Subject-dependent Degrees of Reliability to Solve a Face Recognition Problem Using Multiple Neural Networks

Paolo Sernani, Andrea Claudi, Gianluca Dolcini, Luca Palazzo, Gianluigi Biancucci, Aldo Franco Dragoni

Dipartimento di Ingegneria dell'Informazione (DII),
Università Politecnica delle Marche,
Via Brecce Bianche 60131 Ancona, Italy
{p.sernani, a.claudi, g.dolcini, l.palazzo, g.biancucci, a.f.dragoni}@univpm.it

***Abstract*—The interest towards biometric approach to identity verification is high, because of the need to protect everything that could have a value for some purpose. Face recognition is one of these biometric techniques, having its greater advantage in requiring a limited interaction by user. We present a Face Recognition System (FRS) based on multiple neural networks using a belief revision mechanism. Each network is associated to an a-priori reliability value for each identity stored in database, modelling the specific skill of the modules composing the system with the recognition of a given subject. Every time a network is in conflict with the global response, it is forced to retrain itself, subjecting the system to a continuous learning. The main goal of this work is to carry out some preliminary tests to evaluate accuracy and robustness of FRS with "subject-dependent" reliability values, when some changes can affect the considered features. Tests over digitally aged faces are also conducted.**

***Index Terms*—Face Recognition; Multiple Neural Networks; Bayes Rule; Belief Revision; Aging Process**

I. Introduction

Face recognition is one of the most natural biometric approaches used for identity authentication [1]. Its main advantage over other biometric techniques is that face recognition can be done passively, (i.e. without explicit actions by the user) since face images can be acquired by one or more fixed cameras far from the subject to be identified [2]. There is a high degree of interest of the research community around this field due to the wide range of possible security applications and to the availability of feasible technologies after 30 years of research [3]. An example of a face recognition application for security reasons is video surveillance [4]. Face recognition involves several fields of research: computer vision, pattern recognition, neural networks and psychology [5]. Methods for face recognition can be distinguished in three categories based on the type of features used: holistic methods, local methods and hybrid methods [3]. With holistic methods a single vector, obtained concatenating the grey values of all pixels, is used to represent the whole face image. Local methods uses local features, like geometrical measures or different regions composing the face picture. Hybrid methods use both holistic and local features.

In this paper, we propose a model for a Face Recognition System (FRS) based on a multiple neural networks architecture, following the idea of combining modules together to form a cohesive system more effective than the one composed by a single neural network [6]. Each module is domain specific, performs a subset of the overall task [7] and, compared to the whole system, has a simpler architecture, responding to given input faster than a single complex neural network. [8]. An integrating mechanism has to regulate the combination of the responses of single modules to obtain the global response [7].

The proposed system consists of five neural networks: the task of each one is the recognition of a subject, responding to a given input from a specific face region. We want the system to be resilient to changes that can affect the face to be recognized: for example changes involving beard, haircut, presence or absence of glasses. Another type of change that is considered is that imposed by the natural aging process. Age invariant face recognition is receiving increasing attention in the research community [9], [10]. The basic architecture of the system proposed in this paper is presented in [8]. The model is designed to combine individual responses of each network in critical cases in which there are incompatibilities within them. For this purpose an arbitrary a-priori reliability (i.e. the a-priori probability that the source provides a true information) is assigned to each network. On the basis of the global recognition of the whole group, the reliability values of networks can be re-evaluated producing a-posteriori reliabilities. When conflicts between the outputs of networks occur, a-priori reliabilities are re-calculated using the Bayes Rule, giving the a-posteriori reliabilities. Conflicts arise when there is no global agreement on the recognized subject. The reliabilities calculated with the Bayes Rule are used to make the final choice every time a subject has to be recognized, applying the Inclusion Based Weighted algorithm or another weighted algorithm over the maximally consistent subsets of the outputs of neural networks. These subsets are called *goods* [11]. After the recognition of a subject, the networks that caused the conflicts, i.e. those ones that did not have the chosen subject in their outputs, are forced to retrain themselves. So the system is subjected to a continuous learning. The reliabilities of the retrained networks are then setted to their original values.

