A Neural Network Diagnosis Approach for Analog Circuits

Alessandra Fanni, Alessandro Giua,
Michele Marchesi, Augusto Montisci

DIEE, Università di Cagliari, Piazza d'Armi, 09123 Cagliari, Italy

email: {fanni,giua}@diee.unica.it, tel: +39 070 675.58.92, fax: +39 070 675.59.00

Abstract

This paper presents a neural network system for the diagnosis of analog circuits and shows how the performance of such a system can be affected by the choice of different techniques used by its submodules. In particular we discuss the influence of feature extraction techniques such as Fourier Transforms, Wavelets and Principal Component Analysis.

The system uses several different power supplies and as many neural networks "in parallel". Two different algorithms that can be used to combine the candidate sets produced by each network are also presented. The system is capable of diagnosing multiple faults even if trained on single ones.

Published as:

A. Fanni, A. Giua, M. Marchesi, A. Montisci, "A neural network diagnosis approach for analog circuits," Applied Intelligence, Vol. 11, No. 2, pp. 169-186, September, 1999.

1 Introduction

During the past years, the authors have been involved in several projects on analog circuit diagnosis and quality control of electrical components. The aim of this paper is to present the diagnostic system developed by them and to show how the performance of such a system can be affected by the choice of different techniques used by its submodules.

The system is based on *neural networks*, and is used for *off-line diagnosis* of analog circuits affected by *catastrophic multiple faults*. It may handle linear and nonlinear circuits in transient or steady state behavior.

1.1 Analog circuit faults

Fault diagnosis of analog circuits is a complex problem. Classical solutions require either a huge amount of calculation if parameter identification methods are used, or a great number of simulations of faulty conditions if fault dictionary methods are used [23, 24].

The faults in analog circuits may be *catastrophic faults*, that cause a large and sudden variation of the circuit parameter values, and *deviation faults*, associated to slight variations of the circuit parameter values from their nominal values [2]. Since statistics have shown that 80-90% of analog circuit faults are catastrophic [19], we chose to study faults of this kind, such as short circuits and open circuits between two terminals of a component.

In some applications (regulation systems, nuclear plants, etc.) a prompt fault detection is necessary to avoid damaging the controlled process any further. A diagnostic system capable of detecting a fault during its occurrence performs what is called *on-line diagnosis*.

In many other applications (quality control of circuits, post-mortem diagnosis of electronic boards, etc.) the diagnostic procedure may be applied in an *off-line* fashion, in the sense that the diagnosed device need not be operative. In these cases there is no strict time constraint and even computationally intensive diagnostic systems, such as those based on parameter identification or fault dictionary methods [23, 24], qualitative reasoning [11, 13] model-based and rule-based expert systems [7, 28] etc., may be used. In the case of electric circuits, off-line diagnosis offers an additional advantage: suitable voltage supply configurations may be chosen in order to maximize the observability of the faults [10].

1.2 Diagnosis as pattern recognition

Classically, a pattern recognition system is composed of three modules [12]. A *transducer* acquires data on a physical device and passes them to a *feature extractor* whose purpose is to reduce the data by computing certain features (or properties). These features will be used by a *classifier* to make a final decision on the state of the device.

A circuit diagnostic system is a particular pattern recognition system, in which the physical device is an analog circuit, and the state that must be recognized is the set of faulty components. In particular, in the diagnostic system we have developed, the classifier is a neural network.

This type of diagnostic system offers some advantage over other classical diagnostic methods.

Rule-based systems. These diagnostic systems use compiled sets of rules to associate a symptom to its cause. On the contrary, a neural network automatically derives the symptom-cause correspondence during the training, and does not require an explicit formalization. It is well known that this formalization is the bottleneck of rule-based system technology.

Note that there is a small price to pay for this. A rule-based system has a symbolic-heuristic approach to diagnosis and is generally able to justify its deduction from the rules used to compute the diagnosis. A neural network, on the contrary, has a numerical-algorithmic approach and the knowledge is implicitly memorized in the weights of its synapses. Thus, to justify its deduction a neural network requires additional rule extraction techniques [9].

