

A Receptive Field Based Approach For Face Detection

Bruno J. T. Fernandes, George D. C. Cavalcanti and Tsang I. Ren

Abstract—This paper presents a new neural network method used to perform visual pattern classification. The neural network is called *I-PyraNet* which is a hybrid implementation of PyraNet and the concepts of inhibitory fields. In order to improve the results obtained by this neural network, it is also presented the 2-D Gabor filter. Furthermore, both the neural network and the Gabor filter, are applied in a face detection task and the results are compared to another face detection method based on SVM.

I. INTRODUCTION

The eye is one of the most important organs of the human body. This organ is responsible for the detection of light and for the transformation of this perception in electrical impulses. Moreover, it is from this perception that humans are able to distinguish, to recognize and to understand patterns. The main function of the eye is to perform fast acquisition of information.

Studies in image processing [1] enables researches to reproduce in a computer some behaviors of the human visual system. In this context, neural network plays an important role in the simulation of the human brain. The increasing knowledge about the visual cortex in the brain allowed the creation of models that replicate some of its functionalities. A simple application like face detection which is a very easy task performed by human, on the other hand it is a problem that is not solved yet in image processing.

The work of Hubel [2] presents an important concept of the human visual system. This concept indicated the presence of neurons connected to specific regions of the brain, called receptive fields. Such concept was applied in different techniques, motivating research that uses this concept along with other techniques such as: 2-D Gabor filter [3], [4] and the PyraNet [5], a neural network that presents an excellent ability of recognition.

In this work, we propose a new neural network, called *I-PyraNet*. The *I-PyraNet* is a hybrid implementation that combines the PyraNet and the concepts of inhibitory fields [6]. This neural network is applied in a face detection task. In order to improve the detection ratio, a 2-D Gabor filter is also applied over the images before the classification.

This paper is organized as follows. In Section II the concepts of the receptive and inhibitory fields are presented. In Section III the Gabor filter is presented. In Section IV the *I-PyraNet* classifier is presented. In Section V some

experiments involving the face detection task are shown. Finally, in Section VI some conclusions are presented.

II. RECEPTIVE AND INHIBITORY FIELDS

At the beginning of the 60's, the presence of neurons in the human visual cortex acting as edge detectors in images [2] was found. These neurons send a strong response when they are in a presence of a line in a given orientation and position in the visual field [7].

Later, other kinds of functionalities were revealed in the actuation of these neurons, where the only necessary condition to activate a neuron would be the presence of an appropriate stimulus in a specific region of the visual field. This region is known as the classical receptive field (CRF).

A receptive field is defined in [8] as an area where the presence of a stimulus might lead to a response of a particular neuron. In few words, the receptive field of neuron is a region of affectation that it covers. This kind of neuron has been identified in many parts of the human brain, like the auditory, somatosensory and visual systems.

However, Rizzolati and Camarda [6] showed that simultaneously to the receptive field stimulus, other stimulus outside the receptive field can also affect the neuron. Most of the times, it is an inhibitory stimulus and it is referred as the non-classical receptive field (non-CRF) inhibitory. This kind of stimulus was biologically motivated by the visual cortex of some macaque monkeys and its influence has been shown in the human being as well. The inhibitory field seems to be a common property in biological edge detectors, and it has been applied with successful in the contour detection [9]. The inhibitory field presence can also be seen in the work done by Blakemore and Tobin [10] in which the inhibitory stimulus is measured over the presence of a white bar.

III. GABOR FILTER

The Gabor filter was originally introduced as a multiplication of a cosine/sine wave by a gaussian window [11]. Recently, many works have presented the use of a Gabor function to accomplish the task of texture analysis, due to its optimal localization properties [12].

In another work, Daugman [3] proposed a two-dimensional Gabor filter, developed to model the receptive fields of some mammals. The 2-D complex Gabor filter over the image domain (x, y) is defined as

$$G(x, y) = \exp\left(-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}\right) \times \exp(-2\pi i(u_0(x-x_0) + v_0(y-y_0))), \quad (1)$$

where (x_0, y_0) is the location in the image, with (u_0, v_0) specifying the modulation that has the spatial frequency $\omega_0 =$

Bruno J. T. Fernandes is with the Informatics Center at the Federal University of Pernambuco and with the Academic Unit of Serra Talhada at the Rural Federal University of Pernambuco, Brazil. George D. C. Cavalcanti and Tsang I. Ren are with the Informatics Center at the Federal University of Pernambuco, Brazil. (email: {bjtf, gdcc, tir}@cin.ufpe.br, site: <http://www.cin.ufpe.br/~viisar>).

$\sqrt{u_0^2 + v_0^2}$ and orientation $\theta_0 = \arctan(v_0/u_0)$, and σ_x e σ_y are the gaussian standard deviation along the axis x e y , respectively [12].

