International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 4 (2012) pp. 570-577

A Continuous Learning for Solving a Face Recognition Problem

Aldo Franco Dragoni, Germano Vallesi and Paola Baldassarri

Università Politecnica delle Marche, Italy
{a.f.dragoni, g.vallesi, p.baldassarri}@email.com

***Abstract*: We propose a Hybrid System for dynamic environments, where a "Multiple Neural Networks" system works with Bayes Rule to solve the face recognition problem. One or more neural nets may no longer be able to properly operate, due to partial changes in some of the characteristics of the individuals. For this purpose, we assume that each expert network has a reliability factor that can be dynamically re-evaluated on the ground of the global recognition operated by the overall group. Since the net's degree of reliability is defined as the probability that the net is giving the desired output, in case of conflicts between the outputs of the various nets the re-evaluation of their degrees of reliability can be simply performed on the basis of the Bayes Rule. The new vector of reliability will be used to establish who is the conflict winner, making the final choice (the name of subject), by applying two algorithms, the "Inclusion based weighted" and the "Weighted" one over all the maximally consistent subsets of the global outcome. Moreover the network disagreed with the group and specialized to recognize the changed characteristic of the subject will be retrained and then forced to correctly recognize the subject. Then the system is subjected to continuous learning.**

***Keywords*:** Hybrid System, Multiple Neural Networks, Bayes Rules, Continuous Learning**.**

I. Introduction

Several researches in the field of Artificial Neural Networks indicated that there are problems which cannot be effectively solved by a single neural network [1]. This led to the concept of Multiple Neural Networks systems for tackling complex tasks improving performances w.r.t. single network systems [2]. The idea is to decompose a large problem into a number of subproblems and then to combine the individual solutions to the subproblems into a solution to the original one [1]. This modular approach can lead to systems in which the integration of expert modules can result in solving problems which otherwise would not have been possible using a single neural network [3]. The modules are domain specific and have specialized computational architectures to recognize and respond to certain subsets of the overall task [4]. Each module is typically independent of other modules in its functioning and does not influence or become influenced by other modules. The modules generally have a simpler architecture as compared to the system as a whole, thus a module can respond to given input faster than a complex single system. The responses of the individual modules are simple and have to be combined by some integrating mechanism in order to generate the complex overall system response [4]. The combination of expert modules can be competitive, cooperative or totally decoupled among the individual expert neural networks in a given modular neural network. Generally, in a decoupled approach, individual specialist modules have no information about other modules in the network and the output of the best performing special neural net (according to some criteria) is picked to be the overall output of the modular neural network [1]. The combination of individual responses is particularly critical when there are incompatibilities between them. Such situations may arise for example when the system operates in dynamic environments, where it can happen that one or more modules of the system are no longer able to properly operate [5]. In this context it is necessary to use mechanisms to deal with sets of contradictory information.

In this work we analyze the problem of face recognition and its aim is to propose a model for detecting and solving contradictions into the global outcome. The proposed model consists of a "Multiple Neural Networks" system, where each neural network is trained to recognize a significant region of the face and to each one is assigned an arbitrary a-priori reliability (that may depend on the region of the face that must be recognized). All the networks have a reliability factor that can be dynamically re-evaluated on the ground of the global recognition operated by the overall group. In case of conflicts between the outputs of the various nets the re-evaluation of their "degrees of reliability" can be simply performed on the basis of the Bayes Rule. The conflicts depend on the fact that there may be no global agreement about the recognized subject, may be for s/he changed some features of her/his face. The new vector of reliability obtained through the Bayes Rule will be used for making the final choice, by applying the "Inclusion based" algorithm [3] or another "Weighted" algorithm over all the maximally consistent subsets of the global output of the neural networks. The nets recognized as responsible for the conflicts will be automatically forced to learn about the changes in the individuals characteristics. Networks that do not agree with this choice are required to retrain themselves automatically on the basis of the recognized subject. In this way, the system should be able to follow the changes of the faces of the subjects, while continuing to recognize them even after many years thanks to this continuous process of self training.

II. Theoretical Background

In this section we introduce some theoretical background taken from the Belief Revision (BR) field. Belief Revision occurs when a new piece of information inconsistent with the present belief set (or database) is added in order to produce a new consistent belief system [6].

