Review Article **Bioinformatics**

International Journal of Pharma and Bio Sciences

ISSN
0975-6299

CLASSIFICATION OF DNA BARCODES BASED ON IMAGE PROCESSING TECHNIQUES: A STUDY

S. JYOTHI*[1] AND K. BHARGAVI[2]

[1] Professor, Dept. of Computer Science, Sri Padmavathi Mahila VisvaVidhyalam, Tirupati, A.P
[2]Research Scholar, Dept. of Computer Science, Sri Padmavathi Mahila VisvaVidhyalam, Tirupati, A.P

ABSTRACT

DNA barcoding is a method of identifying the species in biological scientific community without involving the morphological clues. DNA barcoding is a molecular method and which uses DNA from a standardized region of the genome to identify the species. Barcodes are machine readable and capable of storing digital information. The image processing based barcode reading systems started to gain importance which provide more information than laser barcode readers at a time. In this paper, a few traditional barcode methods, computer vision based decoding techniques, image processing based barcode reading system & methods and its performance issues are discussed. Image processing based barcode recognition systems are expected to increase the performance in barcode reading. Preprocessing, segmentation, edge detection and optimization techniques were used in DNA barcode classification.

KEYWORDS: **DNA barcode, Barcode recognition systems, Decoding techniques, Segmentation and Edge detection**

S. JYOTHI
Professor, Dept. of Computer Science, Sri Padmavathi Mahila VisvaVidhyalam, Tirupati, A.P

***Corresponding Author**

INTRODUCTION

Barcode technology is one of the most important parts of Automatic Identification and Data Capture (AIDC). DNA barcoding is a technique, which provides quick identification of species without involving the morphological cues.[1] It uses a relatively small standardized DNA fragment as a tag to define or discover the species. DNA barcode information is used to species identification and species circumscription. Advances in species assignment via DNA barcoding have been achieved, especially through the framework of Bayesian theory. Bayesian method provides the necessary statistical strength to distinguish between well and poorly supported assignments and, most importantly, provide a measure of statistical confidence.2 DNAzymes are DNA molecules with catalytic action also called as DNA enzymes, catalytic DNA or deoxyribozymes[3] . Barcode species allocations are generally only based on DNA information. In the situation of insufficient information concerning the species studied, there are two principal ways to deal with the data: one is the development of sophisticat ed statistical methods, such as the Bayesian methods mentioned above, and another is based on the fuzzy set method. The former has been extensively studied.[4] The barcodes are easy to use and make possible to enter the data more quickly than by manual methods and are highly reliable. Their reliability and speed allow for improving of many operations such as forwarding, reception, packing and manufacturing. DNA is one of the most essential molecules in organisms, containing all the information necessary for organisms to live.5 DNA barcode should be normally a uniform short sequence of DNA (400-800 bp), able to be simply generated and used to characterize all the living organisms. DNA barcoding is very essential for the molecular identification of already described species and the discovery of new species.[6, 7] Numerous biological experiments have demonstrated that DNA is the primary intracellular target of anticancer drugs; interaction between small molecules. A new method of DNA reading was proposed by Nagakavitha et.al.[8] According to the authors' claims it has certain advantages as compared to the Maxam-Gilbert and Sanger methods, which are revealed by automation and rapidity of DNA sequencing. Nevertheless its employment is hampered by a number of biological and mathematical problems. They proposed an algorithm that allows overcoming the computational difficulties occurring in the course of the method during reconstruction of the DNA sequence by its 1-tuple composition. Genomic approaches to tax on diagnosis exploit diversity among DNA sequences to identify organisms.9 In a very real sense, these sequences can be viewed as genetic 'barcodes' that are embedded in every cell. Image process sing based barcode reading software "Softek Barcode Reader" and "DTK Barcode Reader SDK" are two examples developed by DTK Software and Softek Software companies.[10] Many compounds expose their anti tumour activity through binding to DNA and can cause DNA damage.[11]