Derived from this equation, the real 2-D Gabor filter component can be calculated, which is used in [13], [14]:

$$g(x, y; T, \theta) = \exp\left(-\frac{1}{2} \left[\frac{x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2} \right]\right) \cos\left(\frac{2\pi x_\theta}{T}\right), \quad (2)$$

$$x_\theta = x \cos \theta + y \sin \theta, \quad (3)$$

$$y_\theta = -x \sin \theta + y \cos \theta, \quad (4)$$

where θ is the orientation of the Gabor filter and T is the period of the sinusoidal wave.

In Figure 1, taken from [9], it is presented the intensity map of 2-D Gabor filter. The lighter and the darker grayscales, in comparison with the image background, indicate the regions where the Gabor function assume positive and negative values, respectively. The bright ellipse specifies the receptive field boundary. Outside this boundary the Gabor function assumes extremely small values.

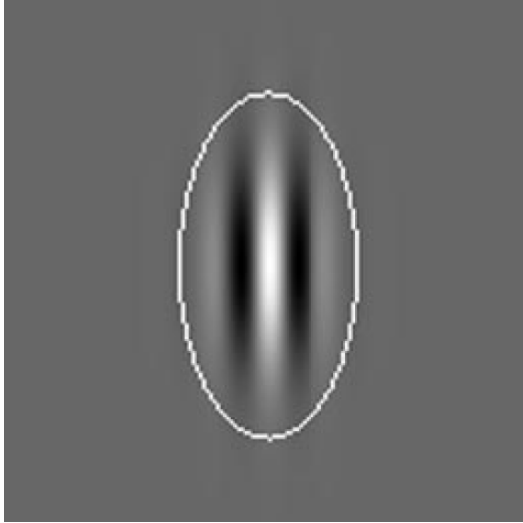


Fig. 1. Intensity map of a 2-D Gabor filter taken from [9]

The Gabor representation of an image is done by the convolution of the image with the Gabor filter [15]. Suppose that $f(x, y)$ is the image intensity in graylevel at the coordinate (x, y) . The image convolution with the Gabor filter $g(x, y; T, \theta)$ is given by

$$\psi(x, y; T, \theta) = f(x, y) \star g(x, y; T, \theta), \quad (5)$$

where \star is the convolution operator.

In the literature, many different parameters in the 2-D Gabor filter configuration have been presented. In the work done by Bhuiyan and Liu [16], the following frequencies, 0.16, 1.0 and 1.4, are used in a 2-D Gabor filter with one of the following orientations, $0, \pi/5, 2\pi/5, 3\pi/5, 4\pi/5$. However, there is no consensus about a heuristic to estimate the Gabor filter parameters.

IV. I-PYRANET

The I-PyraNet [17], [18] is an artificial neural network developed to perform the classification of visual patterns (images). The I-PyraNet is based on the PyraNet [5], which was motivated by the good results obtained by the convolutional neural networks [19] and by the concepts of receptive fields. However, the PyraNet consider only the excitatory effects of the neurons inside a receptive field.

Then, it is proposed here the I-PyraNet, where the neuron might be able to send excitatory or inhibitory stimulus to the neurons in the posteriors layers. The idea behind the use of the inhibitory stimulus is based on the concepts of inhibitory fields that surround a given receptive field. The presence of a neuron in a inhibitory field will affect its output by inverting its signal.

A. I-PyraNet Architecture

The I-PyraNet architecture is very similar to the PyraNet architecture. It is composed by a multilayer network with two kinds of layers:

- 2-D layers: perform the feature extraction and the data reduction and are located at the base of the network. In a 2-D layer, the neurons are arranged in a matrix;
- 1-D layers: perform the image classification and are located at the top of the network.

Like the PyraNet, each 2-D layer inside the I-PyraNet has a parameter that defines the size of its receptive field, given by r and other one that defines the size of the overlap between adjacent receptive fields, given by o . The relation between the receptive field and the overlap is given by $g = r - o$. However, the I-PyraNet has a new parameter: the number of horizontal or vertical inhibitory neurons that surround any given receptive field in a 2-D layer, denoted by h . The neurons in the space defined by h will send negative response when inside of an inhibitory field. In Figure 2, the I-PyraNet architecture is presented.

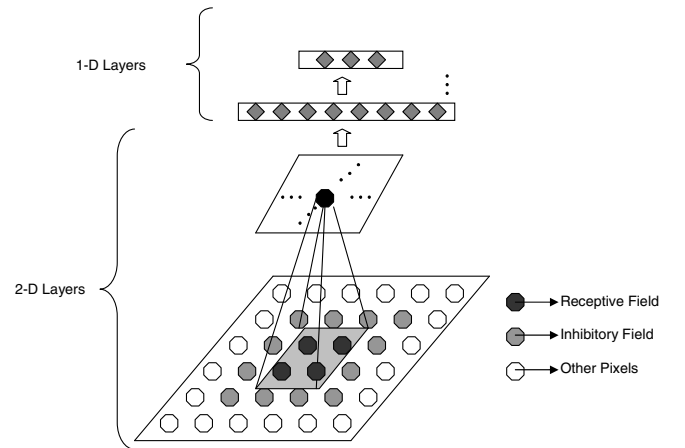


Fig. 2. Arquitetura da I-PyraNet

The entire network is connected in cascade (i.e., the output of one layer works as the input to the next one). The input