Model-based systems. These diagnostic systems usually require the complete knowledge of the circuit scheme and a model of its behavior. Using neural networks it is possible to avoid the problems connected with the calculation of circuit parameters and in general to the modeling.

Fault dictionary method. This method can be used to identify only those faults whose signature has been previously computed and added to the dictionary. Neural networks on the contrary — as reported in several works — may be able to recognize fault configurations not explicitly included in the training set. In [14, 21, 31] neural networks trained to recognize single faults are successfully used to diagnose multiple faults. In [36] neural networks accurately classify previously unseen fault signatures belonging to a deviation fault class known by a few samples.

There have been several works where neural networks have been compared with other pattern classifiers in diagnosis applications. In the domain of single fault diagnosis of circuits, a comparison with Gaussian maximum likelihood and K-nearest neighbors is presented in [26] where neural networks, once trained, are shown to significantly reduce the time of the diagnosis, although they do not offer improvements in the diagnostic accuracy. The same result was independently reported in [36], comparing neural networks and K-nearest neighbors classifiers in the diagnosis of deviation faults.

In the diagnostic system we present, the transducer is an acquisition board that measures the *voltage values in a given set of test points*. Other choices are possible as we will discuss in Section 2. As an example, Spence et al. [34, 35] have used nonintrusive circuit measurements (such as infrared images or magnetic field images); however, nonintrusive measurements have been proved to be very ineffective, in the sense that they can only be used to recognize a limited number of faults. Kirkland and Dean [22] obtained good results using current measurements; however, current measurements are often impractical, since they would require the opening of the circuit, and this is clearly not possible on printed circuits.

We have investigated several feature extraction techniques and have studied their influence on the performance of the diagnostic system. In this paper we compare *Fourier Transforms* [14], *Wavelets* [8], *Principal Components Analysis* [15], and *Sampling*. In [16] *Mean and Root-Mean-Square Values* of the test point voltages were used as features, but due to the large amount of lost information they could only recognize a limited number of faults.

We also observed that the performance of the diagnostic system heavily depends on the choice of power supplies. In particular, it is often the case that a given supply can only lead to the detection of a particular subset of all possible faults. A suitable set of different supplies may be used to build a diagnostic system that combines different diagnoses (one for each supply) dramatically improving the performance of the diagnostic system. In the paper we also present two algorithms that can be used to combine these different diagnoses.

The paper is structured as follows. In Section 2 we recall relevant work on the use of neural networks for circuit diagnosis. In Section 3 we describe the architecture of the proposed diagnostic system and discuss the important issue of simulation versus acquisition. In Section 4 we discuss the choice of power supplies and how this affect the diagnosis. In Section 5 we describe different techniques that can be used by the features extraction module to compactly represent the behavior of the circuit. In Section 6 we present the structure of neural network classifier and show how it is trained. In Section 7 we present two algorithms that can be used to combine the diagnosis computed by different networks. In Section 8 we present statistics of the system performance when diagnosing two different circuits: a board part of a DC motor drive, and an oscillator.

2 Relevant work

Other approaches to the use of neural networks for circuit diagnosis have recently been published.

Keagle *et al.* [21] discuss how networks trained to recognize single faults may be used to detect multiple faults. Tests are performed on a *digital* circuit consisting of nine logical gates affected by stuck-at 1 or stuck-at 0. The paper also presents results on the performance of the diagnostic system as a function of the network architecture.

Meador *et al.* [26] compare feedforward neural network performance with other classifiers: gaussian maximum likelihood and K-nearest neighbors. In each experiment a single parameter deviation fault on an operational amplifier circuit is considered. The classifiers must separate the input patterns corresponding to the correct behavior and to the faulty one.

Parten *et al.* [29] propose using neural networks as part of a model-based expert system for diagnosing lumped parameter devices. The purpose of the net would be that of solving the equations ruling the behavior of the diagnosed device, modeled as a set of interconnected components.

Thompson *et al.* [38] consider the problem of diagnosing an IC board with approximately 60 components, both analog and digital. They use a backpropagation neural network with a modular structure, i.e., each part of the net recognizes a particular fault.

