4.5 The Learning Classifier Systems

In Section 4.5, we introduced several methods that use the RBFN as a classifier. This section introduces another classification method called *the learning classifier system* (*LCS*), proposed by John Holland [62]. The learning classifier system is a rule-based, message-passing, machine learning paradigm designed to process environmental stimuli, much like the input-to-output mapping provided by a neural network. The LCS provides learning through genetic and evolutionary adaptation to changing task environments. Thus, the learning mechanisms are built in directly into paradigm.

The operation of the LCS is centered around a list of rules or *classifiers*. These rules are essentially a set of “if-then” statements, where the “if” part of a rule is called *condition*, and the “then” part is called an *action*. The further details of the LCS will not be given here because they are not the center of this work.

We employed the LCSs to solve our color image classification but they did not give good results. Even though they learn the training samples completely, they could not classify the test samples as well the RBFNs do. This shows that the LCSs over fits the classes without enough generality. The LCS assigns a rule for each sample during the training. When the test samples are applied to the system, it looks for the exactly the same rules in the test images. The LCSs gave 17% correct classification for the test samples. This result can be considered as a random result.

Finally, we conclude that the LCSs are not really proper for this kind of classification problem since they do not have enough ability to generalize. Even though current LCSs cannot provide this feature, if necessary changes are made in the LCSs, they might provide enough generalization. This probably requires another long term research in the area.

4.6 Comments and Discussion on the Methods.

This section gives a summary of the results and comparisons between the methods that use the RBFN. Table 4.1 summarize all methods in terms of their network structure, the types of input feature extraction method, the number of centers in the network, and their performances.

Table 4.1. The summary of the methods with their results.

Method Name	Input Feature Type	# of Centers and RBFs	Network Structure	% Correct Classification
Method I	Average Colors (r, g, b)	8 8	Traditional RBFN	55.06
Method II	Image Partitioning ($r_1, g_1, b_{1.,}\ , r_4, g_4, b_4$)	8 8	Traditional RBFN	66.75
Method III	Average Colors (r, g, b)	23 23	Traditional RBFN	67.79
Method IV	Average Colors (r, g, b)	8 24	O-RBFN	72.20
Method V	Av. Colors and STDs ($r, g, b, {}_r, {}_g, {}_b$)	8 48	O-RBFN	77.66
Method VI	Covariance Matrix (Nine input features)	8 72	O-RBFN	78.12
Method VII	Histogram (24 elements)	8 8	Traditional RBFN	77.14
Method VIII	Histogram with white noise (24 elements)	8 8	Traditional RBFN	78.15
Method IX	Histogram (24 elements)	30 30	RBFN (Fixed Dilation)	77.34
Method X	Histogram (8 elements for each color)	21, 21 19, 19 16, 16	RBFN (Fixed Dilation)	79.74 (red) 80.25 (green) 81.81 (blue)
Method XI	Histogram (Only blue 64 elements.)	80 80	RBFN (Fixed Dilation)	84.41 (blue)

Now, we will compare the results in terms of the network structure, the number of input features in the input vectors, and the number RBFs in the network. First of all, we will compare the results in terms of the type of network structure employed. Figure 4.6 shows how the proposed network structure improved the results. Using the same input features and the same number of RBFs, the O-RBFN structure provides about 17% improvement on the results. While Method I performs a 55.06% correct classification, Method IV performs a 72.2% correct classification.