BARCODE TYPES

The vast majority of modern barcodes fall into two basic categories that are linear (1D) and 2-dimensional (2D).[12] 3D barcodes do exist; they are generally never seen and used in life sciences. The 1D barcode is a ubiquitous labelling technology.[13] UPC codes are 1-D, meaning they only carried information in one direction . One-dimensional codes work fine for carrying small amounts of data like numeric product codes. Many different 1-D coding standards are EAN13,EAN8,EAN128, UPCA, and UPC-E. However, E AN13 barcode standard is the most common one in Turkey and throughout Europe. Linear (also known as 1D or 1-dimensional) barcodes encode data using vertical bars and spaces. The most common format in life science and clinical laboratories is Code 128 since this standard can encode all 128 ASCII characters and requires the smallest amount of space compared to other linear barcode standard.

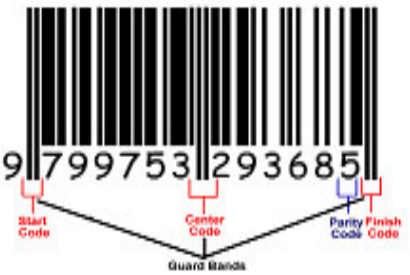

Figure 1
Example of 1D barcode (EAN13 barcode)

2-D barcodes (sometimes called matrix codes) carry information in two directions: vertically and horizontally. The first application for such symbols is unit dose packages in the healthcare industry. Accordingly, 2-D barcodes are capable of holding tens and even hundreds of times as much information as 1-D bar codes.[13, 14] For instance, one of the most popular 2-D barcode formats, Denso Wave's QR Code, can hold more than 7,000 digits or 4,000 characters of text, whereas even the most complex 1-D codes top out around 20 characters. Quick Response code (QR code) is one of the most popular types of two dimensional barcodes dimensional (2-D) barcodes use a grid pattern of data elements to encode information in both

the horizontal and vertical directions. By using both dimensions, 2D barcodes can encode much more information in a smaller area compared to a linear barcode.[16] While most people are probably more familiar with the QR 2D barcode format used by websites.

Figure 2
Example of ID Barcode and 2D Barcode (QR CODE).

The main disadvantage of linear barcodes is size. Since only single dimension is used to encode data, the density of 1D barcodes is dramatically less than 2D barcodes**.** 2D Barcode Disadvantages are High-density encoding of information, more difficult to scan. Even with error correction each element in the barcode encodes a piece of information.

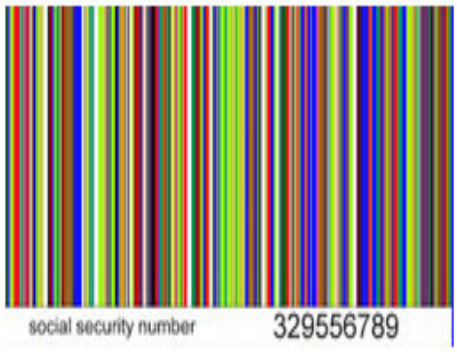

Figure 3
Example of color barcode

3 dimensional barcodes are like in 2D barcodes, can contain different types of information like pricing, height, weight and other product information. The need to this system began to avoid the problems occurred by high temperature, chemicals and solvents that would destroy any barcode as in linear or 2 dimensional barcodes. The 3 dimensional barcodes make it almost unacceptable to change or prohibit the barcode's information.