With this work we introduce a reliability value that is subject-dependent, i.e. an a-priori reliability assigned to every network for each subject to be recognized. The selection algorithms are conveniently modified to use the matrix of reliability values, instead of a single vector. We tested the system not only on changes involving single features of a face (beard, haircut, glasses) but also on changes involving the face as a whole like those caused by the aging process. An aging software is used over a subset of images of the testing database, producing samples at a different aging state (i.e. images corresponding to different ages).

The rest of this paper is organized as follows. In section II the system architecture is described, with particular attention to the selection algorithms. Section III deals with the configuration of experiments, highlighting results. Section IV draws the conclusions and suggests future work.

II. System Architecture

A. Face Recognition System

The basic system architecture is the same proposed in [8]. In the present work there are five recognition networks, specialized in recognizing five face regions: hair, left eye, right eye, nose and mouth. Each region consists in a rectangular block [12] and the recognition technique is based on vectors of grey-level values representing blocks. We used LVQ 2.1 neural networks [13], a variation of Kohonen LVQ [14]. Each network is a source of information because gives one or more outputs to identify a subject. So the selection of a response for the entire system becomes a belief revision mechanism. Belief revision is the process of rearranging a knowledge base to preserve global consistency while accommodating incoming information [11]. On the basis of the conflicts between the sources of information, it is possible to establish maximally consistent subsets of sources, called *goods* and minimally inconsistent subsets of sources, called *nogoods*. In order to make a final choice on the goods and, in our FRS, to recognize a subject, an a-posteriori reliability for each network is calculated with Bayesian Conditioning. Let $S = \{s_1, ..., s_n\}$ be the set of sources, each one associated with an a-priori reliability $R(s_i)$. If the sources are independent the probability that only the sources belonging to $\Phi \in 2^s$, are reliable is:

$$R(\Phi) = \prod_{s_i \in \Phi} R(s_i) \cdot \prod_{s_i \notin \Phi} (1 - R(s_i)). \tag{1}$$

The joint reliability values for each possible Φ have to sum up 1:

$$\sum_{\Phi \in 2^S} R(\Phi) = 1. \tag{2}$$

A Φ containing some sources that give inconsistent information, must have a $R(\Phi)$ equal to zero. So we can calculate the a-posteriori reliability for each source as follows:

- Summing up into $R_{Contradictory}$ the combined reliabilities of all inconsistent sets of sources.
- Putting at zero the reliabilities of all inconsistent sets.
- Dividing the reliability of all consistent set of sources by $1 - R_{Contradictory}$ to obtain the new reliabilities $NR(\Phi)$ (and to assure that constrain (2) is satisfied).
- Calculating the a-posteriori reliability of every source $NR(s_i)$ as the sum of the combined reliabilities $NR(\Phi)$ of those sets containing s_i:

$$NR(s_i) = \sum_{\{\Phi:\, \Phi \in 2^S,\, s_i \in \Phi\}} R(\Phi). \tag{3}$$

In this approach, a subset of networks with an empty intersection of their answers is considered inconsistent. Furthermore in this work we introduce a reliability value that is subject-dependent: we assign to networks a value for each subject to be recognized, modelling the skill of every single network with the identification of a particular subject. The procedure to calculate the a-posteriori reliabilities is the same just described, but applied for all subjects in the goods: we calculate the $R_k(\Phi)$ for each subject k in the goods and then the $NR_k(s_i)$. The selection algorithms are applied using a matrix of reliabilities instead of a single vector. When a winner is selected, only the row of the identified subject is modified and included in the matrix of reliabilities for future recognitions.

Two selection algorithms are applied: the Inclusion Based Weighted (IBW) algorithm and the Weighted algorithm (WA). IBW and WA were presented in [8] and are slightly modified in this work. The Inclusion Based algorithm [12] works as follows:

1) Select all subjects in those goods containing the most reliable source.
2) If the selection returns only one subject stop: that is the identified subject.
3) If the selection contains more than one subject then pop the most reliable source from the list and go to step 1.
4) If there are no more subjects in the selection the ones that were selected at the previous iteration will be returned.

