International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 08 | Aug 2020 www.irjet.net p-ISSN: 2395-0072
National Conference on Recent Advancements in Communication, Electronics and Signal Processing-RACES'20
Organised by: Department of ECE, Velammal Engineering College, Chennai-66

BACTERIA TYPE IDENTIFICATION USING ARTIFICIAL NEURAL NETWORK

Ada Christa C[1], Akshaya D[2], Ankitha M[3]

[1]Assistant Professor, Department of Electronics and Communication Engineering
Prathyusha Engineering College, Chennai
[2,3]UG Scholars, Department of Electronics and Communication Engineering
Prathyusha Engineering College, Chennai

---***---

Abstract: Bacterial type identification plays a vital role to control diseases. Till date biological and microbiological data analysis requires human help which involves lot of time, energy and great cost. To overcome this drawback, to reduce human intervention and to handle increasing volume of data, in this work, efficient method is proposed to identify the type of bacteria automatically. The basic methodology used is pattern recognition technique. The pattern recognition consists of the steps, image segmentation, feature extraction and classification. The desired image is first segmented, then required features are extracted. The back propagation neural network is trained using these training dataset features as input. Then the network is tested with testing dataset features. The efficiency obtained in this proposed work is high compared to other methods.

Keywords: Bacteria classification, pattern recognition, feature extraction.

Introduction

Bacteria is a prokaryotic microorganism, which is small and range in size less than ten micrometres (µm).Bacterial identification is important to understand the type of disease and treat them appropriately. Different techniques are available in microbiology to classify bacteria but the classification is not satisfactory due to increase in bacterial adaptation to human environment. Therefore appropriate and efficient bacterial type diagnosis is required for disease control. The bacterial classification is mainly achieved based on bacterium's shape and cell arrangement. The most ordinary shapes of bacteria include rod,cocci (round) and spiral forms. Cellular arrangement occurred singularly, in series, or in groups. One of the special projections called flagella enables the bacteria to swim and move. The bacterial type identification accomplished using image processing.

According to H.C. Gram bacteria as either Gram positive or negative based on their morphology and differential staining properties [2]. The other types of classification is based on their functions and structures. Till date, human intervention [1] is required for bacterial classification. The main drawback of this methodology is inconsistency in work and also complicated, tedious which requires abundant correlative data [3]. The volume of data available for analysis is also enormous, thus time, energy and cost involvement is very high [1]. Hence human involvement should be reduced. Table 1 shows the classification of bacteria.

Table 1: Classification of Bacteria

Name	Morphology
Staphylococci	Cocci in grape like clusters
Streptococci	Cocci in pairs, chains
Enterococci	Cocci in pairs, chains
Bacilli	Rods, spore forming
Clostridia	Rods, spore formers

Methodology

The automatic bacterial classification pattern recognition techniques. The pattern recognition system includes subsystem to define pattern class subsystem to extract selected features and also subsystem for classification known as classifier. Figure 2 shows the block diagram of pattern recognition system. The main aim of pattern recognition is the classification of patterns and sub patterns in an image.

The three modules are:

- Pre processing
- Feature extraction
- Bacteria classification

Pre-processing

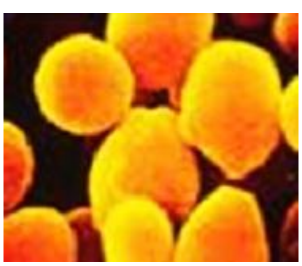
The purpose of pre-processing is to remove the noise from the image.

Feature Extraction

Classification is mainly dependent on features therefore the features extracted should consists of vital data about bacteria [9]. Therefore for bacteria classification, feature extraction is performed to determine the morphology of mycobacterium from their shape. Therefore the geometrical features used are perimeter, area, radii, circularity, compactness, eccentricity, and tortuosity of the samples [10]. Features should remain unchanged even if condition changes and such features are called invariant features [13]. These feature descriptors are stored during training phase and during testing phase, the considered image is compared with all training patterns and best match is selected.

The three main approaches for feature extraction and classification based on the type of features are:

- Statistical approach: This approach is useful for random patterns.
- Syntactic or structural approach: This approach is widely used for complex patterns which gives better results.
- Spectral approach: This approach is less efficient because of textures of spatial frequencies and are evaluated by autocorrelation function of a texture.



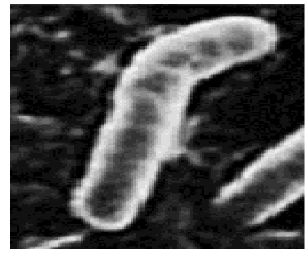


Figure 1: Three classes of bacteria (a) Cocci(b) Vibrio (c) Bacilli

The different features extracted from bacterial image are Relative length, Relative area, Mean, Standard deviation, Entropy, Contrast and Variance

Relative length: The length of each bacterial images is determined by counting the number of pixels in the medial line. It can be determined by normalizing the medial axis using the following equation.

$$l_n = \frac{l_i}{l_t} \quad (1)$$

where l_i is the length of i-th bacterial images
l_t is the total length of 46 bacterial images of one cell

Relativearea: The relative area of the i-th bacterial images can be obtained by counting the pixels of the c bacterial images body and by normalizing the areas using the following equation.

$$A_n = \frac{A_i}{A_t} \quad (2)$$

where A_i is the area of the i-th bacterial images
A_t is the total area

Mean:

$$\mu_{p=\frac{1}{n^2}} \sum_{r=0}^{n-1} \sum_{s=0}^{n-1} p_{r,s} \quad (3)$$

where $p_{r,s}$ is pixel at location r,s.