for the first 2-D layer is the image to be classified. And the input for the first 1-D layer is the output of the last 2-D layer rearranged in a vector. Each neuron in a 2-D layer is connected to a receptive field surrounded by an inhibitory field over the last layer. Neurons in a same layer might share neurons inside their receptive or inhibitory field. It is important to note that in a the 2-D layer the weights are associated to the neuron itself, instead of being associated to the connection between the neurons. Then, it shares not only the output of the neuron but also the weight associated to it.

The output of a 2-D neuron consists in the application of a non-linear activation function applied over the weighted summation of the neurons inside its receptive field subtracted by the weighted summation of the neurons inside its inhibitory field. Then, being (u, v) the position of a neuron in a 2-D layer l , (i, j) the position of the neuron in the previous layer $(l - 1)$ and $b_{u,v}$ the bias of the neuron at (u, v) , the output y of the neuron is given by

$$y_{u,v} = f \left(\underbrace{\sum_{i,j \in R_{u,v}} w_{i,j} \times y_{i,j}}_{\text{Receptive Field}} - \underbrace{\sum_{i,j \in I_{u,v}} w_{i,j} \times y_{i,j}}_{\text{Inhibitory Field}} + \underbrace{b_{u,v}}_{\text{Bias}} \right) \quad (6)$$

where $w_{i,j}$ represents the weight associated to the input position (i, j) to the 2-D layer l , $R_{u,v}$ is the receptive field of the neuron at (u, v) and $I_{u,v}$ is the inhibitory field of the neuron (u, v) .

The neurons at the 1-D layers of the I-PyraNet work like the ones of the PyraNet. The output of a neuron in a 1-D layer is given by the application of a non-linear activation function over the weighted summation of the neurons at the previous layer. In the 1-D layer the weights are associated to the connection between the neurons. Then, the output y of a neuron at the position n of the 1-D layer l is given by

$$y_n^l = f \left(\sum_{m=1}^{N_{l-1}} w_{m,n} \times y_m^{l-1} + b_n^l \right), \quad (7)$$

where N_{l-1} is the number of neurons at the previous layer $l - 1$, $w_{m,n}$ is the synaptic weight of the neuron m at the layer $l - 1$ to the neuron n at the layer l , and b_n^l is the bias associated to neuron n at the 1-D layer l . The output of the last 1-D layer works as the network output.

B. I-PyraNet Training

The I-PyraNet must first be trained to be able to perform visual pattern recognition tasks. As a supervised neural network its objective is to reduce the error obtained through the output desired and the output obtained, and it is made adjusting the synaptic weights in the I-PyraNet. In this work, the approach used to perform this task was the cross-entropy function (CE) [20], where the network outputs estimate the *a posteriori* probability for each known class. The advantages of the use of the cross-entropy function for training neural networks were presented in [21].

Then, being y_n^L the output of the neuron n in the last network layer L for an input image k , the estimated *a posteriori* probability p_n is given by

$$p_n^k = \exp(y_n^{L,k}) / \sum_{i=1}^{N_L} \exp(y_i^{L,k}), \quad (8)$$

where N_L is the number of neurons in the layer L . Therefore, in order to adjust the synaptic weights in the I-PyraNet, the error gradient of the weights must be calculated through the error sensitivity for each neuron.

The error sensitivity δ for each neuron n at the 1-D output network layer L_{1D} , for an input image k is given by

$$\delta_n^{L_{1D},k} = e_n^k f' (s_n^{L_{1D},k}), \quad (9)$$

where e_n^k is the output y_n^k produced by the neuron n at the last 1-D layer L_{1D} minus the desired output d_n^k , then $e_n^k = y_n^k - d_n^k$. And $s_n^{L_{1D},k}$ is the weighted sum input to neuron n at layer L_{1D} and f' is the differential of the activation function f . Then, for the neurons in the others 1-D layers $l_{1D} < L_{1D}$ the error sensitivity is given by

$$\delta_n^{l_{1D},k} = f' (s_n^{l_{1D},k}) \times \sum_{m=1}^{N_{l_{1D}+1}} \delta_m^{l_{1D}+1,k} \times w_{n,m}, \quad (10)$$

where $N_{l_{1D}+1}$ represents the number of neurons in the next layer $l_{1D} + 1$, $w_{n,m}$ is the synaptic weight from the neuron n in the layer l_{1D} to the neuron m at layer $l_{1D} + 1$ and $\delta_m^{l_{1D}+1,k}$ is the error sensitivity of the neuron m at layer $l_{1D} + 1$.

