Cytometry (Communications in Clinical Cytometry) 30:145–150 (1997)

An Artificial Intelligent Diagnostic System on Differential Recognition of Hematopoietic Cells From Microscopic Images

Meral Beksaç,[1] M. Sinan Beksaç,[2]* V. Bahadır Tipi,[3] H. Ali Duru,[3] M. Ümit Karakaş,[3] and A. Nur Çakar[4]

[1]Department of Hematology, Ankara University, Ankara, Turkey

[2]Department of Biomedical Engineering Unit of Obstetrics and Gynecology, Hacettepe University, Ankara, Turkey

[3]Department of Computer Sciences and Engineering, Hacettepe University, Ankara, Turkey

[4]Department of Histology and Embryology, Hacettepe University, Ankara, Turkey

Despite their advantages, none of the automated white blood cell differential counters have replaced the conventional microscopic evaluations of blood and bone marrow slides by hematologists. We have analyzed the smears of 39 patients and 8 control subjects to develop an artificial expert system that recognizes 16 different types of nucleated hematopoietic cells during the stages of differentiation. A charge coupled television camera and a special frame grabber were used for data acquisition, and 247 nucleated cell images were transferred from a microscope to an IBM 386 computer to be processed. One hundred sixty-five and 82 of these images were used for training and testing, respectively. Our system is composed of image processing and analysis (enhancement, thresholding/smoothing, edge detection), pattern recognition (feature extraction and classification with supervised artificial neural network), and expert system development. Image processing and analysis were used to obtain 13 cellular features to be used as the input parameters (neurons) of the artificial neural network. A supervised artificial neural network (back-propagation learning algorithm) was used in the classification of 16 different cells (output neurons of the neural network), which is the second step of pattern recognition. A confusion matrix has been developed to compare the similarities and dissimilarities between the differential recognitions of the hematologist and the expert system. The discriminatory power of the procedure is statistically significant: $Q = (N - n.K)2/N.(K - 1) = 28.2$. The sensitivity and the specificity of the expert system were 71.4% and 90.9%, respectively. Cytometry 30:145–150, 1997. © 1997 Wiley-Liss, Inc.

Key terms: hematopoietic system; blood; image processing/analysis; pattern recognition; artificial neural network

Hematopoietic system disorders (HSDs) are life-threatening problems, and careful examinations and early diagnosis are important tools in the management of patients. There are established conventional approaches in the evaluation of peripheral blood and bone marrow smears (1). However, such an evaluation requires experienced staff to achieve the desired goal. The importance of early diagnosis and follow-up programs obligates screening of the populations at risk for HSDs. However, screening means more patients with less experienced experts and introduces the concept of automation in the field of clinical hematology and cytometry.

Automated cytology may be conceived as a means of identifying human cells in clinical samples (nucleated hematopoietic cells in this study) by computerized intelligent systems (2–6). Various methods used to determine cellular descriptors stand out as particularly promising (7,8). One method uses the traditional information provided by a microscopic image of the object (9,10). The availability of high-capacity microcomputers has faciliated the acquisition and analysis of various images (11,12). Meantime, new generation computer and software engineering technologies have enabled the development of artificial intelligent diagnostic systems for the interpretation of various medical and histocytological images (12–16).

Automated differential blood counters represent the first successful clinical application of automated cytology (2,17). Different investigators have described high-resolu-

*Correspondence to: M. Sinan Beksaç, Biomedical Engineering Unit, Department of Obstetrics and Gynecology, Hacettepe University, 06100 Ankara, Turkey.

Received 10 January 1995; Accepted 9 January 1997

tion scanning systems for leukocyte recognition, and supervised and nonsupervised learning programs have been reported to recognize leukemic cells among normal eritropoietic system cells (3–5). Recently, considerable improvements have been made to hematology analyzers (18,19). The advent of flow cytometry into the modern hematology laboratory has generated the development of a multitude of analyzers capable of performing multiparameter hematologic determinations (20,21). These instruments have the capacity to count and size particles in whole blood and to perform a complete or abridged white blood cell (WBC) differentiation. The automation of the full WBC differential count has proved more difficult (22). Light scatter, coupled with cytochemistries or pattern recognition systems, have been two contrasting technologies employed to provide sufficient WBC differential count (19,23,24). The new trend is to combine artificial intelligent diagnostic programs with automation technology and produce practically available medical systems (6,9,10, 25,26). Several semiautomated and automated systems equipped with control or robotics functions have been described and have brought new concepts to medicine (14–16). Recently, we have developed a semiautomated intelligent diagnostic system based on pattern recognition via neural networks for the evaluation of peripheral blood and bone marrow smears.