Totton and Limb [39] use neural networks to diagnose a circuit board part of a digital telephone exchange. They observed from historical data that failures on four types of components account for more than 85% of all faults. This led them to construct a network whose four outputs signal the presence of a faulty component of a given type, i.e., the network does not pinpoint the faulty component but simply detects what *type* of component is faulty.

Spence *et al.* [34] use a different approach to the single fault diagnosis of printed circuit boards (PCB). The difference between the malfunctioning PCB infrared image and the image of a correctly functioning PCB is interpreted by an artificial neural network to diagnose some types of faulty components. In a subsequent work Spence [35] presents a different test method based on the interpretation of the magnetic field close to the PCB. Although these methods can only recognize a limited number of faults, they have the advantage of requiring nonintrusive measurements.

Rutkowski [31] was the first to suggest the use of neural networks for the diagnosis of multiple faults on analog DC circuits. In this introductory work, the main focus is on testing the capability of the network to generalize from single to double fault diagnosis. In the application example presented in the paper, only a limited number of faults are considered.

Bernieri *et al.* [3] use a neural network for *on-line* analysis of dynamic discrete-time systems whose input/output behavior is ruled by equations of the form:

$$y_k = f(y_{k-1}, \cdots, y_{k-n}, u_k, \cdots, u_{k-m}).$$

The network at the $k-$th instant receives as inputs the value of $y_{k-1}, \cdots, y_{k-n}, u_k, \cdots, u_{k-m}$ and is trained to estimate the value of given parameters that rule the behavior of the system. Parameter deviations over a given threshold are symptoms of faults.

Kirkland and Dean [22] have reported using input *current* measurements as circuit images.

Gu *et al.* [17] combine neural networks and expert systems into a single diagnostic system. To each component is associated a neural network trained to recognize the component's fault. The expert system acts as a coordinator between the different neural networks, supplying suitable inputs to the networks and deriving a diagnosis from the analysis of the networks' output.

Spina and Upadhyaya [36] have considered the problem of diagnosing deviation faults in linear circuits. A white noise source is used to automatically generate test patterns.

Fault signatures are generated associating to a single component a value equal to the nominal value plus 50%. The network can correctly classify previously unseen patterns corresponding to deviation faults of different magnitude.

All these works highlight the prominence in a neural diagnostic system of the aspects related to feature extraction and circuit supplies, thus leading us to a systematic exploration of these issues. The present paper summarizes the results that its authors have obtained throughout a long period of time and that have only partially been presented in the papers referenced in the rest of this section.

In [16] is discussed how networks trained to recognize single faults on analog circuits in dynamic behavior may be used to detect multiple faults. The neural network identifies the faulty components from the mean values of the voltage measurements in a given set of test points. In general it was observed that the network is able to diagnose multiple faults on two and three components, although less sharply than in the single fault case, due to the presence of false alarms. The set of multiple faults was chosen among those single faults well recognized by the network.

In [14] Fourier transforms are used as features of the circuit image, and multiple neural networks were used in parallel by the diagnostic system. This improved the performance of the diagnostic system with respect to the previous one.

In [8] Wavelet transforms are used as features. Wavelets proved to be a good data compression technique when the circuit is studied during a transient. In fact, one can increase the number of wavelets only in particular time intervals depending on the degree of approximation required.

In [15] Principal Component Analysis is used in the feature extraction phase. The main advantage of such a technique lies in the fact that it gives a simple automatic procedure to compress the data.

3 Architecture of the diagnostic system

The architecture of the proposed diagnostic system is shown in Figure 1.

Testing procedure (horizontal path)

Given a circuit to diagnose, we apply a suitable power supply and acquire the voltage signals at a given set of test points, constructing the *circuit image.* We extract significant features, as discussed in Section 5, from the image and use them as inputs to a neural network that has been previously trained to recognize single faults on that circuit. The neural network will generate the *candidate set*, i.e., the set of components recognized as faulty.

In Section 4, we will show that to increase the number of detectable faults it is necessary to use different supplies. Consequently, we will have several neural networks, one for each supply considered. Repeating the procedure described above for all supplies, we obtain several candidate sets. These sets will be combined to derive a single diagnosis using suitable algorithms, as described in Section 7.

Training procedure (vertical path)

The diagnostic system is built training the neural networks that will be used in the testing procedure.