The error sensitivities for the last 2-D layer are calculated using the previous equation but rearranged into a 2-D grid. In the others 2-D layers l_{2D} , the error sensitivity for each neuron at the position (u, v) is given by

$$\delta_{u,v}^{l_{2D},k} = f' (s_{u,v}^{l_{2D},k}) \times w_{u,v} \times \sum_{i=i_l^{max}}^{i_h^{max}} \sum_{j=j_l^{max}}^{j_h^{max}} \gamma_{i,j}^{l_{2D}+1,k}, \quad (11)$$

where $s_{u,v}^{l_{2D},k}$ is the weighted sum input for the neuron (u, v) , $w_{u,v}$ is the weight associate to the neuron (u, v) at layer l_{2D} and $\gamma_{i,j}^{l_{2D}+1,k}$ is given by

$$\gamma_{i,j}^{l_{2D}+1,k} = \begin{cases} \delta_{i,j}^{l_{2D}+1,k} & i_l \leq i \leq i_h, j_l \leq j \leq j_h \\ -\delta_{i,j}^{l_{2D}+1,k} & \text{otherwise} \end{cases}, \quad (12)$$

being $\delta_{i,j}^{l_{2D}+1,k}$ the error sensitivity for the neuron (i, j) at the next layer, and i_l^{max} , i_h^{max} , j_l^{max} , j_h^{max} , i_l , i_h , j_l and j_h were calculated by

$$i_l^{max} = \left\lceil \frac{u - (r_{l+1} + h_{l+1})}{g_{l+1} - h_{l+1}} \right\rceil + 1, i_l = \left\lfloor \frac{u - r_{l+1}}{g_{l+1}} \right\rfloor + 1 \quad (13)$$

$$i_h^{max} = \left\lfloor \frac{u - 1}{g_{l+1} - h_{l+1}} \right\rfloor + 1, i_h = \left\lfloor \frac{u - 1}{g_{l+1}} \right\rfloor + 1 \quad (14)$$

$$j_l^{max} = \left\lceil \frac{v - (r_{l+1} + h_{l+1})}{g_{l+1} - h_{l+1}} \right\rceil + 1, j_l = \left\lfloor \frac{v - r_{l+1}}{g_{l+1}} \right\rfloor + 1 \quad (15)$$

$$j_h^{max} = \left\lfloor \frac{v-1}{g_{l+1} - h_{l+1}} \right\rfloor + 1, j_h = \left\lfloor \frac{v-1}{g_{l+1}} \right\rfloor + 1 \quad (16)$$

Then, the error gradient to the weights and the biases can be derived through the next equations.

- 1-D Weights: the error gradients for the 1-D synaptic weight $w_{m,n}$ from the neuron m at layer $l_{1D} - 1$ to the neuron n at layer l_{1D} for all the images input K , are given by

$$\frac{\partial E}{\partial w_{m,n}} = \sum_{k=1}^K \delta_n^k y_m^{l_{1D}-1,k}. \quad (17)$$

- 2-D Weights: the 2-D synaptic weight $w_{u,v}$ of neuron (u, v) at layer l_{2D} to layer $l_{2D} + 1$ is calculated by

$$\frac{\partial E}{\partial w_{u,v}} = \sum_{k=1}^K \left\{ y_{u,v}^{l_{2D},k} \times \sum_{i=i_l^{max}}^{i_h^{max}} \sum_{j=j_l^{max}}^{j_h^{max}} \gamma_{u,v}^{l_{2D}+1,k} \right\}, \quad (18)$$

- Biases: the error gradients for the bias of neuron n , b_n , at 1-D layer l_{1D} and of neuron u, v , $b_{u,v}$, at 2-D layer l_{2D} are respectively given by

$$\frac{\partial E}{\partial b_n} = \sum_{k=1}^K \delta_n^k, \frac{\partial E}{\partial b_{u,v}} = \sum_{k=1}^K \delta_{u,v}^k \quad (19)$$

The weights were then recalculated through the use of the Gradient Descent method [22] which completes the training phase of the I-PyraNet.

C. I-PyraNet Remarks

It is important to note that the PyraNet can be considered as a special case of the I-PyraNet, where the size of the inhibitory fields are equals to 0.

A neuron inside a 2-D layer in the PyraNet will produce the same input to the neurons in the next layer that contain it inside their receptive field. However, in the I-PyraNet the output of a neuron can be converted in two kinds of inputs to the neurons in the next layer. One kind is a positive input, where the output of the neuron remains the same. The other kind is a negative input, where the output of the neuron is multiplied by -1 .

Finally, it can be said that the main advantage of the use of the inhibitory field concepts is that a same neuron is able to produce two different outputs, depending on its spatial position.

V. EXPERIMENTAL RESULTS

The experiments done in this work were applied over the task of face detection. Face detection consists in the problem of indicating the locations of a face in a given image. This problem might be summarized by the dichotomy task of determine if a given pattern is or is not a face.