MATERIAL AND METHODS

Hardware and Data Acquisition

A microscope (Nikon microphot FXA EPI-FL3 fluorescent microscope) was connected to an IBM-386 by a 64 gray-level-frame grabber (Mini Magiscan IAS25/N25) and charge coupled television camera (Hitachi KP-140 E/K Solid state). Sixty-four single-color 512-×-512 gray-level images were obtained for analysis.

Clinical Material

This study consisted of 39 patients (27 acute nonlymphoblastic leukemia, 4 acute lymphoblastic leukemia, 4 chronic myelogenous leukemia, 1 multiple myeloma, 1 non-Hodgkin's lymphoma, 1 hairy cell leukemia) and 8 control subjects. All pretreatment samples from either peripheral blood or bone marrow were spread onto smears in the conventional manner and stained with Giemsa. The microscopic evaluation of the nucleated cells (16 different types) was made by an expert hematologist and transferred to the computer system simultaneously. One hundred thirty-two microscopic images consisting of 247 cells were selected by the hematologist for data acquisition (46 myeloblasts, 5 monoblasts, 12 normoblasts, 10 proerythroblasts, 42 lymphoblasts, 33 lymphocytes, 12 neutrophils, 13 metamyelocytes, 12 stabs, 27 promyelocytes, 9 myelocytes, 2 basophil, 9 monocyte, 6 eosinophils, 7 plasma cells, and 2 megakaryocytes).

Analysis Program

Our system, called the ANADOLU System is composed of image processing and analysis (enhancement, thresholding and smoothing, edge detection), pattern recognition (feature extraction and classification), and expert system development. The program is written in Borland Pascal version 7.0 (Borland International, Scotts Valley, CA).

