NEURAL NETWORKS AND LETTER RECOGNITION BY HUMANS AND PIGEONS

Carlo Rondinoni

University of Sao Paulo,
Av. Bandeirantes, 3900, 14040-901 Ribeirao Preto - SP - Brazil
crondi@usp.br

ABSTRACT

We used neural networks as a tool to investigate the differences between recognition mechanisms of alphabet letters by humans and pigeons. The most similar pairs of letters from human similarity data and pigeon error data were used as stimuli to a two-layered perceptron. Two additional features were included in our study: lower case letters and biased input defined by a center-surround stimulus reading. After training the network with each letter pair, the weights of the output layer were submitted to cluster analysis. Our results show that humans take into account other factors than the shape (e.g. cursive writing, alphabetical order) when rating similarity between letters, as shown by a regular distribution of distances between clusters. Biased stimulus reading (from center to the borders) for pigeons' pairs resulted in more regularly distributed clusters, giving indications that they may be primed to round feeding stimuli. Support: FAPESP.

1. INTRODUCTION

This is an investigation about the mechanisms of letter recognition by humans and pigeons using a connectionist approach. The idea was to use perceptrons as common tool to compare data from distinct sources, like recognition processes of the same stimuli by different animal species. By introducing novel factors, as lower case letters and bias masks, we used back-propagation neural networks to gain clues about the differences of object recognition processes in the two species.

Data gathered from human subjects' similarity matrices and pigeon's error matrices can be analyzed in terms of Euclidean distances and hierarchies, feature extraction or template matching [1]. Template theories assume that objects are recognized by a point-to-point comparison process, relying on physical overlap between perceived and remembered forms. Significant feature dimensions and hierarchical structure can be detected by multidimensional scaling and clustering analysis. As computed distances seem to be related to the presence of identifiable features in letters, one can abandon distance metric and predict similarity by adding up features. Features as open top or diagonal crossing seem too arbitrary and templates seem too of a mechanistic approach, failing to explain recognition of displaced, rotated or scaled forms. We criticize those points of view by arguing that one should not forget that each letter has a multidimensional identity that pervades any judgment done by humans, that is, data has strong psychological content [2]. It is proposed that letters should be analyzed in terms of stroke number and path length during cursive writing. Other factors as letter frequencies in English language and orthographic contingencies should be taken into account when analyzing human data. For pigeons, we test whether changing stimuli input to a center-surround reading induces a more regular amalgamation clustering schedule, taking into account that these birds are biased to peck round objects [3].

We used perceptrons (Figure 1) as a tool to study visual features and letter recognition. Perceptrons belong to a family of models composed by linear threshold devices at which some connections can be modified by learning. Activations of threshold units can be thought of as visual feature detectors, when inputs are retina-like two-dimensional arrays. The convergence procedure minimizes errors until patterns are correctly classified. The only requirement is that patterns must be linearly separable [4].

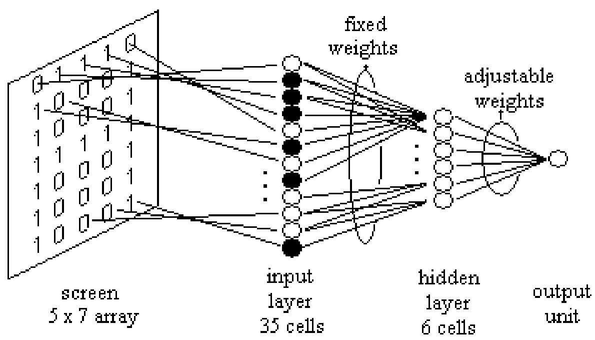

Figure 1. Perceptron architecture used in this study. Figure depicts canonical reading.

2. METHODS

Our rationale is the following: clustering analysis will place most similar pairs of letters at the end nodes of a hierarchical tree. If we stream those letters as training stimuli to a perceptron, the output weights after training will represent the frontier between features of each letter. Introducing new factors to the information pipeline may help to disentangle mechanisms and biases for different observers. We introduced lower-case letters to the set of stimuli as means of investigating factors as global properties (cursive writing path, for instance). Center-surround input reading was introduced to test whether pigeons are biased to peck at round-shaped objects. Once lower-case letters are known only to humans and round object bias happens in birds, we suggest that any regularity in clusterization of output weights must be consequence of internal data consistency. Too short or too long distances between hierarchical branches may indicate lack of stable internal structure in the data.