In other work, Makinen and Raisamo [23] present an evaluation of the different methods for face alignment in order to improve the face detection rates. They had demonstrated that the manual alignment was the only kind of alignment to bring a gain to the detection rate, while the other analyzed methods

did not make any improvement. In all of the evaluated cases, the SVM (Support Vector Machine) [24] obtained the best classification rate. Osuna et al. [25] also showed the advantages of using a SVM approach to perform a face detection task.

We present here the results obtained with the I-PyraNet and the SVM using the Gabor filter in the problem of face detection. The tests were performed using the MIT CBCL Face Database #1 [26] under the domain of the MIT Center For Biological and Computation Learning¹, which contains 2429 face patterns and 4548 non-face patterns for training, and 472 face patterns and 23573 non-face patterns for testing. Examples of the images from this database are shown in Figure 3.

The experimental results are presented in a ROC graph [27]. The ROC curve presents the classification rates of the true positive type for different errors of the false positive type. The true positive rates represent the number of faces correctly classified divided by the total of faces in the database. While the false positive rates represent the number of non-face patterns wrongly classified divided by the total of non-face patterns in the database. It is considered that the result of a ROC curve is as good as the area under the curve. All the tests were realized over a Pentium Dual Core with a 1.73GHz processor and 2-GB RAM.

A. Tests Realized with a Training Subset

The following experiments were realized with 1000 images of each class for training. The I-PyraNet tested in this work has two 2-D layers and one 1-D layer with two neurons, each one giving the probability of an input image belongs or not to a face pattern. Then, in order to find the best receptive field configuration of the I-PyraNet, ROC curves were generated for different sizes of receptive fields with the overlap factor equals to 1 in both 2-D layers and without the presence of inhibitory fields, as shown Figure 4. The receptive fields of size 4 and 3 for the first and the second 2-D layer, respectively, shows the best results. This configuration is used in the following experiments using the I-PyraNet.

Based on the best results obtained varying the receptive fields, the best overlap factors were calculated, Figure 5. It can be seen that some configurations of overlap factors presented the same area under the curve. However, the configuration with the overlap factor equals to 1 for both 2-D layers was used in the following experiments for two reasons. The first one is that this configuration makes the 2-D layers smaller than the other ones, leading to faster results. The second one is that this configuration presented a better true positive rate in the beginning of the ROC curve.

The experiments involving the Gabor filter and the I-PyraNet were also generated. In this task, the images were filtered with the 2-D Gabor filter and then classified with the I-PyraNet. Results obtained in [28] indicated that the cells in the visual cortex of the monkeys had an orientation about 65 degrees. Then, the Gabor filter is tested with $\theta = \pi/3$. In

¹ <http://www.ai.mit.edu/projects/cbcl>



Fig. 3. Image examples of the MIT CBCL Face Database. (a) Training face patterns, (b) testing face patterns, (c) training non-face patterns and (d) testing face patterns.

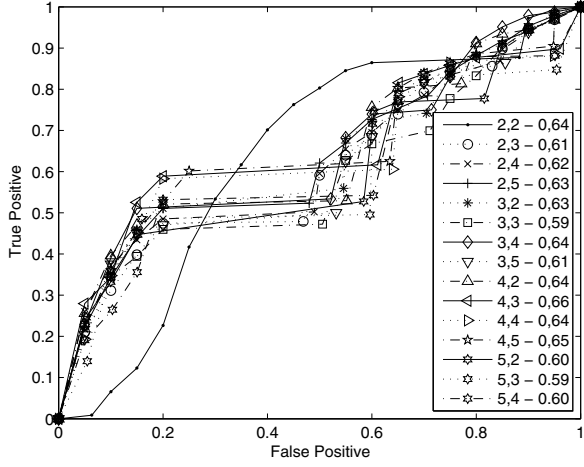


Fig. 4. ROC curves for different receptive fields configuration of the I-PyraNet. The legend presents the size of the receptive field for first and the second 2-D layer, respectively, followed by the area under the curve.

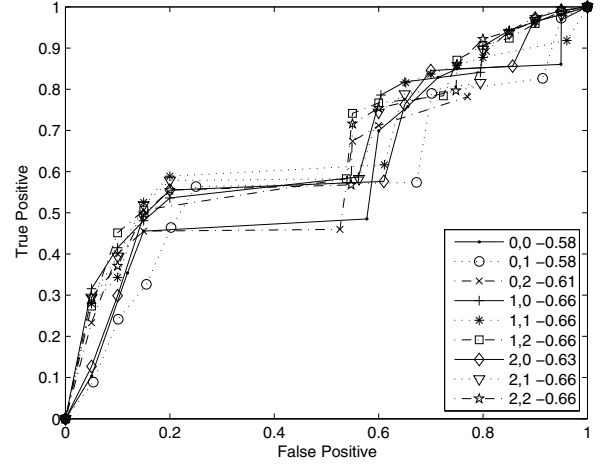


Fig. 5. ROC curves for different overlap factor of the I-PyraNet. The legend presents the size of the overlap for first and the second 2-D layer, respectively, followed by the area under the curve.