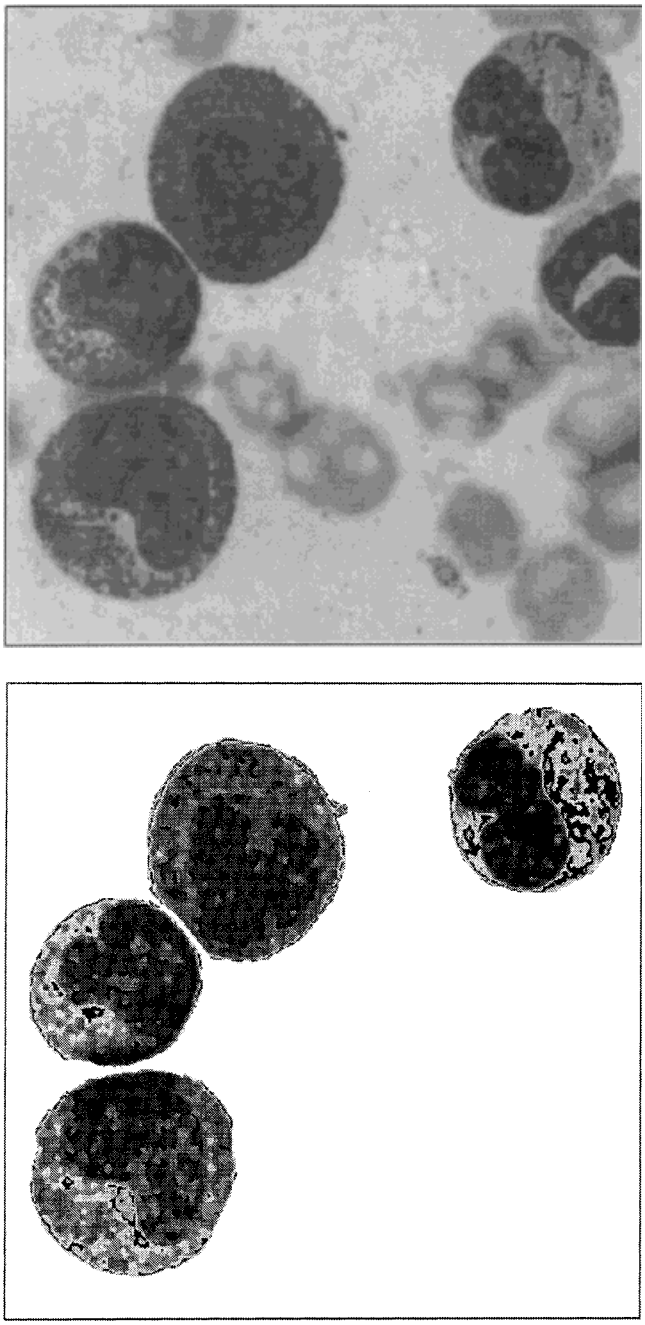

FIG. 1. The original and resulting images after image processing.

Image Processing and Analysis

After the enhancement of the images, various algorithms were used to achieve thresholding, smoothing, and edge detection. Figure 1 shows the original and resulting images obtained after successive processing steps.

Table 1
Mean Values of the Input Vectors (Parameters) of Artificial Neural Networks[a]

Cell type	A	B	C	D	E	F	G	H	I	J	K	L	M
Myeloblast	17,875	30.3	39.4	47.7	12,470	26.4	26.9	7.3	5,405	39.5	70.6	0.8	4.5
Lymphocyte	5,449	24.0	21.9	72.8	3,925	19.4	18.0	5.6	1,524	35.9	72.2	0.1	0.3
Monoblast	21,243	35.2	42.6	72.0	13,433	29.8	27.2	8.8	7,810	44.6	63.7	2.0	1.6
Neutrophil	11,237	35.3	30.3	112.7	5,261	25.8	14.0	23.2	5,975	43.9	48.9	1.3	0.7
Erytroblast	6,237	25.3	22.8	67.2	3,572	19.2	16.4	10.2	2,665	33.4	58.6	0.2	0.1
Proerythroblast	13,585	31.0	33.7	42.6	10,157	27.9	24.0	7.7	3,428	38.8	72.8	0.7	8.5
Metamyelocyte	11,927	29.8	32.0	76.1	6,878	23.1	18.8	10.8	5,049	39.2	58.5	1.9	0.9
Stab	11,059	32.2	29.9	80.9	6,614	25.8	17.3	15.7	4,446	41.7	60.8	1.0	2.3
Promyelocyte	19,529	31.4	41.3	37.2	14,231	28.6	25.9	8.5	5,298	38.9	73.5	2.0	7.4
Myelocyste	17,520	26.9	39.9	38.2	11,203	23.6	21.4	7.0	6,317	33.7	66.6	4.6	3.1
Lymphoblast	8,962	28.8	27.1	41.3	6,801	26.0	20.7	8.2	2,161	38.0	76.2	0.3	3.8
Monocyte	21,692	33.9	41.3	115.1	12,264	25.7	22.7	18.8	9,428	44.8	56.7	1.8	0.7
Basophil	13,380	30.5	34.5	117.0	7,126	21.5	16.5	21.5	6,254	41.0	53.7	7.0	0.5
Eosinophil	13,246	31.7	31.5	62.8	7,831	26.3	17.0	15.8	5,415	39.8	61.7	3.5	2.2
Plasma cell	17,085	31.6	39.0	54.0	12,043	27.6	21.3	14.0	5,041	40.9	69.5	1.3	2.7

[a]A, area of the cell; B, average color of the cell; C, ratio of cell area to circumference; D, homogeneity of the cell; E, area of nucleus; F, average color of nucleus; G, ratio of nucleus area to circumference; H, homogeneity of the nucleus; I, area of cytoplasm; J, average color of cytoplasm; K, ratio of nucleus area to cell area; L, number of cytoplasmic granules; M, number of nucleoluses.

Pattern Recognition

In the first step, we selected 13 cell characteristics (features) to be extracted after edge detection. Feature extraction was used to obtain input parameters (input neurons) of the artificial neural network (ANN) of the system. Table 1 shows the list of feature vector components (input parameters) used in this system and the results of the measurements performed on different types of nucleated hematopoietic system cells. Figure 2 shows 16 different types of nucleated cells after image processing and analysis ready to be used for feature extraction.