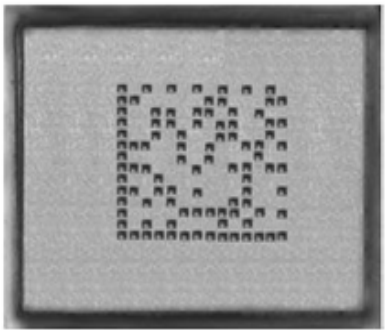

Figure 4
Example 3 D Barcode

BARCODE RECOGNITION SYSTEM

To decode barcode information, widths of the dark and light regions on the barcode must be known. Peak locations of the dark and light regions are found with statistical pattern recognition.[17] Barcode information is deciphered based on the distance between the peak locations simple thresholding applied to all pixels.[18] If the pixel value is less than threshold, it is replaced by zero.[19] Then the number of zeros is counted to obtain widths of the barcode lines. Detection of the transition between the two barcode regions, transition values are obtained by zero-crossings of the second derivative of the zarcode.[20] Then, distance measurement is performed between the light and dark barcode lines and barcode is decoded using the distance information. The decoding process requires the widths of the barcode lines to be known.[21, 22] In order to do this, the midpoints of the dark bars and light spaces could be used since the images at hand may be noisy or low quality. However, quite often barcode area obtained via barcode localization may not include only the barcode itself. There may be some areas which are not actually a part of the barcode or that area may not completely enclose the barcode. The extracted barcode area must be checked if start and finish are included. In the following, various decoding techniques are discussed.

DECODING TECHNIQUES

Image Approximation with B-Spline Smoothing

A B-spline function consists of polynomial pieces on subintervals joined together with certain continuity conditions.[23] Spline Functions with equally spaced knots are called Cardinal Splines. B-spline smoothing based estimation is applying to get the derivative of barcode.[24, 25] This result is inaccurate in the calculation of midpoint values. In order to overcome this problem, a threshold is used.

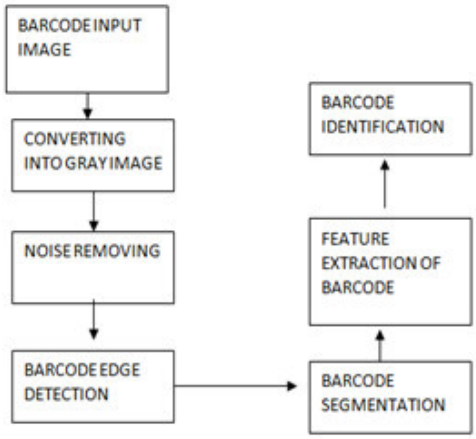

Figure 5
Block Diagram of Barcode Detection Process

Then we detect the point where the line intersects zero. Hence, the midpoints of the bars and spaces on the barcode are found. Midpoints are used to detect bar/space borders, which are later used to reconstruct a higher quality barcode sample by least squares approximation. Finally, barcode data is deciphered by use of bar/space widths and match filtering. The barcode recognition block takes in the barcodes and tries to match up the barcode with the numbers of pixels generated from the bar detection.

Decoding EAN-13 Barcodes from Images Captured By Digital Cameras

A decoding technique of vision-based technique is used to locate and decode EAN-13 barcodes from images captured by digital cameras.[26] The proposed algorithm identifies parallel line patterns at block level to determine the location of the barcode within the given input image. Once the barcode image region is identified, the decoding process is carried out according to the EAN-13 specification.

SVD (Singular Value Decomposition) method

An alternative sequence encoding method by adopting the Latent Semantic Indexing (LSI) analysis is used in the field of information retrieval and information filtering.[27,28] The sequence encoding schema is used to convert molecular sequences (character strings) into input vectors (numbers) of the neural network classifier.[29]

Quick Response (QR) Decoding

QR codes are now used over much wider range of applications, including commercial tracking, transport ticketing, website Uniform Resource Locator (URL), and identity verification. QR code is a kind of matrix 2D (two-dimensional) barcode designed by Denso-Wave Corporation of Japan in 1994. Its benefits are high-speed recognition, strong error-correcting capability, effectual expression for Chinese characters and all-directional recognition. A two-dimensional bar code, QR (Quick Response) code, is usually used to specify the URL of the Internet web services. This QR Code was standardized as the JIS standard (X 0510) in 2000 and revised in 2004.[30] First the version of QR code is selected by considering the coding capability of the data information and the smallest number is chosen for information accommodation.