Figure 6, the results for different frequencies, denoted by f ($f = 1/T$), of the Gabor filter using $\theta = \pi/3$, $\sigma_x = 4$ and $\sigma_y = 4$ are presented. It is easy to see that $f = 8$ achieved the best results, with an area of 0.71.

Then, in Figure 7, the ROC curves for different values of σ_x and σ_y for the Gabor filter are presented. The application of the Gabor filter using $\sigma_x = 4$ and $\sigma_y = 8$ achieved the best results, with an area of 0.8.

It is important to see that the images in the database were collected under different environment and illumination conditions. Then, in Figure 8, the results between the I-PyraNet over the not pre-processed images and over the images with the histogram equalized [1] are compared. It is easy to see that the application of the histogram equalization brought gains to the results.

B. Tests Realized with the Complete Training Set

The following experiments were realized with all the images of the training set. The I-PyraNet used in the following experiments has a receptive field of size 4 and 3 for the first and the second 2-D layer, respectively, with an overlap factor of 1 for both layers. And the Gabor filter tested has the following parameters: $\theta = \pi/3$, $f = 8$, $\sigma_x = 4$ and $\sigma_y = 4$.

First, it was tested the best configurations for the inhibitory fields of the I-PyraNet using the Gabor filter over images with the histogram equalized, Figure 9. The use of the inhibitory field of size 0 in the first 2-D layer and 1 or 2 in the second 2-D layer achieved the best results, with an area under the ROC curve of 0.83.

On the other hand, it was also generated the results using the SVM classifier. The SVM tested uses a 3rd degree polynomial as its kernel function [26]. In Figure 10, it can be seen that the use of the SVM with the Gabor filter

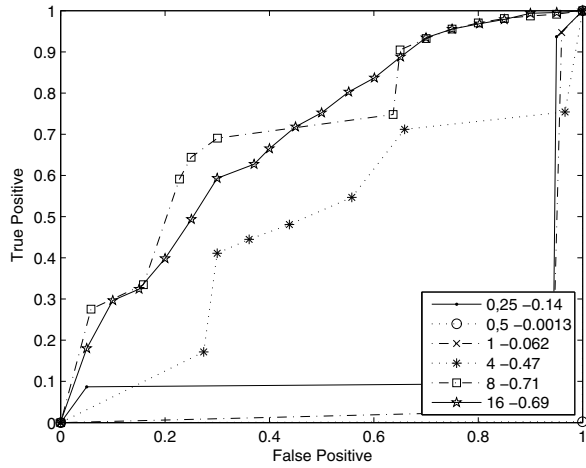


Fig. 6. ROC curves for different frequencies of the Gabor filter. The legend presents the value of the frequency followed by the area under the curve.

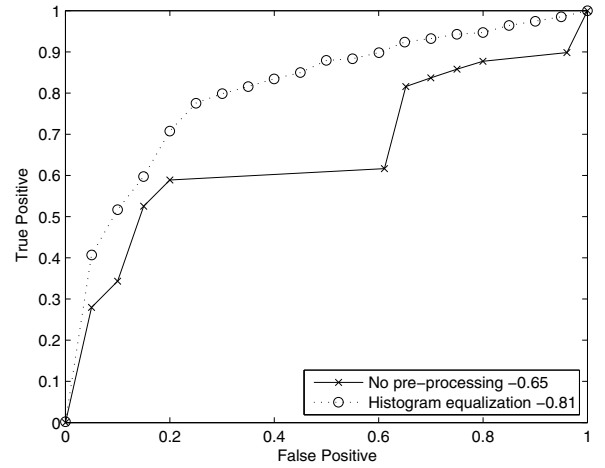


Fig. 8. Comparison between the I-PyraNet over not pre-processed images and images with the histogram equalized.

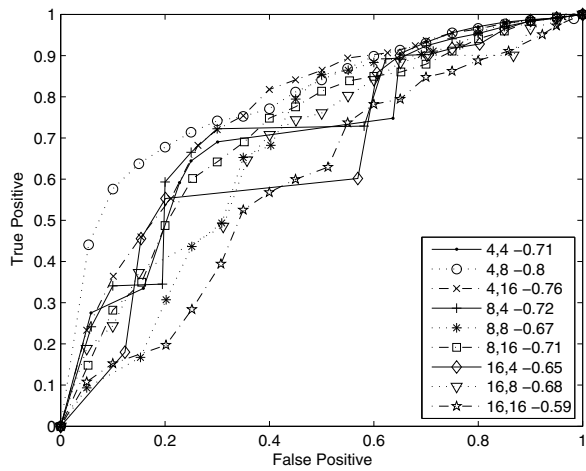


Fig. 7. ROC curves for different values of σ_x and σ_y for the Gabor filter. The legend presents the value of σ_x and σ_y followed by the area under the curve.

