Automatic character recognition based on graph theory. A new approach to automation

EDINEI PERES LEGASPE
Power and Electric Automation Engineering Department
Escola Politécnica da Universidade de São Paulo
Av. Professor Luciano Gualberto, Trav. 3, 158, Room A2-06
BRAZIL
edineilegaspe@hotmail.com

WELLINGTON SPERANDIO SILVA
Power and Electric Automation Engineering Department
Escola Politécnica da Universidade de São Paulo
Av. Professor Luciano Gualberto, Trav. 3, 158, Room A2-06
BRAZIL
wellington.sperandio@usp.br

CAIO FERNANDO FONTANA
Power and Electric Automation Engineering Department
Escola Politécnica da Universidade de São Paulo
Av. Professor Luciano Gualberto, Trav. 3, 158, Room A2-06
BRAZIL
caioffontana@uol.com.br

EDUARDO MARIO DIAS
Power and Electric Automation Engineering Department
Escola Politécnica da Universidade de São Paulo
Av. Professor Luciano Gualberto, Trav. 3, 158, Room A2-06
BRAZIL
emdias@pea.usp.br – http://www.gaesi.eng.br

Abstract: The aim of this paper is to present a new method Optical Character Recognition (OCR). For it will be used for projections of images and the technique of Dijkstra's shortest path in graph-based structures can To validate the model presented in this work is implemented using software techniques commented above in the knowledge base organized around the Latin alphabet. We made several tests on the software and thus validating the proposed model. The presented method has the benefit that the training time of knowledge is much lower when compared to a neural network that performs the same operation.

Keywords: OCR, graphs, Dijkstra, Image recognition.

1. Introduction

The automation process is already something very widespread today, being used in various applications.

One area that is being privileged to automate processes in recent times is the area of ITS (Intelligent Transportation System).

And one of those needs in automation is the automatic recognition and optical character, and the area of ITS may be mentioned applications:

a. Radars
b. Vehicular
c. Reading shipping container

It is important that the recognition system is accurate and fast.

There are various software in the market that do the optical character recognition [1].

There are several techniques for recognition, but rather a technique is used today is the neural networks [2] [3] [4].

But the training of these neural networks in the vast majority using the *Backpropagation algorithm*, which is quite slow at this stage of training, often lingering in the range of seconds or even minutes [5] [6].

The purpose of this paper is presenting a new recognition method based on the techniques of graph theory and algorithm of Dijkstra's shortest path aiming to lower training time.

To demonstrate the validity of the current solution, we developed a prototype software to validate the same.

2. Problem formulation

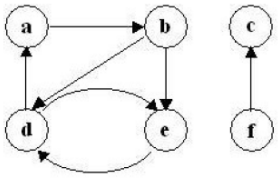
Because the problem is relatively complex, it is necessary to divide it into smaller parts and then address each item separately. In the next subsections will present the techniques for forming optical recognition used in this work.

2.1. Graph theory

The Graph Theory (GT) is a branch of mathematics that began in the eighteenth century with the mathematician Leonard Euler (1707-1783). The development of TG had more momentum with applications focused on optimization problems in the second half of the twentieth century, along as the advent of computers, without which many of the applications of TG would be impossible [7] [8].

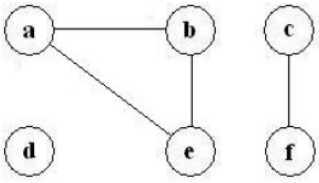
There are several definitions of what a graph. There are some differences in definitions and terminology presented in the literature. Despite these variations in a directed graph G is a pair (N, A), where N is a finite set of nodes or vertices that make up G [7] [8].

The binary relation that can exist between nodes is called an arc or edge. This finite set of arcs gives rise to another finite set A, whose elements are the arcs or edges are valid for a graph G [8]. Figures 1 and 2 show examples of directed and undirected graphs.