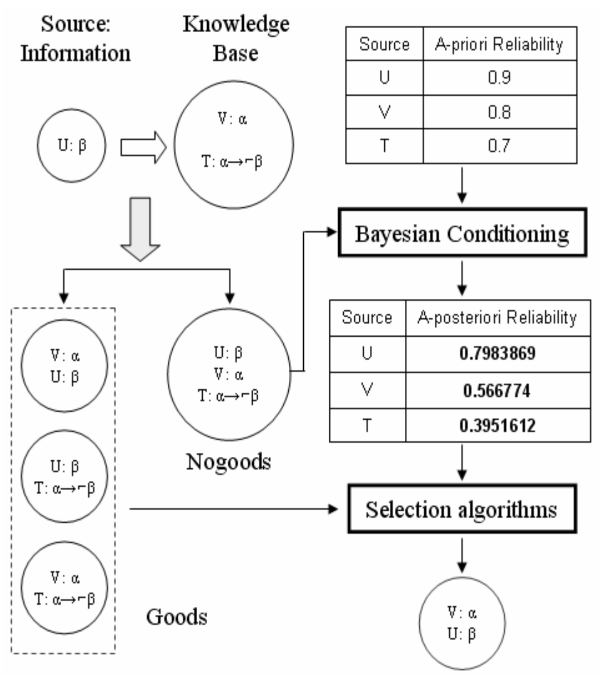


Figure 1. "Belief Revision" mechanism

In Figure 1, we see a Knowledge Base (KB) which contains two pieces of information: the information α, which come from source V, and the rule "If α, then not β" that comes from source T. Unfortunately, another piece of information β, produced by the source U, is coming, causing a conflicts in the KB. To solve the conflicts we have to found all the "maximally consistent subsets", called Goods, inside the inconsistent KB, and we choose one of them as the most believable one. In our case (Figure 1) there are three Goods: {α, β}; {β, α →¬β}; {α, α→¬β}. Maximally consistent subsets (Goods) and minimally inconsistent subsets (Nogoods) are dual notions. Given an inconsistent KB finding all the Goods and finding all the Nogoods are dual processes. Each source of information is associated with an a-priori "degree of reliability", which is intended as the a-priori probability that the source provides correct information.

In case of conflicts the "degree of reliability" of the involved sources should decrease after "Bayesian Conditioning" which is obtained as follows. Let $S = \{s_1, ..., s_n\}$ be the set of the sources, each source si is associated with an a-priori reliability $R(s_i)$. Let ϕ be an element of 2^S. If the sources are independent, the probability that only the sources belonging to the subset $\phi \subseteq S$ are reliable is:

$$R(\phi) = \prod_{s_i \in \phi} R(s_i) * \prod_{s_i \notin \phi} (1 - R(s_i)) \quad (1)$$

This combined reliability can be calculated for any ϕ providing that:

$$\sum_{\phi \in 2^S} R(\phi) = 1 \quad (2)$$

Of course, if the sources belonging to a certain ϕ give inconsistent information, then $R(\phi)$ must be zero. Having already found all the Nogoods, what we have to do is:

- Summing up into $R_{Contradictory}$ the a-priori reliability of *Nogoods*.
- Putting at zero the reliabilities of all the contradictory sets, which are the *Nogoods* and their supersets.
- Dividing the reliability of all the other (no-contradictory) set of sources by $1 - R_{Contradictory}$ we obtain the new reliability (*NR*).

The last step assures that the constrain (2) is still satisfied and it is well known as "Bayesian Conditioning". The revised reliability $NR(s_i)$ of a source si is the sum of the reliabilities of the elements of 2^S that contain si. If a source has been involved in some contradictions, then $NR(s_i) \leq R(s_i)$, otherwise $NR(s_i) = R(s_i)$.

For instance, the application of this Bayesian conditioning to the case of Figure. 1 is showed in Table 1 and 2.

A. Selection Algorithms

These new "degrees of reliability" will be used for choosing the most credible Goods as the one suggested by "the most reliable sources". There are three algorithms to perform this task:

- Inclusion based (IB) This algorithm works as follows
  1. Select all the Goods which contains information provided by the most reliable source.
  2. If the selection returns only one Good, STOP, that's the searched most credible Good.
  3. Else, if there are more than one Good then pop the most reliable source from the list and go to step 1.
  4. If there are no more Goods in the selection, the ones that were selected at the previous iteration will be returned as the most credible ones with the same degree of credibility.
- Inclusion based weighted (IBW) is a variation of Inclusion based [7]: each Good is associated with a weight derived from the sum of Euclidean distances between the neurons of the networks (i.e. the inverse of the credibility of the recognition operated by each net). If IB select more than one Good, then IBW selects as winner the Good with a lower weight.
- Weighted algorithm (WA) combines the a-posteriori reliability of each network with the order of the answers provided. Each answer has a weight $1/n$ where $n \in [1; N]$ represents its position among the N responses. Every Good is given a weight obtained by joining together the reliability of each network that supports it with the weight of the answer given by the network itself, as shown

in the following equation, where W_{Goodj} is:

$$W_{Good_j} = \sum_{i=1}^{M_j} \left(\frac{1}{n_i} * \mathrm{Rel}_i \right) \quad (3)$$

W_{Goodj}: Weight of $Good_j$.
Rel_i: Reliability of network i ($i \in Good_i$).
n_i: Position in the list of answers provided by the network i.
M_j: Number of network that compose $Good_j$.
If there are more than one Good with the same reliability then the winner is the Good with the highest weight.

