improved by selecting a suitable distance metric for the dataset. There are various types of distance metrics to determine the distance for *k*-NN classifiers such as

Euclidean distance, city block metric, cosine distance and correlation distance. The following formulas are used to compute each of the distance metrics [14].

Euclidean distance[14]*:*

$$d^2{}_{st} = (x_s - y_t)(x_s - y_t)$$

City Block Metric[14]:

$$d_{st} = \sum_{j=1}^{n} | x_s - y_t |$$

Cosine distance[14]:

$$d_{st} = \left( 1 - \frac{x_s y_t}{\sqrt{(x_s x_s)(y_t y_t)}} \right)$$

Correlation distance[14]:

$$d_{st} = \left( 1 - \frac{(x_s - \bar{x}_s)(y_t - \bar{y}_t)'}{\sqrt{(x_s x_s)(x_s x_s)} \sqrt{(y_t y_t)(y_t y_t)}} \right)$$

## 6.5. Neural network

The neural network processes consist of steps which are the learning and testing process. The back propagation technique is used for neural network training process[10]. It is commonly used for pattern recognition, image classification, and medical analysis. Training process can obtain optimum weight; therefore neural network can operate successfully and produce targeted results Multi layer perception (MLP) is the most familiar neural network model, consisting of succeeding linear transformations followed by processing with non-linear activation tasks [10,25].

The learning algorithm for multilayer can be conveyed using generalized Delta Rule and gradient descent since they are having non-linear activation tasks [25]. Radial basis function (RBF) is a network for approximating functions which consists ofmodeling an input-output mapping as a linear groupingof radially symmetric tasks. RBFnetwork has rapidtraining and simple system and does not necessitate a heavy iterative training process asin MLP [10].

General regression neural network estimates any random function involving input and output vectors, then drawing the function estimated directly from the training data [24]. It is related to the radial basis function network and is based on a standard statistical procedure called kernel regression [24]. For the probabilistic neural network, Gaussian distribution is assumed. However, the assumption of normality does not always be securely justified. When the distribution is not known and the exact distribution diverges considerably from the assumed one, the conventional statistical methods normally run into major classification problems thus results in high rate of misclassification [19].

A research made by Riries (2002) used neural network to classify tuberculosis bacteria in which result in an excellent output [10]. Among of the severe diseases which can be treated when early diagnosed are tuberculosis and other mycobacteriosus diseases. Traditionally, in order to identify tuberculosis bacteria, sputum images is analysed to detect the presence of bacilli. This method however consume a lot of time and requires specialists to avoid large errors. Therefore, neural network is used resulting in a mean square error of 0.000368 and zero error classification when a

number of new data is being classified [10]. Figure 4 shows the result of tuberculosis classification using neural network [20].

| Image | Geometric Features | NN Classification Result |
|---|---|---|
| | • Circularity:      0.798063<br>• Compactness: 15.7461<br>• Eccentricity: 1.00187<br>• Tortuosity:      0.295393 | TUBERCULUS BACTERIA |
| | • Circularity:      0.804989<br>• Compactness: 15.6106<br>• Eccentricity: 1.00014<br>• Tortuosity:      0.294254 | TUBERCULUS BACTERIA |
| | •Circularity:      0.801031<br>• Compactness: 15.6877<br>• Eccentricity: 1.00237<br>• Tortuosity:      0.29493 | TUBERCULUS BACTERIA |
| | • Circularity:      0.73098<br>• Compactness: 17.1911<br>• Eccentricity: 1.02675<br>• Tortuosity:      0.303604 | NOT TUBERCULUS BACTERIA |
| | • Circularity:      0.742655<br>• Compactness: 16.9209<br>• Eccentricity: 1.02937<br>• Tortuosity:      0.305133 | NOT TUBERCULUS BACTERIA |

Figure 4. Result of tuberculosis classification using neural network[10]

## 7. LIMITATIONS AND CHALLENGES

Traditional techniques of bacterial classification rely on phenotypic identification using gram staining, biochemical process as well as culture methods. However, these techniques of bacterial identification have two major problems.  They can be used only for organisms which are cultivated in vitro and show unique biochemical attributes that are not suitable for patterns that have been used as a characteristic of any known groups of microorganism.In the recent years, molecular procedures have been proven to be auspicious in defeating the limitations of traditional phenotypic methods for the detection and classification of bacteria. Real time PCR and microarrays is also among of the most commonly used molecular methods. However, it is highly sensitive besides permitting bacteria quant at species level.