In the second step of pattern recognition, we used supervised ANN for classification (back-propagation learning algorithm) (27–29). Sixteen different types of nucleated hematopoietic cells made up the output neurons of the ANN. Figure 3 shows the schematic representation of our ANN structure. The decision-making logic of this system includes individual recognition of 16 cell types and differentiation of immature cells from the mature ones. In this study, 165 nucleated cell images were used for the training of ANN, and the remaining 82 were used subsequently to test the performance of the system.

RESULTS

The nucleated hematopoietic cells of normal peripheral blood includes neutrophils, lymphocytes, monocytes, eosipnophils, and basophils. Promyelocytes, myelocytes, metamyelocytes, stabs, and plasma cells may be considered normal in bone marrow in certain amounts. However, myeloblasts, monoblasts, proerythroblasts, mormoblasts, and lymhoblasts are considered as immature blastic cells.

Figure 2 shows 16 nucleated hematopietic cells used in this program after image processing and analysis ready to be used for feature extraction. Lymphocytes, neutrophils, monocytes, eosinophils, basophils, promyelocytes, myelocytes, metamyelocytes, stabs, and plasma cells are grouped as mature cells, and myeloblasts, monoblasts, proerythroblasts, normoblasts, and lymphoblasts are grouped as immature cells.

In this study, we used 82 cells to test the performance of the system. Table 2 shows the confusion matrix of this ''testing'' for each individual nucleated cell types.

The confusion matrix was used to compare the similarities and dissimilarities among the decisions of the hematologist and the analysis program. The discriminatory power of the discrimination procedure has been established by this matrix to see whether the discriminator is really useful. Table 3 shows the confusion matrix for mature and immature nucleated cells groups. The matrix provides a convenient method of summarizing the number of correct and incorrect classifications. Diagonal elements of the confusion matrix indicate the number of correct classifications, and the off-diagonal elements indicate the number of incorrect classifications. In this study, the discriminatory power of the procedure was statistically significant: $Q = (N - n.K)2/N.(K - 1) = 28.2$. Thus, it can be concluded that both the hematologist and the analysis program classify the cells as mature and immature in a similar way. Our findings showed that sensitivity and specificity of this system are 71.4% and 90.9%, respectively.

DISCUSSION

Microscopic evaluation of peripheral blood or bone marrow smears has been the task of hematologists, pathologists, and cytologists. Development of automated cell counter and flow cytometry has transferred this time-consuming job from human subjects to automated systems (17–21). Nevertheless, these systems have their disadvantages and application restrictions (22–24).

Transference of microscopic images to computers has gained popularity during the past few years (12–15). An ideal computerized system has not yet been developed for differential recognition of nucleated hematopoietic system cells. In this study, we have developed a semiautomated artificial intelligent diagnostic system for the recognition

