

'Mind Reading': Hitting Cognition by Using ANNs to Analyze fMRI Data in a Paradigm Exempted from Motor Responses

José Paulo Marques dos Santos^{1,2,3}, Luiz Moutinho⁴ and Miguel Castelo-Branco⁵

¹University Institute of Maia, Av. Carlos Oliveira Campos, Castelo da Maia
4475-690 Avioso S. Pedro, Portugal

²NECE - Research Unit in Business Sciences, University of Beira Interior,
Estrada do Sineiro, 6200-209 Covilhã, Portugal

³Department of Experimental Biology, Faculty of Medicine, University of Porto
Al. Prof. Hernâni Monteiro, 4200 - 319 Porto, Portugal

⁴Adam Smith Business School, University of Glasgow, West Quadrangle,
Gilbert Scott Building, Glasgow G12 8QQ, Scotland, U.K.

⁵IBILI (Institute for Biomedical Research on Light and Image), University of Coimbra
Azinhaga Santa Comba, Celas, 3000-548 Coimbra, Portugal

Abstract. The main goal of the present study is to launch the foundations of a pipeline for fMRI-based human behavior classification, addressing however some particularities of cognitive processes. While studying cognition, much of the experiments with fMRI use devices to record subjects' responses, which recruits the participation of the motor cortex. Although the influence of this aspect may be reduced in subtractive univariate analyses methods, it may negatively interfere in multivariate methods. The fMRI data here used is exempted of motor responses. Subjects were asked to form impressions about persons, objects, and brands, but their thoughts were not recorded by devices. The feed-forward backpropagation artificial neural network was used. With this procedure it was possible to correctly classify above randomness. The analysis of the hidden nodes reveals the extensive participation of the fusiform gyri and lateral occipital cortex in this cognitive process, corroborating the critical participation of these structures during classification in the natural brain.

1 Introduction

Although some classifiers have been being proposed for fMRI (functional magnetic resonance imaging) data analysis [1], and the advantages of such methods have been already addressed, especially in the study of cognitive processes [2, 3], ANNs have been missing this trend, with sporadic cross-talks [4-6]. However, ANNs' advantages are well known (e.g. modeling non-linear systems) which may be useful in the study of cognitive processes by modeling and classifying functional data. The present study's main goal is to contribute to this developing field.

Since its practical implementation in the mid 90s, fMRI has been extensively used to better understand human cognition. GLM (General Linear Model) is the mostly used method to analyze the fMRI signal. However, GLM has some limitations. Its

univariate nature is one of them. GLM procedure analyzes the signal in a voxel by voxel basis, i.e. in one voxel independently of any activity in the remaining voxels of the brain. Nonetheless, there is compelling evidence that processes in the brain unfold in interconnected networks [7], and neurons' interdependencies have to be assumed in order to fully understand cognition. One advantage of multivariate-based methods is that they consider the activity of all voxels included in the model, which may emulate brain function closely, at least better than the GLM approach. Comparisons of these methods may be found elsewhere [8, 9].

The present study explores the applicability of ANNs. ANNs were already successfully used with fMRI. Misaki and Miyauchi [5] used ANNs to model signal, although in a procedure similar to the GLM analysis. Recently ANNs were used to detect Resting State Networks (RSNs) at the individual level [10], acknowledging that RSNs probably are the most important brain networks discovered with neuroimaging techniques [11]. Therefore, ANNs may be a suitable method for fMRI signal analysis, especially in the cognitive domains.

ANNs have a potential advantage in comparison with other multivariate methods. ANNs' structure includes nodes in hidden layers which may emulate similar regularities that exist in the decision process. This facet may be explored in order to model cognitive processes, especially the complex and multi-stepped ones. Therefore, the present study uses a proven and effective backpropagation feedforward ANN with one single hidden layer, in order to favor, more the interpretation of the results (mainly the psychological interpretation), than the ANN performance.

One particularity disclosed in [12] is the probable over performance introduced by motor components when the objective of the study focus on earlier cognitive stages. Although counterbalancing may reduce motor influences in subtractive univariate methods, this may not happen in ANNs and the neural activity produced in motor responses may introduce biases that inflate successful hits. Hence, in this study ANNs are used to model a cognitive task, but the task is exempted of motor components.

