

AN INTELLIGENT MODULAR NEURO - FUZZY BASED APPROACH FOR AUTOMATED CLASSIFICATION OF CHROMOSOMES

Mousami V. Munot

Associate Prof. Dept of E&TC, Pune Institute of Computer Technology, Pune, India

Abstract: The manual process of ordering of chromosomes (karyotyping) is long, tiresome and prone to errors. Automated classification of chromosomes in one of the appropriate 24 classes has therefore emerged as a thrust area of research. Researchers have studied the important features of the chromosomes which play a significant role in karyotyping of chromosomes and reports success of neural networks as promising classifiers. Perceptive vagueness in the texture of the chromosomes, however, encourages use of fuzzy classifier for efficient classification. This research initially revalidates the performance of neural network and fuzzy system individually, to examine the dominant features for each and every Denver group and sub-group and further presents a two-stage classification process to resemble manual approach of generating a karyogram. The proposed novel and intelligent approach initially identifies the best features for every class / subgroup and further utilizes them to build a modular Neuro-Fuzzy classifier for organizing the chromosome as per a standard karyogram. The proposed algorithm is validated using a publically available database comprising of 90 metaphase images (4320 segmented chromosomes). The results conforms the efficacy of the algorithm by achieving improvement by 2 % in overall performance of the proposed hybrid classifier and thus offers a practically deployable solution to the issue of manual karyotyping.

Keywords: karyotyping, chromosome classification, neural networks, fuzzy systems

I. INTRODUCTION

Chromosomes are the genetic carriers of the human body [1, 2]. Karyotyping is the procedure of envisaging and sorting the chromosomes within a cell during the metaphase stage of cell division [3, 4]. It is an indispensable assignment in cytogenetics for the diagnosis of genetic disorders [5]. Manual process of classification of chromosomes is tedious and prone to human errors [6]. It still faces lot of challenges and therefore is a long standing issue in the domain of medical image processing [7]. The task of generating the karyogram can actually be envisioned as a pattern recognition problem demanding recognition of the features and further classification of chromosome in its respective classes(24 classes with seven main groups called as Denver groups: A-G and 24 subgroups) [8, 9]. Literature describes overall prevailing features of the chromosomes considering the entire classification problem. Various approaches based on Hidden Markow Model [HMM] [8], pairing of chromosomes [9], Rough Set theory [10], incremental learning [11, 12] are explored in literature for the task of chromosome segmentation and classification [13]. Moreover, fuzzy based system [14, 15] and neural network based approaches are also reported in the literature with encouraging results. It is however also essential to identify the specific prevailing features for every class. This is imperative when designing an automated system which duplicates manual behavior and approach of generating a karyogram. The classifier designed considering this approach is expected to intelligently perform the task of karyotyping in two stages. The first stage will involve classifying the chromosomes in major Denver Groups and next stage will classify the chromosomes in the respective subgroups. A hybrid approach combining both best results of neural networks and fuzzy is proposed to exploit the fuzziness in the features of the subgroup as described in section II. Section III details the experimental results and the best dominant features selected for the proposed two stage classifier. Paper concludes with the discussion section.

II. PROPOSED METHODOLOGY

Literature reports various algorithms for feature extraction [16, 17] of features like band profile [18, 19], dominant points [20,21] and shape representation [22] for the polarity free [23, 24]classification of chromosomes. A rule based classification system [25,2] has delivered encouraging results and is in line with the process of manual classification. Literature reports use of many features and feature extraction methodologies to identify the class of the chromosomes. This experimentation presented in this research considers individual features and various combinations of all the features to finally identify the features that achieve highest classification accuracies. The features considered as listed :

- F1 : Centromeric Index*
- F2 : Length of Chromosome*
- F3 : Total number of bands*
- F4 : Intensity*
- F5 : Area of the Chromosomes*

- F6: Perimeter of chromosome
 F7: Number of Bands above centromere
 F8: Number of bands below centromere
 F9: Wavelet Coefficient - 1
 F10: Wavelet Coefficient - 2
 F11 Ratio of larger arm to the shorter arm

Multilayer Perceptron model with back propagation algorithm is used for the neural network based classifier. A Fuzzy based classifier with Mamdani model is used and Gaussian membership function is used. Implementation of Gaussian membership function requires specification of mean and deviation. The mean and the deviation value of the feature vector were decreased to half of the original values. The next section details the experimental results of individual classifiers and results of the proposed modular approach using both NN and Fuzzy classifiers.