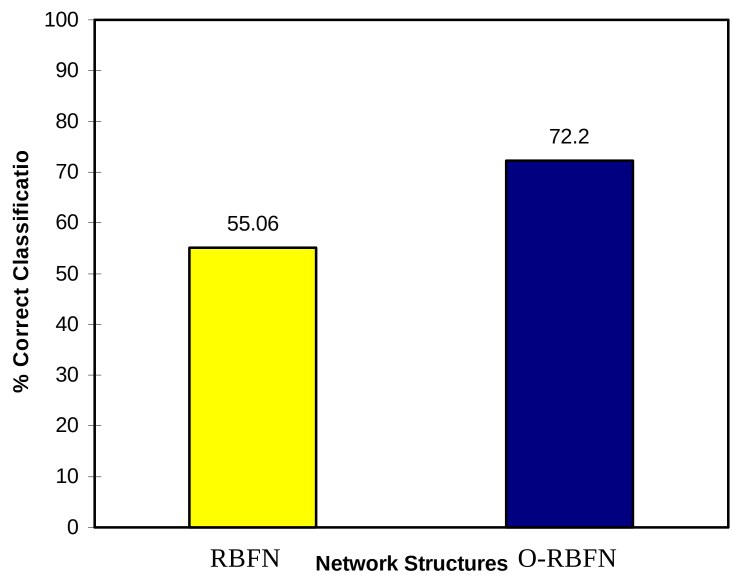

Figure 4.6. The comparison between the RBFN structure and O-RBFN structure.

Next, we will see how the number of input features in the input vectors improves the results with traditional RBFN structure and O-RBFN structure. As can be seen in Figure 4.7, Method I performs a 55.06% correct classification with three input features while Method II performs a 66.75% correct classification with 12 input features in the input vectors. Method I and Method II use the traditional RBFN structure. Method IV performs a 72.2% correct classification with three input features in the input vectors while Method V performs a 77.66% correct classification with six input features. Finally, Method VI performs a 78.12 correct classification with nine input features in the input vectors. As can be seen , the more input features we used in the input vectors, the more the correct classification the network performs.

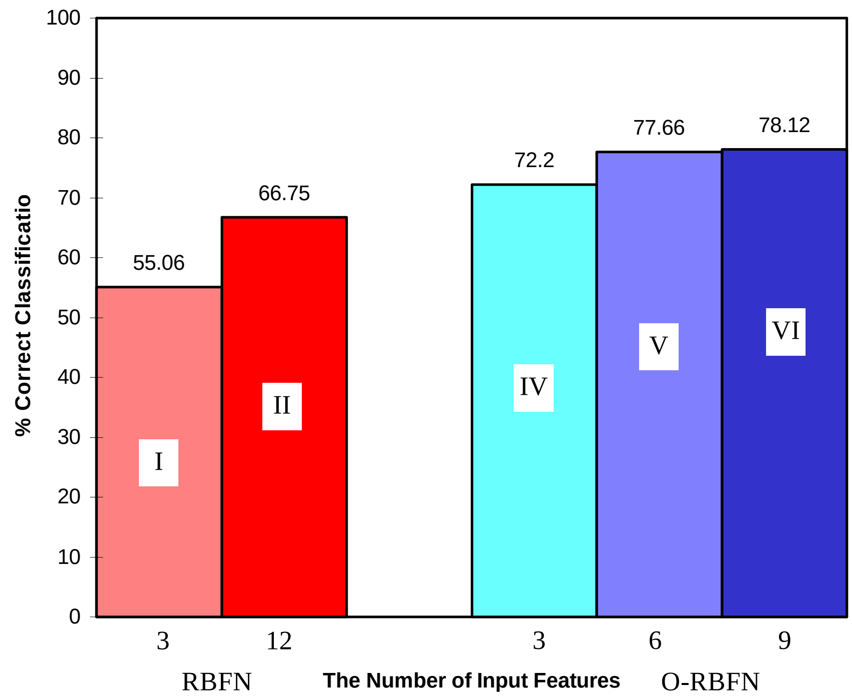

Figure 4.7. The comparison of the methods in terms of the number of input features in the input vectors.

The number of RBFs employed in the network is another parameter that affects the results. Figure 4.8 shows how the number of RBFs used in the network affects the results. We should consider Method I and Method II together since they have the same type of network structure. Similarly, Method IV, V, and VI are considered together. As can be seen in the Figure 4.8, there is almost 13% improvement in the results since the number of RBFs is increased from 8 to 23. Similarly, the results are improved while we are increasing the number of RBFs in the network in Methods IV, V, and VI. There is an about 6% improvement in the results while the number of RBFs is increased from 24 to 72 in Method IV through Method VI.