## 8. CONCLUSION

Further research in classification of bacteria using machine learning would be of great help in microbiological field. It introduces a new and efficient method to categorize bacteria classes to replace the inefficient old procedure which is tedious and time consuming. This new method which uses pattern recognition would feasibly help to accurately tell the classes of bacteria which have caused many types of life threatening diseases and infections.  This would directly help to save lives of millions of people, should the detection of their diseases is done at an early stage.

## REFERENCES

[1] S.Prabakar, et al., "Development of Image Processing Scheme for Bacterial Classification Based on Optima Discriminant Feature",Proceedings of the 2010 International Conference of Bioinformatics and Biometrics Engineering, 2010.

[2] L.Frank, "Bacterial Classification, Structure and Function", 2009, pp.1-2.

[3] Z. Ying, et al., "Bacteria Classification Using Neural Network", Institute of Biology Gansu Academy of Sciences, 2011.

[4] O.Richard, et al., "Pattern Classification", 2nd ed., Wiley-Interscience, 2000.

[5] Y. Huiyong, et al., "Research on Insect IdentificationBased on Pattern Recognition Technology", Taiyuan Extension Station of Application Technology, 2010.

[6] S. Sapna, et al., "Comparative Study of Image Segmentation Techniques and Object Matching Using Segmentation", Methods and Models in Computer Science, 2009.

[7] T. Wenbing, J.Hai, "Color Image Segmentation Based on Mean Shift and Normalized Cuts", Man and Cybernetics, 2007.

[8] S.Mohsen, et al., "A Classified and Comparative Study of Edge Detection Algorithms", Information Technology:Coding and Computing, 2002.

[9] R.Pekka, et al.,"Efficient Automated Method for Image-Based Classification of Microbial Cells", Department of Chemistry and Bioengineering Tampere University of Technology, 2008.

[10] R.Riries,"Automatic Classification of Tuberculosis Bacteria Using Neural Network", International Conference on Electrical Engineering and Informatics,2001.

[11] J.Mohammed,et al., "Investigating the Performance of Naive- Bayes Classifiers and K-Nearest Neighbor Classifiers",Department of Electrical and Computer Engineering University of Windsor, 2007.

[12] K.Roy, G.Subrahmanya, "Combining Feature Descriptors for Efficient Classification and Recognition of Images", 5th International Congress, Image and Signal Processing (CISP), 2012.

[13] N.Y. Khan, et al., "SIFT and SURF Performance Evaluation Against Various Image Deformations on Benchmark Dataset",Digital Image Computing Techniques and Applications (DICTA), 2011.

[14] Z. Supardi, et al., Classification of Blasts in Acute Leukemia Blood Samples Using K-Nearest Neighbour", Universiti Malaysia Perlis, 2012.

[15] B.Oren, "In Defense of Nearest-Neighbor Based Image Classification", The Weizmann Institute of Science. 2008.

[16] M.Nittaya, et al., "Classification Students with Learning Disabilities Using Naive Bayes Classifier and Decision Tree", Faculty of Science and        Engineering, Kasetsart University.

[17] P. Mahesh, M. Paul, "A Comparison of Decision Tree and Backpropagation Neural Network Classifiers for Land Use Classification",  Geoscience and Remote Sensing Symposium, 2002.

[18] S.Hojijatoleslami, J.Ecittler, "Detection Of Clusters Of Microcalcificatioin Using A K-Nearest Neighbour Classifier", IEEE, 1996.

[19] T. Yildirim, Classification of Escherichia Coli Bacteria by Artificial Neural, First International IEEE Symposium "Intelligent Systems",2002, p.14.