III. EXPERIMENTAL RESULTS OF CHROMOSOME CLASSIFICATION USING HYBRID CLASSIFIER:

The database used for experimentation consists of segmented chromosome from publically available Biolab database. Matlab R2012b was used for the experimentation. The results of the neural network and fuzzy engine are summarized below.

(i) Classification using NN networks:

Multilayer Perceptron model with back propagation algorithm is used for the neural network based classifier. The experimentation involves approximating the various factors of NN to identify the best feature vector. A two stage classifier is build to initially identify the Denver group of the segmented chromosome at the input and further has seven neural networks to categorize the segmented chromosome in its approximate sub-group. Table I details the various combination of the features and the neurons experimented to achieve the best results. Fig. 1 depicts the best results achieved for each group and sub-group and number of neurons in the hidden layer. A target of 100 epochs was set along with other parameters.

Table I : Neural Network based classifier for automated karyogram generation
 {GA : Group A ; GB : Group B ; GC : Group C ; GD : Group D ; GE : Group E ; GF : Group F ; GG: Group G}

	Denver Group	GA	GB	GC	GD	GE	GF	GG
Features (F1 to F11)	<ul style="list-style-type: none"> F1 to F11 F1 to F6 Excluding F9, F10, F11 	<ul style="list-style-type: none"> F1 to F11 F1 to F6 Excluding F7, F8, F9 	<ul style="list-style-type: none"> F1 to F11 F1 to F6 Excluding F7, F8, F9, F10, F11 	<ul style="list-style-type: none"> F1 to F11 Excluding F8, F9 Excluding F1 F1 to F6 Excluding F1, F6, F11 	<ul style="list-style-type: none"> F1 to F11 Except F8 to F11 Except F1 	<ul style="list-style-type: none"> F1 to F11 Excluding F8, F9, F10, F11 Excluding F8, F9 	<ul style="list-style-type: none"> F1 to F11 F1 to F6 Excluding F8 to F11 Excluding F11 	<ul style="list-style-type: none"> F1 to F11 F1 to F6 Excluding F8 to F11 Excluding F11
Neurons in hidden layer	5, 20, 25, 30, 35, 37, 40, 100	20, 30	20, 25, 30, 35	20, 30, 50, 100	20, 30, 50, 100	20, 30, 50	20, 30, 50, 100	20, 30, 50, 100



Figure 1: Best results using neural network based approach

(ii) Classification using Fuzzy Systems

A fuzzy engine is developed using a Mamdani model and Gaussian function as a membership function to develop a rule-base classifier that exploits the fuzziness of the overlapping feature of the chromosomes. The various combination of the features is explored even to test the developed fuzzy classifier. The results signify the usability of only first six features F1 to F6 for efficient decision making and classifying the chromosome in its respective class. The best results are depicted in Figure 2.