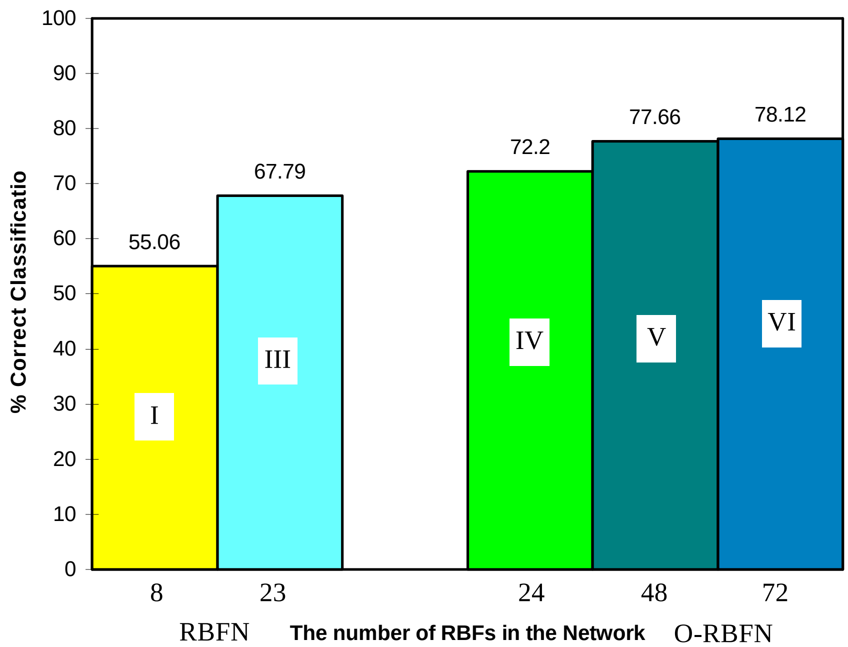

Figure 4.8. The comparison of the methods in terms of the number RBFs in the network

At last, we will examine how the usage of histogram data improves the results. We should consider Method VII and Method VIII together because their network structure differ from the network structure that Methods IX, X, and XI employ. While Method VII and Method VIII use the traditional RBFN structure with adjusted dilations, Methods IX, X, and XI use the same structure with fixed dilations. Figure 4.9 illustrates the results of these methods. Method VII and Method VIII differ from each other in the sense that Method VIII has white noise in the histogram vectors. These two methods perform almost the same but Method VIII performs around 1% better than Method VII. While we are examining the Methods IX, X,, and XI, we need to consider Method IX and Method X together because they have the same number of input features in their input vectors. These two methods differ from each other in the sense that the Method IX applies each color's histogram separately to the network. This improves the results even though the network uses fewer input features. Finally, Method XI uses only the blue histogram with 64 bins. The input vectors in this method have 64 features. Using more input features improved the results about 3% more.