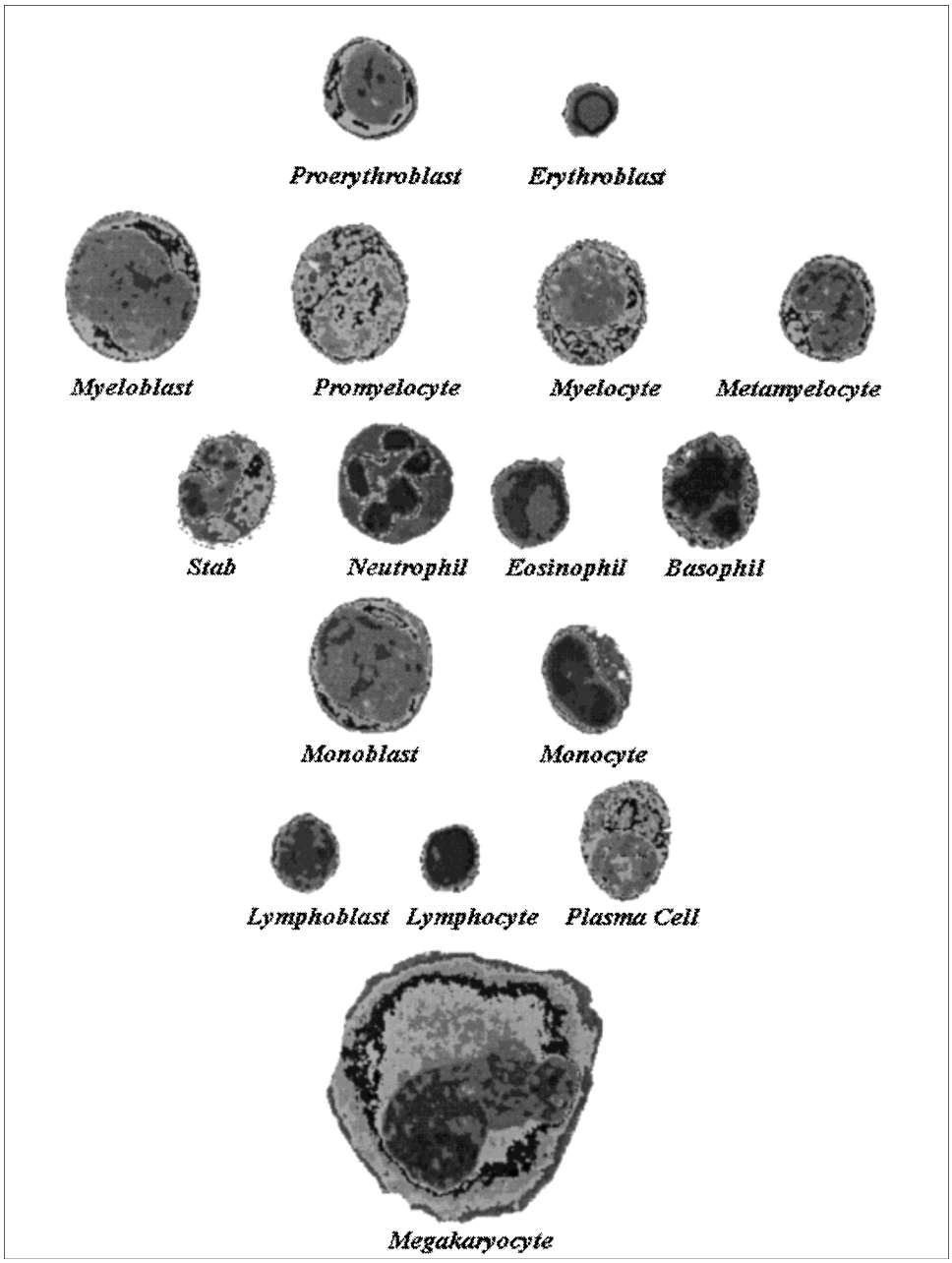

FIG. 2. Computerized outcome of 16 different types of hematopoietic system nucleated cells after image processing and analysis, ready to be used for feature extraction.

of both peripheral and bone marrow nucleated cells to obtain standardization and to prevent problems arising from visual analysis of an inexperienced staff. The decision-making structure of the system is designed to recognize each individual nucleated hematological cell and differentiate mature and immature cells that are to be interpreted subsequently. It has been demonstrated that the sensitivity and specificity of this system are 71.4% and 90.9%, respectively. The consistency between the hematologist and the computerized system is 79.3%. The main discrepancy exists during the evaluation of immature cells, which may be diagnosed as normal (false negativity = 28.5%). However, false positivity is quite rare (9.1%). The system currently enables more successful screening than differential counting of peripheral or bone marrow specimens.

The advantage of this system is in handling physical characteristics of the cell images but not the densitometric specifications of the cells. However, there are still some problems to be solved in the case of overlapping and touching cells. For this reason, we prefer the semiautomated approach. We believe that texture analysis and some other automated cell separation techniques will be

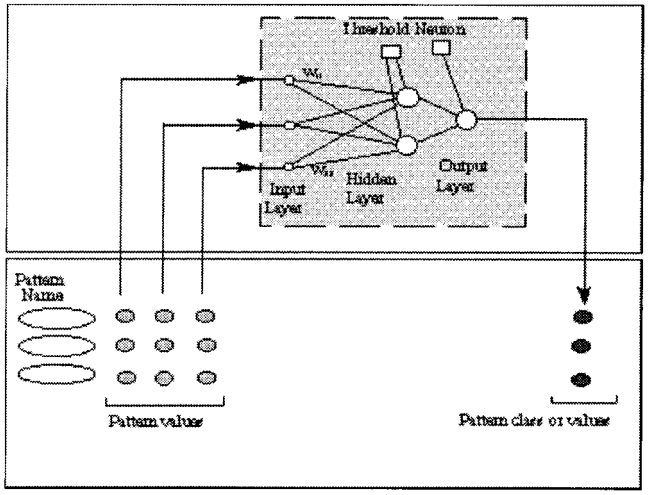

FIG. 3. Schematical representation of our ANN structure.

