Automatic Image Segmentation Colors by neocognitron neural network

Abstract

This paper contain algorithm of a neocognitron neural network that has been trained to recognize segmentation of Image from color image. Corners and Edge contour extraction are segmentation by models of complex and end-stopped cells. Detection of corners and local edge maxima is performed by selection of local maxima in both edge and corner sub image. The advantage of the proposed model over other models is that the same low constant thresholds for corner and local edge maxima detection for each color are used for different images.

A neocognitron neural network prove success 100% for each sub image segment color.The system can segments the image in any number of parts choices and process with 16 x 16 or 24 x 24 bits for each process we can used any types of image with any size.

Keywords: Neural network, neocognitron neural network, image segmentation ,Hybrid image segmentation techniques ,Edge color detection

1. Introduction

Computer images are extremely data intensive and hence require large amounts of memory for storage. As a result, the transmission of an image from one machine to another can be very time consuming[2]. By using data compression techniques, it is possible to remove some of the redundant information contained in images, requiring less storage space and less time to transmit. Neural nets can be used for the purpose of image segmentation.

The existing automatic image segmentation techniques can be classified into four approaches
1) threshold techniques,
2) boundary-based methods,
3) region- based methods, and
4) hybrid techniques[1].

1.1 Threshold techniques are based on the assumption that adjacent pixels whose value (grey level, color value, texture etc) lies within a certain range belong to the same class[2]. Threshold techniques can obtain good segmentation of images that include only two opposite components.

1.2 Boundary-based techniques use the assumption that pixel values change rapidly at the boundary between two regions [5], [6]. Edge detectors used in these techniques can be simple ones such as the Sobel or Roberts operators, or more complex ones such as the Canny operator.

The output of most existing edge detectors can only provide candidates for the region boundaries, because these obtained color edges are normally discontinuous or over-detected[5].

1.3 Region-based techniques rely on the assumption that adjacent pixels in the same region have similar visual features such as grey level, color value, or texture[3]. The performance of this approach largely depends on the selected homogeneity criterion. High-level knowledge of the image components can be exploited through the choice of seeds. This property is very attractive for semantic object extraction toward content-based image database applications.

1.4 Hybrid techniques which integrate the results of boundary detection and region growing, are expected to provide more accurate segmentation of images. Neural networks can be viewed as massively parallel computing systems consisting of an extremely large number of simple processors with many interconnections.

Neural network models attempt to use some organizational principles (such as learning, generalization, adaptively, fault tolerance and distributed representation, and computation) in a network of weighted directed graphs in which the nodes are artificial neurons and directed edges (with weights) are connections between neuron outputs and neuron inputs. The main characteristics of neural networks are that they have the ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to the data[6].we used this type for our system because Neural networks can be viewed as massively parallel computing systems and contains the effective results.

2- Segmentation Colors Algorithm

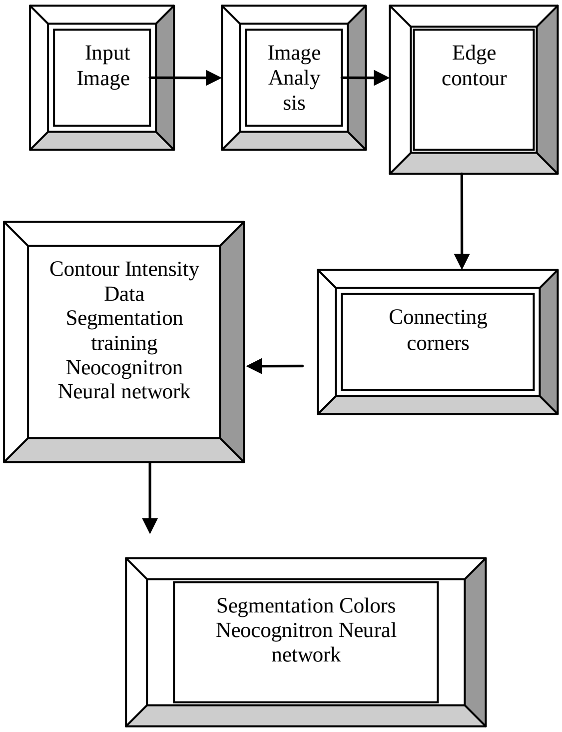
Our system contain 4 stages that are shown in figure below :

1-Image Analysis

2- Edge contour extraction for each color

3- Connecting corners and training Neocognitron Neural network

4-Testing Segmentation Colors Neocognitron Neural network



Figure(1):Block Diagram for algorithm system

2-1 Image Analysis

Image analysis process can be broken down into three stages :

1- Preprocessing :

Is used to determine the noise and eliminate irrelevant visually information , noise is unwanted information can be result form the image acquisition process , also in this stage we can find regions of interest for father process .