Each neural network is trained using a set of patterns corresponding to all possible single-faults, as detailed in Section 5 and 6. The training patterns are constructed from the faulty circuit images using the same feature extraction technique that will be used in the testing.

It may be possible to obtain each faulty circuit image using an acquisition board. One has to produce, one by one, all single faults on the circuit and then has to acquire the corresponding faulty circuit image. This procedure is not practical in many cases. Thus we resorted to PSpice simulation of the circuit behaviour in faulty conditions.

On the contrary, when constructing the circuit image in fault-free condition, both real acquisitions and PSpice simulations are possible. As we will later discuss, several real acquisitions will be used to estimate the magnitude of the measurement noise.

Our results showed that if the circuit PSpice model is accurate enough, there is no difference between a network trained with "simulated" patterns and a network trained with "acquired" patterns. In fact, the distance between a simulated and an acquired pattern has the same order of magnitude of the distance (due to measurement noise and component parameter tolerance) between two patterns acquired during the same fault condition.

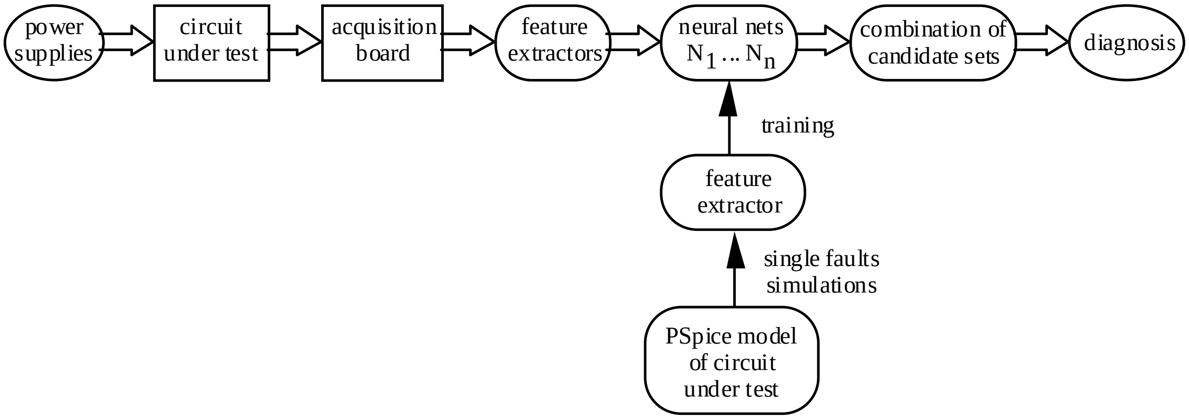


Figure 1: The proposed diagnostic system architecture.

4 Power supplies

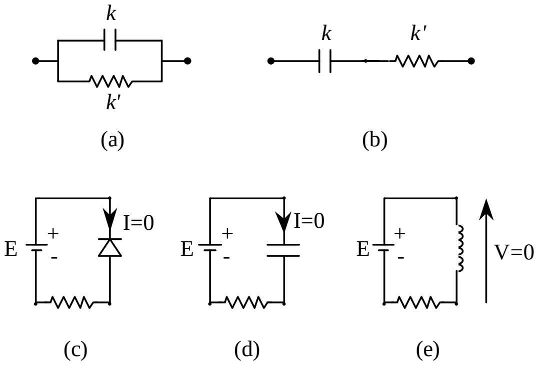
One of the main problems in the diagnosis of circuits is the presence of undistinguishable and undetectable faults.

Consider two (or more) components, say k and k' in *parallel* as in Figure 2.(a). Clearly the behavior of the circuit is the same whenever component k or component k' is short circuited. The same problem appears when we consider open circuit faults of *series* components as in Figure 2.(b). Faults of this kind are called *undistinguishable*, in the sense that they produce the same voltage configuration at the available test points.

A similar problem may arise when a fault is *undetectable*. In this case, the measured behavior of the fault-free circuit is the same as the measured behavior of the faulty circuit.

The presence of undistinguishable and undetectable faults may have different causes.

