EURASIP Journal on Applied Signal Processing 2005:15, 2536–2545

Astrophysical Information from Objective Prism Digitized Images: Classification with an Artificial Neural Network

Emmanuel Bratsolis

Département Traitement du Signal et des Images, École Nationale Supérieure des Télécommunications, 46 rue Barrault, 75013 Paris, France
Email: bratsoli@tsi.enst.fr

Section of Astrophysics, Astronomy and Mechanics, Department of Physics, University of Athens, 15784 Athens, Greece
Email: ebrats@cc.uoa.gr

Received 28 May 2004; Revised 14 December 2004

Stellar spectral classification is not only a tool for labeling individual stars but is also useful in studies of stellar population synthesis. Extracting the physical quantities from the digitized spectral plates involves three main stages: detection, extraction, and classification of spectra. Low-dispersion objective prism images have been used and automated methods have been developed. The detection and extraction problems have been presented in previous works. In this paper, we present a classification method based on an artificial neural network (ANN). We make a brief presentation of the entire automated system and we compare the new classification method with the previously used method of maximum correlation coefficient (MCC). Digitized photographic material has been used here. The method can also be used on CCD spectral images.

Keywords and phrases: objective prism stellar spectra, classification, artificial neural network.

1. INTRODUCTION

Large surveys are concerned with two things. The first is finding unusual objects. Once detected, these unusual objects must always be analyzed individually. The second one is to do statistics with large numbers of objects. In this case, we need an automated classification system.

High-quality film copies of IIIa-J (broad blue-green band) plates, taken with the 1.2 m UK Schmidt Telescope in Australia, have been used. The spectral plates are with dispersion of 2 440 Å/mm at H_γ and spectral range from 3 200 to 5 400 Å. The photographic material has been digitized at the Royal Observatory of Edinburgh using the SuperCOSMOS machine.

Stellar classification with ANNs as a nonlinear technique has been used by many other researchers in the last decade [1, 2, 3, 4, 5]. These methods were utilized for different databases and different spectral dispersion images. In this work, we use wide-field images from the 1.2 m UK Schmidt Telescope in Australia with an objective prism P1. In this case, we can work directly on the image making detection, extraction, classification, and testing of population synthesis. The main contribution here is that there is an automated method, useful to study the spatial distribution of stars (we have the stellar coordinates from the detection method) in groups with the same spectral type (from the classification method). It is useful in astrophysics because we can have a spatial distribution of stellar groups with the same age (grosso modo) and we can study them separetly (morphology, mixture of different populations, etc.).

The final aim of this automated method is to study the stellar population synthesis of Magellanic cloud regions. The detection procedure gives the stellar coordinates on the prism plate [6]. Here we test an ANN based on the classical backpropagation learning procedure.

2. IMAGE REDUCTION

Our test image contains, in pixel size, a region of 3 200 (EW) × 3 150 (SN) of the small Magellanic cloud. The scanning pixel size of the SuperCOSMOS measuring machine is 10 μm and the plate scale is 67.11 arcsec/mm. So our image is centered $\mathrm{RA}_{2\,000} = 1^h16^m$ and $\mathrm{DEC}_{2\,000} = -73^\circ 20'$ and contains a region of 35.8 arcmin (EW) × 35.2 arcmin (SN) of the SMC (Figure 1).

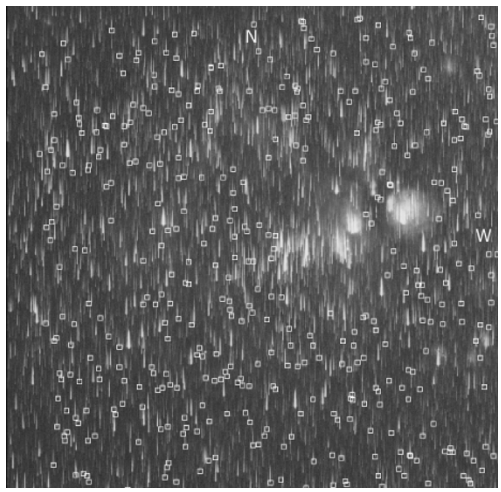


Figure 1: Low-dispersion objective prism image of a region of 35.8 arcmin (EW) × 35.2 arcmin (SN) of the small Magellanic cloud (inner wing). The squares correspond to the positions of the selected spectra for test.

Table 1: Details for the features on objective prism P1. Every spectrum has a length of 128 pixels. The *zero-point* (or detection point) corresponds to the pixel number 10.

Feature	$\lambda(\AA)$	Distance (mm)	Pixel no.
Zero-point	5 400	0.000 ± 0.005	10 ± 1
TiO	5 000	0.100 ± 0.005	20 ± 1
H_β	4 861	0.150 ± 0.005	25 ± 1
TiO	4 800	0.160 ± 0.005	26 ± 1
$H_\gamma + G$	4 340, 4 300	0.320 ± 0.005	42 ± 1
CaI	4 227	0.340 ± 0.005	44 ± 1
H_δ	4 101	0.430 ± 0.005	53 ± 1
$H + H_\epsilon$	3 970	0.500 ± 0.005	60 ± 1
CaII + K	3 936, 3 934	0.520 ± 0.005	62 ± 1
MgI + FeI blend	3 820	0.570 ± 0.005	67 ± 1
FeI+H blend	3 730	0.640 ± 0.005	74 ± 1
FeI blend	3 580	0.740 ± 0.005	84 ± 1

Spectral plates taken with Schmidt-class telescopes contain thousands of spectra. Our initial objective is to detect these spectra and to extract the basic information. The detection algorithm takes as input an image frame, divides the frame in subframes, and applies a signal processing method.

The processing of detection (DETSP) is carried out in four sequential stages [6].

(1) *Image frame preprocessing.* The whole image frame is filtered by a sequence of median and smoothing filters. A grid of subframes is fixed on the filtered image, according to the overlapping mode.