III. Face Recognition System: an example

In this section we will apply the theoretical background to the problem of recognizing faces by means a "Multiple Neural Networks" system. The sources will be neural nets and the pieces of information will be the outputs. The conflict will be a simple disagreement. Face recognition is a biometric approach that employs automated methods to verify or recognize the identity of a living person based on his physiological characteristics [7]. Many methods of face recognition have been proposed during the past 30 years. Face recognition problem has attracted several fields of research: psychology, pattern recognition, neural networks, computer vision, and computer graphics [8]. These methods are broadly classified in three categories, according to the types of features used by various methods: Holistic methods, Local methods and Hybrid methods [9]. In the Holistic methods each face image is represented as a single highdimensional vector by concatenating the grey values of all the pixels in the face; Local methods use the local facial features for recognition, and finally Hybrid methods use both local and holistic features to recognize a face. We focus the attention on the Local methods that provide flexibility to recognize a face based on its parts. Local methods are classified into two main categories: local features-based method and local appearance-based method. The first method is based on the geometrical measures, while the second method divides the face image in different regions. The simplest and the most widely-used region shape is rectangular blocks [10].

ϕ	R(U)	R(V)	R(T)	R(ϕ)	NR(ϕ)
	0.1	0.2	0.3	0.006	0.0120967
T	0.1	0.2	0.7	0.014	0.0282258
V	0.1	0.8	0.3	0.024	0.048387
VT	0.1	0.8	0.7	0.056	0.1129032
U	0.9	0.2	0.3	0.054	0.1088709
UT	0.9	0.2	0.7	0.126	0.2540322
UV	0.9	0.8	0.3	0.216	0.4354838
UVT	0.9	0.8	0.7	0.504	0
				$\sum_{\phi \in 2^S} R(\phi) = 1$	$\sum_{\phi \in 2^S} NR(\phi) = 1$

Table 1. Conflict table.

In this work we also consider a recognition technique based on the use of whole image grey-level templates. So each person is represented by a database of a set of four rectangular masks representing eyes, nose, mouth and hair [11]. In the simplest version of multiple template matching, each template of the face to recognize is compared using a suitable metric (typically the Euclidean distance) with the corresponding template of each image belonging to the database.

ϕ	NR(ϕ)	NR(U $\in$ S)	NR(V $\in$ S)	NR(T $\in$ S)
	0.0120967	0	0	0
T	0.0282258	0	0	0.0282258
V	0.048387	0	0.048387	0
VT	0.1129032	0	0.1129032	0.1129032
U	0.1088709	0.1088709	0	0
UT	0.2540322	0.2540322	0	0.2540322
UV	0.4354838	0.4354838	0.4354838	0
UVT	0	0	0	0
		NR(U)=0.7983869	NR(V)=0.566774	NR(T)=0.3951612

Table 2. The revised reliability.

Instead, in the present work a number of independent recognition modules, such as neural networks, are specialized to respond to individual template of the face. In order to solve the problem to recognize the face even if partially changes occurred it is necessary to introduce a system in which expert modules can be adapted to the new situation. Unlike the Euclidean distance neural networks are better able to upgrade themselves in presence of changes in the input pattern. We propose a modular system consisting of four neural networks, for example four Self Organizing Maps of Kohonen [12], in a way that each network is specialized to perform a specific task: eyes recognition (E network), nose recognition (N network), mouth recognition (M network) and, finally, hair recognition (H network). Considering a simple theoretical example, we suppose that during the testing phase, the system has to recognize the face of four persons: Andrea, Franco, Lucia and Paolo. So, we suppose that each network has the following possible codified outputs: A output of each network is the subject Andrea, F output of each network is the subject Franco, L output of each network is the subject Lucia and finally, P output of each network is the subject Paolo. According to the value of the weights of each trained network, each net will provide in output a list of names of subjects, ordered from the most probable to the least one, by considering the nearest one in terms of Euclidean distance. For the purpose of this example, we considered to take into account only the first two outputs as threshold (i.e. let's limit the uncertainty to the first two most probable names). Let's suppose that, after the testing phase, the outputs of the networks are as follows: E gives as output "A or F", N gives "A or P", M gives "L or P" and, finally H gives "L or A", how showed in Table 3.

E	N	M	H
A	A	L	L
F	P	P	A

Table 3. Networks Outputs.

So, the 4 networks do not agree in the choice of the subject since there is no individuals in the intersection of the four outputs (intersection is void). Now the problem is to establish the most credible individual corresponding to the contradictory outputs. To solve this problem we adopt the method described in Section 2. First of all we need to give an a-priori reliability factor of degree of each network. Then we have to find Goods and Nogoods. Considering the Table 3 we can detect three Goods, that are the largest subsets of

{E,N,M,H} which agree in the choice of at least one subject; these Goods are: {E,N,H} corresponding to Andrea, {N,M} corresponding to Paolo and, finally {M,H} corresponding to Lucia. Besides, we identify two Nogoods, that are the smallest subsets of {E,N,M,H} which have no subject in common; these Nogoods are: {N,M,H} and {E,M}.

Now we have to choose the most credible Good, i.e. the one "provided by the most reliable networks". However the reliability of the networks are changed due to the fact they felt in conflict. Starting from an undifferentiated a-priori reliability factor of 0.9, and applying the method described in the previous Section we get the following new degrees of reliability for each network, how showed in Table 4.