N = {a,b,c,d,e,f}
A = {(a,b), (b,d), (b,e), (d,a), (d,e), (e,d), (f,c)}

Figure 1 - directed graph



N = {a,b,c,d,e,f}
A = {(a,b), (a,e), (f,c)}

Figure 2 - undirected graph

In Figure 1, one can observe a sense of direction of the arcs in this graph is

given the name of directed graph [7] [8].

These structures based on graphs can represent various models as special structures, map of cities, routes, among other applications.

A possible computer implementation of graph structures are the adjacency lists or adjacency matrices [8].

2.2. Dijkstra shortest path

The shortest path algorithm was suggested by Edsger Dijkstra in 1959, from the need to find the path of shortest distance between two points x and y. Today this need is present in determining routes between cities, technical logistics for load distribution and products, computer networks, among others [8].

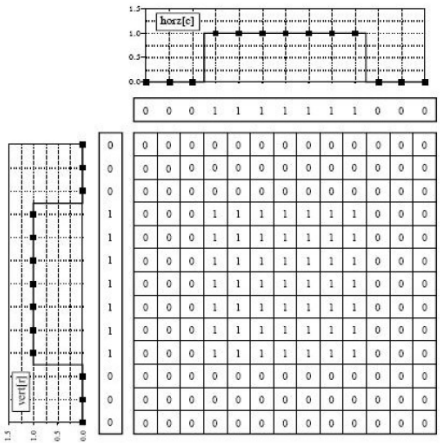
According to the Dijkstra algorithm aims at finding the path whose sum of these weights in a weighted connected graph, is the smallest among the vertices x and y.

The proposed algorithm is demonstrated by Dijkstra (bellow).

The algorithm part of an initial estimate for the minimum cost and will successively adjusting this estimate. He believes that a vertex will be closed when it was obtained a minimum cost path from vertex taken as the search root to it. Otherwise he said to be open [5].

2.3. Projection

The technique of convolution by separability is a rapid technique that allows for a separate image, a two-dimensional signal in two-dimensional signs, treating these signs as projections, has a vertical projection and horizontal projection of an image. Figure 4 illustrates the vertical and horizontal projections [9].



1. Assign zero to the estimated minimum cost of the vertex s (the root of the search) and the other infiniteestimates;
2. Assign a value to any previous (the precedent of a vertex t is the vertex that precedes t in the least cost path from s to t);
3. As long as there apex open:
the vertex k is still open which estimate is the lowest of all the open vertices;
close to the vertex k
o For every vertex j that is still open successor to k do:
♣ some estimate of the vertex k with the cost of the arc that joins kaj;
♣ If this sum is better than the previous estimate for the vertex j, replace it and write k as a precedent for j.

For automatic recognition of characters is necessary to develop a knowledge base. This knowledge base will be provided by horizontal projections (topic covered in this work item 2.3). Figure 5 displays the projection of the letter E.

Imagem da letra "E"									∑
2	2	2	2	2	2	2	2	2	20
2	2	1	1	1	1	1	2	2	14
2	2	1	1	1	1	1	2	2	14
2	2	1	1	2	2	2	2	2	18
2	2	1	1	1	1	1	2	2	14
2	2	1	1	1	1	1	2	2	14
2	2	1	1	2	2	2	2	2	18
2	2	1	1	1	1	1	2	2	14
2	2	1	1	1	1	1	2	2	14
2	2	2	2	2	2	2	2	2	20

Figure 5 - Example of projection of the letter E

This case was adopted for the blank pixel value and the two black pixels in a value, making the flat sum has the value of each item in the projection.

The procedure should be done with all the characters of the alphabet, thus generating a set of projections that will be used as a knowledge base structured in a graph.

3.2. Construction of the knowledge graph

In Figure 6, displays the array of projections which are demonstrated in three different projections of three figures (ie, three possible characters for the knowledge base).

11	10	9	8	7	6
11	9	9	7	8	6
11	9	7	10	8	6

Figure 6 - Matrix projections of three different figures

Having the array of projections is possible to construct the graph that represents the knowledge base.

To construct the graph of knowledge must first be created nodes: start and end, which are respectively called here S and E.