2- Data Reduction :

Involve either reducing the data in spatial domain or transforming it into another domain called the frequency domain and the extracting features for the analysis process .

3- Feature analysis :

The feature extracted by the data reduction process are examined and evaluated for use in application .

2-2 Edge contour extraction for each color

The second step in the algorithm is contour extraction. Local edge maxima and corners should be treated differently. Two contours can start and in order to avoid missing any, a contour will be followed in every direction. contour is always passing through a local maximum a corner should have been marked there. Hence at a local maximum two opposite directions will be selected to be followed.

Contour following is based on a greedy algorithm: walking from one top,

local edge maximum, to another and trying to keep as high as possible. The latter is done by selecting always that coordinate (within certain constraints) with the highest Call σ -response. In the layer of linking cells this means that the neighbor with strongest input response is activated and linked by firing synchronously with the other activated neurons.

```
1  Function ExtractContour (C_σ^all, C, M, c, n)
2    p := n; contour := p; pp := c
3    repeat
4      if (NotAtBorderImage(p) and NoMember(contour, p) and C_σ^all(p) ≥ T and p ∉ C ∪ M)
5        p1 := Neighbor of p to pp at 180°
6        p2 := Neighbor of p to pp at 135°
7        p3 := Neighbor of p to pp at 225°
8        pp := p;  p := pj, where C_σ^all(pj) = max(C_σ^all(pi)) ∀i; i, j ∈ {1,2,3}
9        Add (contour, p)
10     else stop
11   until stop
12   if (p ∈ C ∪ M or Length (contour) ≥ Len)
13     return contour
```

Figure(2):Algorithm for extraction of a single contour for each color segmentation

In this figure variable *p* denotes the current point or coordinate and *pp* previous point, which is initially a corner or local edge maximum. At every step there are three possible neighbors of *p* that can be selected. These neighbors form a 135, 180, and 225 degree angle with *p* and *pp* (lines 5-7). From these three neighbors the one with the highest Call σ -response is selected (line 8).

3- Connecting corners and training Neocognitron Neural network

Third step is Connecting corners to contours necessary because a contour does not necessarily have to pass through a corner, to get a fully symbolic representation, an additional step is taken that connects corners to contours if the closest distance between corner and contour is less than $d\sigma$. This value is derived from the single end-stopped operator.

The amplitude for each of the grey value, red-green, and blue-yellow channels yields the edge operator[2]

$$c_{\sigma}^{all} = \sqrt{c_{\sigma}^{2} + (\frac{1}{2} c_{\sigma}^{r,g})^{2} + (\frac{1}{2} c_{\sigma}^{b,y})^{2}} \quad \text{..............................(1)}$$

Where: c=color r=red ,g=green, b=blue, y=yellow

at a single scale. The corner operator is similar to (1), with the exception that for C one should substitute E.

We found the averaging the responses over arrange of frequencies yield much more robust corner detection operator:-

$$\Sigma_{avg}^{all}(x,y) = \frac{1}{s} \sum_{j=0}^{s-1} \Sigma_{j=0}^{all}(x,y) \quad \text{...(2)}$$

We are interested in determine all the different ways in which a given graph may be colored using at most (m) colors .

Adjacency matrix GRAPH (1:n,1:n) , where GRAPH (I,J) =true if (I,J) is an edge of G and otherwise GRAPH (I,J) = false . Then used Neocognitron Neural networks which contains:

3.1. Preprocessing

Preprocessing is used to determine the noise and eliminate irrelevant visually information , noise is un wanted information can be result form the image acquisition process , also in this stage we can find regions of interest for father process[7] .