Neural Nets	A-priori R(ϕ)	A-posteriori NR(ϕ)
E	0.9	0.7684
N	0.9	0.8375
M	0.9	0.1459
H	0.9	0.8375

Table 4. A-priori and A-posteriori Reliability.

The networks N and H have the (same) highest reliability, and by applying a selection algorithm (Section 2.1) it turns out that the most credible Goods is {E,N,H}, which corresponds to Andrea. So Andrea is the response of the system.

Figure 2 shows a schematic representation of this Face Recognition System (FRS). Which is able to recognize the most probable individual even in presence of serious conflicts among the outputs of the various nets.

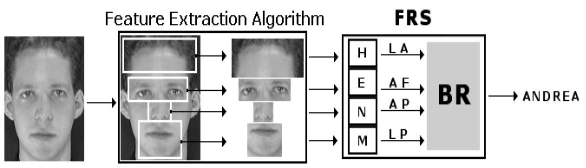


Figure 2. Schematic representation of Face Recognition system (FRS)

IV. Face Recognition System in a dynamical environment

As seen in the previous Section, one or more networks may fail to recognize the subject, there can be two reasons for the fault of the net: either the task of recognizing is objectively harder, or the subject could have recently changed something in the appearance of his face (perhaps because of the grown of a goatee or moustaches).

The second case (showed in Figure 3) is very interesting because it shows how our FRS could be useful for implementing Multiple Neural Networks able to follow dynamic changes in the features of the subjects. In a such dynamic environment, where the input pattern partially changes, some neural networks could no longer be able to recognize the input. However, if the changes are minimal, we guess that most of the networks will still correctly recognize the face. So, we force the faulting network to re-train itself on the basis of the recognition made by the overall group. This is an evolutionary system. On the basis of the a-posteriori reliability and of the Goods, our idea is to automatically re-train the networks that did not agree with the others. The network that do not support the most credible Good are forced to re-train themselves in order to "correctly" (according to the opinion of the group) recognize the changed face. Each iteration of the cycle applies Bayesian conditioning to the a-priori "degrees of reliability" producing an a-posteriori vector of reliability. To take into account the history of the responses that came from each network, we maintain an "average vectors of reliability" produced at each recognition, always starting from the a-priori degrees of reliability. This average vector will be given as input to the two algorithms, IBW and WA, instead of the a-posteriori vector of reliability produced in the current recognition. In other words, the difference with respect to the BR mechanism described in Paragraph 2 is that we do not give an a-posteriori vector of reliability to the two algorithms (IBW and WA), but the average vector of reliability calculated since the FRS started to work with that set of subjects to recognize.

Figure 3. Partial canges of the face

Figure 3 shows the behaviour of the system when the testing image partially changes. Now the subject has moustaches and goatee, while, when the system is trained, the subject did not have them. So O_M network (specialized to recognize the mouth) is no longer able to correctly indicate the tested subject. Since all the others still recognize Andrea, O_M will be retrained with the mouth of Andrea as new input pattern, as showed in Figure 4.

In order to consider only the single stable changes we used an approach based on the concept of temporal window. For this purpose, we consider each changing for each feature and simultaneously we analyze if the same feature maintains its change. The re-learning procedure is only in the case of the changing is longer than the previously fixed temporal window (windowlength). So we avoid the re-learning for a subject with a very variable feature.

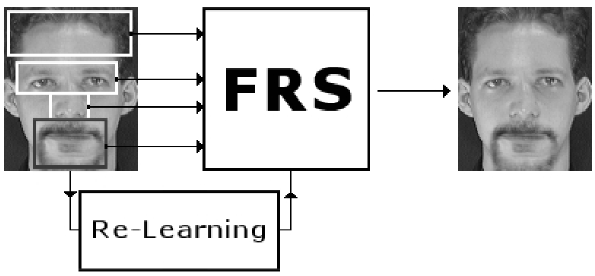


Figure 4. Re-learning to the System when the input is partilly changed

We define imm_i the portion of the image containing the feature analyzed by the network r_i; S the subject identified by the synthesis function of the FRS; s_{ik} is the subject i in the k-th position of the list ordered on the base of the distance of the LVQ output.

So the re-learning procedure consists of the following steps:

1. For each network r_i the system compares S and s_{ik} used to find the Good. If $S \neq s_{ik}$ $\forall$ k then in the

temporary directory temp(S_i) of the subject S related to the network i is saved the imm$_i$ portion, as showed in Figure 5. On the contrary if S = s_{ik} for one k the temporary directory temp(S_i) is empitied.

2. If in temp(S_i) there are windowlength samples, the temp(S_i) images are transferred in riadd(S_i) removing its old images, then the retraining of the r_i network begins using the riadd(S_i) images for S and the most recent images for all other subjects.