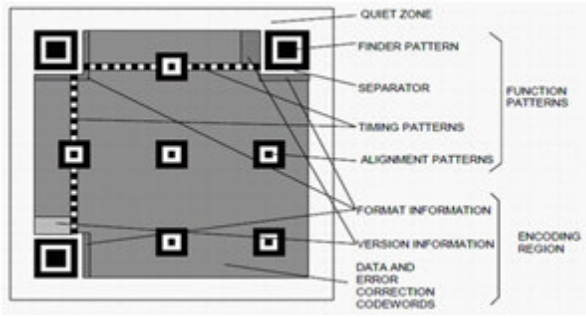

Figure 6
The Structure of QR Code

Fuzzy-based identification

Most methods of species assignment make binary decisions regarding classification of unknown samples, i.e. the sample belongs or does not belong to taxon X.[31, 32]. DNA barcoding uses a single gene (or combination of gene) that constitutes incomplete information to assign an unknown specimen to a species. Species membership, therefore, could be viewed in probabilistic terms, i.e. understood on a continuous scale. Fuzzy set theory allows for a scaled evaluation of membership in which species assignment takes a value between 0 and 1, referred to as a fuzzy membership function (FMF).[33]

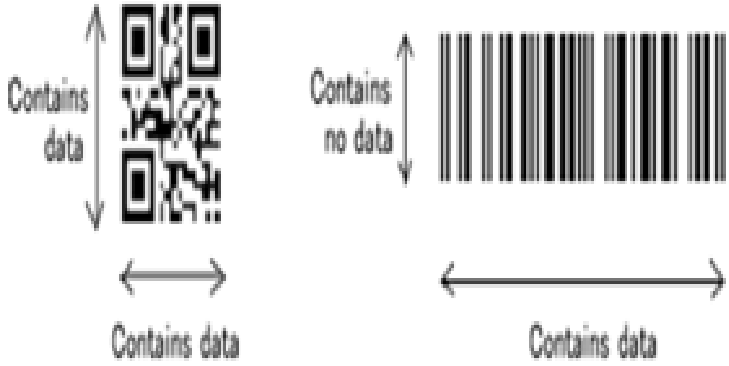

Figure 7
a) QR code b) Barcode

Analogous to a barcode gap analysis, fuzzy membership is evaluated using two parameters, the maximum intra-specific genetic distance (θ1) and minimum inter-specific genetic distance (θ2), both calculated as K2P distances directly from each potential species, when possible, and not averaged across the dataset, i.e. parameters are computed using all individuals from the known species against which a query sequence is being compared, along with all members of the nearest neighbouring species in the dataset. When a species is limited to just one sample, however, θ1 is estimated as an average among species. Thus, the fuzzy method is also able to assign singleton samples to their own group, which helps alleviate the problem of false positive assignments.[34]

Statistic Assignment Package (SAP)

This method uses a Bayesian phylogenetic approach to species assignment. For every query sequence a database, first a search is performed using BLAST to find homologous DNA sequences in GenBank.[35] Only homologues with a BLAST score equal to or greater than half of the best matching homologue are used, except when fewer than 50 best hits are recovered, in which case less similar sequences are included until a set of 50 potential homologues is obtained. Sequences are then aligned using ClustalW and a large set of phylogenetic trees is sampled using Bayesian Markov chain Monte Carlo sampling of the posterior distribution of trees.[36, 37]. Lastly, every query sequence is assigned a posterior probability of belonging to nested monophyletic groups corresponding to the taxonomic hierarchy. In the case of the Panama dataset the analysis was performed using default parameters and the program took about 2 weeks to finish. Eleven sequences in particular caused the program to crash. Fortunately, incomplete runs can be restarted without losing prior results. These troublesome sequences were eliminated from the input in order to continue with the analysis, resulting in 261 out of the 272 sequences used as query sequences. The analysis was also performed using only COI sequences for the Panama dataset, in which case the analysis took 17 h. For the Malagasy dataset the analysis was performed using 52 sequences and the program took 210 min to finish. All analyses were performed on a Mac OSX 2.66 GHz Intel Core 2 Duo.