(2) *Subframe signal processing.* Each one of the fixed subframes is processed by applying the detection algorithm based on a signal processing method. The detected spectral positions are saved in table (file).

(3) *Detection table processing.* There are possible double detections of spectra near the edges of neighboring subframes. For this reason, the table of detected spectra is now processed to remove the doubling. It is sorted as well.

(4) *Detection fine adjustment.* The signal processing approach is used again. Now, as many subframes are fixed as the number of detected spectra. The subframes are narrower and each one includes a particular detected spectrum image. This leads to fine adjustment of the position. The adjusted position table is finally sorted.

One of the advantages of the SuperCOSMOS machine is that it scans the plates with a direction parallel to the longitudinal axis of the spectra. Thus, our spectra are parallel to a coordinate axis. The success of DETSP procedure is that it detects all the spectra at the same common-wavelength zero-point at 5 400 Å. This zero-point (0.000 mm) corresponds to our pixel scale (1–128) at 10 pixels.

After the spectral detection, a new procedure starts, responsible for the extraction of spectra (EXTSP) in one-dimensional streams containing all the basic information. The spectral length contains 128 pixels. These are the zero-point plus 118 pixels on the right of zero-point plus 9 pixels on the left of zero-point. For a better signal-to-noise ratio, the actual extraction of the spectrum is performed by means of rectangular weighted "slit" sliding on data. Its width and shape are either fixed or determined by the average fit on the transversal sections of the spectrum.

Our detected zero-point defined by DETSP at 5 400 Å on the dispersion curve of the objective prism P1 helps us to define the distance measurements for various features. The results are shown in Table 1.

The extracted spectra are stored in a two-dimensional file $n \times 128$, where n is the number of detected spectra. Every row of this file is an independent normalized spectrum with length 128 pixels. The maximum number of spectra used for testing here is $N = 426$. The low-dispersion objective prism P1 allow us to classify the stellar spectra only in six classes (OB, A, F, G, K, M). Although the number of classes is limited, the method is useful to study the spatial distribution of stars in groups with the same spectral type.

3. CLASSIFICATION BY USE OF ANN

The objective of classification is to identify similarities and differences between objects and to group them. These groups (classes) are motivated by a scientific understanding of the objects. From spectral energy distributions, we take useful informations about the intrinsic properties of stars like the mass, age and abundances or the related to these like the radius, effective temperature and surface gravity.

An ANN is designed to solve a particular problem by completing two stages: training and verification. During the training stage, a proposed network is provided with a set of examples (input with desired output) of the relationship to be learned, and by implementing specific algorithms,

usually iterative in nature, the network becomes able to reproduce these examples. Once the training stage has been completed, the verification stage, can begin. During the verification stage, a set of new examples, not contained in the training set, are presented to the network. If the network is unable to generalize the new set, then some redesign steps involving addition of more examples and/or modifications in topology of network must be accomplished and the two stages are repeated until satisfactory results are achieved.

ANNs are connectionist systems consisting of many primitive units (artificial neurons) which are working in parallel and are connected via directed links. The general neural unit u_i has M inputs. Each input is weighted with a weight factor w_{ij}, so that input information is $x_i = \sum_{j=1}^{M} w_{ij}u_j$. The main processing principle of these units is the distribution of activation patterns across the links similarly to the basic mechanism of a biological neural network. The knowledge is stored in the structure of the links, their topology and weights which are organized by training procedures. The link connecting two units is directed, fixing a source and a target unit. The weight attributed to a link transforms the output of a source unit to an input on a target unit. This is a supervised learning. Depending on the weight, the transmitted signal can take a value ranging from highly activating to highly forbidding.

The basic function of a unit is to accept inputs from units acting as sources, to activate itself, and to produce one output that is directed to units-targets. Based on their topology and functionality, the units are arranged in layers. The layers can be generally divided into three types: input, hidden, and output. The input layer consists of units that are directly activated by the input pattern. The output one is made by the units that produce the output pattern of the network. All the other layers are hidden and directly inaccessible.

Supervised learning proceeds by minimizing a cost (or error) function with respect to all of the network weights. The cost function J of the network is given by

$$J = \frac{1}{2}\|\mathbf{e}\|^2 = \frac{1}{2}\|\mathbf{t} - \mathbf{y}\|^2, \tag{1}$$

where $\mathbf{t}$ is the desired output vector and $\mathbf{y}$ the response vector of the network to the training pattern.

The activation function f of the unit u_i is given by the sigmoid function

$$y_i = f(x_i) = \frac{1}{1 + \exp\left(-\sum_{j=1}^{M} w_{ij}u_j\right)}. \tag{2}$$

A neuron i in layer l has an output $y_i^{(l)}$ that is given by

$$y_i^{(l)} = f\left(\mathbf{w}_i^{(l)}\mathbf{y}^{(l-1)} - b_i^{(l)}\right), \tag{3}$$

where $\mathbf{w}_i^{(l)}$ is the weight vector of the connections between the neurons from the previous layer $l-1$ and neuron i in layer l, $\mathbf{y}^{(l-1)}$ is the output vector of neurons in layer $l-1$, and $b_i^{(l)}$ is a bias term for the neuron i in layer l [7].

The network training is a nonlinear minimization process in W dimensions, where W is the number of weights in the network. As W is typically large, this can lead to various complications. One of the most important is the problem of local minima. To help avoid local minima, a momentum term is added in the weight update equation.

Weights and biases have been initialized by real random numbers between -1 and 1 and adjusted layer by layer backward, according to the enhanced back-propagation learning rule given by

$$\Delta\mathbf{w}_i^{(l)}(n+1) = -\gamma\nabla J\left(\mathbf{w}_i^{(l)}\right) + \alpha\Delta\mathbf{w}_i^{(l)}(n), \tag{4}$$

where γ is the learning rate of the network, α is a momentum parameter, and n is the number of cycles. A delta learning algorithm (δ_i is the local gradient for the neuron i) has been used for error minimization [8].