Standard deviation:

$\sigma_p =$ (4)

Entropy: It is a measure of randomness

$$e=-\sum_{b=0}^{L-1} p(b) log_2 p(b) \quad (5)$$

where p(b)=N(b)/n^2 for {0<=b<=L-1},
where L is the number of different values which pixels can adopt
N(b)=number of pixels of amplitude (b) in the pixel window of size 'n×n'.

Contrast: Contrast is defined as

$$\sum_{n=0}^{N_g} n^2, \quad (6)$$

where n=$\sum_{i-1}^{N_g} \sum_{j-1}^{N_g} p_{i,j}$|i-j|
where n= number of gray levels.

Variance: Variance is the measure that tells how much the gray level are varying from the mean

$$\text{Variance}=\sum_i \sum_j (i,j) p_{i,j} - \mu^2 \quad (7)$$

Bacterial images classifier

In the classification phase the unprocessed data is categorized into certain classes according to the features [11]. The main objective of classification is to make a decision on the dataset based on the information obtained during testing phase.Including neural network in classification consists of two steps as learning and testing phase. The back propagation is used for training phase [10], which learns efficiently during pattern recognition, image classification and medical analysis [10, 25]. In general Gaussian distribution is assumed but when the distribution is not known prior then the conventional statistical methods will result in high misclassification rate [19]. The research work by Riries in 2002, classified tuberculosis bacteria very efficiently using neural network.

Results and Discussion

The input to the feature extraction algorithm is the bacterial images. The pattern vectors extracted from the images is given as an input to the classifier. In this system the sample of 180 bacterial images are collected and 120 images are used for training phase and remaining for testing. The classification efficiency is obtained by

$$\text{Accuracy}=\frac{totalno.of bacterialimages - no.of misclassified}{TotalNumberof Bacterialimages} \quad (8)$$

The total number of images considered for evaluation of the model is 100. Thus in this total dataset, 75 dataset is considered for training and 25 dataset is considered for testing. The classification efficiency obtained for training dataset is 92% and classification efficiency obtained for testing dataset is 88%. The classification efficiency obtained for combination of training and testing dataset is 90%. Table 2 tabulates the result of bacterial type identification efficiency.

Table 2: Bacterial type identification classification efficiency

	No. of images as input	Correctly classified data	Classification efficiency
Training dataset	75	68	92%
Testing dataset	25	22	88%
Average	100	90	90%

Conclusion

In this work, an effective methodology to identify the bacterial type automatically with less intervention of human being is proposed. The input images is segmented for the required input and seven features are extracted from these images. These features are given as input for the back propagation algorithm. During testing phase the efficiency obtained is 92% and during training phase the efficiency obtained is 88%. Thus on an average the classification efficiency obtained is 90%.

References

[1] Trattner S, Greenspan H, Tepper G, Abboud S., "Automatic identification of bacterial types using statistical imaging methods", IEEE transactions on medical imaging. 2004; 23(7):807-820.
[2] Blackburn N, HagstroÈm Å, Wikner J, Cuadros-Hansson R, Bjørnsen P K., "Rapid determination of bacterial abundance, biovolume, morphology, and growth by neural network-based image analysis",

Applied and Environmental Microbiology, 1998; 64(9):3246-3255.

[3] Perner P., "Classification of HEp-2 cells using fluorescent image analysis and data mining", International Symposium on Medical Data Analysis. Springer; 2001. p. 219-224.

[4] Hiremath P, Bannigidad P., "Automated gram-staining characterization of digital bacterial cell images", Proc. IEEE Int'l. Conf. on Signal and Image Processing", 2009. p. 209-211.

[5] De Bruyne K, Slabbinck B, Waegeman W, Vauterin P, De Baets B, Vandamme P, "Bacterial species identification from MALDI-TOF mass spectra through data analysis and machine learning", Systematic and applied microbiology. 2011; 34(1):20-29.

[6] Forero MG, CristoÂbal G, Desco M. Automatic identification of Mycobacterium tuberculosis by Gaussian mixture models", Journal of microscopy. 2006; 223(2):120±132.

[7] Liu J, Dazzo FB, Glagoleva O, Yu B, Jain A K., "CMEIAS: A computer-aided system for the image analysis of bacterial morpho types in microbial com-munities", Microbial Ecology. 2001; 41(3):173-194.

[8] Ahmed WM, Bayraktar B, Bhunia AK, Hirleman ED, Robinson JP, Rajwa B. Classification of bacterial contamination using image processing and distributed computing. IEEE journal of biomedical and health informatics. 2013; 17(1):232-239.

[9] Cortes C, Vapnik V. Support-vector networks. Machine learning. 1995; 20(3):273-297.

[10] Holmberg M, Gustafsson F, HoÈrnsten EG, Winquist F, Nilsson LE, Ljung L, "Bacteria classificationbased on feature extraction from sensor data", Biotechnology techniques. 1998; 12(4):319-324.

[11] Ates H, Gerek O N, "An image-processing based automated bacteria colony counter", Computer andInformation Sciences", 2009. ISCIS 2009. 24th International Symposium on. IEEE; 2009. p. 18-23.

[12] Sommer C, Gerlich D W., "Machine learning in cell biology teaching computers to recognize phenotypes", J Cell Sci. 2013; 126(24):5529-5539.