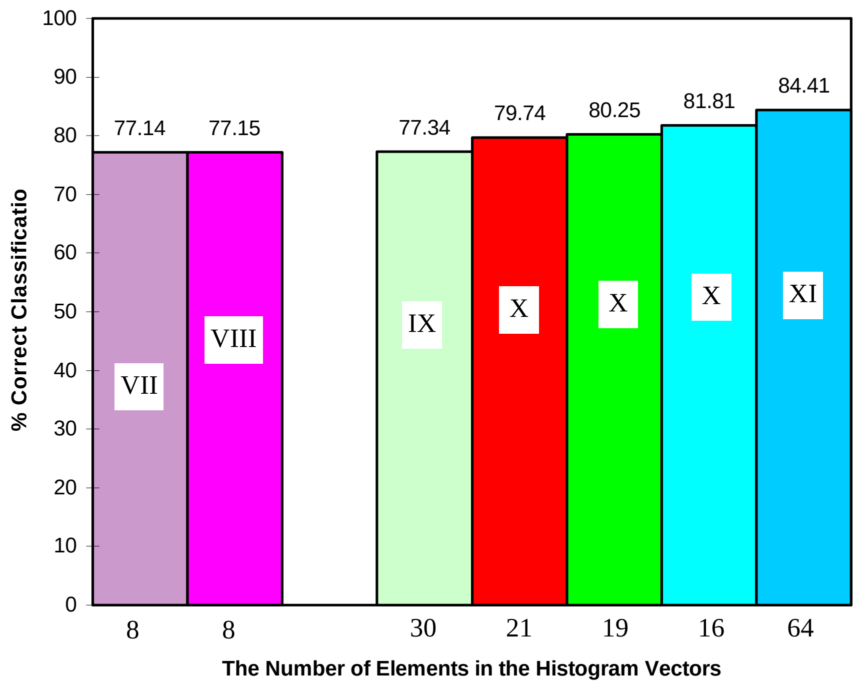

Figure 4.9. The Comparison of the methods in terms of the network structure and the number of elements in the histogram vectors. Methods IX, X, and XI have fixed dilations in their RBFs. Method VII and Method VIII have adjusted dilations in their RBFs.

Now, we compare our results with the results in [1]. Figure 4.10 illustrates our results and the results in [1] together. The methods compared here have the same input features with methods in [1]. This allows us to make a fair comparison between our methods and the methods in [1], the minimum distance classifiers.

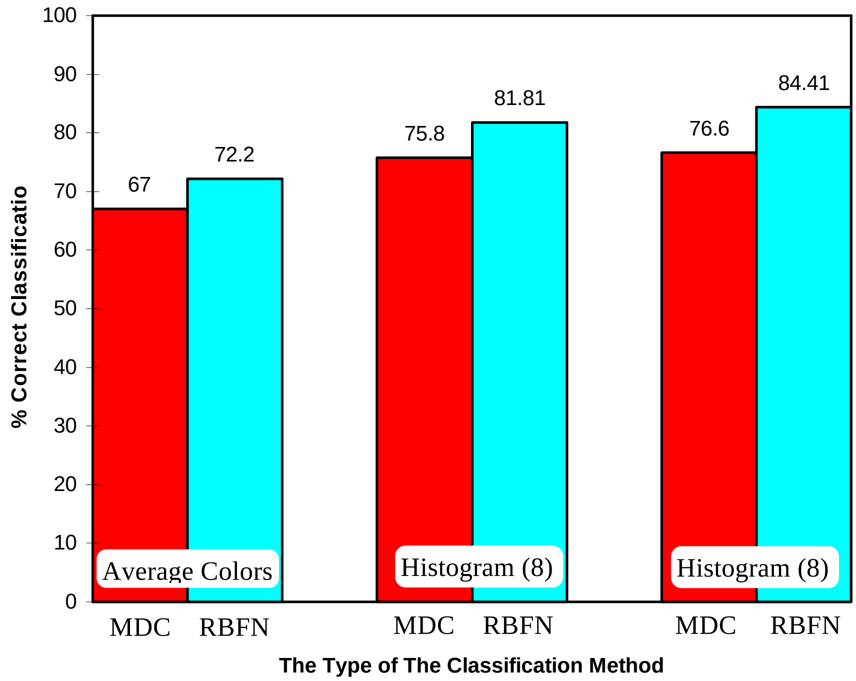

Figure 4.10. The comparisons of the minimum distance classifier (MDC) method and radial basis function network (RBFN) method.

The first two bars show the performance of the methods with three input features, the average values of red, green, and blue in the images. The next two bars show the results of the methods that use the histogram vectors created with eight bins for each color. Finally, the last two bars show the results of the methods that use the histogram vectors created with 64 for each color. In our methods, we used the uniform quantization to create the histogram vectors, explained in Section 4.3.4. Three dimensional uniform quantization was used in [1]. In this case, we have fewer elements in the histogram vectors than in the histogram vectors in [1]. Even though we have fewer elements in the histogram vectors, the RBFN performed better than the minimum distance classifiers did.

To summarize, the radial basis function networks performed better than both the LCS method and the minimum distance classifier method. The RBFNs can define the classes more accurately than the other two methods with a great deal of generalization property. Although the RBFNs gave better results that the minimum distance classifiers, 97% correct classification was obtained using different color quantization method in [1]. In this work, since we are focused on the classification method we compared our results with the results obtained by minimum distance classifiers. A new approach to color quantization was used to have the input features in [1]. If the same input features obtained by this method are used the RBFN methods might give almost 100% correct classification because theRBFNs gave better results with other input features.