According delta rule, the synaptic weights of the network in layer l are

$$\begin{aligned} w_{ij}^{(l)}(n+1) = {} & w_{ij}^{(l)}(n) + \gamma\delta_i^{(l)}(n)u_j^{(l-1)}(n) \\ & + \alpha\left[w_{ij}^{(l)}(n) - w_{ij}^{(l)}(n-1)\right] \end{aligned} \tag{5}$$

and the δ's for the backward computation are

$$\delta_i^{(L)}(n) = e_i^{(L)}(n)y_i(n)\left[1 - y_i(n)\right] \tag{6}$$

for neuron i in output layer L and

$$\delta_i^{(L)}(n) = u_i^{(l)}(n)\left[1 - u_i^{(l)}(n)\right]\sum_k \delta_k^{(l+1)}(n)w_{ki}^{(l+1)}(n) \tag{7}$$

for neuron i in hidden layer l.

Every input causes a response to the neurons of the first layer, which in turn cause a response to the neurons of the next layer, and so on, until a response is obtained at the output layer. The response is then compared with the target response, and the error difference is calculated. From the error difference at the output neurons, the algorithm computes the rate at which the error changes as the activity level of the neuron changes. Here is the end of forward pass. Now, the algorithm steps back one layer before the output layer and recalculates the weights between the last hidden layer and the neurons of the output layer so that the output error is minimized. The algorithm continues calculating the error and computing new weight values, moving layer by layer backward, toward the input. When the input is reached and the weights do not change, the algorithm selects the next pair of

input-target patterns and repeats the process. Although responses move in a forward direction, weights are calculated by moving backward, hence the name back-propagation. As the patterns are chosen randomly, the complete name of this method is "stochastic back-propagation with momentum."

The learning algorithm can be applied as follows.

(1) Initialize the weights to small random values.

(2) Choose a training example pair of input-target $(\mathbf{x}, \mathbf{t})$.

(3) Calculate the outputs $y_i^{(l)}$ from each neuron i in a layer l starting with the input layer and proceeding layer by layer toward the output layer.

(4) Compute the $\delta_i^{(l)}$ and the $w_{ij}^{(l)}$ for each input of the neuron i in a layer l starting with the output layer and backtracking layer by layer toward the input.

(5) Repeat steps (2)–(4) until the termination criterion.

4. CLASSIFICATION BY USE OF MCC

Let $D_{ij} = D_i(\lambda_j)$, $j = 1, \ldots, N$, be the normalized value for the ith stellar spectrum and let $S_{kj} = S_k(\lambda_j)$, $j = 1, \ldots, N$, be the normalized value of the kth class standard stellar spectrum with $k = 1, \ldots, 6$. For $k = 1, \ldots, 6$, the standard stellar spectra are OB, ..., M. The correlation coefficient for the ith stellar spectrum for the kth class is

$$r_{ik} = \frac{\sum_{j=1}^{N} (D_{ij} - \bar{D}_i)(S_{kj} - \bar{S}_k)}{\sqrt{\sum_{j=1}^{N} (D_{ij} - \bar{D}_i)^2} \sqrt{\sum_{j=1}^{N} (S_{kj} - \bar{S}_k)^2}}, \quad (8)$$

with $\bar{D}_i$ being the mean value (over the j variable) of the ith spectrum, $i = 1, \ldots, 426$, and $\bar{S}_k$ the mean value of the kth class standard spectrum.

The correlation coefficient r_{ik} for the ith spectrum for the class k was calculated with displacement ± 3 pixels to predict a possible displacement from the detection algorithm caused by the local background. For these seven correlation coefficients for every class k, the maximum value was chosen. The final classification was given by the maximum value of the coefficient r_i of all the r_{ik} coefficients as

$$r_i = \arg\left(\max r_{ik}\right), \quad k = 1, \ldots, 6. \quad (9)$$

Only spectra with correlation coefficients $r_{ik} > 0.95$ were accepted. In other case, stellar spectra were overlapped or saturated.

5. EXPERIMENTS AND RESULTS

A neural network of three layers and 72 input units, 32 hidden units, and 6 output units has been chosen here. The input units are normalized pixel value units with pixel positions from 11 to 82 corresponding to the central part of digitized spectra. The output units are units corresponding to the six different classes of low-dispersion stellar spectra (OB, A, F, G, K, M). The O and B stars present one class, the OB. The "back-propagation" learning procedure has been used with a "training" mode in which the network learns to associate inputs and desired outputs which are repeatedly presented to it (supervised learning) and a "verification" mode in which the network simply responds to new patterns according to prior training.

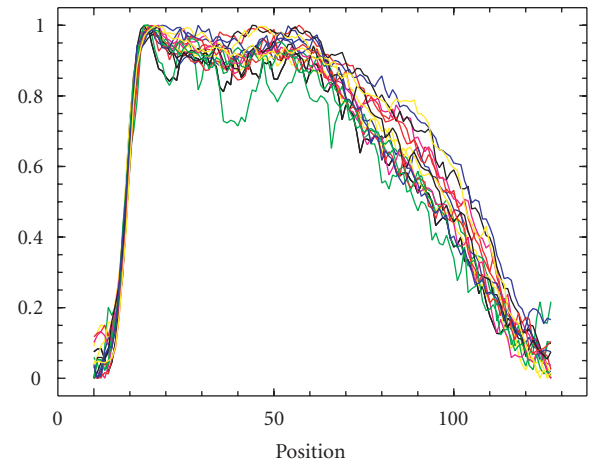


Figure 2: Some of the OB spectra used for training of the ANN.