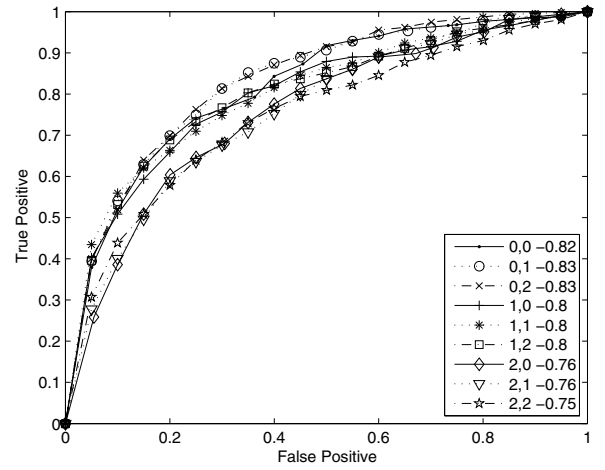


Fig. 9. ROC curves for different inhibitory fields configuration of the I-PyraNet. The legend presents the size of the inhibitory field for first and the second 2-D layer, respectively, followed by the area under the curve.

over images with the histogram equalized achieved the best results, with an area of 0.9.

Finally, in Figure 11, the comparison between the I-PyraNet, the PyraNet and the SVM is presented. The I-PyraNet had better results than the PyraNet, but the SVM presented the biggest area under the curve. However, the use of the I-PyraNet might be justified by its short time to perform the classification task. While the I-PyraNet takes only 0.04 milliseconds to classify a pattern, the SVM takes 7 milliseconds. Which means that the I-PyraNet with the configuration calculated during the experiments is 175 times faster than the SVM. It is also important to note that the application of the Gabor filter takes 0.3 milliseconds to generate the Gabor representation of an image.

The fact of the I-PyraNet is much faster than the SVM might motivate its use in embedded systems, where some

processing restrictions are critical.

On the other hand, the results of the I-PyraNet can be improved by repeating the classification step with different instances of the I-PyraNet for a same face many times. The application of the I-PyraNet 25 times takes only 1 millisecond, which is still seven times faster than the SVM. Then, the results obtained by each classification step must be combined in order to decide if a given pattern is or is not a face.

VI. CONCLUSIONS AND DISCUSSIONS

In this work, a new neural network, called I-PyraNet, was presented. The I-PyraNet is a hybrid implementation between the recently described PyraNet and the concepts of inhibitory fields. Also, it was presented here the 2-D Gabor filter that was motivated by the concepts of receptive fields. Both, I-PyraNet and 2-D Gabor filter, were applied over a

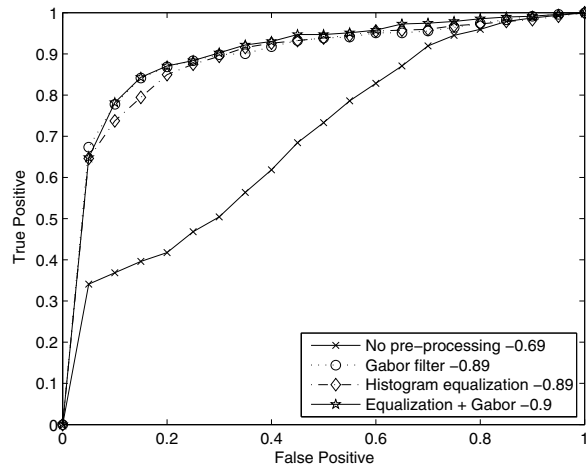


Fig. 10. ROC curves for the SVM classifier using not pre-processed images, images with the histogram equalized and the Gabor filter.

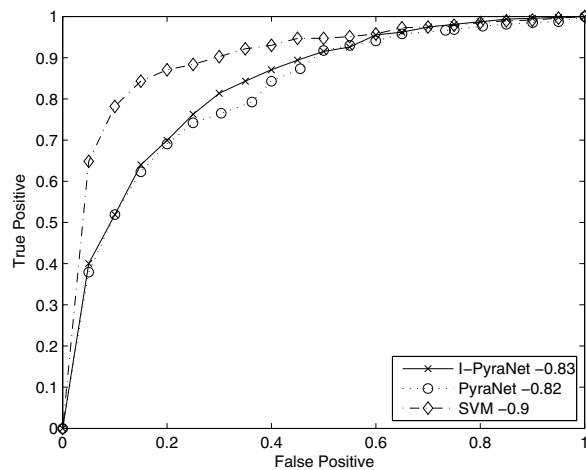


Fig. 11. Comparison between the I-PyraNet, the PyraNet and the SVM classifier.

face detection task and showed very good results. The I-PyraNet was compared to the regular PyraNet and to the SVM. In the experiments, the MIT CBCL which is a well-know database, was used. The experiments showed that the I-PyraNet achieved better results than the regular PyraNet. On the other hand, SVM achieved the best classification rates among the three classifiers. However, it is important to note that the I-PyraNet works 175 times faster than the SVM, this fact can be good reason to motivate the use of the I-PyraNet, since in certain applications such as in embedded systems the can be critical.