Analyzing the matrix of projections in Figure 6, checks if there is a triplicity in the first column with the number 11, so it generated only one node in the graph, with the value 11. Moving in the next column of Figure 6, has the values 10 and 9 in the amount of 1 and 2 respectively, then two nodes must be created with these values in the graph. The procedure should be done with all the columns of the matrix of projections, with a directed graph as shown in Figure 7.

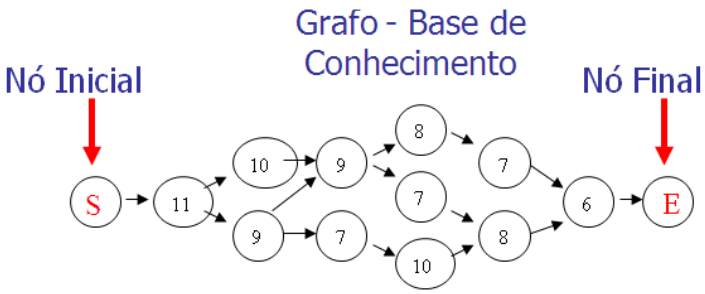


Figure 7 - Graph knowledge base

The shape of the directed graph is determined by the columns of the projection matrix.

The conversion of the matrix of projections of Figure 6 in a directed graph can be seen in Figure 7.

With the graph constructed in memory, you have the knowledge base for a recognition.

ISBN: 978-1-61804-056-5

3.3. Recognition

In this the first time will be considered the same recognition with an image that is trained in the graph.

For the graph shown in Figure 7, is placed with the second projection of Figure 6, for the recognition. Figure 8 displays the path to the previously mentioned entry.

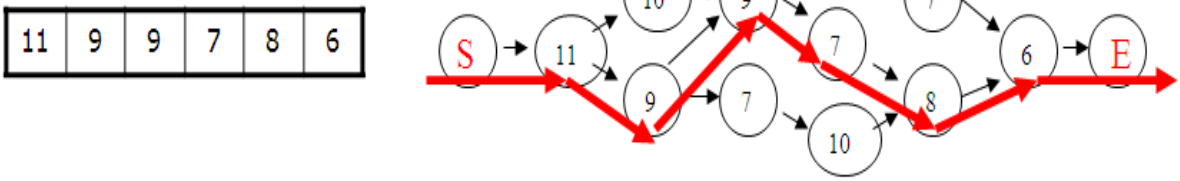


Figure 8 - Projected path input and path in the graph

Making each node in the graph has an extra information that defines that original projections are connected with it, then the path followed by the graph is given an input it is possible to determine with whom she most closely resembles the projection.

In the real world input to the recognition is not identical to the contents trained, so the system or apply a similar approach.

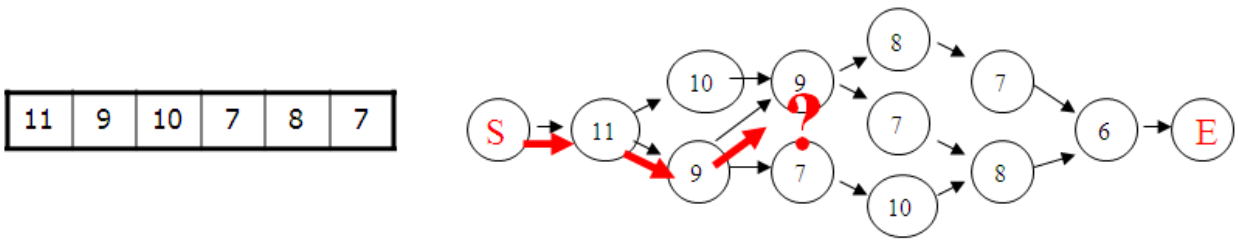
For this problem it is proposed to use a factor of error calculation, shown in equation 1.

$$E_i = \left| 1 - \frac{P_i}{G_i} \right| \quad (1)$$

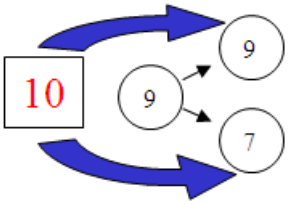
The error added to each of the values of the graph is derived from the values of the projection that refuse to recognize. This error value is calculated using equation 1, where i is the index of the element of the projection P and the level of the graph G.