Table 2
Confusion Matrix of the Test Results of the System and the Medical Expert

Anadolu v. 1.0 evaluates	Medical expert evaluates														
	a	b	d	c	e	n	j	h	i	f	g	l	m	k	o
a. Myeloblast	15			2	2					4					
b. Monoblast		2			1										
d. Normoblast			3		1	1									
c. Proerytroblast										1					
e. Lymphoblast	1				5				2	1					
n. Lymphocyte	1		1		4	9									
j. Neutrophil							5		1						
h. Metamyelocyte			1					3	1					1	
i. Stab									1	1					
f. Promyelocyte	2	1		1	2					9					1
g. Myelocyte	1									2	3				
l. Basophil												1			
m. Monocyte															
k. Eosinophil														1	
o. Plasma cell															1

of benefit for the solution of these problems. The other important issue is to increase the number of cells used in the learning/training phase to increase the negligibility of the system. One disadvantage in this study is the limitation of the hardware facilities, which can be overcome by using more sophisticated data-acquisition and computer systems.

We have used neural network application in the classification of cell types and believe that it is the appropriate methodology for obtaining reasonable results. At the beginning, we tried unsupervised neural network approaches but could not obtain satisfactory results (25,30,31). We later used the supervised ANN structure to achieve our goal (28,29). However, a hybrid ANN structure may be tried by using Kohonen's clustering algorithms to obtain a better performance (25,31).

Our future aim is to combine and adapt our system to a motorized cell finder. This improvement will enable us to broaden our spectrum in clinical applications. Thus, we will be able to quantify and make automated differential counts.

Table 3
Comparison of the Test Results of the System and the Medical Expert

Anadolu v. 1.0 evaluates	Medical expert evaluates	
	Immature	Mature
Immature	35	3
Mature	14	30

In conclusion, this study (Anadolu system) is the first artificial intelligent diagnostic system using neural networks in the recognition of peripheral blood and bone marrow nucleated cells. We believe that further investigations are necessary to achieve our final goal.

ACKNOWLEDGMENTS

We thank Dr. Haluk Koç, Dr. Hamdi Akan, and Dr. Osman Ilhan for giving us their patients' slides.

LITERATURE CITED

1. Kjeldsberg CR: Principles of hematologic examination. In: Wintrobe's Clinical Hematology, Ed. 9, Vol. 1. Lee GR, Bithell TC, Foerster J, Athens JW, Lukens JN (Eds.) Lea and Febiger, Philadelphia, 1993, pp 7–37.
2. Prewitt JMS, Mendelsohn ML: The analysis of cell images. Ann NY Acad Sci 128:1035–1043, 1966.
3. Jarkowski TL, Layton JM, Bahr GF, Wied GL, Bellamy JC, Bartels PH: Computer recognition of cells from asymptomatic lymphocytic leukemia. I. Methodologic study. Acta Cytol 14:147–153, 1971.
4. Brenner JF, Gelsema ES, Necheles TF, Neurath PW, Selles WD, Vastola E: Automatic classification of normal and abnormal leukocytes. J Histochem Cytochem 22:647–706, 1974.
5. Vastola EF, Hertzberg R, Neurath P, Brenner J: Description by computer of texture in Wright-stained leukocytes. Acta Cytol 18:231–247, 1974.
6. Miller AS, Blott BH, Hames TK: Review of neural network applications in medical imaging and signal processing. Med Biol Eng Comput 30:449–464, 1992.
7. Bartels PH, Bahr GF, Bibbo M, Wied GL: Objective cell image analysis. J Histochem 20:239–254, 1972.
8. Bartels PH, Wied GL: Computer analysis and biomedical interpretation of microscopic images: Current problems and future directions. Proc IEEE 65:252–262, 1977.
9. Errington PA, Graham J: Application of artificial neural networks to chromosome classification. Cytometry 14:627–639, 1993.
10. Eskiizmirliler S, Erkmen AM, Çakar N, Beksaç MS: A computerized system for the classification of chromosomes based on pattern recognition and image analysis techniques (Çankaya system). Turk J Med Res 11:4–8, 1993.
11. Beksaç MS, Gören C, Kartal G, Erkmen A: An image processing for umbilical artery blood flow velocity waveforms. Turk J Med Res 10:298–300, 1992.
12. Beksaç MS, Eskiizmirliler S, Erkmen AM, Çakar AN: Çankaya system: An expert system for automated cytogenetics using neural networks. Gynecol Obstet Reprod Med 1:72–77, 1995.
13. Başaran F, Beksaç MS, Erkmen AM: An automated system for the evaluation of antepartum blood flow velocity waveforms: Feature extraction. Turk J Med Sci 18:227–232, 1993.
14. Graham J: Automation of routine clinical chromosome analysis. I: Karyotyping by machine. Anal Quant Cytol Histol 9:383–390, 1987.
15. Lundsteen C, Gerdes T, Granum E, Philip J: Automatic chromosome analysis. II. Karyotyping of banded human chromosomes using band transition sequences. Clin Genet 19:26–36, 1981.
16. Piper J, Baldock R, Tower S, Rutovitz D: Towards a knowledge-based chromosome analysis system. In: Automation of Cytogenetics, Lundsteen C, Piper J (eds). Springer Verlag, Berlin, 1989, pp 275–294.
17. Preston K Jr: Machine techniques for automatic leukocyte pattern analysis. ANN NY Acad Sci 92:482–490, 1962.