The paradigm is largely inspired in the work of Mitchell, Macrae and Banaji [13]. In their study participants made impressions of persons and objects, two stimuli classes that have been extensively used in cognitive neuroscience. In the present study, a third class was added: brand logos. However, in order to explore the discriminative power of ANNs, this class is split in two: preferred brands and indifferent brands.

2 Methods

2.1 Paradigm

The structure of the present study relies on the work of Mitchell et al., [13]. It also includes the same two classes of stimuli, photographs of human faces and objects, but it adds a third new one: brands' logos. However, two subclasses of brands' logos are considered: preferred and indifferent brands. To disentangle between the two subclasses, subjects performed a preliminary session where they assessed 200 logos using the PAD (Pleasure – Arousal – Dominance) scale [14], and

the SAM (Self Assessment Manikin) [15]. Fig. summarizes all the assessments.

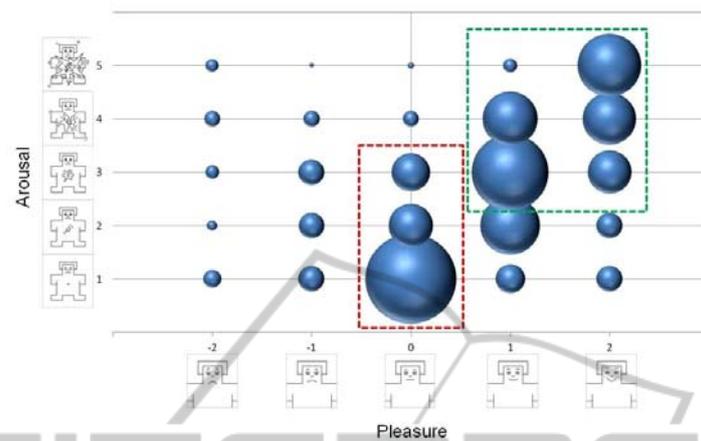


Fig. 1. Subjects' assessments in the preliminary session plotted in the Pleasure - Arousal matrix; the two rectangles define the selection criteria: the green outline define preferred brands (high pleasure, high arousal), and the red define indifferent brands (null pleasure, low arousal).

As in [13], the images with stimuli are accompanied with a caption. The caption includes some information about the subject depicted in the image (person / brand / object). Participants were instructed to covertly form an impression of the person, brand, or object taking into account the information in the caption. During the interstimuli interval participants fixated a cross.

2.2 Data Analysis and Preprocessing

Due to the difficulty in dealing with huge amounts of input data, ANNs have been used with ROIs (regions of interest) for fMRI signal analysis. ICA (Independent Component Analysis) may be used in order to previously reduce data dimensionality, which then allows whole brain analyses [12].

fMRI data pre-processing was carried out using FEAT (FMRI Expert Analysis Tool) version 5.98, and also using probabilistic independent component analysis (PICA) [16] as implemented in MELODIC (Multivariate Exploratory Linear Decomposition into Independent Components) version 3.10, both part of FSL - FMRIB's Software Library, [17].

Fifteen subjects were randomly assigned to the train group, and the remaining seven subjects were allocated to the test group.

The fMRI data of the train group entered the PICA analysis for dimension reduction, which output 173 ICs (independent components). Features were then extracted from each of the 173 time courses. The strategy adopted was to average the second and third signals after stimulus onset. By this way, the average time distance from the onset was 5000 ms, i.e. the signals considered were consistently in the neighborhood of the hemodynamic response peaks. At the end of this stage the result is a matrix with 2399 rows (each corresponding to an epoch with the corresponding

event), and 173 columns (each corresponding to one IC) plus one more column with the event code. This matrix is the training set.

The fMRI data of the test group was preprocessed in FEAT. The following pre-statistics processing was applied: motion correction using MCFLIRT [18]; slice-timing correction using Fourier-space time-series phase-shifting; non-brain removal using BET [19]; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with $\sigma=30.0s$). No spatial smoothing was applied. Registration to high-resolution structural and/or standard space images was done using FLIRT [18, 20]. All acquisitions were previously registered to a standard brain (MNI152) in order to make comparisons between subjects possible.