In Figure 9, an entry is placed not identical to recognize when it is applied to the third column in the projection graph of knowledge, it is noted that there is a similar amount, so it is necessary to calculate the error in both cases.

Figure 9 - Example of the non-identical



The error calculation is done according to procedure described above based on an equation. Figure 10 is shown the error calculation for the case shown in Figure 9.



Erro_1 = | 1 - 10 / 9 | = 0.11

Erro_1 = | 1 - 10 / 7 | = 0.42

Figure 10 - Calculation of error for the nodes

As Figure 10, the smallest error is the way to the top of 0.11.

Using the error factor in determining degrees of similarity between nodes and the Dijkstra shortest path algorithm that uses the error factor to determine the best degree of similarity or words you enter the best way to generate an error, it is possible to recognize characters with a knowledge base pre-defined.

In order to validate the solution proposed was implemented software that performs the character recognition.

3.4. Implementing the Solution

The proposed solution was built in C + +. He was born as a knowledge base set of pictures representing the letters of the alphabet.

Figure 11 displays the test software, with the load set of letters to the formation of projections to the knowledge base.

Figure 11 - Recognition Software

For the tests, we used figures on the size of 50x50 pixels as the training base.

In the figure 11 is shown the creation time plus the time of the projections of the graph construction of knowledge, getting around 31 ms.

In figure 12 we show the values of all the projections, in this figure one can see the values of each of the projections loaded.

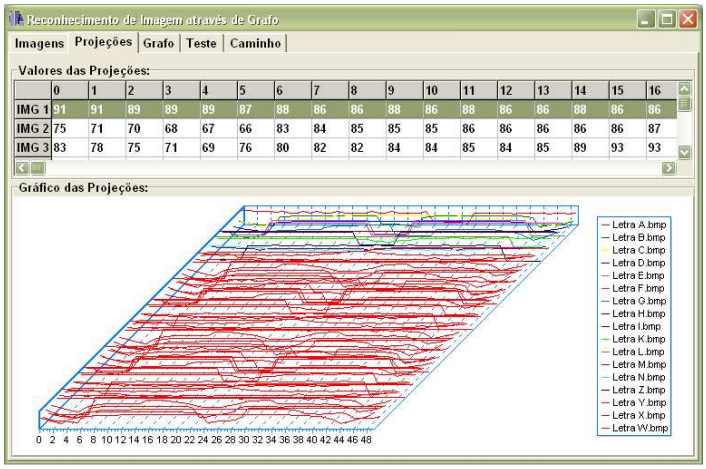
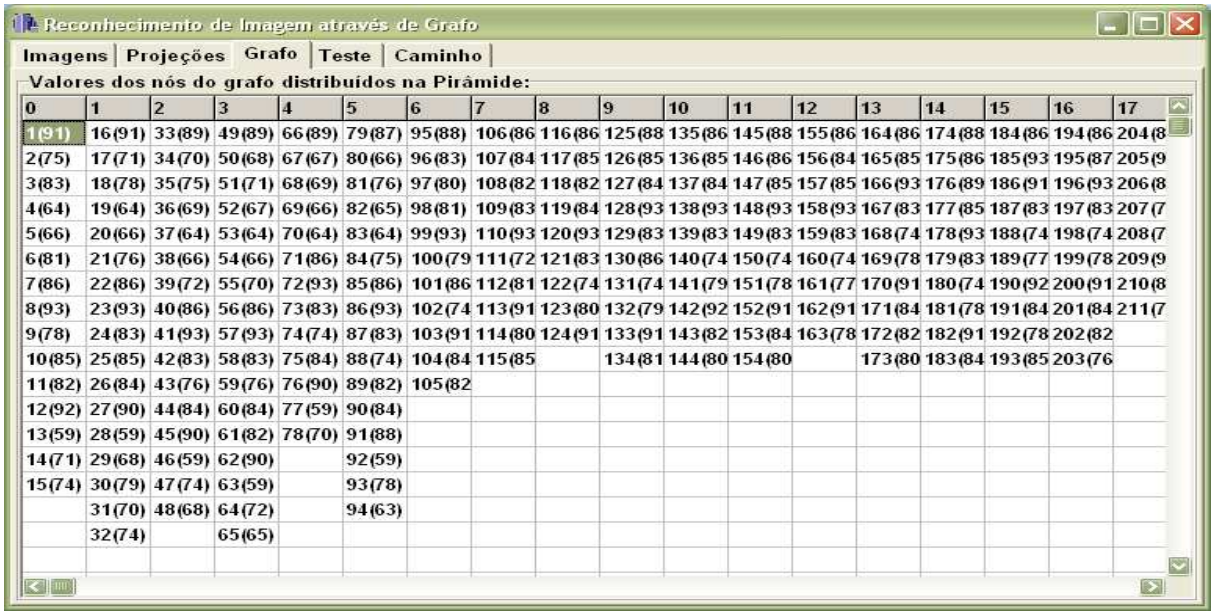