In IBW each answer in a good is associated to a weight equal to the sum of Euclidean distances between the codebooks and the input pattern:

$$P_k = \sum_{1}^{\#Nets} d_{ik} \tag{4}$$

where d_{ik} is the euclidean distance between the codebook of the i-th network associated to the k-th class and the input pattern. When two or more modules has the same reliability IBW considers first the network with the lower q_i, obtained as follows:

$$q_i = \min\{P_k : i \in good_k\}. \tag{5}$$

In this work we consider a reliability value for each subject. However IBW selects the most reliable network considering a reliability value for each network. So, in order to apply IBW, the average of all reliability values for each network is used.

Unlike IBW, WA always takes into account the ranking of the outputs of a single network. Each answer has a weight $1/n$, where $n \in [1, N]$ represents its position among the N outputs. Let M_k be the subset of sources who gives subject k as a possible output, i.e. the good containing k. Each answer has a weight:

$$W_k = \sum_{s_i \in M_k} \frac{1}{n} \cdot R_k(s_i). \tag{6}$$

The selected subject will be the one with the highest weight.

B. Example

Suppose that the FRS has to discriminate between 4 subjects: A, B, C and D. Taking into account the first two subjects identified by each network (N1, N2, N3, N4 and N5) , table I contains the answers, ordered from the most probable to the least probable one. So N1 gives “A or B”, N2 gives “A or B”, etc.

Table I
NETWORK OUTPUTS

N1	N2	N3	N4	N5
A	A	D	A	C
B	B	B	D	B

The maximally consistent subsets of the five networks with respect to these answers are {N1, N2, N4}, {N1, N2, N3, N5} and {N3, N4}. The intersection of answers of networks included in the goods are respectively A, B and D. Table II shows the reliabilities of all networks, with respect to all subjects. Table III reports the euclidean distance between the codebooks of networks and the respective class.

Table II
NETWORK RELIABILITIES

	N1	N2	N3	N4	N5
A	0.7875	0.1456	0.2832	0.4820	0.1960
B	0.4445	0.9000	0.8960	0.4726	0.4726
C	0.8717	0.8307	0.8940	0.1186	0.6957
D	0.7218	0.9000	0.8984	0.4732	0.4732
AVG	0.7064	0.6941	0.7429	0.3866	0.4594

Table III
EUCLIDEAN DISTANCE

	N1	N2	N3	N4	N5
A	1.8814	1.8824	1.9560	1.9028	1.9632
B	1.9050	1.9120	1.9234	1.9844	1.9550
C	1.9366	1.9953	1.9804	1.9799	1.9036
D	1.9734	1.9722	1.9031	1.9301	1.9933

Considering the average reliabilities of networks, as showed in the last line of table II, the order of networks to apply IBW is {N3, N1, N2, N5, N4}. So, on the first iteration IBW selects the subjects in those goods containing N3 (i.e. {N1, N2, N3, N5} and {N3, N4}). On the second iteration IBW selects the goods including N1 between the remaining ones. Only the good corresponding to subject B survives: B is the global answer of the FRS. With WA the winner is A, corresponding to the highest weight W. Table IV summarizes the weights calculated for each subject resulting from goods.

Table IV
RESULTS

Goods	Subjects	W	P
{N1,N2,N4}	{A}	0.2830	9.5858
{N1,N2,N3,N5}	{B}	0.2713	9.6798
{N3,N4}	{D}	0.2270	9.7721

Observing table III, it can be noted that with only one answer, IBW selects D as most credible winner, WA indicates A. With three outputs there is only one good (composed by all networks), so there are no conflicts and the identified subject is A. The following section highlights the recognition performances also considering different number of responses.