The 173 brain activation maps obtained with the train group were used as masks to average the individual time courses in the test group. The same procedure for feature calculation was adopted, i.e. the second and third acquisitions after stimulus onset were averaged (average time distance from the onset was 5000 ms, equal to the training set). Finally, the 1119 epochs obtained were normalized for each subject. At the end of this stage the result is a similar matrix with 1119 rows and 173 columns (each corresponding to one IC) plus one more column with the event code (that is used to assess the ANN calculations). This matrix is the training set.

In order to only include input nodes containing critical information for the classification [21], the 173 ICs were screened. For each IC a GLM was applied. The timecourse of the IC was the independent variable, and the stimuli onsets convolved with a gamma function were the explanatory variables. The parameters were estimated with least mean squares and z statistics computed. The ICs that survived the screening were those where at least one of the four z was superior to 2.3. Thus, the ICs screened out had not correlations with the stimuli. This procedure reduced the quantity of ICs to 82.

2.3 Parameters of the Artificial Neural Networks

The AMORE package [22] implemented in R [23] was used to design and perform the necessary calculations of the backpropagation feedforward ANN. Exploratory analyses yielded a global learning rate of 0.07 and a global momentum of 0.8. It was considered a hidden layer with six nodes. The selected activation function for the hidden nodes was “tansig”, while for output neurons the function was “sigmoid”.

In order to investigate possible bias derived from the network structure, the ANN was also fed with a matrix similar to the test set, but now including random values from a normal distribution. This procedure was completed for 10,000 times in order to have a large distribution.

3 Results

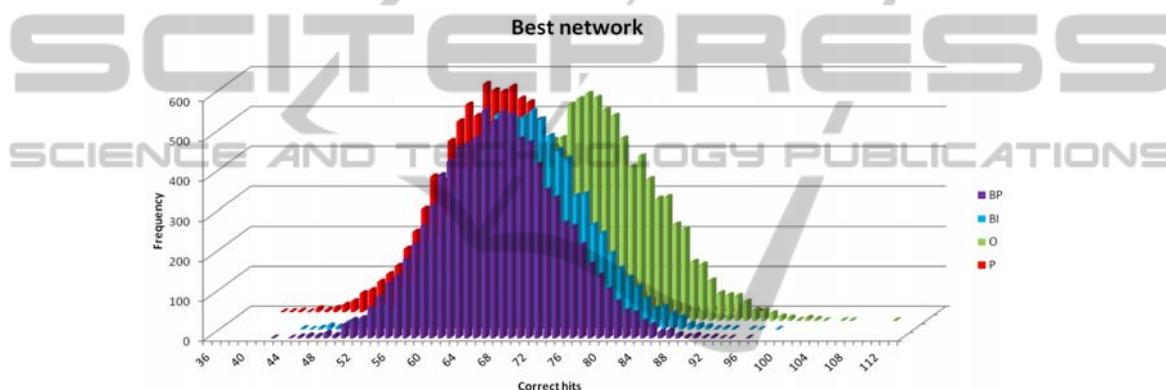
The results of the ANN with the best performance (more global correct hits) are represented in Table 1, with the respective accuracies and precisions.

Table 1. Confusion matrix with the predictions of the ANN.

Class		Predicted assessment				Total
		BP	BI	O	P	
Real assessment	BP	100	80	45	55	280
	BI	91	88	71	30	280
	O	49	61	134	36	280
	P	59	33	46	141	279
Total		299	262	296	262	1119
Accuracy		33,4%	33,6%	45,3%	53,8%	
Precision		35,7%	31,4%	47,9%	50,5%	

BP brands preferred; BI brands indifferent; O objects; P persons.

Fig. 1 depicts the results of feeding the network with random values from a normal distribution. Table 2 represents the probability values of the predictions in Table 1.

**Fig. 1.** Plot of the results of feeding the network with random values of a normal distribution (10,000 feeds).**Table 2.** Probabilities values of the predictions of the ann based on the distribution obtained after feeding the network with random normal values for 10,000 times.

Class		Predicted assessment			
		BP	BI	O	P
Real assessment	BP	0.000	0.085	1.000	0.942
	BI	0.002	0.008	0.794	1.000
	O	0.998	0.901	0.000	1.000
	P	0.914	1.000	1.000	0.000

BP brands preferred; BI brands indifferent; O objects; P persons.