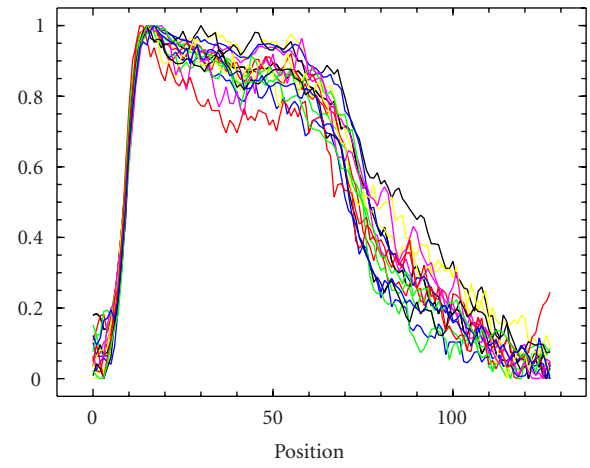


Figure 3: Some of the A spectra used for training of the ANN.

The experimental database consisted of 426 digitized spectra. This allowed us to initialize, update, and train the ANN. No more than 2000 cycles were needed to stabilize the learning with learning rate $\gamma = 0.05$ and momentum parameter $\alpha = 0.1$. Some of the training spectra are presented in Figures 2, 3, 4, 5, 6, and 7. 85 spectra have been used for training, 85 for verification, and 426 for "test."

The results are presented in Tables 2, 3, 4, 5, 6, and 7. The ANN column has numbers with an integer and a decimal part. The integer part corresponds to the accepted-by-the-system class and the decimal part to the percentage gravity of the accepted-by-the-system class. For example, ANN 1.86

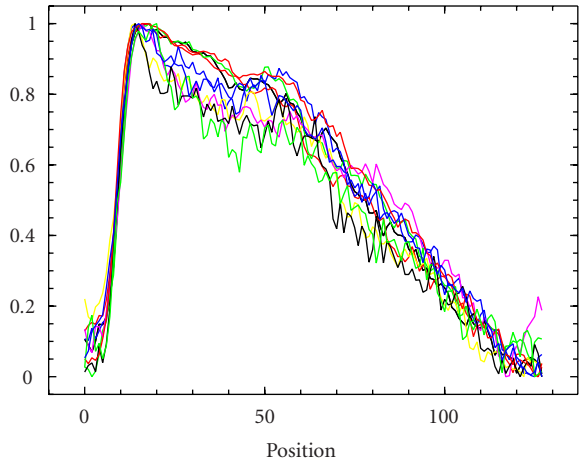


Figure 4: Some of the F spectra used for training of the ANN.

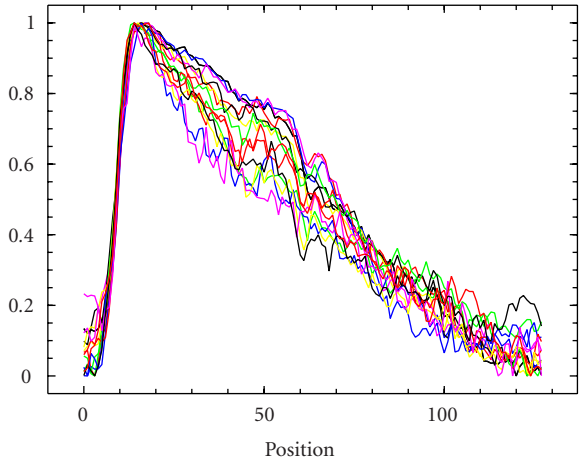


Figure 5: Some of the G spectra used for training of the ANN.

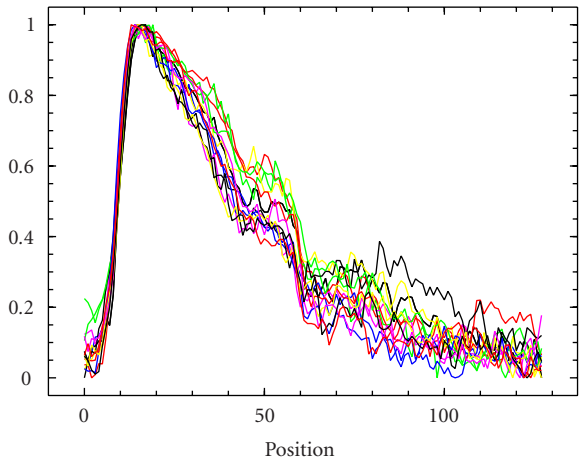


Figure 6: Some of the K spectra used for training of the ANN.

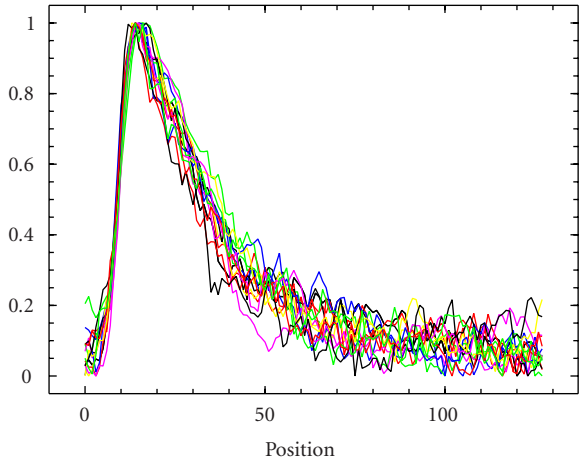


Figure 7: Some of the M spectra used for training of the ANN.