A Skew Detection Method for 2D Barcode Images Based on the Least Square Method

The fast and robust algorithm for skew detection in 2D barcode images is based on the least Square Method.[38,39,40] Unlike the methods based on Hough Transform that are computationally expensive, it quickly obtains skew angles making it applicable to real-time applications and this method includes two processes; the segmenting process searches for barcode region and the line fitting process fits the barcode line and obtains the skew angle. Thus the Experimental results show this method reduces the running time. Line information of the barcodes borders can search for the border's skew angle, even when the detected image has high noise.

Support Vector Machines (SVM)

Support Vector Machines (SVMs) have long been considered as powerful classification tools and often yield better results than traditional neural networks.[41, 42.] Constructs a hyper plane which is used to separate classes in the feature space has the largest distance (margin) to the nearest training data points.[43] A spectrum kernel for DNA barcoding sequences analysis used the statistics of short substrings of length *k* (*k*-mers) contained in the original sequence to transform the arbitrary long sequences into fixed-length representations.[44]

CONCLUSION

This study focuses on barcode recognition using image processing techniques. In this study, barcodes classification, decoding techniques and barcode methods are discussed. Their performance metrics are also compared individually. Barcode recognition is highly accurate, powerful, fast, and reduces manual work.

REFERENCES

1. Kopila Pariyar. Process of Detecting Barcodes Us ing Image Processing. International Journal of Sci entific Engineering and Research (IJSER).[Internet].2014 [cited 2014 July]; 2(7):2347-3878. Available from: http://www.ijser.in/archives/v2i7/SjIwMTMzMjY=.p df
2. Lou M, Golding G.B. Assigning sequences to species in the absence of large inter specific differences. Mol. Phylogenet. [Internet].2010[cited 2010 Jan]:56: Available from:187-194.http://www. ecbol.org/docs/Publications/Lou_2010.pdf
3. Babunageswararao K, Ramakrishna R, Kamal Kumar P.D and Syed Sameer. Dnazymes as Novel Therapeutics. International Journal of Pharma and Bio Sciences.[Internet].2010[cited 2011 Oct]: 2(4): 161-173. Available from:http://www.ijpbs.net/archive-issue. php?issueid=16
4. Abdo Z, Golding G.B. A step toward barcoding life: a model-based, decision theoretic method to assign genes to pre-existing species groups. Systematic Biology. [Internet].2007[cited 2007]: 56(1): 44–56. Available from: http://sysbio.oxfordjournals.org/co ntent/56/1/44.full.pdf+htmlDOI:10.1080/10635150 601167005
5. Chandrasekhar S, Moksharagni B and Hussain R eddy K.Synthesis, Spectral Characterization And DNA Binding Studies Of Dinuclear Copper(II) Complexes With 1, 10-Phenanthroline. International Journal of Pharma and Bio Sciences.[Internet]2015 [cited 2015 Oct] 6(4): 25-39.Available from: http://www.ijpbs.net/c ms/php/upload/4577_pdf.pdf
6. Cortes C, Vapnik V.Support-vector network. Machine Learning.1995[cited 1995]; 20:273–297.Available from: http://image.diku.dk/imagecan on/material/cortes_vapnik95.pdf
7. Altschul S.F, Madden T.L, Schaffer A.A, Zhang J, Zhang Z, Miller W, Lipman D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research.[Internet].1997 Sep[cited 1997 June20]; 25(17): 3389 3402.Avail able from: http://www.ncbi.nlm.nih.gov/pmc/article s/PMC146917/pdf/253389.pdf
8. Nagakavitha D and Hussain Reddy K. Synthesis, Spectral Characterization and DNA Binding Properties of Copper(II) Complexes of Functionalized Hydrazones. Resea rch Journal of Pharmaceutical, Biological and Che mical Sciences [Internet].2014 [cited 2014 May];5(3):1086-1085.Available from: http://www.rjpbcs.co m/pdf/2014_5(3)/[189].pdf
9. Kuroki M, Yoneoka T, Satou T, Takagi Y, Kitamura T, Kayamori N. Bar-code Recognition S ystem Using Image. IEEE Xplore DigitalLibrary. [Internet].1997[cited 1997 Sep]:56: Available from: 568-572. http://ieeexplore.ieee.org /xpls/abs_all.jsp?arnumber=616333DOI:10.1109/ ETFA.1997.616333
10. Joanne M. Stempak, Kyoung-Jin Sohn,,En-Pei Chiang , Barry Shane. Cell and stage of transformation-specific effects of folate deficiency on methionine cycle intermediates and DNA methylation in an in vitromodel. Oxford Journals.[Internet]. 2005 Jan [cited 2004 Oct];26(5): 981 - 990.Available from: http://carcin. oxfordjournals.org/content/26/5/981.long
11. Raman N, Sobha S and Selvaganapathy M. (2012), Probing The DNA Binding Mode Of Transition Metal Based Biologically Active Compounds: Validation By Spectroscopic Methods. International Journal of Pharma and Bio Sciences.2012; 3(1):251-268. Available from: http://www.ijpbs.net/archive-issue.php?issueid=17
12. Normand N,Viard Gaudin C. A Two dimensional b arcode reader. . IEEE Xplore Digital Library. [Inter net].1994 [cited 1994Sep]:3:201-203.Available from: http://ieeexplore.ieee.org/xpls /abs_all.jsp?arnumber=577158.DOI:10.1109/ICP R.1994.577158.