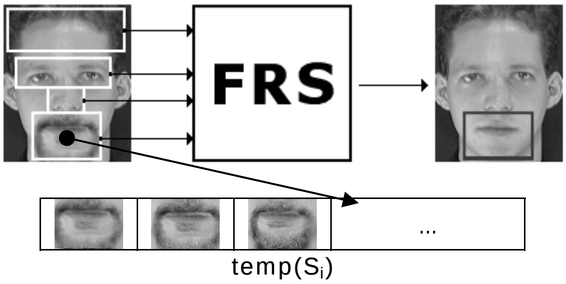


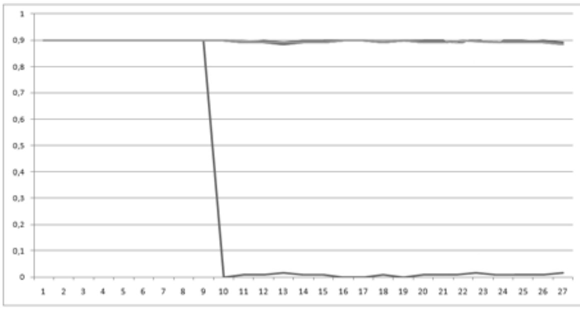
Figure 5. Creatino of the temporal window

We have to highlight that the windowlength chosen strongly depends on the variability of subjects, and so on the database used for the testing. It is important also to note that there will always be a limit to the size of the windowlength beyond which for any dataset the system will be able to filter all the changes, to a value beyond this limit the system behaves as a system without re-learning.

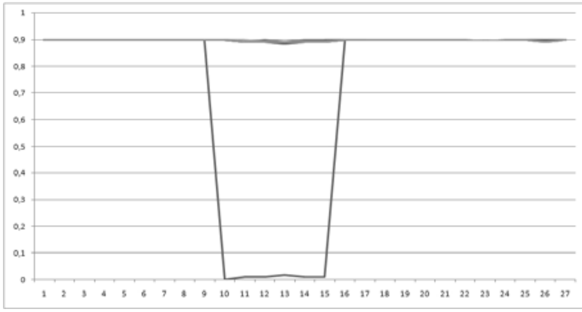
If not recognized by the networks, the introduction of re-learning in the facial recognition system allows, that the networks maintain higher reliability values than in the case without re-learning, as shown in Figures 6.a and 6.b. This because, the network now can recognizes a feature that could not recognize with the original knowledge acquired during the first training of all networks. If a network is no longer able to recognize one feature can not contribute to the final choice. Moreover if other networks do not recognize the subject, but indicate the same wrong subject, the whole system fails. In this case there would be a wrong Good that could be the most likely for the system but associated to the incorrect subject.

Figure 6 shows the a-posteriori reliability trend related to the five expert neural networks concerning a particular subject. Observing the two graphs in Figure 6 we can see that until the networks are agree, the reliability maintains high values. While, if one of the networks (eg mouth) comes into conflict with the others giving in output another subject, since it changed its appearance, then the reliability goes down.

In Figure 6.a, we can see how this conflict will bring the net loser to have a low reliability, which can not be recovered even when the network agrees with the others. Conversely, in Figure 6.b, we can see that if the network does not recognize the subject for a consecutive number of times corresponding to the windowlength samples the re-learning begins, after which the a-posteriori reliability back to its highest level. Thus allowing the neural network to return to fully functioning.



(a)



(b)

Figure 6. Performance of the a-priori reliability without and with Re-learning

V. Experimental results

In this work we compared the results of two groups of neural networks: the first consisting of four networks and the second of five networks (the additional network is obtained by separating the eyes in two separate networks). This in order to see if increasing the number of networks we obtain significant changes in the final results. All the networks are LVQ 2.1 [13], a variation of Kohonen's LVQ [14], each one specialized to respond to individual template of the face. For each data point (x,y) from the training set S = $\{(x_i, y_i)\}_{i=1}^{N}$, LVQ 2.1 first selects the two nearest prototypes θ_l, θ_m according to the Euclidean distance. If the labels c_l and c_m are different and if one of them is equal to the label y of the data point, then the two nearest prototypes are adjusted according to:

$$\begin{aligned} \theta_l(t+1) &= \theta_l(t) + \alpha(t)(x-\theta_l), \qquad c_l = y \\ \theta_m(t+1) &= \theta_m(t) - \alpha(t)(x-\theta_m), \qquad c_m \neq y \end{aligned} \quad (4)$$

If the labels c_l and c_m are equal or both labels differ from the label y of the data point, no parameter update is done. The prototypes, however, are changed only if the data point x is close to the classification boundary, i.e. if it falls into a window:

$$\min\left(\frac{d(x,\theta_m)}{d(x,\theta_l)}, \frac{d(x,\theta_l)}{d(x,\theta_m)}\right) > s \quad \text{where } s = \frac{1-\omega}{1+\omega} \qquad (5)$$

of relative width $0 < \omega \leq 1$. This "window rule" had to be introduced, because otherwise prototype vectors may diverge. Learning rate used is shown in the following equation:

$$\alpha(t) = \eta e^{(-\beta t)} \quad (6)$$

where $\alpha(t)$ decreases monotonically with the number of iterations t (η=0.25 and β=0.001, values obtained after a series of tests to optimize networks).