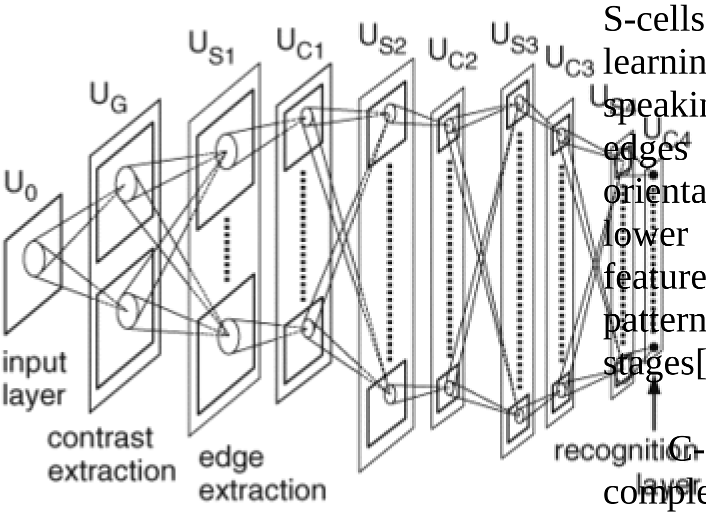


Figure (3): A typical architecture of the neocognitron

Figure (4) shows a typical architecture of the network. The lowest stage is the input layer consisting of two-dimensional array of cells, which correspond to photoreceptors of the retina. There are retinotopically ordered connections between cells of adjoining layers. Each cell receives input connections that lead from cells situated in a limited area on the preceding layer. Layers of "S-cells" and "C-cells" are arranged alternately in the hierarchical network.

S-cells work as feature-extracting cells. They resemble simple cells of the primary visual cortex in their response. Their input connections are variable and are modified through learning. After having finished learning, each S-cell come to respond selectively to a particular feature presented in its receptive field. The features extracted by S-cells are determined during the learning process. Generally speaking, local features, such as edges or lines in particular orientations, are extracted in lower stages. More global features, such as parts of learning patterns, are extracted in higher stages[8].

C-cells, which resembles complex cells in the visual cortex, are inserted in the network to allow for positional errors in the features of the stimulus. The input connections of C-cells, which come from S-cells of the preceding layer, are fixed and invariable. Each C-cell receives excitatory input connections from a group of S-cells that extract the same feature, but from slightly different positions. The C-cell responds if at least one of these S-cells yield an output. Even if the stimulus feature shifts in position and another S-cell comes to respond instead of the first one, the same C-cell keeps responding. Thus, the C-cell's response is less sensitive to shift in position of the input pattern. We can also express that C-cells make a blurring operation, because the response of a layer of S-cells is spatially blurred in the response of the succeeding layer of C-cells.

In other words, the connections to a cell-plane have a translational symmetry. As a result, all the cells in a cell-plane have receptive fields of an identical characteristic, but the locations of the receptive fields differ from cell to cell. The modification of variable connections during the learning progresses also under the restriction of shared connections.

3-2 errors of local features Image

Since small amounts of positional errors of local features are absorbed by the blurring operation by C-cells, an S-cell in a higher stage comes to respond robustly to a specific feature even if the feature is slightly deformed or shifted.

Thus, tolerating positional error a little at a time at each stage, rather than all in one step, plays an important role in endowing the network with the ability to recognize even distorted patterns.

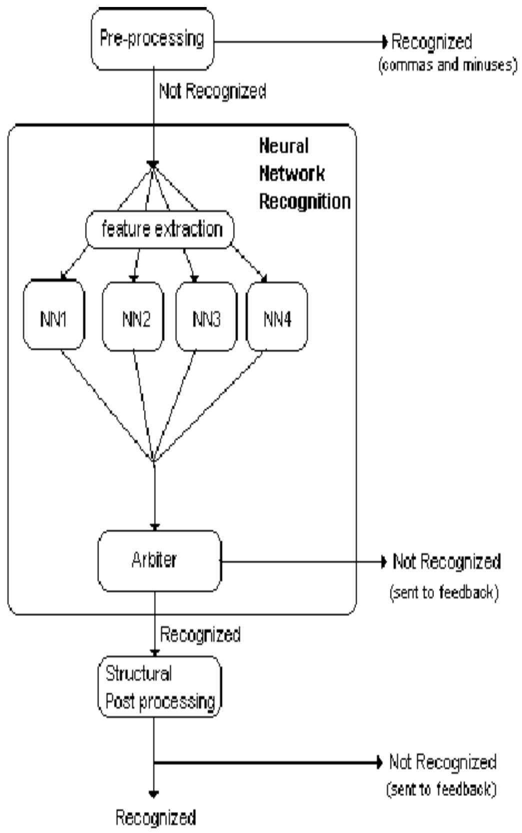
The C-cells in the highest stage work as recognition cells, which indicate the result of the pattern recognition. Each C-cell of the recognition layer at the highest stage integrates all the information of the input pattern, and responds only to one specific pattern. Since errors in the relative position of local features are tolerated in the process of extracting and integrating features, the same C-cell responds in the recognition layer at the highest stage, even if the input pattern is deformed, changed in size, or shifted in position. In other words, after having finished learning, the neocognitron can recognize input patterns robustly, with little effect from deformation, change in size, or shift in position.