13. Howlett R.J, Berthier S, and Awcock.J.G Determining the location of industrial bar-codes using neural networks. IEEE Xplore Digital Library .[Internet].1997 [cited 1997 Jul]:2:511-515. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=615578DOI:10.1049/cp:19970946
14. Pavel Šimurda. Bar code localization using texture analysis Information Sciences and Technologies Bulletin of the ACM Slovakia, Special Section on the ACM Student Project. [Internet].2011[cited 2011]: 3(4): 55–56. Available from: http://acmbulletin.fiit.stuba.sk/vol3num4/simurdaSPY.pdf
15. Japanese Industrial Standards.Two Dimensional Symbol QR Code Basic Specification.[Internet] JIS X 0510.2004.[Updated 2004 Nov 20] Available from: http://www.webstore.jsa.or.jp/webstore/Com/FlowControl.jsp?lang=en&bunsyoId=JIS%20X%200510%3A2004&dantaiCd=JIS&status=1&pageNo=1
16. Kato, H., & Tan, K. Pervasive 2D Barcodes for Camera Phone Applications.[Internet].2011[cited 2011]6(4): 76-85. Available from: http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=2506&context=ecuworks
17. Joseph E, Pavlidis T.Bar code waveform recognition using peak locations. ACM Digital Library .[Internet].1994[cited 1994 Jun]:16(6):630-640.Available from: http://dl.acm.org/citation.cfm?id=628614. DOI:10.1109/34.295907
18. Oktem R.Bar code localization in wavelet domain by using binary morphology.IEEE Xplore Digital Library .[Internet].2004 [cited 2004 Apr]:499-501.Available from: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=1338574DOI: 10.1109/SIU.2004.1338574
19. Tanee Wiputtikul, Sartid Vongpradhip. A Technique to Add Error Detection of QR CodeDecoding by Using Micro QR Code.International Journal of Electrical Energy. [Internet].2013[cited 2013 Sep]; 1(3): 1-5.Available from: http://www.ijoee.org/uploadfile/2013/0912/20130912055005793.pdf
20. Stephen K. Reed. Pattern recognition and categorization. Cognitive Psychology.Science Direct. [Internet].1972 [cited 1972 July];3(3):382-407.Available from: http://www.sciencedirect.com/science/article/pii/001002857290014X
21. Hungwen Li, Mark A Lavin, Ronald J Le Master .Fast Hough transform: A hierarchical approach. Computer Vision, Graphics, and Image Processing.Scince Direct. [Internet].1996[cited1996 Nov];36(2-3):139-161.Available from: http://www.sciencedirect.com/science/article/pii/0734189X86900733
22. Electron J.Super-resolution of bar codes.Imaging. [Internet].2001 [cited 2000 Mar];10(1):213-220.Available from: http://electronicimaging.spiedigitallibrary.org/article.aspx?articleid=1097649%20DOI:10.1117/1.1329337
23. Orhan Bulan.High Capacity Color Barcodes: Per Channel Data Encoding via Orientation Modulation in Elliptical Dot Arrays. IEEE Transactions on Image Processing. [Internet].2011[cited 2011 May]; 20(5):1337-1350.Available from: http://www.ece.rochester.edu/~gsharma/papers/BulanColorBarcodesTIPMay2011.pdf. DOI :10.1109/TIP.2010.2092437