The Training Set is composed of 20 subjects (taken from FERET database [15]), for each one 4 pictures were taken for a total of 80. During the learning phase, networks were trained using three different epochs: 3000, 4000 and 5000.

We use two methods to find Goods and Nogoods from the networks responses:

1. Static method: the number of responses provided by the networks to be used for the construction of Goods and Nogoods is fixed a priori. In our case we have chosen values from 1 to 5. With the value 1 are taken only the first responses provided by each network, while the value of 5 were considered the top 5 answers for each network.
2. Dynamic method: the number of responses provided by the networks to be used for the construction of Goods and Nogoods changes dynamically according to a threshold. This threshold is the minimum number of Goods to be reached by the networks. We start with one answer for each network and increase by 1 until we get a number of Goods equal to the threshold. In our case we have chosen values from 1 to 5.

In the next step we applied the Bayesian conditioning [16], [17], depending from Goods obtained with these two techniques. In this way we obtain the new reliability for each network. These new "degrees of reliability" will be used for choosing the most credible Good (then the name of subject). We use two selection algorithms (showed in subsection 2.1) to perform this task:

1. Inclusion based weighted (IBW).
2. Weighted algorithm (WA).

To test our work, we have taken 488 different images of the 20 subjects and with these images we have created the Test Set. Figure 7 shows the correct recognition rate for this Test Set, obtained by considering the hybrid system consisting of four or five neural networks, and applying the selection algorithms IBW and WA on Goods identified by the static method. Figure 8 shows however, the rate of correct recognition obtained for the same Test Set using the same set of selection algorithms and neural networks, but with Goods obtained through the dynamic method. Figure 7 and 8 shows also, how WA selection algorithm is better than IBW for all four cases. The best solution for WA is achieved with five neural networks and 5000 epochs in both the methods (Static and Dynamic).

The results of Figure 9 show how the union of the Dynamic method with the selection algorithm WA and five neural networks gives the best solution to reach a 79.39% correct recognition rate of the subjects. Figure 9 also shows as using only one LVQ network for the entire face, we obtain the worst result. In other words, if we consider a single neural network to recognize the face, rather one for the nose, one for the mouth and so on, we have the lowest rate of recognition equals to 66%. This is because a single change in one part of the face makes the whole image not recognizable to a single network, unlike the hybrid system.