III. EXPERIMENTAL RESULTS

A. Methodology

We tested our system considering 5 and 10 subjects (taken from ORL database [15]). The training set is composed of 4 images per subject. For the testing phase we used 10 pictures per subject from the original database and a variable number of pictures (between 5 and 10 per subject) obtained by aging one of the original images. We considered for the training phase a maximum of 10 000 epochs, having obtained the best results on the tests over every single network between 5 000 and 10 000 epochs.

In order to find maximally consistent subsets of networks, we take into account both a fixed number of answers for each network and a fixed minimum number of subjects resulting from goods (i.e. a minimum of subjects competing to be the recognized identity). We called the first *static method* and the latter *dynamic method.* We tested both IBW and WA, using a matrix of reliabilities, as illustrated in the example in section II. In all tests a network is forced to retrain itself when its outputs do not include the winner identity: in this case, the reliabilities of the network are setted to their original values.

B. Results

Fig. 1 illustrates the rate of correct recognition using the static method. These tests highlights that, using IBW, the accuracy rate decreases when the number of answers per network increases: this is due to the reduction of the number of conflicts. There is a conflict when two or more networks do not include in their responses one or more common answers; thus a greater number of outputs causes less conflicts. In general, with only one answer a network could retrain itself more often because it is more probable it does not agree with the response of the whole system. WA avoids this problem using a ranking of answers to recognize a subject. In both the experiments, with one answer per network the rate of correct recognition is high, while with IBW the accuracy is lower with three answers per network.

Fig. 2 shows the rate of correct recognition with the dynamic method: the obtained values are comparable to those of static method and indicate a similar trend. Considering a minimum of one good is the same to consider one answer per network. In the other cases the number of answers is not fixed, but changes for each image given as input: when the minimum number of subjects to be considered is not reached, increasing the number of response introduces more subjects as candidate to be the recognized identity, but less goods. So the belief revision mechanism generates less conflicts and the networks are retrained less often.

These preliminary tests underline that belief revision mechanism, with subject-dependent reliability values, gives its better results with one answers per network. In this configuration IBW is slightly better than WA (with an average rate of correct recognition of 89.87% compared to 88.79%). However WA seems to improve its performances when the number of answers increases, indicating that the ranking of answers better balances the decreasing of conflicts. These early results are promising, at least to permit the use of FRS in a context where a small number of subjects has to be recognized.