After the learning, the S-cell acquires the ability to extract a feature of the stimulus presented during the learning period. Through the excitatory connections, the S-cell receives signals indicating the existence of the relevant feature to be extracted. If an irrelevant feature is presented, the inhibitory signal from the V-cell becomes stronger than the direct excitatory signals from the C-cells, and the response of the S-cell is suppressed.

4-Testing Segmentation Colors Neocognitron Neural network

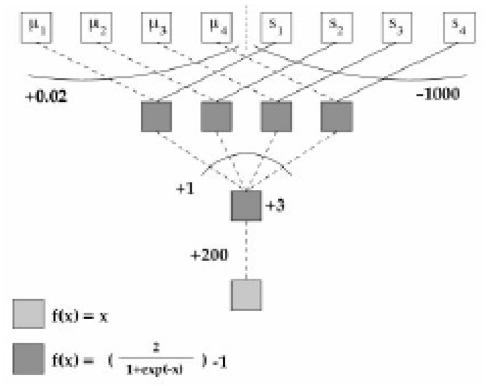
Once an S-cell is thus selected and has learned to respond to a feature, the cell usually loses its responsiveness to other features. When a different feature is presented, a different cell usually yields the maximum output and learns the second feature. Thus, a "division of labor" among the cells occurs automatically.

The second principle for the learning is introduced in order that the connections being strengthened always preserving translational symmetry, or the condition of shared connections. The maximum-output cell not only grows by itself, but also controls the growth of neighboring cells, working, so to speak, like a seed in crystal growth.



Figure(4): The flow of the recognition module for neocognitron neural network

To be more specific, all of the other S-cells in the cell-plane, from which the “seed cell” is selected, follow the seed cell, and have their input connections strengthened by having the same spatial distribution as those of the seed cell.



Figure**(5)** The module for selecting the right mean. The top layer is
the input layer.

S~-neuron is trained on the base of function of weight coefficients change

$$w(t+1) = w(t) + \frac{1}{t+1} \cdot \left[u(t+1) - w(t)\right]$$
……(4)

where *t* is number of training iteration, *w* is value of weight coefficient, *u* is value of activity for the neuron on an input of link to be trained.
The activation function of *S*~-neuron corresponds to the radial basic function.

$$U_{Sl}(n,k) = \exp\left(\sqrt{\frac{\sum_{p=1}^{P_{Cl-1}} \sum_{v \in Al} (\tilde{U}_{Cl}\ (n+v,p) - w(v,k))^2}{N}} \right)$$
……(5)

where l is the serial number of a layer; k is the number of a trained plane; n - two-dimensional index of a neuron in k -th plane; w - weight coefficient of link; v – two-dimensional shift of entry link in a subgroup of links Al ;
Al - two-dimensional value describing a size of a receptor subgroup in previous C - sublayer, N , p , - total amount of entry links and serial number of sub layer to be joined with a trained neuron correspondingly.

The argument $\tilde{U}_{Cl}$

$$\tilde{U}_{Cl}\ (n+v,p) = U_{Cl}\ (n+v,p) - \min_{\forall v}(U_{Cl}(n+v,p))$$
…..(3)

Is entered into a relation with the purpose of exception in
receptor a subgroup of a constant component influence.

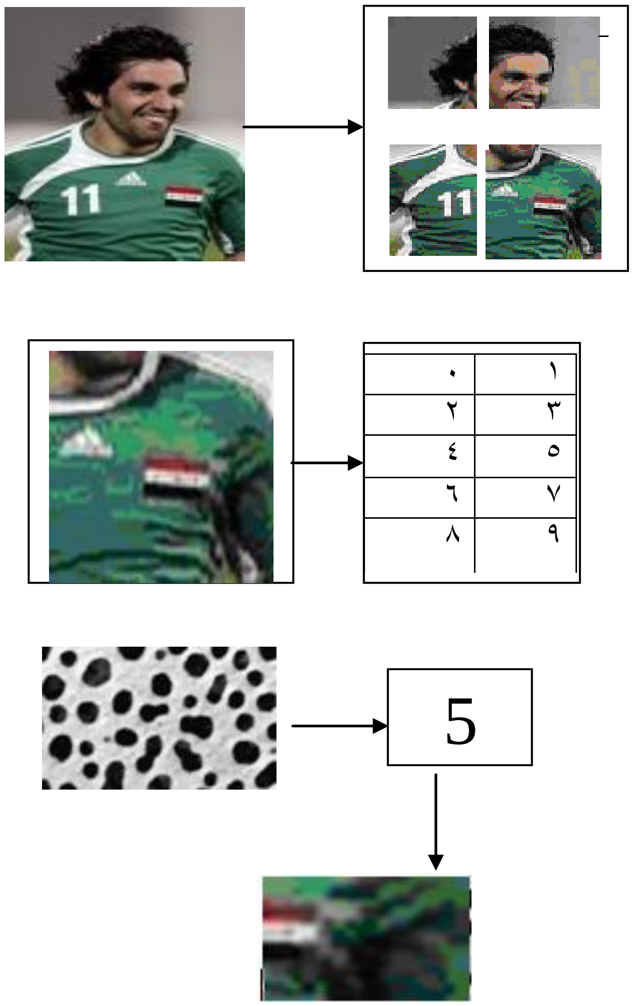
This is necessary because the neural network mast be invariant to change of brightness for the recognized image.