- *Topology of the circuit*, as in the examples discussed above.
- *Limited number of test points*, that may not allow detection of an abnormal behavior of the circuit.
- Components whose measured behavior is the same when faulty or correctly functioning. We recall some of the possible causes.
  - *Operating point of the component.* Consider the diode in Figure 2.(c). It is



(a) (b)

(c) (d) (e)

Figure 2: Examples of undistinguishable and undetectable faults.

reverse biased and thus for all practical purposes its behavior is the same when the diode is functioning well or when it is affected by an open circuit fault.

- *Frequency content of the supplies.* Some frequency components may not be suitable for highlighting a given fault. In DC steady state, for instance, capacitors behave as open circuits and inductors behave as short circuits, as shown in Figure 2.(d) and Figure 2.(e), respectively.
- *Protection subcircuits.* The behavior of the protection components is not supposed to affect the overall behavior unless other faults are present.

There is little we can do to resolve the ambiguity due to the topology of the circuit or due to the choice of test points. However, we may try to resolve the ambiguity due to the behavior of the circuit by an appropriate choice of power supplies.

As an example, a different choice of supply, such a high frequency square voltage, force the diode in Figure 2.(c) to alternatively switch from reverse to forward bias, and the capacitor and inductance in Figure 2.(d),(e) to work in AC.

This problem has also been discussed by Dague *et al.* [10]. These authors add an external *stimulation* in suitable points so as to disturb the circuit operating conditions.

We will train different networks to process the data collected for each different supply configuration. Thus, our diagnostic system is composed of several neural networks, each one specialized in detecting a given set of faults. When the system is used to diagnose a circuit, each network will produce a set of candidates, i.e., of possibly faulty components. The overall diagnosis can be computed by means of different algorithms, given in Section 7.

5 Feature extraction techniques

We assume that the information on the circuit behavior, i.e., the *circuit image*, is given by the voltage measurements in a set of available test points. These points are usually given by the circuit board manufacturer and cannot be arbitrarily chosen.

Since the voltage signal at each test point is a function of time, we need to extract significant features to compactly represent the circuit behavior. Extensive experimental studies showed the influence of the particular feature chosen. The feature used in [16] was the mean value (MV). The diagnostic system performances improved when root-mean-square values (RMSV) or a combination of MV and RMSV were used. When MV or RMSV are used, all the information on the dynamic behavior is lost. Thus other feature extraction techniques are required. We discuss here four different techniques: *Fourier Transforms*, *Wavelets*, *Principal Components Analysis*, and *Sampling.*

During the *training*, the goal of the feature extraction procedure is to construct an $(s \times r)$ matrix X. Each row of this matrix represents the circuit behavior during one of the s acquisitions and each column represents the value of a particular feature. Each row of X is use as a training pattern input for the neural network, hence there will be r nodes in the network input layer, and s training patterns, as discussed in Section 6.

During the *testing* of a circuit, the same feature extraction procedure is used to derive the inputs that will be given to the neural network.

In this section, we mainly discuss the feature extraction module as used during the training.

Consider a circuit with n components and a given set of m test points.

The voltage of all test points is measured on a *real circuit* by an acquisition board during p acquisitions *in the absence of faults.* These measurements will be used to estimate the magnitude of measurement noise.

On the contrary, the faulty circuit images, i.e., the voltage of all test points in presence of a fault, are constructed via *PSpice simulation*. We consider two single faults for each bipolar component: open circuit and short circuit. We also considered faults on components with more than two terminals. As an example, in the circuit shown in Figure 4, there are trimmers and operational amplifiers. We considered two possible faults on a trimmer (cursor stuck up and cursor stuck down) and just one single fault on an operational amplifier (it was made inoperative by feeding with exceedingly high voltage).

In general, let s' be the number of the single faults taken into account; then one needs to collect $s = p + s'$ circuit images.

5.1 Fourier transforms

A simple technique for compacting the information given by the circuit image without losing the dynamics of the system is given by the Fourier analysis that converts the signals into frequency components [37].

We compute the Fast Fourier Transform (FFT) of the sampled voltage signal measured at each test point. If we have t voltage samples, we obtain — for each test point — $q = t/2$ frequency components and we take the amplitude of each component.