Figure 12 - Values of the projections used in training

Figure 13 displays a textual version of the graph constructed as knowledge base.

ISBN: 978-1-61804-056-5



0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1(91)	16(91)	33(89)	49(89)	66(89)	79(87)	95(88)	106(86	116(86	125(88	135(86	145(88	155(86	164(86	174(88	184(86	194(86	204(8
2(75)	17(71)	34(70)	50(68)	67(67)	80(66)	96(83)	107(84	117(85	126(85	136(85	146(86	156(84	165(85	175(86	185(93	195(87	205(9
3(83)	18(78)	35(75)	51(71)	68(69)	81(76)	97(80)	108(82	118(82	127(84	137(84	147(85	157(85	166(93	176(89	186(91	196(93	206(8
4(64)	19(64)	36(69)	52(67)	69(66)	82(65)	98(81)	109(83	119(84	128(93	138(93	148(93	158(93	167(83	177(85	187(83	197(83	207(7
5(66)	20(66)	37(64)	53(64)	70(64)	83(64)	99(93)	110(93	120(93	129(83	139(83	149(83	159(83	168(74	178(93	188(74	198(74	208(7
6(81)	21(76)	38(66)	54(66)	71(86)	84(75)	100(79	111(72	121(83	130(86	140(74	150(74	160(74	169(78	179(83	189(77	199(78	209(9
7(86)	22(86)	39(72)	55(70)	72(93)	85(86)	101(86	112(81	122(74	131(74	141(79	151(78	161(77	170(91	180(74	190(92	200(91	210(8
8(93)	23(93)	40(86)	56(86)	73(83)	86(93)	102(74	113(91	123(80	132(79	142(92	152(91	162(91	171(84	181(78	191(84	201(84	211(7
9(78)	24(83)	41(93)	57(93)	74(74)	87(83)	103(91	114(80	124(91	133(91	143(82	153(84	163(78	172(82	182(91	192(78	202(82	
10(85)	25(85)	42(83)	58(83)	75(84)	88(74)	104(84	115(85		134(81	144(80	154(80		173(80	183(84	193(85	203(76	
11(82)	26(84)	43(76)	59(76)	76(90)	89(82)	105(82											
12(92)	27(90)	44(84)	60(84)	77(59)	90(84)												
13(59)	28(59)	45(90)	61(82)	78(70)	91(88)												
14(71)	29(68)	46(59)	62(90)		92(59)												
15(74)	30(79)	47(74)	63(59)		93(78)												
	31(70)	48(68)	64(72)		94(63)												
	32(74)		65(65)														

Figure 13 - Graph of textual knowledge

In figure 14 loads a value test, where a figure that resembles with the letter A, and then created a projection for this entry.