24. Boor, Carlde. A Practical Guide to Splines [Internet]. New York: Springer-Verlag ; 1978 [cited 1978]. XVIII, 348. Available from: http://www.springer.com/us/book/9780387953663
25. Reinsch C.H. Smoothing by Spline Functions.Numerische Mathematica. [Internet].1967 [cited 1967 Oct]; 10(3): 177-183.Available from: http://link.springer.com/article/10.1007%2FBF02162161?LI=true DOI: 10.1007/BF02162161
26. Schoenberg I.J. Cardinal Spline Interpolation. [Internet].Philadelphia, PA: Soc. Indus Apllied Math. 1973.[cited 1973].121.Availble from: https://books.google.co.**in**/books?isbn=089871009X
27. Ohbuchi E, Hanaizumi H, Hock L.A.Barcode readers using the camera device in mobile phones. ACM Digital Library. [Internet].2004 [cited 2004]:260-265.Available from: http://dl.acm.org/citation.cfm?id=1033123. DOI: 10.1109/CW.2004.23
28. Canny J. A Computational Approach to Edge Detection. ACM Digital Library. [Internet].1986 [cited 1986 Jun];8(6):679-698. Available from: http://dl.acm.org/citation.cfm?id=11275.DOI:10.1109/TPAMI.1986.4767851.
29. Nageswara Rao PV, Uma Devi, Dsvgk Kaladhar, Sridhar Gr, Allam Appa Rao. A Probabilistic Neural Network Approach For Protein Superfamily Classification. International Journal of Application or Innovation in Engineering & Management (IJAIEM). [Internet].2009[cited 2009]; 6(1):101-105.Available from: http://www.jatit.org/volumes/research-papers/Vol6No1/11Vol6No1.pdf
30. Ashna Thomas, Remya Paul. An Effective Method for Removing Scratches and Restoring Low -Quality QR Code Images.IJARET. [Internet].2013[cited 2013 June]; 1(V): 1-5.Available from: http://www.ijoee.org/uploadfile/2013/0912/20130912055005793.pdf
31. Valentina Boeva, Mireille Regnier, Dmitri Papatsenko,Vsevolod Makeev . Short fuzzy tandem repeats in genomic sequences, identification, and possible role in regulation of gene expression.Bioinformatics, Oxford Journals. [Internet].2006[cited 2006 Jan]; 22(6):676-684.Available from:http://bioinformatics.oxfordjournals.org/content/22/6/676.short DOI:10.1093/bioinformatics/btk032
32. Mansoori E.G, Zolghadri M.J, Katebi S.D, Mohabatkar H, Boostani R, Sadreddini H,et al.Generating Fuzzy Rules for Protein Classification. Iranian Journal of Fuzzy Systems. [Internet].2008 [cited 2008 June];5(2): 21-33.Available from: http://www.sid.ir/en/VEWSSID/J_pdf/90820080206.pdf
33. Lihong Ren, Yongsheng Ding, Hao Ying, Shihuang Shao. Emergence of Self-Learning Fuzzy Systems by a New Virus DNA–Based Evolutionary Algorithm. International Journal of Intelligent Systems.[Internet].2003 [cited 2003 Sep];18: 339-354.Available from: http://onlinelibra