Table 3 lists the weights of the axons that link hidden nodes to the output nodes. For the sake of space, the weights of the axons that link input to hidden nodes are not here fully reported. However, Fig. 2 depicts two axial slices of IC2 and IC7. IC2 has important positive weights with hidden nodes 1 (10.73) and 5 (22.33) and important negative weights with hidden node 3 (-10.60) and 4 (-39.56); also IC2 encompasses voxels in the occipital and temporal occipital fusiform gyrus, and lateral occipital

cortex, all bilaterally. IC7 has important positive weights with hidden nodes 1 (13.11), 3 (14.17), and 5 (11.72); IC7 includes voxels from the lateral occipital cortex, inferior temporal gyrus (temporooccipital part) all bilaterally.

Table 3. Weights of the axons that link hidden to output nodes. The most important positive weights have blue background, and red for the most important negative weights.

Output nodes	Hidden nodes					
	1	2	3	4	5	6
BP	1.569	0.811	1.224	0.288	-1.314	0.679
BI	-0.411	-0.680	1.154	-0.838	1.302	0.998
O	-0.184	0.983	-1.387	-1.478	1.081	-1.305
P	-1.783	-0.797	-1.151	1.714	-0.966	0.353

BP brands preferred; BI brands indifferent; O objects; P persons.

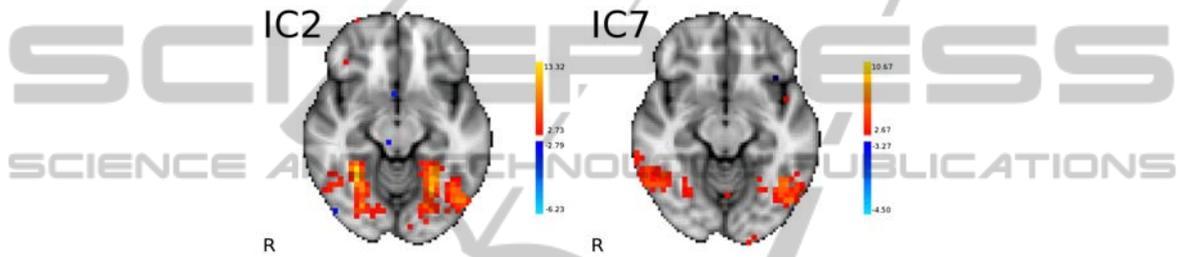


Fig. 2. Two axial slices ($z = -12$) of IC2 and IC7; MNI152 coordinates; radiological convention.

4 Discussion

Being 280 the quantity of cases of each class presented to test the network, it would be expectable that random choice is around 70 (25%) because there are four classes. In fact the peaks of the distributions in Fig. 1 fall around this value (69 for BP, 71 for BI, 76 for O, and 63 for P). In Table 2 all the p-values in the diagonal are zero or close to. The main conclusion of this study is that the basic backpropagation feedforward ANN with one hidden layer is correctly predicting much above random choice, i.e. the ANN is extracting critical information from brain data in order to correctly predict behavioral responses.

It is important to highlight the conditions of this study. Subjects never performed actions, just made mental impressions about the stimuli. Thus, the ANN is extracting neural information in the pre-motor stages, supposedly during the perception / decision periods, which are the most interesting for studies on cognition.

The analysis of the hidden nodes reveals interesting aspects. Hidden node 3 has important positive weights for preferred and indifferent brands and important negative weights for objects and persons. Hidden node 5 has important positive weights for indifferent brands and objects and negative weights for preferred brands and persons. These two nodes alone are sufficient to discriminate between the four classes.

However, considering hidden node 4, each class (brand, object, and person) is discriminated. It is possible to conclude that this node is able to successfully segregate among classes, which is also supported by the data in Table 1 and Table 2. In fact, the cells with correct hits (grey background) concentrate the majority of the assessments and have the lowest probability values.

Nonetheless, discriminating between preferred and indifferent brands it is not so good. The values of the four cells that involve preferred and indifferent brands in Table 1 are approximate, and in the case of indifferent brands, the network has more tendency to classify as preferred brands. The analysis of the same four cells in Table 2 confirms this observation. The reason for such has to be explored. The problem may be intrinsic to the stimulus, because of its low salience, or the method has to be improved in order to attain such refinement.