18. Bollinger PB, Drewinko B, Brailas CD, Smmeton NA, Truzillo JM: The Technicon H*1™—An automated hematology analyzer for today and tomorrow. Complete blood count parameters. Am J Clin Pathol 87:71–78, 1987.
19. Warner BA, Reardon DM: A field evaluation of the Coulter STKS®. Am J Clin Pathol 95:207–217, 1991.
20. Corash I: Automated leukocyte differential analysis by flow cytochemistry. Lab Med 14:503–508, 1983.
21. Drewinko B, Bollinger P, Rountree M, Johnston D, Corrigan G, Dalton WT, Trujillo JM: Eight-parameter automated hematology analyzer: Comparison of two flow cytometric systems. Am J Clin Pathol 78:738–747, 1982.
22. Dacie JV, Lewis SM: Basic haematological techniques (differential leucocyte count). In: Practical Haematology, Dacie JV, Lewis SM (eds). Churchill Livingstone, Edinburgh, 1984, pp 39–41.
23. Rowan RM. Automated examination of the peripheral blood smear. In: Automation and Quality Assurrance in Haematology, Rowan RM, England JM (eds). Blackwell Scientific, Oxford, 1986, pp 129–177.
24. Barnard DF, Barnard SA, Carter AB, Patterson KG, Yardumian A, Machin SJ: An evaluation of the Coulter VCS differential counter. Clin Lab Haematol 11:255–266, 1989.
25. Beksaç MS, Özdemir K, Erkman A, Karakaş Ü: Assessment of antepartum fetal heart rate tracings using neural networks. In: A Critical Appraisal of Fetal Surveillance, van Geijn HP, Copray FJA (eds). Elsevier, Amsterdam, 1994, pp 354–362.
26. Beksaç MS, Özdemir K, Karakaş Ü, Yalçın S, Karaağoğlu E: Development and application of a simple expert system for the interpretation of the anepartum fetal heart rate tracings (version 88/2.29). Eur J Obstet Gynecol Reprod Biol 37:133–141, 1990.
27. Hopfield JJ: Neural networks and physical systems with emergent collective computational abilities. Proc Nat Acad Sci USA 79:2254–2258, 1982.
28. Rumel Hart DE, Hinton E, Williams RJ: Learning representations by backpropagating errors. Nature 323:533–536, 1986.
29. Sitting DF, Orr JA: A parallel implementation of the backward error propagation neural network training algorithm: Experiments in event identification. Comp Biomed Res 25:547–561, 1992.
30. Kohonen T: Self Organization and Associative Memory, Ed. 3. Springer Verlag, Berlin, 1988.
31. Kohonen T: Correlation matrix memories. IEEE Trans Comput C-21: 353–359, 1970.