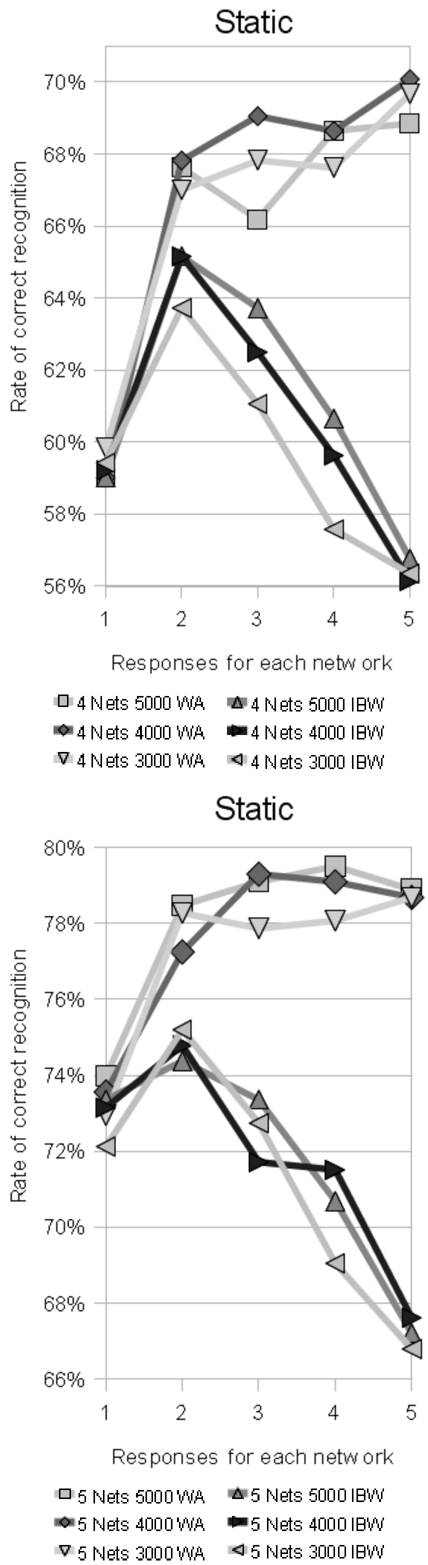


Figure 7. Rate of correct recognition with Static method

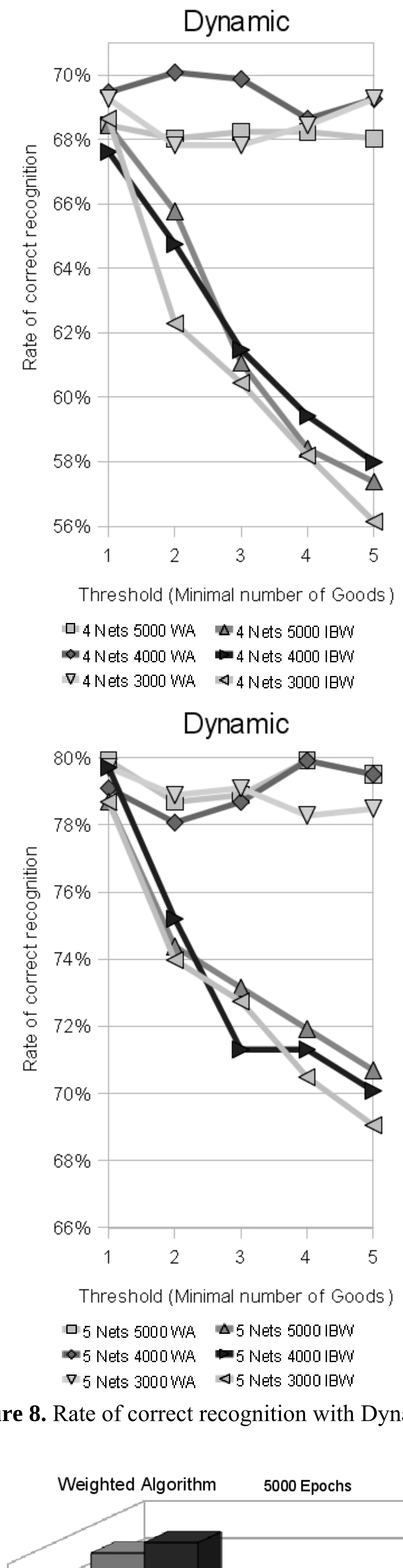
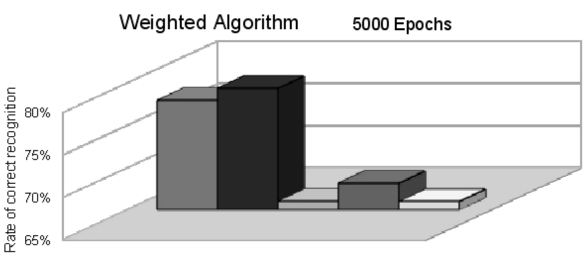


Figure 8. Rate of correct recognition with Dynamic method



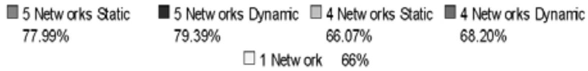


Figure 9. Average rate of correct recognition without Re-learning

In Figure 10, we can see the comparison between the average rate of correct recognition of the following cases:

- Hybrid system with retraining (Dynamic method, WA selection algorithm), 89,25 %
- Hybrid system without retraining (Dynamic method, WA selection algorithm), 79,39 %
- Only one neural network for the entire face, 66 %

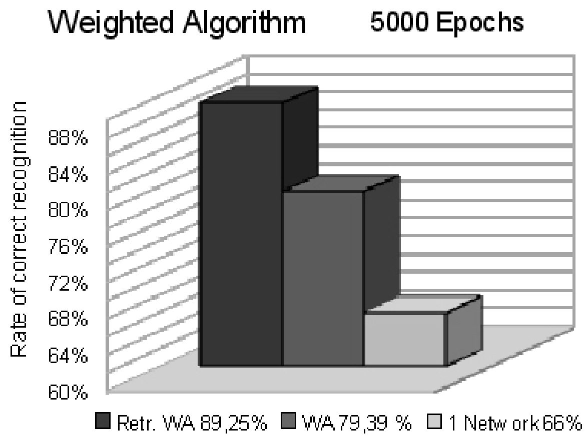


Figure 10. Average rate of correct recognition with Re-learning

VI. Conclusion

Our hybrid method integrates multiple neural networks with a symbolic approach to Belief Revision to deal with pattern recognition problems that:

- Require the cooperation of multiple neural networks specialized on different topics.
- The individuals to recognize change dynamically some of their features so that some nets occasionally fail.

We tested this hybrid method referring to a face recognition problem, training each network to recognize a specific region of the face: eyes, nose, mouth, and hair. Every output unit is associated with one of the persons to be recognized. Each net gives the same number of outputs. We consider a constrained environment in which the image of the face is always frontal, lighting conditions, scaling and rotation of the face being the same. We accommodated the test so that changes of the faces are partial, for example the mouth and hair do not change simultaneously, but one at a time. Under this assumption of limited changes, our hybrid system ensures great robustness to the recognition. The system assigns a reliability factor to each neural network, which is recalculated on the basis of conflicts that occur in the choice of the subject. The new "degrees of reliability" are obtained through the conflicts table and Bayesian Conditioning. These new "degrees of reliability" can be used to select the most likely subject. When the subject partially changes its appearance, the network responsible for the recognition of the modified region comes into conflict with other networks and its degree of reliability will suffer a sharp decrease. The networks that do not agree with the choice made by the overall group will be forced to

re-train themselves on the basis of the global output. So, the overall system is engaged in a never ending loop of testing and re-training that makes it able to cope with dynamic partial changes in the features of the subjects. To maintain high values of the reliability for all the networks is very important since the choice of the right subject strongly depends on the credibility of all the experts.
In future works we want to test the system differentiating the a-priori reliability of the experts depending on the subject and the portion of the face to identify.