IC2 and IC7 (depicted in Fig. 2) are two important sources of data for successful classification. Interestingly these two brain networks encompass brain regions from visual and visual associative areas. This is in line with the findings of Hanson, Matsuka and Haxby [4], which also found in fusiform gyri sources of cognitive data for accurate classification.

References

1. Pereira, F., Mitchell, T.M., Botvinick, M.: Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45, S199-S209 (2009)
2. Mitchell, T. M., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M., Newman, S.: Learning to Decode Cognitive States from brain images. *Machine Learning* 57, 145-175 (2004)
3. Haynes, J.-D., Rees, G.: Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience* 7, 523-534 (2006)
4. Hanson, S. J., Matsuka, T., Haxby, J. V.: Combinatorial Codes in Ventral Temporal lobe for Object Recognition: Haxby (2001) revisited: is there a "face" area? *Neuroimage* 23, 156-166 (2004)
5. Misaki, M., Miyachi, S.: Application of artificial neural network to fMRI regression analysis. *Neuroimage* 29, 396-408 (2006)
6. Sona, D., Veeramachaneni, S., Olivetti, E., Avesani, P.: Inferring Cognition from fMRI Brain Images. In: Marques de Sá, J., Alexandre, L., Duch, W., Mandic, D. (eds.) *Artificial Neural Networks – ICANN 2007*, vol.4669, pp.869-878. Springer Berlin/ Heidelberg, 2007
7. Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America* 102, 9673-9678 (2005)
8. Pereira, F., Botvinick, M.: Information mapping with pattern classifiers: a comparative study. *Neuroimage* 56, 476-496 (2011)
9. Misaki, M., Kim, Y., Bandettini, P.A., Kriegeskorte, N.: Comparison of multivariate classifiers and Response Normalizations for Pattern-information fMRI. *Neuroimage* 53, 103-118 (2010)
10. Hacker, C.D., Laumann, T.O., Szrama, N.P., Baldassarre, A., Snyder, A.Z., Leuthardt, E.C., Corbetta, M.: Resting state network estimation in individual subjects. *NeuroImage* 82, 616-633 (2013)
11. Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D.A., Shulman,

- G. L.: A default mode of brain function. *Proceedings of the National Academy of Sciences* 98, 676-682 (2001)
12. Santos, J. P., Moutinho, L.: Tackling the cognitive processes that underlie brands' assessments using artificial neural networks and whole brain fMRI acquisitions. 2011 IEEE International Workshop on Pattern Recognition in NeuroImaging (PRNI), pp. 9-12. IEEE Computer Society, Seoul, Republic of Korea (2011)
 13. Mitchell, J. P., Macrae, C. N., Banaji, M. R.: Forming impressions of people versus inanimate objects: social-cognitive processing in the medial prefrontal cortex. *Neuroimage* 26, 251-257 (2005)
 14. Russell, J. A., Mehrabian, A.: Evidence for a three-factor theory of emotions. *Journal of Research in Personality* 11, 273-294 (1977)
 15. Morris, J. D.: Observations: SAM: The self-assessment manikin - an efficient cross-cultural measurement of emotional response. *Journal of Advertising Research* 35, 63-68 (1995)
 16. Beckmann, C.F., Smith, S.M.: Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging* 23, 137-152 (2004)
 17. Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M.: Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 Suppl 1, S208-S219 (2004)
 18. Jenkinson, M., Bannister, P.R., Brady, J.M., Smith, S.M.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825-841 (2002)
 19. Smith, S.M.: Fast robust automated brain extraction. *Human Brain Mapping* 17, 143-155 (2002)
 20. Jenkinson, M., Smith, S.M.: A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* 5, 143-156 (2001)
 21. Walczak, S., Cerpa, N.: Heuristic principles for the design of artificial neural networks. *Information and Software Technology* 41, 107-117 (1999)
 22. Limas, M. C., Meré, J. B. O., Marcos, A. G., Ascacibar, F. J. M. d.P., Espinoza, A. V. P., Elías, F. A.: AMORE: A MORE flexible neural network package. León (2010)
 23. R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2010)