C -sublayer consists from *C*~ -neurons, that generalize one feature from previous sublayer. The generalization means, that if in receptor field of *C*~-neuron is found out even one active neuron, then *C*~-neuron too passes in active state, i.e. the fuzzy logic function “or” on all neuron receptor field is fulfilled and its value is assigned to an output signal of a neuron.

$$U_{Cl}(n,k) = \max(U_{Sl-1}(n+v,p)), \forall v \in Dl, \forall p \in P$$
……..(6)

where l is serial number of a layer; p , P – serial number of a plane and set of planes from the previous S -layer accordingly; Dl - two-dimensional value describing sizes of C~-neuron for a receptor subgroups; v – two-dimensional index of link inside these subgroups.

when the networks are trained a little further as a whole with a low learning rate (0.1), the algorithm improves rapidly: after only 100–500 learning cycles training can be stopped. The final effects of network are shown in Figure below.



Figure(6):The original image and its segmentations (for each process give to network) were converted to floating point images, with grey values in the range [20.5, 0.5].

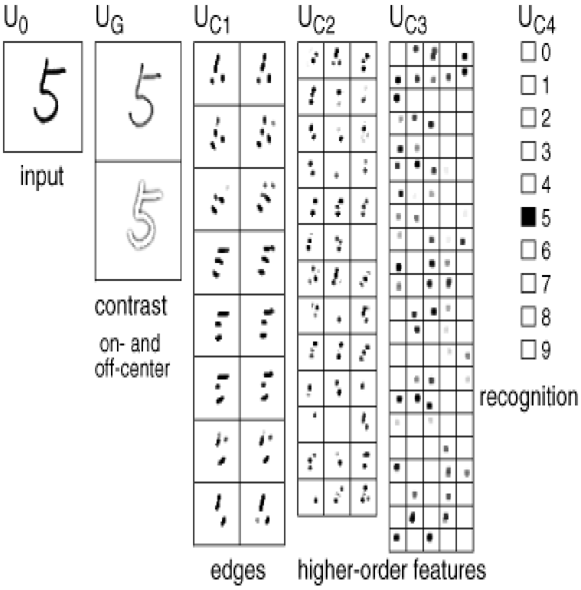


Figure (7): The response of a neocognitron that has been trained to recognize Image.

The procedure for save values image in buffer array is:

```
for I:=0  to
image1.picture.height-1
do
  for j:=0  to
image1.picture.width-1
do
   begin
z1:=image1.canvas.pixel
s[i,j];
  for k:=1 to 24 do
     begin
       if z1 mod 2=0 then
          begin
            if k <= 8 then
               r[i,j]:=r[i,j]+'0'
            else if (k <= 16)
then

g[i,j]:=g[i,j]+'0'
                 else
b[i,j]:=b[i,j]+'0';
          end
       else
          begin
                if k <= 8
then
             r[i,j]:=r[i,j]+'1'
           else if  (k<=16)
then

g[i,j]:=g[i,j]+'1'
                 else
b[i,j]:=b[i,j]+'1';
          end;
     z1:=z1 div 2;
   end;
   s1:=r[i,j]+g[i,j]+b[i,j];
   end;
```

5-Conclusion

This paper proposed a recognition module to Image Segmentation Colors by neocognitron neural network. The segments produced are normalized in order to simulate uniformity between sub image that should be recognized as the same value. This is required because the major difficulty in image segmentation is the variation of image style. Variations in width, height, slant and stylistic differences in image all make segmentation extremely difficult. Therefore, segments are normalized to a 16x16 input or 24 x 24 for the segmentation module.