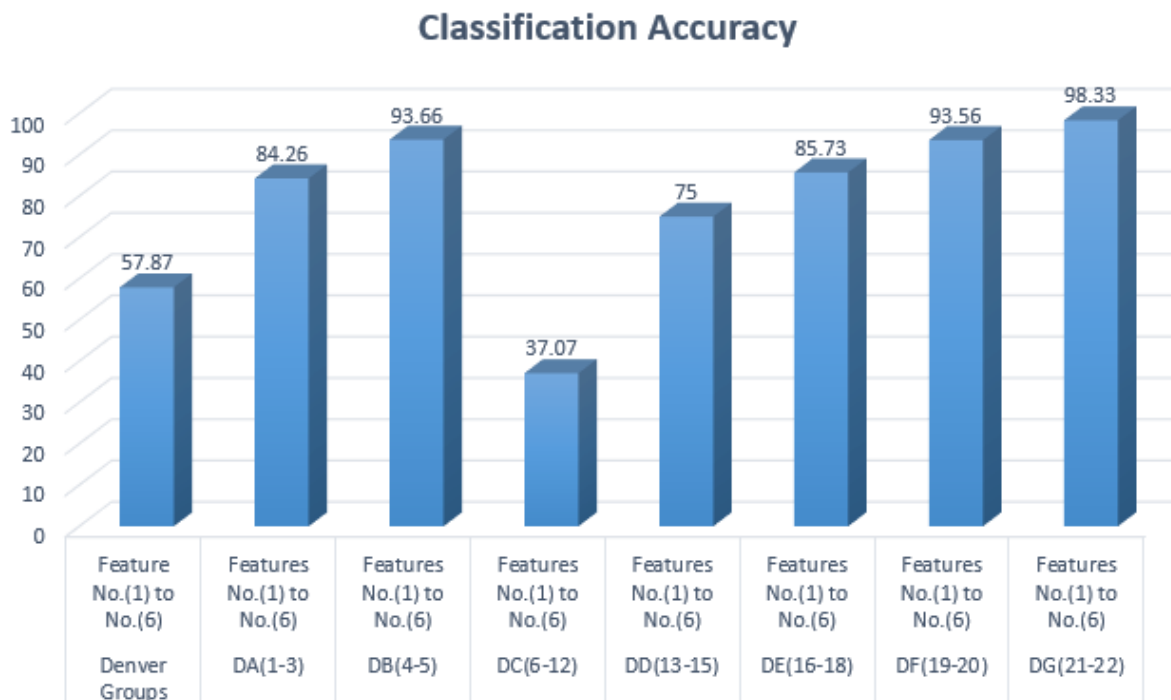


Figure 2: Classification results using fuzzy classifier (Mamdani Model and Gaussian Membership function)

(iii) Classification using Hybrid approach (Neuro - Fuzzy)

The performance of the neural network and fuzzy classifier vary for each subgroup or class. This research proposes a novel approach combining the best classifiers for the overall procedure of automated generation of a karyogram. Figure 3 indicates the results of the proposed hybrid classifier. A fusion of the NN and fuzzy modules at classifier levels is unique contribution of this research which delivers encouraging and improvement in the results. The proposed approach demonstrates increase in the true positive value and the classification accuracy by 2 % as depicted in Figure 3. The proposed novel modular neuro-fuzzy classifier offers an encouraging approach for AKS. Table II details the achieved enhancement in classification accuracies achieved using developed hybrid classifier