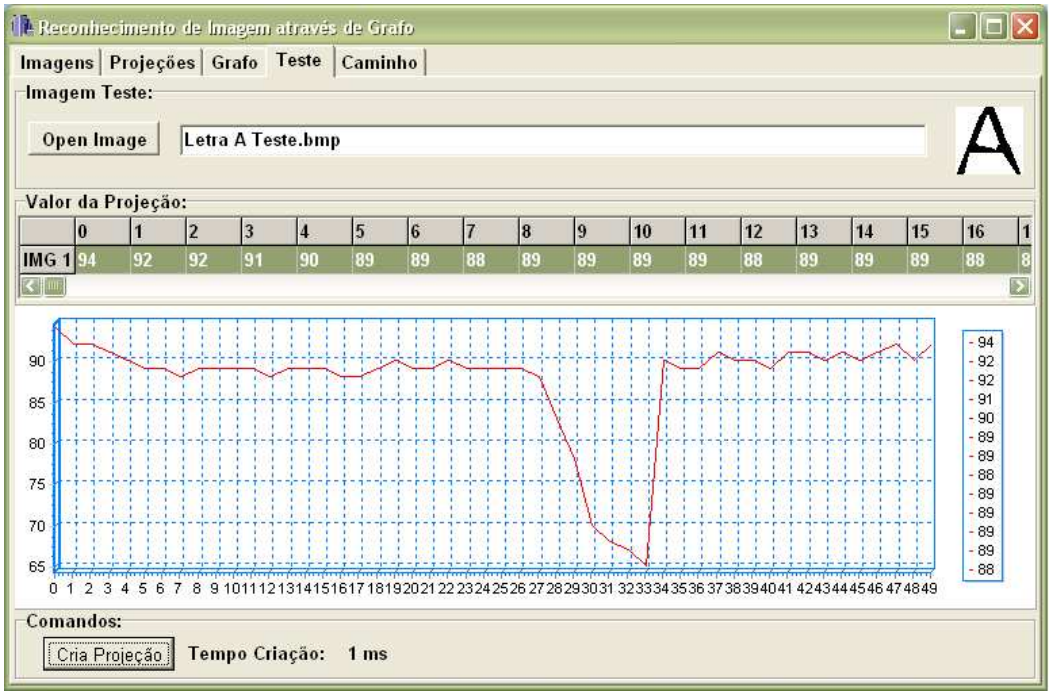


Figure 14 - Projection of the input value similar to letter A

By submitting this entry in the graph knowledge base and applying the techniques and error approach by Dijkstra's shortest path is noted that in figure 15 was recognized that input as the letter A.

Figure 15 is displayed in addition to the recognized image, also the path as well as errors adopted to determine the best way, ie with the smallest error.