ry.wiley.com/doi/10.1002/int.10097/pdf.DOI 10.1002/int.10097

34. Yohsikawa T,Furuhashi T, Uchikaawa Y. DNA coding method and a mechanism of development for acquisition of fuzzy control rules.I EEE Xplore Digital Library. [Internet].1996 [cited 1 996 Sep]; 3: 2194-2200.Available from:http://ieee xplore.ieee.org/xpls/abs_all.jsp?arnumber=55354 0 DOI:10.1109/FUZZY.1996.553540
35. Buldyrev S.V, Goldberger A.L, Havlin S, Mantegna R.N, Matsa M.E, Peng C-K,,et al. Long range correlation properties of coding and noncod ing DNA sequences: GenBank analysis.PHYSICA L REVIEW E.Internet].1995 [cited 1995 May];51(5): 5084 .Available from: http://journals.aps.org/pre/ abstract/10.1103/PhysRevE.51.5084. DOI:http://d x.doi.org/10.1103/PhysRevE.51.5084
36. Stephen F. Altschul, Thomas L. Madden, Alejandr o A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller,David J. Lipman,et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research.Oxford Journals.[Internet] .1997 [cited 1997 June];25(17): 3389-3402.Availa ble from: http://nar.oxfordjournals.org/content/25/1 7/3389.full. DOI: 10.1093/nar/25.17.3389.
37. Kuroki M, Yoneoka T, Satou T, Takagi Y, Kitamura T, Kayamori N. Bar-code Recognition S ystem Using Image Processing. Sixth Internationa l Conference on Emerging Technologies and Factory Automation,1997: 568-572.
38. Ying-Hong Liang, Zhi-Yan Wang. A Skew Detection Method for 2D Bar Code Images Based on the Least Square Method.IEEE Xplore Digital Library .[Internet].2006 [cited 2006 Aug]:39 74-3977499.Available from: http://ieeexplore.ieee. org/xpls/abs_all.jsp?arnumber=4028766DOI:10.1 109/ICMLC.2006.258793
39. Xiao-Wei Xu , Zhi-Yan Wang,Yan-Qing Zhang,Ying-Hong Liang. A Skew Distortion Correction Method for 2D Bar Code Images Base d on Vanishing Points.IEEE Xplore Digital Library .[Internet].2007 [cited 2007 Aug]:1652-1656.Avail able from: http://ieeexplore.ieee.org/xpls/abs_all.j sp?arnumber=4370412.DOI:10.1109/ICMLC.200 7.4370412
40. Ersoy Öz, Hüseyin Kaya. Support vector machines forquality control of DNA sequencing.Jo urnal of Inequalities and Applications.[Internet].20 13 [cited 2012 Dec];18: 339-354.Available from: h ttp://link.springer.com/article/10.1186/1029-242X-2013-85.DOI: 10.1186/1029-242X-2013-85
41. Terrence S. Furey,Nello Cristianini,Nigel Duffy, D avid W. Bednarski, Michèl Schummer David Haussler. Support vector machine classification and validation of cancer tissue samples using microarray expression data.Bioinfo rmatics. Oxford Journals .[Internet].2007 [cited 20 07 Aug]; 16(10):906-914. Available from: http://bi oinformatics.oxfordjournals.org/content/16/10/906 .short