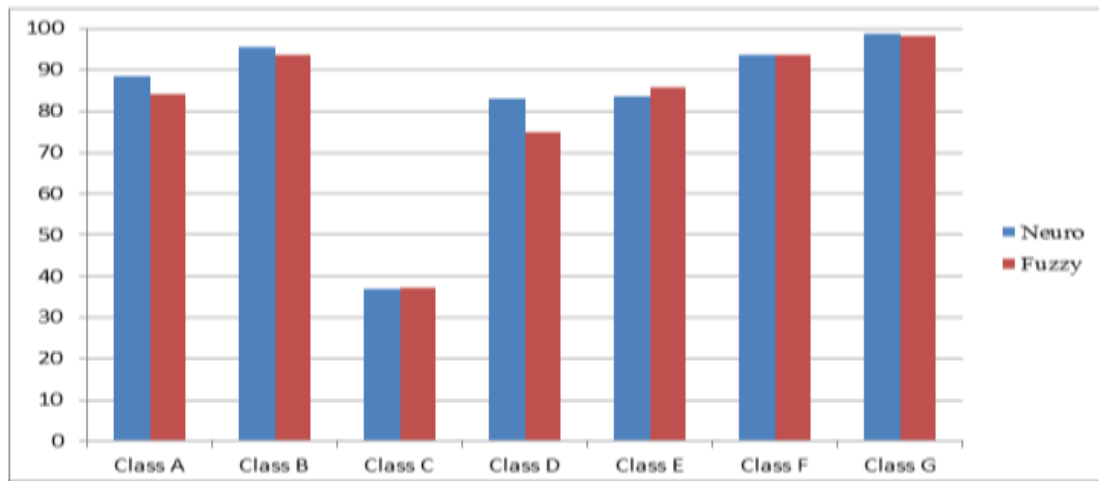


Figure3. Comparison of ANN and Fuzzy Classifiers

Table II Confusion Matrix obtained using Neuro-Fuzzy approach

Class	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	XXI	XXII
I	138	24	1	12	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
II	2	140	3	29	16	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
III	0	25	100	36	27	3	4	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0
IV	0	3	5	143	8	9	11	10	5	3	1	1	1	0	0	0	0	0	0	0	0	0
V	0	3	6	8	133	0	4	6	12	12	6	1	1	2	0	0	0	0	0	0	0	0
VI	0	2	0	42	14	55	45	29	11	1	0	0	5	0	1	0	0	0	0	0	0	0
VII	0	0	1	17	19	12	47	57	37	4	0	0	1	1	6	0	0	0	0	0	0	0
VIII	1	0	0	9	5	3	32	68	44	17	2	0	4	8	13	1	0	0	0	0	0	0
IX	0	0	0	8	10	1	0	16	68	26	9	0	3	6	24	1	2	0	0	0	0	0
X	0	0	1	2	8	1	0	10	55	70	17	1	5	8	22	0	0	0	0	0	0	0
XI	0	0	0	11	6	0	4	8	20	51	44	9	11	14	26	2	0	0	1	0	0	0
XII	0	0	0	3	6	1	1	7	23	48	51	21	13	19	10	0	0	2	0	0	0	0
XIII	0	0	0	1	0	1	4	5	6	2	4	1	142	26	3	0	2	12	0	1	0	0
XIV	0	0	0	0	0	1	4	2	1	11	10	2	9	129	11	2	5	17	0	1	0	0
XV	0	0	0	2	0	2	1	0	3	12	4	3	3	29	92	6	20	25	4	1	0	0
XVI	0	0	0	0	0	0	0	0	1	2	1	1	8	9	20	87	30	0	13	12	0	0
XVII	0	0	0	0	0	0	0	2	2	1	0	0	3	5	15	9	106	7	24	18	0	2
XVIII	0	0	0	0	0	0	0	3	1	1	0	0	8	11	20	1	14	89	7	15	0	0
XIX	0	0	0	0	0	0	0	0	0	0	0	0	5	3	4	12	11	16	118	9	1	34
XX	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2	20	4	3	10	131	3	12
XXI	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	6	6	3	14	17	87	2
XXII	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	6	2	34	37	1	82

IV. DISCUSSION AND CONCLUSION

A two level classification approach is presented in this research work for automated karyotyping. Artificial neural network and fuzzy classifier have been explored in this research. Literature reports various features for the organization of the chromosomes. However, this research presents rigorous experimentation to examine the usability of all the features for each Denver group and sub-group and

identifies the dominant features responsible for the decision about the ordering of the chromosomes in a karyogram and further uses it in two-level classification approach. First level of classification involves grouping of the chromosomes in the main Denver Group (macro level) and the next stage classifies the chromosomes in the sub-groups (micro level). The true positive (TP) value achieved for the Denver class is 71 %, whereas the value of TP ranges from 92% to 99 % for sub-group classification. This two level classification process resembles the procedure of manual karyotyping and thus leads to increased classification accuracies. The novel hybrid approach combining the neural network and fuzzy classifiers based on the best accuracies obtained for the respective groups, outperforms the individual results, leading to encouraging increase in overall TP by 2% (84 correctly classified chromosomes).