ISBN: 978-1-61804-056-5

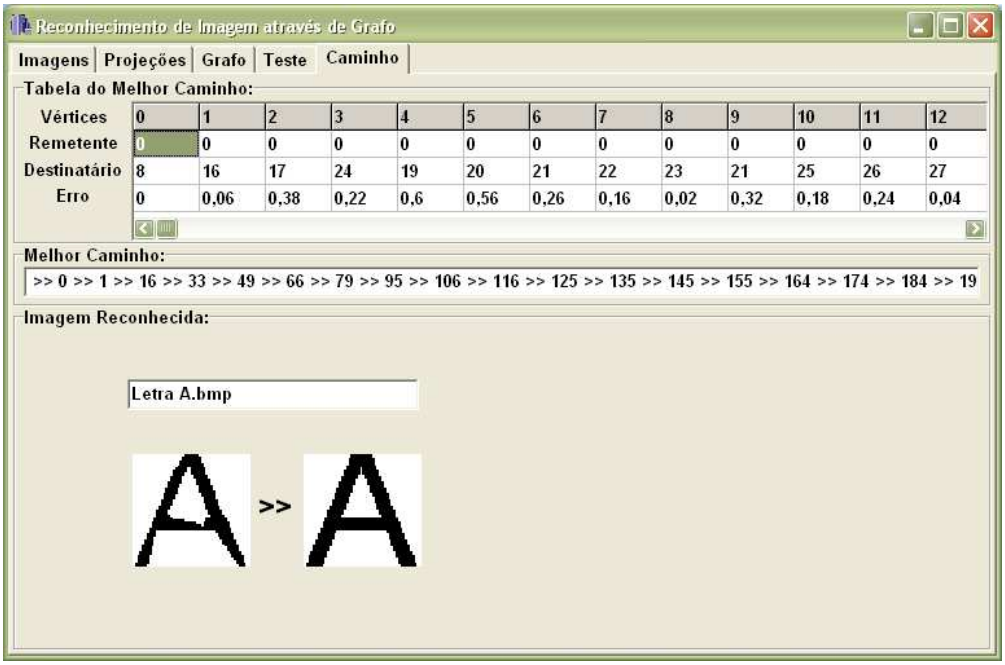


Figure 15 - The Letter of Recognition

4. Conclusions

The results of applying the system of automatic recognition of characters in a set of images that make up the basic Latin alphabet (A to Z) of varying sizes, were viable in the automatic identification of letters.

Tests were conducted to show that the system is able to classify the vowels even in the presence of small deformations and noise that may exist, if considered a real recognition of a character present in an image, respecting the limitations required for recognition.

Another technique which allowed the determination of the shortest path to adoption was the error that allowed reaching a decision which way to go, considering the difference between the values belonging to the graph and the projection of the image you want to recognize.

The algorithm showed a good result when inserted into the system was adapted from the shortest path algorithm proposed by Edsger Dijkstra, the moment that determined the way that more resembled the desired projection, computing a large number of possibilities, sometimes to be considered impractical, if used techniques ride classic rides like the depth and width.

During the tests have also been detected limitations of the system, usually applied to non-Latin characters and input characters and deformed, a possible solution to this problem would be to apply the technique of projection and also vertically in the figures using the same technique discussed here for both projections horizontal and vertical.

Also found that training time was 31ms in order, ie, well below the standards of time running on a neutral network in training do.

A future work is to develop a library for reuse of this software and use in various

ISBN: 978-1-61804-056-5

applications, such as in the field of ITS to the management of cargo containers.

References

[1] DAY, IN; Piker, JRC; PEREIRA, SL; MANTOVANI, E. *Use of Vision's Weighing Systems for Control of Railway Cars in Port Areas: An Automatic Identification Proposal.* WSEAS Transactions on Systems, Vol.4 No 1, p.74-78, 2005

[2] Gonzales, R. & WOODS, R.. Digital Image Processing. 3rd ed. United States: Prentice Hall, 2007.

[3] FEDOROVICI, L.; VOISAN, E. ; DRAGAN, F. & IERCAN, D., *Improved neural network based on OCR preprocessed blob classes.* International Joint Conference on Computational Cybernetics and Technical Informatics (CONTI-ICCC), 2010

[4] Shah, P. ; KARAMCHANDANI, S. ; NADKAR, T. ; GULECHHA, N. , Koli, K. & LAD, K., *OCR-based chassis number recognition using artificial neural networks.* IEEE International Conference on Vehicular Electronics and Safety (ICVES), 2009

[5] Werber, *PJ.* *Backpropagation through time: what it does and how to do it.* Proceedings of the IEEE. Vol 78, No 10, p. 1550-1560, 1990.

[6] Stoica, M. ; Calang, GA & Sisak, F., *Measuring the time needed for training the neural network based on the number of training steps.* 19th IEEE International Workshop on Robotics, 2010

[7] KALICHARAN, N., Data Structures in C. 1st ed. United States: CreateSpace, 2008.

[8] Tenenbaum, A. M, et Al, Data Structures Using C and C + +. 2nd ed. United States: Prentice Hall, 1995.

[9] SMITH, SW, Digital Signal Processing: A Practical Guide for Engineers and Scientists. 2nd ed. United States: Newnes, 2002.