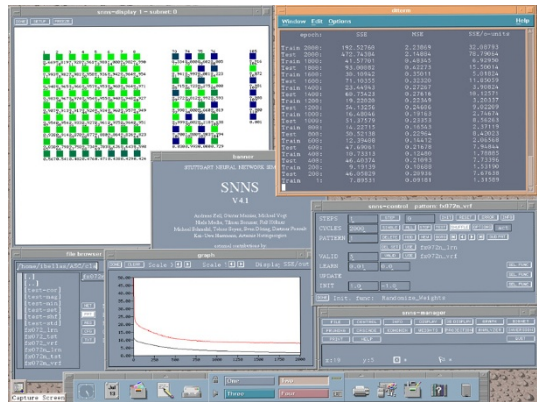


Figure 8: Training by using Stuttgart neural network simulator (72 input units, 32 hidden units, and 6 output units).

means that this spectrum has been classified as 1 (or OB) with gravity 86%. *No* is the number of spectrum (from the sample of 426 spectra), *Sp* the spectral type (class) which is 1 for OB, 2 for A, 3 for F, 4 for G, 5 for K, and 6 for M. *Us* the quality of spectrum which denotes 1 for only recognizable, 2 for good, and 3 for very good spectra. Spectra denoted as 10, 20, and 30 are the corresponding recognizable, good, and very good used for the training of ANN. The ANN has been developed with the freely distributed Stuttgart neural network simulator (Figure 8).

The results for the MCC and ANN methods are presented in Tables 8 and 9. Confusion matrices for two methods, compared with a human expert (HE) results, are presented in Tables 10 and 11. It is evident that the ANN method is better than the MCC. There is an exception for the extreme classes, OB and M, in which the MCC method gives better results. This means that for some specific cases, like the detection of OB stars, the MCC method gives very good results [9].

To quantify the degree of agreement between different classification methods, we have calculated the mean error

Table 2: OB-type spectra from the sample of 426 selected for test. *No* is the sample number of spectrum, *Sp* is the spectral class, *Us* is the quality of spectrum, and *ANN* is the artificial neural network classification. For the *ANN* column, the integer part corresponds to the accepted-by-the-system class and the decimal part to the percentage gravity of the accepted-by-the-system class.

No	*Sp*	*Us*	*ANN*	*No*	*Sp*	*Us*	*ANN*	*No*	*Sp*	*Us*	*ANN*	*No*	*Sp*	*Us*	*ANN*
2	1	2	1.86	72	1	2	1.74	187	1	2	1.68	284	1	2	1.38
3	1	1	1.46	74	1	30	1.70	191	1	2	1.20	287	1	2	1.83
5	1	1	1.74	76	1	1	1.22	192	1	2	1.82	290	1	1	1.88
7	1	2	1.78	77	1	1	1.28	193	1	1	1.28	293	1	2	1.27
9	1	2	1.85	78	1	1	1.70	194	1	1	1.49	299	1	2	1.49
10	1	2	1.90	79	1	1	1.22	201	1	2	1.77	302	1	2	1.17
11	1	3	1.73	80	1	1	1.49	202	1	2	1.74	304	1	2	1.67
12	1	1	1.44	81	1	2	1.75	203	1	2	1.87	305	1	2	2.18
14	1	1	1.29	82	1	1	1.41	207	1	2	1.84	306	1	2	2.34
15	1	10	1.76	89	1	10	1.84	208	1	1	1.50	311	1	2	1.65
17	1	2	1.77	93	1	2	1.87	209	1	2	1.87	313	1	2	1.81
21	1	2	1.88	94	1	2	1.85	211	1	2	1.35	314	1	2	1.26
22	1	3	1.57	96	1	2	1.69	213	1	1	1.68	318	1	2	1.84
24	1	1	1.10	99	1	2	1.75	215	1	1	1.64	320	1	2	1.18
33	1	1	1.89	102	1	1	1.28	216	1	2	1.81	323	1	2	1.81
34	1	2	1.84	103	1	2	1.61	217	1	1	1.76	327	1	1	1.47
35	1	3	1.88	104	1	1	2.18	220	1	2	1.38	335	1	2	2.52
36	1	3	1.86	110	1	1	1.23	221	1	30	1.67	337	1	1	2.26
37	1	2	1.72	118	1	20	1.78	223	1	2	3.18	342	1	2	1.89
40	1	2	1.63	120	1	1	3.22	225	1	1	1.42	344	1	2	1.86
43	1	2	1.76	131	1	2	1.72	227	1	2	1.75	347	1	2	1.19
46	1	20	1.87	139	1	1	1.78	228	1	2	1.17	354	1	2	3.16
47	1	1	1.43	141	1	30	1.58	236	1	2	3.25	358	1	1	1.28
49	1	1	1.68	146	1	1	1.59	238	1	1	1.48	359	1	2	1.40
51	1	1	1.85	149	1	1	1.62	240	1	2	1.83	360	1	30	1.62
53	1	10	1.67	154	1	20	1.82	246	1	2	1.61	367	1	2	1.82
56	1	2	1.45	156	1	2	1.71	251	1	2	1.61	368	1	2	1.52
59	1	2	1.46	157	1	1	2.43	252	1	2	1.38	369	1	10	1.83
60	1	10	1.80	158	1	1	1.54	254	1	2	1.87	377	1	2	3.16
62	1	1	1.87	167	1	20	1.69	256	1	2	1.46	378	1	3	1.70
63	1	10	1.73	172	1	2	1.79	260	1	2	1.87	381	1	2	1.71
65	1	10	1.78	174	1	2	1.62	261	1	1	1.57	383	1	2	1.37
67	1	1	1.53	175	1	1	1.86	273	1	2	1.80	385	1	2	1.38
68	1	30	1.89	178	1	1	1.67	280	1	2	1.52	387	1	1	1.33
69	1	2	1.88	180	1	1	1.72	281	1	2	1.86	390	1	2	1.79
70	1	20	1.82	181	1	1	1.42	283	1	2	2.19	401	1	2	1.61

ME_{HEMCC} between human expert and maximum correlation coefficient classification and ME_{HEANN} between human expert and artificial neural network and the corresponding dispersions σ_{HEMCC} and σ_{HEANN}. C^i_{HE} is the i classified spectrum by the human expert, C^i_{MCC} by the maximum correlation coefficient method, and C^i_{ANN} by the artificial neural network

Table 3: A-type spectra from the sample of 426 selected for test. *No* is the sample number of spectrum, *Sp* is the spectral class, *Us* is the quality of spectrum, and *ANN* is the artificial neural network classification. For the *ANN* column, the integer part corresponds to the accepted-by-the-system class and the decimal part to the percentage gravity of the accepted-by-the-system class.