We are now ready to construct the *input pattern matrix*. The matrix has initially s rows, one for each acquisition, and $m \cdot q$ columns, one for each feature, i.e., frequency component computed at each test point. Thus the input pattern matrix takes the form $X' = \{x'_{i,j} \mid 1 \leq i \leq s, 1 \leq j \leq m \cdot q\}$. The first p rows of X' are associated to the fault-free acquisitions.

Matrix X' is still unusable because of its high dimensionality. Domain dependent knowledge may be used to further reduce its number of columns [37].

The data reduction algorithm we propose, requires two phases.

1. *Remove features that give no information.* We compute for each column j the difference Δ_j between its maximum and minimum element. We also compute the difference δ_j between the maximum and minimum element in the first p rows of the column: this is an index of the numerical uncertainty associated to the value of feature j during the p different fault-free acquisitions. Fix a threshold $\theta > 1$. If $\Delta_j \leq \theta \delta_j$ then the variation of the feature j has the same order of magnitude of the numerical uncertainty and column j will be removed. We used a value of $\theta = 10$.

2. *Scale the inputs.* To improve separability between patterns we scale the columns of the input pattern matrix in the interval $[-1, 1]$.

3. *Select a subset of significant features.* The idea is to keep only those columns that are necessary to distinguish between different patterns. Fix a threshold $\sigma < 1$. If $\mid x'_{i,j} - x'_{i',j} \mid \leq \sigma$ then the variation of feature j is not large enough to distinguish pattern i from pattern i'. We used a value of $\sigma = 0.1$.

   We proceed as follows.

```
begin
let the initial set of significant features be $S := \emptyset$;
for $i := 2, s$                    (* compare each row $i$ of $X'$ with all previous ones *)
   for $i' := 1, i - 1$
   begin
   • let $S_{i,i'} := \{j : \mid x'_{i,j} - x'_{i',j} \mid > \sigma\}$ be the set of those features, i.e., columns,
     that may be used to distinguish between patterns $i$ and $i'$;
   • if $S_{i,i'} \cap S = \emptyset$ and $S_{i,i'}$ is not empty then add to $S$ the most significant feature,
     i.e., feature $j$ such that $\mid x'_{i,j} - x'_{i',j} \mid \geq \mid x'_{i,j'} - x'_{i',j'} \mid$ for all $j, j' \in S_{i,i'}$;
   end
end
```

We thus obtain a new matrix X of order $(s \times r)$ with $r \leq m \cdot q$.

The data reduction algorithm we use with FFT falls into the category of *unsupervised* feature extraction methods [4], i.e., methods that do not use information on the target data. Note, however, that the data reduction is performed opportunistically, by projecting the features onto a subspace that still contains all information required to separate the input patterns.

5.2 Wavelets

The origins of Wavelets date back to 1909, when Haar proposed them as a viable solution to function decomposition problems. In fact Fourier series, as stated in its original formulation, show a non-uniform convergence even for particular continuous functions. Wavelets approach is more suitable than Fourier one, especially when signals are non-stationary. Both “time-frequency” and “time-scale” wavelets are suited to signal analysis ranging from "quasi-stationary" to fractal structure type. Mathematicians speak of "atomic decomposition" of signals, where wavelets are the elementary constituents. The various wavelets are obtained from a single wavelet by scaling and shifting operations.

There are several definition of wavelets. One possible is the following [27]: a wavelet is a function $y(x)$ in $L^2(I\!\!R)$ such that $2^{j/2}y(2^j x - k)$, is an orthonormal basis for $L^2(I\!\!R)$. The most frequently used wavelets are the Grossmann-Morlet wavelets, that are also similar to Daubechies wavelets and to Gabor-Malvar wavelets. The last algorithm is of time-frequency type, while the former is a time-scale algorithm.

In the wavelet theory [30, 25] any signal of finite energy can be represented as a linear combination of wavelets whose coefficients represent the features we want to extract, and indicate how close the signal is to a particular basis function.

Discrete wavelet transform (DWT) is a relatively recent method whose biggest potential has been found to be signal compression. The two major advantages of the wavelet transformation are that it can zoom in time discontinuity and that it is possible to construct an orthonormal basis, localized in time and frequency.