The perceptron was programmed in E-Basic environment (E-Prime, PST. Inc.). The basic algorithm for one training trial ran as follows:

1. randomly choose a letter among upper- and lower-case of each most similar pairs from each hierarchy;
2. read the pixels from the screen;
3. activate retina;
4. activate a 6-cells hidden layer;
5. calculate output response;
6. test if observed and desired categories match;
7. update output weights proportional to the learning rate (η = 0,57) if no match;
8. test mean square error and trial number;
9. end procedure if error is low or maximum trial number is reached;
10. go to step 1.

Each block consisted of at least 10 epochs presenting the four letters (two pairs of lower- and upper-case letters) and criterion to end procedure was mean squared error below 0,005.

Upper-case letters were the same as used in another study [5], defined by a 5x7 pixels bitmap files. Lower-case letters were intuitively created and always touched the matrix borders. At each trial, one of the four possible letters was displayed on the screen. Stimulus input was done in the presence of noise (mean 0.52). The pixels were read and fed forward into the input layer (retina) with values 1 for white and -1 for black pixels. We tested two ways of input to the 6-cells hidden layer. Always in the presence of noise and receptive field superposition, the canonical condition reads the pixels from top to bottom and from left to right (Figure 2a). Stimulus reading in the biased condition started on the central cell and then in a spiral-like path clock-wise toward the borders (Figure 2b).

Four blocks were ran, two with canonical reading and two with center-surround reading for both human and pigeon end node letter pairs. Output weights for each letter pair were clustered in STATISTICA Software (StatSoft Inc.).

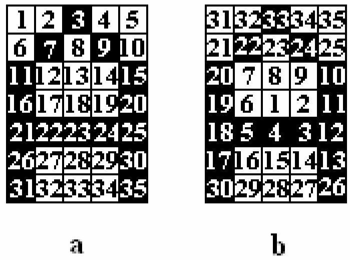

Figure 2. Input reading sequence. a. Canonical stimulus reading; b. Center-surround reading.

3. RESULTS

The table below shows the output weights found by the perceptron after training, with most similar letters from pigeon cluster trees [1,5]. Table 2 shows results for humans. Mean epoch numbers for canonical reading was 15 and for center-surround reading was 28. Figures 3 – 6 show hierarchical tress for each condition.

Ee x Ff	Ii x Tt	Xx x Yy	Nn x Ww	Uu x Vv	Dd x Oo	Cc x Gg	Bb x Pp	Aa x Rr
a								
11	**13**	**23**	**11**	**12**	**11**	**11**	**11**	**11**
-3,2	6,0	2,4	0,6	1,6	2,5	0,7	-1,8	-2,5
-1,2	-3,0	-4,4	6,9	-0,7	-2,4	0,8	-2,9	-7,0
0,0	-2,9	-5,4	8,2	-2,3	-0,7	-3,7	-1,5	-1,9
6,9	-1,1	-2,4	0,3	0,5	-1,9	-2,3	5,6	2,9
4,9	7,9	4,3	-6,1	2,9	3,1	-2,4	6,7	7,5
3,7	7,8	5,3	-7,3	4,4	1,4	2,1	5,3	2,3
b								
12	**24**	**11**	**27**	**12**	**11**	**11**	**38**	**13**
-0,3	-7,7	7,8	-6,2	3,2	7,8	-4,3	11,2	-6,0
7,2	8,8	-2,2	14	-1,9	5,5	-8,5	3,7	12
3,0	1,2	6,0	-11	-6,1	-3,5	-6,6	1,5	-2,7
6,8	-0,1	-4,7	9,6	4,3	-0,4	-5,7	-4,4	7,3
-0,7	1,1	5,3	-10	9,3	1,9	-1,4	3,2	10
3,5	-9,0	-2,9	14	14	11	-3,4	5,3	4,0

Table 1. Number of training epochs (bold fonts) and output weights found by the perceptron for pigeon similar pairs. **a.** Canonical stimulus reading; **b.** Center-surround reading.