References

[1] Azam, F., "Biologically inspired modular neural networks", *PhD Dissertation*, Virginia Tech., 2000
[2] Shields, M.W., and Casey, M.C., "A theoretical framework for multiple neural network systems", *Neurocomputing*, vol. 71(7-9), pp. 1462-1476, 2008
[3] Sharkey, A.J., "Modularity, combining and artificial neural nets", *Connection Science*, vol. 9(1), pp. 3-10, 1997
[4] Li, Y., and Zhang, D., "Modular Neural Networks and Their Applications in Biometrico", *Trends in Neural Computation*, vol. 35, pp. 337-365, 2007
[5] Guo, H., Shi, W., and Deng, Y., "Evaluating sensor reliability in classification problems based on evidence theory", *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, vol. 36(5), pp. 970-981, 2006
[6] Gärdenfors, P., "Belief Revision", *Cambridge Tracts in Theoretical Computer Science*, vol. 29, 2003
[7] Brewka, G., "Preferred subtheories: an extended logical framework for default reasoning", Proc. 11th Inter. *Joint Conf. on Artificial Intelligence*, pp. 1043-1048, 1989
[8] Tolba, A.S., El-Baz, A.H., and El-Harby, A.A., "Face Recognition a Litterature Review", *International Journal of Signal Processing*, vol. 2, pp. 88-103, 2006
[9] Zhao, W., Chellappa, R., Phillips, P.J., and Rosenfeld, A., "Face Recognition: a Litterature Survey", *ACM Computing Surveys*, vol. 35(4), pp. 399-458, 2003
[10] Tan, X., Chen, S., Zhou, Z.H., and Zhang, F., "Face Recognition from a Single Image per Person: a Survey", *Pattern Recognition*, 39:1725-1745, 2006
[11] Tan, X., Chen, S., Zhou, Z.H., and Zhang, F., "Recognizing Partially Ocluded, Expression Variant Faces from Single Training Image per Person with SOM and Soft kNN Ensemble", *IEEE Transactions on Neuraal Networks*, 16(4):875-886, 2005
[12] Brunelli, R., and Poggio, T., "Face recognition: features versus template", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15(10), pp. 1042-1052, 1993
[13] Kohonen, T., *Self-Organizing maps*, Springer-Verlag, Berlin, 1997
[14] Kohonen, T., Kangas, J., Laaksonen, J., and Torkkola, K., "LVQ pak: A program package for the correct application of learning vector quantization algorithms", *IEEE*, pp. 725-730, 1992
[15] Kohonen, T., "Learning vector quantization", in *Self-Organizing Maps*, Springer Series in Information Sciences, vol. 30, Berlin, 2001
[16] Philips, P.J., Wechsler, H., Huang, J., and Rauss, P., "The FERET Database and Evalutation Procedure for Face-Recognition Algorithms", *Image and Vision Computing J.*, vol. 16, no. 5, pp. 295-306, 1998
[17] Dragoni, A.F., "Belief revision: from theory to practice", *The Knowledge Engineering Review*, vol. 12, no. 2, pp. 147-179, 1997
[18] Dragoni, A.F., Animali, S., "Maximal Consistency and Theory of Evidence in the Police Inquiry Domain", in *Cybernetics and Systems: An International Journal*, vol. 34, no. 6-7, pp. 419-465, Taylor & Francis, 2003

Author Biographies

Aldo Franco Dragoni was born in Ascoli Piceno (Italy), the 22nd of june 1961. He received a degree in Electronics Engineering from the University of Ancona, discussing a thesis in the field of Artificial Intelligence on "Plan Recognition from Visual Information". Currently he is in charge as Associate Professor at the "Università Politecnica delle Marche", where he teaches "Foundamentals of Computer Science", "Artificial Intelligence" and "Real Time Systems". He served as Program Committee and Reviewer for several International Conferences and Journals on Artificial Intelligence.

Germano Vallesi Ph.D. He received a degree in Electronics Engineering from the University of Ancona (Italy) in 2005 and the Ph.D degree in "Electronics Engineering, Informatic and Telecommuncations" from the "Università Politecnica delle Marche" Ancona (Italy) in 2008. Him research interests include the study of neural networks, biometric systems, hybrid architectures, robotics, vision and image processing by neural networks.

Paola Baldassarri Ph.D. She received a degree in Electronics Engineering from the University of Ancona (Italy) in 1999 and the Ph.D degree in "Artificial Intelligent Systems" from the University of Ancona (Italy) in 2005. Her research interests include the study of hybrid architectures for reactive robotic localization vision based, unsupervised learning theory for analysis of electromyography signals, vision and image processing by neural networks.