An important issue of wavelet analysis is the choice of the proper type of wavelet and of the methodology to use, i.e., time-scale, time-frequency or a combination of the two.

In our diagnostic system, Haar wavelets are chosen to realize data compression of circuit-image information. Decomposition proposed by Haar results as follows:

$$s_n(t) = \langle f, h_0 \rangle \; h_0(t) + \cdots + \langle f, h_n \rangle \; h_n(t)$$

where $\langle f, h_i \rangle = \int_0^1 f(t) \; h_i(t) \; dt$, and $s_n(t)$ is the $n-$th order summation which uniformly converges to the signal $f(t)$, and *Haar wavelets* are defined as:

$$h_n(t) = 2^{j/2}\mathbf{H}(2^j \; t - k); \qquad \begin{array}{l} j \geq 0; \\ 0 \leq k < 2^j; \\ n = 2^j + k. \end{array}$$

Here, the scaling factor is a power of 2, and k defines the time shift with respect to the basic wavelet $\mathbf{H}$, that is the unit square window function. The various wavelets $(n > 0)$ are obtained starting from the basic wavelet $(n = 0)$ by combining, scaling and shifting operations. It is important to note that the time range has to be limited in the $[0, 1]$ interval. This is not limiting because real signals always have a finite time length and this will become the new time unit. It is also possible to realize a suitable time windowing of the signal.

Thus, it is possible to project the time signal onto a set of mutually orthonormal wavelets. The number of the wavelets may be arbitrary, depending on the required approximation in reconstruction or, as in the present case, on the amount of information to extract from the signal.

Because a circuit image results from a set of digital acquisition, signals are not continuous in time, but discrete due to sampling. Hence, a discrete transform has to be used and particular care is required to compute the inner products.

The construction of the input matrix X using wavelets follows the same procedure presented in Section 5.1 for Fourier transforms and will not be repeated here.

5.3 Principal components analysis

Principal Component Analysis (PCA) is another unsupervised feature extraction method. Compression by means of PCA is accomplished by projecting each data vector along the directions of the individual orthonormal eigenvectors of the covariance matrix of data. As the first few eigenvalues of the covariance matrix contain most of the signal energy, the dimensionality of the data can be greatly reduced without losing much information on the input data.

It may happen that the information associated to the discarded PC subspace is important for the subsequent classification phase [4] and in this case PCA is not suitable. However, PCA is a potentially useful method because it works in many applications. In [1] PCA is used for terrain classification, and it is shown that it can lead to a significant improvement in the classifier performance. In [6] there is a comparison between Gabor filters and PCA as feature extraction methodologies applied to SAR images segmentation with neural networks.

Let s be the number of the circuit behaviors taken into account, and t be the number of samples for each test point voltage. Each circuit image is represented by $m \cdot t$ values. We have a $(s \times m \cdot t)$ data matrix X' which could be used as input for the neural network.

As previously mentioned, preprocessing is necessary to extract from these data the salient features. We would like to reduce the number of columns of this matrix from $m \cdot t$ to $r \ll m \cdot t$, with acceptable loss of information. Using PCA [20] this compression is accomplished projecting the s circuit images along the directions of the principal eigenvectors of the covariance matrix of $(X')^T$.

Given the data matrix $(X')^T$, whose $i-th$ column $\vec{x}_i$ $(i = 1, \cdots, s)$ represents a circuit image, the covariance matrix of these data is the $(m \cdot t \times m \cdot t)$ matrix

$$C = \sum_{i=1}^{s} \vec{x}_i \cdot \vec{x}_i^T - \frac{1}{s^2} \left(\sum_{i=1}^{s} \vec{x}_i \right) \cdot \left(\sum_{i=1}^{s} \vec{x}_i^T \right)$$

The eigenvectors of this matrix form an orthonormal basis, and any vector $\vec{x}_i$ can be represented with respect to this basis by means of a coefficient vector with $m \cdot t$ elements.