Uu x Vv	Oo x Qq	Cc x Gg	Ee x Ff	Bb x Pp	Ii x Tt	Xx x Kk	Mm x Nn	Aa x Hh
a								

12	**11**	**11**	**11**	**11**	**13**	**16**	**54**	**11**
1,6	2,7	0,7	-3,2	-1,8	6,0	-0,1	3,6	8,7
-0,7	5,4	0,8	-1,2	-2,9	-3,0	-3,4	2,4	-1,7
-2,3	-4,2	-3,7	0,0	-1,5	-2,9	1,5	-3,5	-1,9
0,5	-0,9	-2,3	6,9	5,6	-1,1	0,8	0,3	-3,8
2,9	-3,5	-2,4	4,9	6,7	7,9	4,1	1,4	6,5
4,4	6,1	2,1	3,7	5,3	7,8	-0,8	7,3	6,7

b

16	**14**	**11**	**15**	**69**	**32**	**11**	**18**	**68**
9,7	-11	-8,2	1,1	15,0	-5,2	-0,1	-3,0	-2,3
4,1	-16	-14	11,5	2,0	-12	1,3	11,6	10,5
-3,5	-13	-13	5,8	5,1	2,0	-0,7	6,6	-4,5
2,0	-5,5	-9,3	7,4	-6,2	-0,6	-1,0	-1,5	3,3
7,6	-1,2	-3,3	-2,9	6,8	6,1	-2,4	-16	-9,5
15,3	-4,2	-4,4	2,8	3,7	-7,7	-0,3	-11	5,6

Table 2. Number of training epochs (bold fonts) and output weights found by the perceptron for human similar pairs. **a.** Canonical stimulus reading; **b.** Center-surround reading.

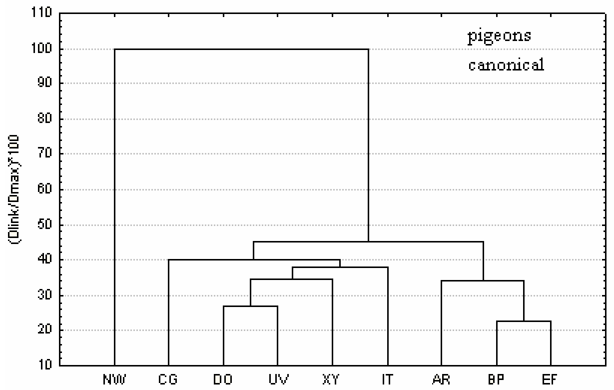

Figure 3. Hierarchy tree found from cluster analysis of output weights of Table 1a.

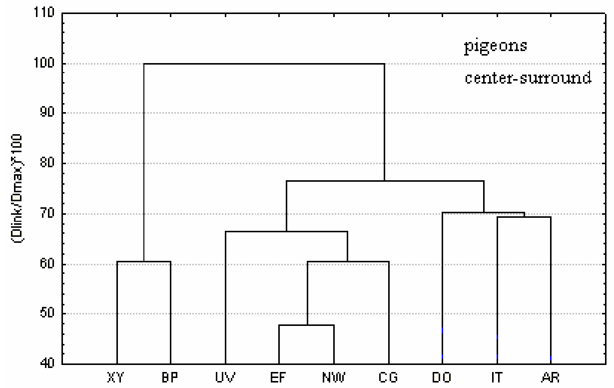

Figure 4. Hierarchy tree found from cluster analysis of output weights of Table 1b.

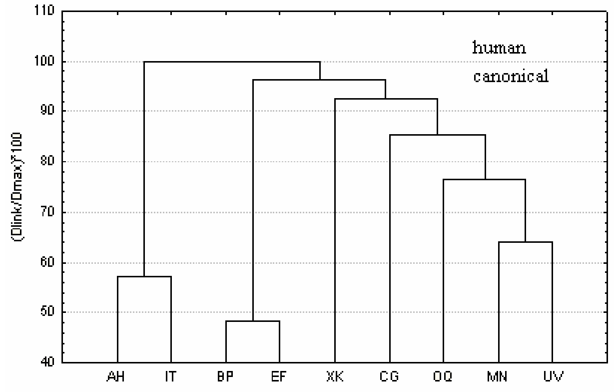

Figure 5. Hierarchy tree found from cluster analysis of output weights of Table 2a.

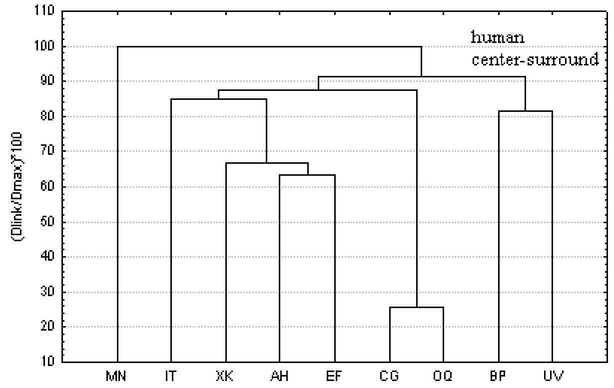

Figure 6. Hierarchy tree found from cluster analysis of output weights of Table 2b.