Finally, spite the fact that the I-PyraNet was only used here for a face detection problem, the proposed method can also be applied other different type of visual recognition problems, like object detection and word recognition.

REFERENCES

- [1] R. C. Gonzales and R. E. Woods, *Digital Image Processing*. Prentice-Hall, 2007.
- [2] D. H. Hubel, "The visual cortex of the brain," *Scientific American*, no. 209, pp. 54–62, 1963.
- [3] J. Daugman, "Two-dimensional analysis of cortical receptive field profiles," *Vision Research*, vol. 20, pp. 846–856, 1980.
- [4] J. Jones and P. Palmer, "An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex," *Journal of Neurophysiology*, vol. 58, pp. 1233–1258, 1987.
- [5] S. L. Phung and A. Bouzerdoum, "A pyramidal neural network for visual pattern recognition," *IEEE Transactions on Neural Networks*, vol. 18, no. 2, 2007.
- [6] G. Rizzolatti and R. Camarda, "Inhibition of visual responses of single units in the cat visual area of the lateral suprasylvian gyrus (clare-bishop area) by the introduction of a second visual stimulus," *Brain Res.*, vol. 88, no. 2, pp. 357–361, 1975.
- [7] C. Grigorescu, N. Petkov, and M. A. Westenberg, "The role of non-crf inhibition in contour detection," *Journal of WSCG*, vol. 11, no. 1, 2003.
- [8] M. Levine and J. Shefner, *Fundamentals of sensation and perception*. Oxford University Press, 2000.
- [9] C. Grigorescu, N. Petkov, and M. A. Westenberg, "Contour detection based on nonclassical receptive field inhibition," *IEEE Transactions on Image Processing*, vol. 12, no. 7, pp. 729–739, 2003.
- [10] C. Blakemore and E. Tobin, "Lateral inhibition between orientation detectors in the cats visual cortex," *Exp. Brain Res.*, vol. 15, 1972.
- [11] D. Gabor, "Theory of Communication," *Journal of the Institute of Electrical Engineers*, vol. 93, no. 26, pp. 429–457, 1946.
- [12] J. Yang, L. Liu, T. Jiang, and Y. Fan, "A modified Gabor filter design method for fingerprint image enhancement," *Pattern Recognition Letters*, vol. 24, pp. 1805–1817, 2003.
- [13] A. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.
- [14] L. Hong, Y. Wan, and A. Jain, "Fingerprint image enhancement: Algorithm and performance evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 777–789, 1998.
- [15] M. Zhou and H. Wei, "Face verification using Gabor Wavelets and AdaBoost," *18th International conference on Pattern Recognition*, pp. 404–407, 2006.
- [16] A.-A. Bhuiyan and C. H. Liu, "On Face Recognition using Gabor Filters," *Proceedings of the World Academy of Science, Engineering and Technology*, vol. 22, pp. 51–56, 2007.
- [17] B. J. T. Fernandes, G. D. C. Cavalcanti, and T. I. Ren, "Classification and segmentation of visual patterns based on receptive and inhibitory fields," *Eighth International Conference on Hybrid Intelligent Systems*, pp. 126–131, 2008.
- [18] B. J. T. Fernandes and G. D. C. Cavalcanti, "A pyramidal neural network based on nonclassical receptive field inhibition," *20th IEEE International Conference on Tools with Artificial Intelligence*, 2008.
- [19] Y. Lecun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computing*, vol. 1, no. 4, pp. 541–551, 1989.
- [20] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 2007.
- [21] D. M. Kline and V. L. Berardi, "Revisiting squared-error and cross-entropy functions for training neural network classifiers," *Neural Computing and Applications*, vol. 14, pp. 310–318, 2005.
- [22] D. Rumelhart, G. Hinton, and R. Williams, "Learning internal representations by backpropagation," *Nature*, vol. 323, pp. 533–536, 1986.
- [23] E. Makinen and R. Raisamo, "Evaluation of gender classifications methods with automatically detected and aligned faces," *IEEE. Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 3, pp. 541–547, 2008.
- [24] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 1995.
- [25] E. Osuna, R. Freund, and E. Girosit, "Training support vector machines: an application to face detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 130–136, 1997.
- [26] B. Heisele, T. Poggio, and M. Pontil, "Face detection in still gray images," Tech. Rep. 1687, Center for Biological and Computational Learning, MIT, Cambridge, MA, 2000.

- [27] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006.
- [28] R. D. Valois, W. Yund, and N. Hepler, "The orientation and direction selectivity of cells in macaque visual cortex," *Vision Research*, vol. 22, pp. 531–544, 1982.