References

- [1] EneaPoletti, Enrico Grisan, Alfredo Ruggeri, "A modular framework for the automatic classification of chromosomes in Q-band images", Elsevier journal on computer methods and programs in biomedicine, pp 1-11, 2011.
- [2] Xingwei Wang, Bin Zheng, Shibo Li, John J. Mulvihill, Marc C. Wood, Hong Liu, "Automated classification of metaphase chromosomes: Optimization of an adaptive computerized scheme", Elsevier Journal of Biomedical Informatics 42, pp 22-31, 2009.
- [3] Delie Ming, Jinwen Tian, "Automatic Pattern Extraction and Classification for Chromosome Images", *J Infrared MilliTerahz Waves* 31, pp 866-877, 2010.
- [4] Mousami Munot, Prachi Joshi, Madhuri Joshi, Parag Kulkarni, "Efficient Pairing of Chromosomes in Metaphase Image for Automated Karyotyping", IEEE-EMBS Conference on Biomedical Engineering and Sciences, IECBES - 2012, Langkawi, Malaysia, Dec-2012, pp:916-921.
- [5] Artem Khmelinskii, Rodrigo Ventura, and Joao Sanches, "A Novel Metric for Bone Marrow Cells Chromosome Pairing", *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp 1420-1429, June 2010.
- [6] Mousami Munot, Madhuri Joshi, "Towards Automated Karyotyping of Curved Chromosomes", Coimbatore Institute of Information Technology, CiiT International Journal of Artificial Intelligent Systems and Machine Learning, January 2012.
- [7] Mukul Joshi, Mousami Munot, Madhuri Joshi, Kruti Shah, Ketan Soni, "Automated Detection of the Cut-points for the Separation of Overlapping Chromosomes", IEEE-EMBS Conference on Biomedical Engineering and Sciences, IECBES - 2012, Langkawi, Malaysia, Dec-2012, pp:820 - 825
- [8] John M. Conroy, Tamara G. Kolda, Dianne P. O'Leary, and Timothy J. O'Leary, "Chromosome Identification Using Hidden Markov Models: Comparison with Neural Networks, Singular Value Decomposition, Principal Components Analysis, and Fisher Discriminant Analysis" Laboratory Investigation, vol. 80, no. 11, pp 1629-1641, November 2000.
- [9] Mousami Munot, Prachi Joshi, Madhuri Joshi, Parag Kulkarni, "Efficient Pairing of Chromosomes in Metaphase Image for Automated Karyotyping", IEEE-EMBS Conference on Biomedical Engineering and Sciences, IECBES - 2012, Langkawi, Malaysia, Dec-2012, pp:916-921.
- [10] Munot M., Joshi M., Mitra P., "Genetic Algorithm Incorporates with Rough Set Theory: Application to Automated Karyotyping", Indian International conference on artificial intelligence, IICAI 2011.
- [11] Prachi Joshi, Mousami Munot, Parag Kulkarni, Madhuri Joshi, "Efficient Karyotyping of Metaphase Chromosomes Using Incremental Learning", The IET Journal of Science Measurement and Technology, Vol.7, Issue 5, September 2013, pp. 287-295.
- [12] Mousami Munot, Prachi Joshi, Madhuri Joshi, Parag Kulkarni, "An Incremental Approach for Efficient Karyotyping Systems", Journal of Medical Imaging and Health Informatics, Med. Imaging Health, Inf. 6, February 2016, pp. 221-225.
- [13] Mousami Munot, Jayanta Mukhopadhyay, Madhuri Joshi, "A Novel Approach for Efficient Extrication of Overlapping Chromosomes in Automated Karyotyping", Springer Journal of Medical and Biological Engineering and Computing, August 2013, pp. 1-14.
- [14] Liang Wang, John Yen, "Extracting fuzzy rules for system modelling using a hybrid of genetic algorithms and Kalman filter", *Fuzzy Sets and Systems* 101, pp 353-362, 1999.
- [15] Mohammad Ali Akbari and Masayuki Nakajima, "Improving the Automatic Karyotyping Accuracy of the Unrefined Chromosome Features using Fuzzy Logic", *IEEE Region 10 Conference on TENCN*, vol.3, pp 616-619, 2004.
- [16] Mehdi Moradi, S. Kamaledin Setarehdan, "New features for automatic classification of human chromosomes: A feasibility study", *Pattern Recognition Letters* 27, pp 19-28, 2006.
- [17] Akila Subasinghe Arachchige, Jagath Samarabandu, Joan Knoll, Wahab Khan, Peter Rogan, "An Image Processing Algorithm for Accurate Extraction of the Centerline from Human Metaphase Chromosomes", *IEEE International Conference on Image Processing*, pp 3613-3616, September 2010.
- [18] Jau-hong Kao, Jen-hui Chuang, Tsaipei Wang, "Chromosome classification based on the band profile similarity along approximate medial axis", *Pattern Recognition* 41, pp 77 - 89, 2008.
- [19] Shadab Khan, Joao Sanches and Rodrigo Ventura, "Robust Band Profile Extraction Using Constrained Nonparametric Machine-Learning Technique", *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp 2587-2591, October 2010.
- [20] Seung Yun Ryu, Jong Man Cho, SeungHyo Woo, "A Study for the Feature Selection to Identify Giemsa-Stained Human Chromosomes Based on Artificial Neural Network", *IEEE International Conference on EMBS*, Istanbul, Turkey, pp 691-692, 25-28 October 2001.
- [21] Gunter Ritter, GernotSchreib, "Using dominant points and variants for profile extraction from chromosomes", *Pattern Recognition* 34, pp 923-938, 2001.
- [22] Gunter Ritter and GernotSchreib, "Profile and feature extraction from chromosomes", *IEEE Conference on Pattern Recognition*, vol.2, pp 287-290, 2000.
- [23] L. V. Guimaraes, A. Schuck Jr and A. Elbern, "Chromosome Classification for Karyotype Composing applying Shape Representation on Wavelet Packet Transform", *International Conference of the IEEE EMBS*, Vol.1, pp 941-943, September 17-21, 2003.
- [24] Gunter Ritter, Christoph Pesch, "Polarity-free automatic classification of chromosomes", *Elsevier Science of Computational Statistics & Data Analysis* 35, pp 351-372, 2001.
- [25] EneaPoletti, Enrico Grisan, and Alfredo Ruggeri, "Automatic classification of chromosomes in Q-band images", *IEEE Conference on Engineering in Medicine and Biology Society (EMBS)*, pp 1911-1914, 2008.
- [26] Xingwei Wang, Bin Zheng, Shibo Li, John J. Mulvihill, Hong Liu, "A rule-based computer scheme for centromere identification and polarity assignment of metaphase chromosomes", Elsevier *Journal of computer methods and programs in biomedicine* 89, pp 33-42, 2008.