From tables 1 and 2, we can note that it takes more epochs to correctly classify letters with diagonal strokes (like M, N and W) or with redundant shapes between upper and lower case letters (like P and B or A and H). In most of the cases, classification converged vary fast, taking only the minimum number of ten epochs to reach square error criterion.

Visual inspection of hierarchical trees shows that canonical stimulus reading for pigeons' similar letter pairs promotes poor separation of branches (Figure 3), as indicated by mean distance between them, around 0,30. The main cluster in this hierarchy is composed by the pairs CG, DO, UV, XY, and IT and it can be labeled as the radial symmetry cluster, considering that radial symmetry includes letters which have strokes along the radius of a circle or along its perimeter. Vertical symmetry is the characteristic of the other group, composed by AR, BP, and EF, that is, letters that can be mirrored either vertically or horizontally and are still recognized. As stated above, those features defined as symmetry should be taken as a loose criterion and not a mechanistic point-to-point relationship. The pair with diagonal lines NW appears as an outer group, consequence of a longer convergence time or epochs to converge. Figure 4 shows already another tendency: center-surround stimulus reading for pigeons' letter pairs result in a more regular distribution, with relative Euclidean distances around 0,60. No specific feature can be identified as defining a cluster.

Canonical reading for human-defined pairs resulted in clusters with relative distances varying from 0,50 to 0,90, also a well balanced and distributed amalgamation schedule. Vertical symmetry seems to be the predominant feature which defines similarity. Center-surround input for humans (Figure 6), on the contrary, induces the formation of very loose branches, that is, too distant clusters. Letters with diagonal lines (M and N) are again classified as an out-group, as is the case in Figure 3.

4. DISCUSSION

In their general discussion, Podgorny and Garner [5] state that they remain uncommitted to any descriptive system for characterizing properties of alphabetic stimuli that lead to confusion matrices from letter pair. This must be the case either because the statistical tools are not enough powerful to identify them or those features were not plausibly chosen. The latter possibility must be the case. Feature-based approaches, as they are presented in the literature, consider letters as any other simple visual stimuli, relying only upon their very physical traits, like general symmetry, open loops and line orientations. The point is whether such an arbitrary conceptualization of letter formation would once match the patterns of natural similarity ratings or decision reaction times. By introducing two distinct information biases, which stem from the human cognitive domain from one side (lower case letters) and pigeon perceptive domain from the other (center-surround input), we could show that features should be chosen by taking into account not only physical properties but also ideogramatic, calligraphic, and orthographic properties of letters. As stated in another study [2], multidimensional scaling and features identification should serve as an exploratory tool rather than an end in itself.

In the case of pigeon data, it was argued that preference for round objects would be a strong biasing factor [3]. Here the author describes experiments where more than 1000 chicks were tested by presenting objects with graded angularity, from a sphere to a pyramid. Pecks on the sphere were 10 times oftener that on the pyramid. Tri-dimensional shapes as the sphere were also preferred instead of flat disks. Our results show that it can be indeed the case, once we found good regularity among clusters when stimuli were read in a center-surround manner.

5. CONCLUSION

In this study we used neural networks to investigate internal consistencies of recognition mechanisms by humans and pigeons. By introducing stimuli that were not present in the data presented to real subjects and modifying the way those stimuli were fed into the artificial pattern classifier, we found that humans are biased by other factors than physical properties of letter stimuli while pigeons show preference for round objects or, at least, rely mostly upon their global characteristics, with no feature extraction. The present results can be used in a further step of constructing Optical Character Recognition software with cognitively more plausible mechanisms.

6. REFERENCES

[1] D. S. Blough, Discrimination of Letters and Random Dot Patterns by Pigeons and Humans. *J. Exp. Psych.: Animal Behavior Processes,* vol. 11, no. 2, pp. 261-280, 1985.

[2] Keren, G. and Baggen, S. Recognition models of alphanumeric characters. *Perception & Psychophysics*, vol. 29. no. 3, pp. 234-246, 1981.

[3] Fantz, R. L. The origin of form perception. *In:* Perception: Mechanisms and Models. Readings from Scientific American, pp. 334-340, 1972.

[4] Ellis, R. and Humphreys, J.W., *Connectionist Psychology*. Psychology Press, 1999.

[5] Podgorny, P. and Garner, W.R. Reaction time as ma measure of inter- and intraobject similarity: Letters of the alphabet. *Perception & Psychophysics*, vol. 26. no. 1, pp. 37-52, 1979.