No	*Sp*	*Us*	*ANN*	*No*	*Sp*	*Us*	*ANN*
4	2	30	2.66	182	2	1	2.40
13	2	1	3.21	206	2	1	2.52
19	2	1	3.31	218	2	2	1.29
25	2	1	2.28	231	2	1	2.21
31	2	1	2.70	241	2	2	2.77
38	2	1	3.16	263	2	1	2.50
42	2	30	2.85	264	2	2	2.78
44	2	2	2.76	269	2	2	2.49
45	2	2	2.53	277	2	2	2.57
50	2	1	1.10	286	2	1	2.35
71	2	10	2.76	296	2	3	2.11
73	2	2	2.31	308	2	2	2.68
83	2	1	3.21	309	2	2	2.79
98	2	30	2.61	312	2	20	2.85
107	2	2	2.77	316	2	1	2.37
109	2	30	2.88	329	2	20	2.59
111	2	1	2.57	338	2	10	2.67
113	2	20	2.88	341	2	2	2.85
119	2	10	2.42	346	2	30	2.84
122	2	2	2.72	357	2	2	2.48
125	2	1	2.11	389	2	1	2.49
129	2	1	2.83	398	2	2	3.20
138	2	1	1.51	405	2	1	2.11
159	2	20	2.85	406	2	1	1.14
161	2	1	2.56	411	2	20	2.31
164	2	1	2.80	414	2	10	2.29
168	2	1	3.36	416	2	1	2.82
169	2	1	2.30	421	2	1	2.59
177	2	20	2.93	424	2	30	2.80

Table 4: F-type spectra from the sample of 426 selected for test. *No* is the sample number of spectrum, *Sp* is the spectral class, *Us* is the quality of spectrum, and *ANN* is the artificial neural network classification. For the *ANN* column the integer part corresponds to the accepted-by-the-system class and the decimal part to the percentage gravity of the accepted-by-the-system class.

No	*Sp*	*Us*	*ANN*
30	3	1	3.32
54	3	20	3.35
61	3	20	3.31
91	3	20	3.55
133	3	1	2.20
140	3	2	3.16
152	3	1	3.20
153	3	20	3.22
183	3	2	3.52
186	3	1	2.28
188	3	1	3.10
198	3	1	3.47
212	3	10	3.24
214	3	10	3.48
219	3	1	3.21
235	3	1	1.27
258	3	2	3.17
262	3	2	3.17
265	3	10	3.24
271	3	1	3.24
285	3	10	3.39
294	3	2	3.33
298	3	1	2.24
331	3	10	3.34
349	3	1	3.28
386	3	20	3.47
426	3	2	3.30

method:

$$\begin{aligned} \mathrm{ME}_{\mathrm{HEMCC}} &= \frac{1}{426}\sum_{i=1}^{426}\left|C_{\mathrm{HE}}^{i}-C_{\mathrm{MCC}}^{i}\right| = 0.15, \\ \sigma_{\mathrm{HEMCC}} &= \sqrt{\frac{1}{426}\sum_{i=1}^{426}\left(C_{\mathrm{HE}}^{i}-C_{\mathrm{MCC}}^{i}\right)^{2}} = 0.43, \\ \mathrm{ME}_{\mathrm{HEANN}} &= \frac{1}{426}\sum_{i=1}^{426}\left|C_{\mathrm{HE}}^{i}-C_{\mathrm{ANN}}^{i}\right| = 0.13, \\ \sigma_{\mathrm{HEANN}} &= \sqrt{\frac{1}{426}\sum_{i=1}^{426}\left(C_{\mathrm{HE}}^{i}-C_{\mathrm{ANN}}^{i}\right)^{2}} = 0.37. \end{aligned} \tag{10}$$

Tables 12 and 13 give the global statistical properties between different classification methods.

Figure 1 shows the region N83-84-85 which belongs to the inner wing of the SMC and is of interest because of its OB associations and nebulae. This region has been studied by different astronomers in the past. It is evident that there is a correlation between associations like NGC 456, 460a,b, and 465 with the nebulae of ionized gas.

There are groups of stars with age variations of 4–10 Myr and spatial scales of 30–400 pc. There is also an extended region containing N83-84-85 with a diameter of more than 500 pc and sequential star formation on a scale of 10^7 yr which seems to be part of a supergiant shell.

We focus on this region because it seems to show a feedback between OB star formation and the physical properties

Table 5: G-type spectra from the sample of 426 selected for test. *No* is the sample number of spectrum, *Sp* is the spectral class, *Us* is the quality of spectrum, and *ANN* is the artificial neural network classification. For the *ANN* column, the integer part corresponds to the accepted-by-the-system class and the decimal part to the percentage gravity of the accepted-by-the-system class.