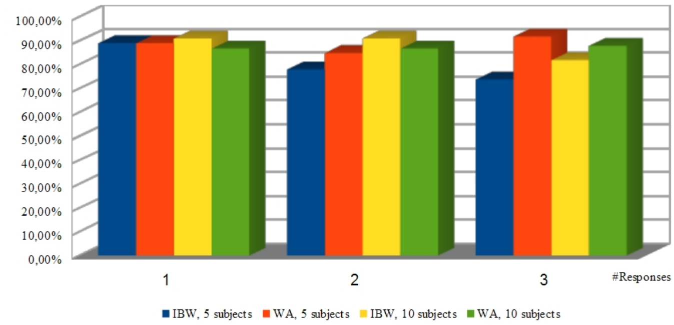

Figure 1. Rate of correct recognition with IBW and WA: static method

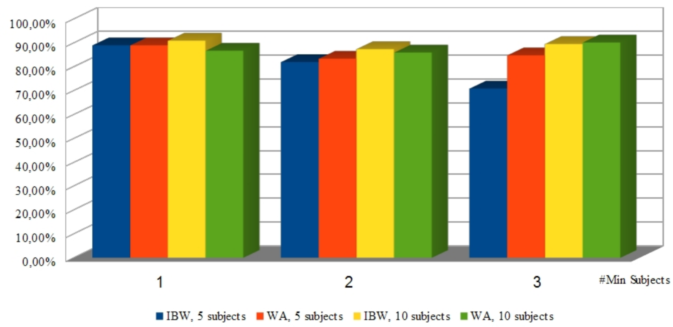

Figure 2. Rate of correct recognition with IBW and WA: dynamic method

IV. Conclusions

We described a Face Recognition System composed by Multiple Neural Networks using a Belief Revision mechanism to recognize a subject when a face image is presented as input. FRS has to be resilient to changes involving a feature and to changes involving the face as a whole like those imposed by the natural aging process. The system is composed by five LVQ 2.1, one for each face region that we considered: hair, left eye, right eye, nose and mouth. Each network has a degree of reliability for each subject to be recognized, modelling the capacity of that network with the recognition of a particular face. The reliability factors are recalculated using Bayesian Conditioning when conflicts occur between networks. So for subsequent inputs the network which recognized the wrong subject on previous inputs will be more accurate. We conducted some preliminary tests to verify accuracy and robustness of our approach. Using a database of face images from ten subject, taken at different times, varying lighting, facial expressions and details [15] and simulating an aging process, we obtained an average recognition rate near 87%. This result is promising and could support the use of FRS in those cases in which few people are to be identified.

As future work, further tests will be carried out, with larger and different databases of subjects and with more pictures in the aged version to obtain a comparison with other face recognition systems, with particular attention to changes involving features. To highlight results regarding the aging process, a specific database including images corresponding to different ages will be used. Other types of neural network could be considered and also different selection algorithms could be applied on the goods in order to recognize a subject using reliability values.

Acknowledgements

The work reported in this paper is supported, in part, by OPENMOB srl. The images used for testing are taken from ORL database, provided by AT&T Laboratories Cambridge. The authors would like to thank Aprilage Inc. for allowing the use of their aging software.

References

[1] A. Khairwa, K. Abhishek, S. Prakash, and T. Pratap, "A comprehensive study of various biometric identification techniques," in *Computing Communication Networking Technologies (ICCCNT), 2012 Third International Conference on*, pp. 1–6, 2012.

[2] R. Jafri and H. Arabnia, "A survey of face recognition techniques," *Journal of information processing systems*, vol. 5, no. 2, pp. 41–68, 2009.

[3] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face recognition a literature survey," *ACM Computing Surveys*, vol. 35(4), pp. 399–458, 2003.

[4] A. Claudi, F. Di Benedetto, G. Dolcini, L. Palazzo, and A. Dragoni, "Marvin: mobile autonomous robot for video surveillance networks," in *Computer Modeling and Simulation (EMS), 2012 Sixth UKSim/AMSS European Symposium on*, pp. 21–26, 2012.

[5] A. Tolba, A. El-Baz, and A. El-Harby, "Face recognition a literature review," *International Journal of Signal Processing*, vol. 2, pp. 88–103, 2006.

[6] M. Shields and M. Casey, "A theoretical framework for multiple neural network systems," *Neurocomputing*, vol. 71(7-9), pp. 1462–1476, 2008.

[7] Y. Li and D. Zhang, "Modular neural networks and their applications in biometrico," *Trends in Neural Computation*, vol. 35, pp. 337–365, 2007.

[8] A. Dragoni, G. Vallesi, and P. Baldassarri, "A continuous learning for solving a face recognition problem," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 4, pp. 570–577, 2012.

[9] Z. Li, U. Park, and A. Jain, "A discriminative model for age invariant face recognition," *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 3, pp. 1028–1037, 2011.

[10] U. Park, Y. Tong, and A. Jain, "Age-invariant face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 5, pp. 947–954, 2010.

[11] A. Dragoni, "Belief revision: from theory to practice," *The Knowledge Engineering Review*, vol. 12, no. 2, p. 147–179, 1997.

[12] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognition*, vol. 39, no. 9, pp. 1725 – 1745, 2006.

[13] T. Kohonen, *Self-organizing maps*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1997.

[14] T. Kohonen, J. Kangas, J. Laaksonen, and K. Torkkola, "Lvq pak: a program package for the correct application of learning vector quantization algorithms," pp. 725–730, IEEE, 1992.

[15] F. Samaria and A. Harter, "Parameterisation of a stochastic model for human face identification," in *Applications of Computer Vision, 1994, Proceedings of the Second IEEE Workshop on*, pp. 138–142, 1994.