Advantages of using multi-parameter solutions are:

• solutions are relatively easily obtained, using general methods to obtain thez

parameters

• no simplifications of the models are required

Disadvantages of multi-parameter solutions are:

• solutions maybe inefficient, often more parameters are used than is necessary

• little knowledge is gained about the solution itself

• the quality of the solution can often only be determined experimentally

Using only the three colors (per pixel) the object identification is pretty good, but not perfect. There are quite a few pixels that are miss-classified . The obvious extension of using just color would be to use other features as well.

For corner and contour detection we used low constant thresholds for different images. In this respect our method is very robust and outperforms the other edge detection methods, where setting a threshold often strongly influences the results. The extraction algorithm guarantees contours without gaps and avoids global threshold settings.

The success of neural net architecture does not simply depend on the type of network used, but rather on the methods used to tailor the total architecture to the environment that it will operate. In this application, the segmentation has accomplished this task through different methods. First, redundancy is utilized through an array of networks to both increase recognition rate and reduce false positive rates. Second, a version of the image feature extraction is used, tuning the networks to different information from the input segments and proving to increase the discrimination and generalization power of the respective neural nets and thus the overall architecture. Finally, the networks have been trained on segments gathered from the application's segment or itself.

The segmentation of image information is a difficult process. In order to improve the performance of any segmentation system, constraints on the use or the environment should be used. Given this, the segmentation module presented has utilized standard techniques in a unique way and has successfully taken advantage of the environment that it operates within.

We found that the Neocognitron is highly dependent on the selection of
the relevant filter sets. Failure to select good filters can lead to biased results. We also applied this understanding to provide an additional layer in the structure to handle sub image with various color as confirmation of our study.

In conclusion, we emphasize that proper selection of filters is the key to the success of the Neocognitron.

6-**Further research**

A number of subjects, which should be investigated further, were only shortly addressed
in this work. They include items concerned with fast and reliable learning in neural networks like efficiency of learning rules, problems with local minima and stopping criteria, quality of training sets, optimal network architectures and clever initialization of connection weights

Several architectures for image filters were only very briefly addressed. An interesting architecture for image filtering that was mentioned, is an architecture with symmetric weights that takes the whole image as input. The advantages of such an architecture
are that the storage requirements are not excessive, while in contrary to the local neighborhood based image filtering approach, it is able to use global image features.

A disadvantage is the increased complexity.

The research presented in this work also creates several possibilities and challenges for further and new research.

• Applying AVR for design and evaluation of image filters and other image processing operations segmentation

• Determining cost functions for specific applications by analyzing the effect

of errors in the output of an image processing operation on the output of the total image processing system
• Generation of test and training images using 3D scene and imaging models
• Training neural networks for minimum average risk
• Analysis of internal representations of neural networks using decomposition in basic components
• Using neural network image filters in image processing systems
• Investigating applicability of neural networks for other image processing operations.

٧-References

[1] Cottrell M, Fort J C and Pages G 1994," Two or three things that we know about the Kohonen algorithm", Proc. Eur. Symp. on Artificial Neural Networks (Brussels: D facto) pp 235–44.

[2]C. Murat and V.T. Paschos. The probabilistic minimum coloring problem. In Proc. Workshop on Graph Theoretic Aspects ofComputer Science, Lecture Notes in Computer Science. Springer Verlag, 2003

[3]Erwin E, Obermayer K and Shulten K 1992 ,"Self-organizing maps: ordering, convergence properties and energyfunctions", Biol. Cyber. 67 47–55

[4] K. Fukushima, N. Wake. Handwritten alphanumeric character recognition by the Neocognitron, IEEE Trans. on Neural Networks 2 (3) (1991). p. 355-365.

[5]K. Fukushima: "Neocognitron for handwritten digit recognition", Neurocomputing, 51, pp. 161-180 (April 2003).

[6]K. Fukushima: "Restoring partly occluded patterns: a neural network model", Neural Networks, 18[1], pp. 33-43 (2005).

[7]Kunihiko Fukushima (2007)," Neocognitron". Scholarpedia, 2(1):1717, Created: 23 June 2006, reviewed: 18 January 2007, accepted: 19 January 2007

[8]Martinetz T M, Berkovich S G and Schulten R J 1993, "Neural-Gas network for vector quantization and its applicationto time-series", prediction IEEE Trans. Neural Networks 4 558–69