No	*Sp*	*Us*	*ANN*	*No*	*Sp*	*Us*	*ANN*
6	4	2	4.61	189	4	1	4.36
8	4	10	4.54	222	4	30	4.65
23	4	1	4.59	244	4	2	4.33
27	4	1	4.51	249	4	1	4.50
57	4	2	4.23	250	4	2	4.41
64	4	2	4.28	253	4	2	4.56
66	4	2	4.65	259	4	2	4.19
86	4	20	4.36	270	4	30	4.57
87	4	30	4.58	278	4	1	4.65
95	4	2	4.54	279	4	2	4.42
100	4	1	4.64	288	4	2	4.51
106	4	30	4.42	291	4	1	3.45
112	4	1	4.61	297	4	20	4.70
115	4	1	4.56	303	4	2	4.36
117	4	2	4.47	315	4	1	4.59
128	4	1	4.40	322	4	2	4.56
137	4	1	4.38	332	4	2	4.39
143	4	20	4.63	352	4	30	4.56
145	4	2	4.57	362	4	20	4.65
150	4	1	5.74	370	4	2	4.37
160	4	10	4.29	388	4	30	4.58
162	4	1	4.60	394	4	2	4.51
170	4	2	4.46	396	4	30	4.69
173	4	20	4.58	403	4	1	4.68
176	4	2	4.66	413	4	2	4.32
185	4	2	4.58	422	4	1	4.59

Table 6: K-type spectra from the sample of 426 selected for test. *No* is the sample number of spectrum, *Sp* is the spectral class, *Us* is the quality of spectrum, and *ANN* is the artificial neural network classification. For the *ANN* column, the integer part corresponds to the accepted-by-the-system class and the decimal part to the percentage gravity of the accepted-by-the-system class.

No	*Sp*	*Us*	*ANN*	*No*	*Sp*	*Us*	*ANN*
1	5	30	5.80	234	5	2	5.27
18	5	30	5.82	237	5	2	5.58
26	5	10	5.83	239	5	2	5.81
32	5	2	5.56	242	5	2	5.44
39	5	2	5.61	255	5	2	5.65
41	5	1	4.23	266	5	3	5.69
58	5	20	5.75	267	5	2	5.85
75	5	20	5.44	268	5	2	4.48
85	5	30	5.63	275	5	2	5.64
88	5	1	4.41	292	5	2	5.68
90	5	1	5.70	300	5	2	5.87
97	5	3	4.59	325	5	2	5.36
124	5	1	5.77	334	5	2	5.85
134	5	2	5.29	336	5	2	5.44
135	5	2	5.60	339	5	2	5.86
136	5	1	5.83	340	5	2	5.79
142	5	20	5.54	355	5	2	5.58
144	5	20	5.87	361	5	2	4.33
147	5	20	5.59	366	5	2	4.37
155	5	2	5.52	371	5	2	4.27
163	5	2	5.73	376	5	2	5.88
165	5	20	5.74	379	5	30	5.68
195	5	2	5.67	380	5	2	5.29
196	5	2	4.60	384	5	30	5.79
197	5	2	4.40	395	5	2	5.73
199	5	2	5.43	397	5	3	5.72
200	5	2	5.42	399	5	2	5.84
204	5	10	5.67	402	5	2	5.74
205	5	2	4.40	410	5	1	5.63
210	5	1	4.53	412	5	1	5.60
224	5	2	5.29	415	5	2	5.72
226	5	20	5.86	417	5	2	4.36
230	5	1	5.68	419	5	2	5.55
233	5	2	5.60	420	5	1	5.62

of the interstellar medium. It suggests that star formation and interstellar medium properties probably are self-regulated. Our automated method helped us to show some morphological characteristics of these OB associations and to explain the triggered star formation by possible supernova explosions.

We have to note that the reddening for the used SMC stellar spectra is considered negligible and the training of the ANN has been made directly by the normalized measured spectra of our region.

Initially the field of view of CCD images was much smaller than photographic plates, but digital detectors have now caught up with the development of CCD mosaic cameras. Our algorithm can be used also on CCD spectral images. It is evident that if the resolution of digitized plates is the same as the resolution of CCD spectral images, the method is exactly the same. In the other case, we have to modify the method to the new resolution.

This method is used for the moment as classification tool only at the University of Athens.

6. CONCLUSIONS

In this paper, we have described an automated method of classification for digitized low-dispersion objective prism stellar spectra by using an ANN system. This method has

Table 7: M-type spectra from the sample of 426 selected for test. *No* is the sample number of spectrum, *Sp* is the spectral class, *Us* is the quality of spectrum, and *ANN* is the artificial neural network classification. For the *ANN* column, the integer part corresponds to the accepted-by-the-system class and the decimal part to the percentage gravity of the accepted-by-the-system class.

No	*Sp*	*Us*	*ANN*	*No*	*Sp*	*Us*	*ANN*
16	6	2	6.33	295	6	2	6.82
20	6	30	6.94	301	6	20	6.94
28	6	30	6.96	307	6	2	6.90
29	6	30	6.60	310	6	2	6.94
48	6	3	6.92	317	6	2	6.70
52	6	20	6.90	319	6	2	6.93
55	6	30	6.93	321	6	2	6.79
84	6	1	6.33	324	6	2	6.45
92	6	10	6.96	326	6	2	6.88
101	6	2	5.31	328	6	2	6.95
105	6	30	6.87	330	6	2	6.93
108	6	30	6.95	333	6	2	6.93
114	6	3	6.97	343	6	2	5.84
116	6	2	6.74	345	6	2	6.93
121	6	1	6.54	348	6	2	6.47
123	6	1	5.48	350	6	2	6.24
126	6	20	6.76	351	6	2	6.91
127	6	3	6.93	353	6	2	6.92
130	6	1	6.71	356	6	2	6.32
132	6	2	5.54	363	6	2	6.87
148	6	10	6.75	364	6	2	5.18
151	6	2	6.87	365	6	2	6.95
166	6	10	6.81	372	6	2	6.91
171	6	20	6.91	373	6	2	6.96
179	6	2	6.75	374	6	2	6.93
184	6	10	6.97	375	6	2	6.78
190	6	2	6.92	382	6	2	6.91
229	6	2	6.94	391	6	2	5.63
232	6	2	6.95	392	6	2	6.24
243	6	2	6.58	393	6	3	6.96
245	6	2	6.95	400	6	2	6.86
247	6	1	5.50	404	6	10	6.87
248	6	3	5.76	407	6	2	6.54
257	6	2	6.79	408	6	3	6.97
272	6	2	6.37	409	6	2	6.91
274	6	2	6.95	418	6	2	6.66
276	6	2	6.93	423	6	2	6.84
282	6	2	6.92	425	6	2	6.65
289	6	2	5.89				

Table 8: Results for the MCC after the test with 426 stellar spectra.