[20] G.Manuel, et al., "Identification of Tuberculosis Bacteria Based On Shape and Color", Real-Time Imaging, Volume 10, Issue 4, 2004, pp. 251–262.

[21] Z.Z. Htike, S.L. Win "Recognition of Promoters in DNA Sequences Using Weightily Averaged One-dependence Estimators", Procedia Computer Science, Volume 23, 2013, Pages 60-67, ISSN 1877-0509.

[22] Z.Z. Htike, S.L. Win "Classification of Eukaryotic Splice-junction Genetic Sequences Using Averaged One-dependence Estimators with Subsumption Resolution", Procedia Computer Science, Volume 23, 2013, Pages 36-43, ISSN 1877-0509.

[23] Z.Z. Htike, et al., "A Monocular View-Invariant Fall Detection System for the Elderly in Assisted Home Environments," 7th International Conference on Intelligent Environments (IE), 2011, pp.40-46, 25-28 July 2011.

[24] M.Tatlier, et al., Artificial neural network methods for the estimation of zeolite molar compositions that form from different reaction mixtures, Computers and Chemical Engineering, Volume 30, Issue 1, 15 November 2005, pp.137-146.

[25] L. Ozyilmaz, T. Yildirim, "Diagnosis of thyroid disease using artificial neural network methods" Proceedings of the 9th International Conference on Neural Information Processing, 2002vol.4, pp.2033-2036 vol.4, 18-22 Nov. 2002.

[26] K.S. Woods, et al., "Comparative Evaluation of Pattern Recognition Techniques for Detection of Microcalcifications in Mammography", Series in Machine Perception and Artificial Intelligence: Volume 9, 1994, pp. 213-231.

[27] S. Hotta, et al., "Pattern recognition using average patterns of categorical k-nearest neighbors",Proceedings of the 17th International Conference on Pattern Recognition, 2004. vol.4, pp.412-415 23-26 Aug. 2004.

[28] Z.Z. Htike, "Can the future really be predicted?," Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 2013 , pp.360-365, 26-28 Sept. 2013.

[29] Z.Z. Htike, "Multi-horizon ternary time series forecasting," Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 2013, pp.337-342, 26-28 Sept. 2013.

[30] S.L. Win, et al., "Cancer Recurrence Prediction Using Machine Learning", International Journal of Computational Science and Information Technology (IJCSIT), 2014, Vol. 6, Issue 1, ISSN: 0975-3826.

[31] S.N.A. Hassan, et al., "Vision Based Entomology: A Survey", International Journal of Computer science and engineering Survey (IJCSES), 2014, Vol. 5, Issue 1, ISSN: 0976-2760.

[32] E.M. Azhari, et al., "Brain Tumor Detection and Localization in Magnetic Resonance Imaging", International Journal of Information Technology Convergence and services (IJITCS), 2014, Vol. 4, Issue 1, ISSN 2231-1939.

[33] N.A. Mohamad, et al., "Bacteria Identification from Microscopic Morphology Using Naïve Bayes", International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), 2014, Vol. 4, Issue 1, ISSN 2231-3117.

[34] S.L. Win, et al., "Cancer Classification from DNA Microarray Gene Expression Data Using Averaged One-Dependence Estimators", International Journal on Cybernetics & Informatics (IJCI), 2014, Vol. 3, Issue 1, ISSN: 2320-8430.

[35] E.M. Azhari, et al., "Tumor Detection in Medical Imaging: A Survey", International journal of Advanced Information technology (IJAIT), 2014, Vol. 4, Issue 1, ISSN 2231-1548.

[36] S.N.A. Hassan, et al., "Vision Based Entomology – How to Effectively Exploit Color and Shape Features", Computer Science & Engineering: An International Journal (CSEIJ), 2014, Vol. 4, Issue 1, ISSN: 2231-3583.

[37] S.L. Win, et al., Cancer Survivability Prediction Using Semi-Naïve Bayesian Models, International Journal on Bioinformatics & Biosciences, 2014, Vol. 4, Issue 1, ISSN: 1839-9614.