MCC Class	Human expert classification					
	OB	A	F	G	K	M
−3	0	0	0	0	0	0
−2	0	0	2	0	0	1
−1	0	8	7	2	9	4
0	141	40	16	45	52	72
+1	1	9	2	5	7	0
+2	2	1	0	0	0	0
+3	0	0	0	0	0	0
Bad class	3	18	11	7	16	5
All used	144	58	27	52	68	77

Table 9: Results for the ANN after the test with 426 stellar spectra.

ANN Class	Human expert classification					
	OB	A	F	G	K	M
−3	0	0	0	0	0	0
−2	0	0	1	0	0	0
−1	0	4	3	1	12	9
0	132	48	23	50	56	68
+1	7	6	0	1	0	0
+2	5	0	0	0	0	0
+3	0	0	0	0	0	0
Bad class	12	10	4	2	12	9
All used	144	58	27	52	68	77

Table 10: Confusion matrix for human expert and MCC combination. The results are expressed in percentages.

	OB	A	F	G	K	M
OB	97.92	0.69	1.39	0.00	0.00	0.00
A	13.79	68.97	15.52	1.72	0.00	0.00
F	7.41	25.92	59.26	7.41	0.00	0.00
G	0.00	0.00	3.85	86.54	9.61	0.00
K	0.00	0.00	0.00	13.24	76.47	10.29
M	0.00	0.00	0.00	1.30	5.19	93.51

Table 11: Confusion matrix for human expert and ANN combination. The results are expressed in percentages.

	OB	A	F	G	K	M
OB	91.67	4.86	3.47	0.00	0.00	0.00
A	6.90	82.76	10.34	0.00	0.00	0.00
F	3.70	11.11	85.19	0.00	0.00	0.00
G	0.00	0.00	1.92	96.16	1.92	0.00
K	0.00	0.00	0.00	17.65	82.35	0.00
M	0.00	0.00	0.00	0.00	11.69	88.31

been compared with the previously used MCC method and gave better results. The automated classification is a part of a fully automated method, developed for stellar detection, extraction of basic information, and classification from low-dispersion objective prism images. The detected objects with their coordinates are stored in tables (files). The method is useful because we can study the spatial distribution of stars in groups with the same spectral type.

Table 12: Statistical properties for the different classification methods: human expert, maximum correlation coefficient, and artificial neural network.

Spectral-type class	OB	A	F	G	K	M
	1	2	3	4	5	6
HE	144	58	27	52	68	77
MCC	148	48	32	61	62	75
ANN	137	58	35	62	66	68

Table 13: Statistical comparison between the different classification methods: human expert-maximum correlation coefficient, and human expert-artificial neural network.

Test	Mean error (ME)	Dispersion (σ)
HEMCC	0.15	0.43
HEANN	0.13	0.37

ACKNOWLEDGMENT

The author is grateful to I. Bellas-Velidis for fruitful discussions.

REFERENCES

[1] T. von Hippel, L. J. Storrie-Lombardi, M. C. Storrie-Lombardi, and M. J. Irwin, "Automated classification of stellar spectra—I. Initial results with artificial neural networks," *Monthly Notices of the Royal Astronomical Society*, vol. 269, no. 1, pp. 97–104, 1994.

[2] R. K. Gulati, R. Gupta, P. Gothoskar, and S. Khobragade, "Stellar spectral classification using automated schemes," *Astrophysical Journal*, vol. 426, no. 1, pp. 340–344, 1994.

[3] E. F. Vieira and J. D. Ponz, "Automated classification of IUE low-dispersion spectra. I. Normal stars," *Astronomy and Astrophysics Supplement Series*, vol. 111, pp. 393–398, 1995.

[4] H. P. Singh, R. K. Gulati, and R. Gupta, "Stellar spectral classification using principal component analysis and artificial neural networks," *Monthly Notices of the Royal Astronomical Society*, vol. 295, no. 2, pp. 312–318, 1998.

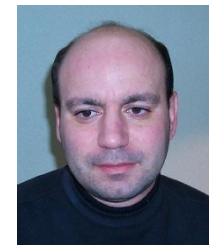
[5] C. A. L. Bailer-Jones, M. Irwin, and T. von Hippel, "Automated classification of stellar spectra—II. Two-dimensional classification with neural networks and principal components analysis," *Monthly Notices of the Royal Astronomical Society*, vol. 298, no. 2, pp. 361–377, 1998.

[6] E. Bratsolis, I. Bellas-Velidis, E. Kontizas, F. Pasian, A. Dapergolas, and R. Smareglia, "Automatic detection of objective prism stellar spectra," *Astronomy and Astrophysics Supplement Series*, vol. 133, no. 2, pp. 293–297, 1998.

[7] D. E. Rumelhart and J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing*, vol. 1, MIT Press, Cambridge, Mass, USA, 7th edition, 1988.

[8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan, New York, NY, USA, 1994.

[9] E. Bratsolis, M. Kontizas, and I. Bellas-Velidis, "Triggered star formation in the inner wing of the SMC. Two possible supernova explosions in the N83-84-85 region," *Astronomy and Astrophysics*, vol. 423, no. 3, pp. 919–924, 2004.



Emmanuel Bratsolis was born in Greece. He recieved a B.S. degree in physics from the University of Athens (UA), Greece, an M.S. degree in astrophysics and space technology from the University of Paris VII, France, and an M.S. degree in signal and image processing from École Nationale Supérieure des Télécommunications (ENST), Paris, France. He also recieved a Ph.D. degree in astrophysics (UA) and a Ph.D. degree in image processing (ENST). He has been a researcher in different projects. His research interests include image and signal processing, remote sensing, and data analysis in astrophysics.