To reduce the data dimension, it is possible to consider only those eigenvectors associated to the dominant eigenvalues of C. Fix a threshold $c \in [0,1]$, and let $\{\lambda_1, \lambda_2, \cdots, \lambda_{m \cdot t}\}$ be the ordered set of eigenvalues of C, i.e., $\lambda_j \leq \lambda_{j+1}$. We say that there are r dominant eigenvalues if $\frac{\sum_{j>r} \lambda_j}{\sum_j \lambda_j} < c$. If $\vec{v}_1, \cdots, \vec{v}_r$, are the eigenvectors associated to the dominant eigenvalues, we may use as compressed representation of a vector $\vec{x}_i$ the coefficient vector: $\vec{a}_i = \left(\vec{x}_i^T \cdot \vec{v}_1, \cdots, \vec{x}_i^T \cdot \vec{v}_r \right)^T$. We used a value of $c = 0.999$.

Thus, the data matrix X' is reduced to a $(s \times r)$ matrix X. The same compression technique will be used on subsequent circuit images acquired during the test phase.

5.4 Sampling

Given the circuit image (i.e., the sampled voltage signals at all test points) one may compact the data retaining just a limited number q of the t samples.

Experimental results [16] showed that this is not a viable technique if the circuit is in AC steady-state or if there are many test points. In fact, this leads to a neural network with too many nodes in the input layer, i.e., too many features. This may reduce the performance of the classification system and leads to a higher computational cost of the training.

However, this technique was effective when studying short transients on circuits with a limited number of test points. The choice of the samples to retain must be opportunistic, and depends on the signal variation pattern.

6 Neural model

As proposed in most of the literature discussed in Section 2, we use a three level neural network with sigmoid activation functions and backpropagation learning with generalized delta rule.

6.1 Fault coding

The network has r input nodes, i.e., as many as there are columns in the input pattern matrix X derived with any of the different feature extraction procedures previously described. The output nodes of the network are as many as the number of circuit components n.

We construct the s input-output patterns that will be used to train the neural network for the diagnosis of the circuit as follows. Each pattern is given by a pair $(\vec{x}_i, \vec{y}_i)$. The vector $\vec{x}_i$ is the $i-th$ row of matrix X while the associated vector $\vec{y}_i$ is defined as follows:

$$y_i(k) = \begin{cases} 0 & \text{if component } k \text{ is not faulty} \\ & \text{during the } i-th \text{ acquisition;} \\ 1 & \text{if component } k \text{ is faulty} \\ & \text{during the } i-th \text{ acquisition.} \end{cases}$$

This general scheme must be altered to take into account undistinguishable faults.

Topologically undistinguishable faults are easy to deal with. From an inspection of the circuit a list of all sets of parallel components is made. Then, a single short circuit fault acquisition for each set $\mathcal{C}_i$ of parallel components is considered. There will be a single training pattern $(\vec{x}_i, \vec{y}_i)$ for such a fault. The vector $\vec{y}_i$ is such that $y_i(k) = 1$ for all $k \in \mathcal{C}_i$, while all other components have a 0 value. A dual procedure takes care of sets of series components.

Two faults i and i' are *behaviorally* undistinguishable if

$$\| \vec{x}_i - \vec{x}_{i'} \|_\infty = \max_j | x_{i,j} - x_{i',j} | \leq \sigma,$$

where σ is the threshold introduced in section 5.1. A fault i is behaviorally undetectable if the condition $\| \ \vec{x}_i - \vec{x}_0 \|_\infty \leq \sigma$ is satisfied for all input vectors $\vec{x}_0$ obtained in the faulty free condition. We combine the patterns of behaviorally undistinguishable faults (as we did for topologically undistinguishable faults) and remove from the training set the patterns associated to undetectable faults.

The fault coding here described is different from the one presented in [16], that defined the vector $\vec{y}_i$ as follows:

$$y_i(k) = \begin{cases} 0 & \text{if component } k \text{ is short circuited} \\ & \text{during the } i-th \text{ acquisition;} \\ 0.5 & \text{if component } k \text{ is not faulty} \\ & \text{during the } i-th \text{ acquisition;} \\ 1 & \text{if component } k \text{ is open circuited} \\ & \text{during the } i-th \text{ acquisition.} \end{cases}$$

The new coding gives sharper identification of the faulty component and is more robust when diagnosing multiple faults because the values of interest (0 and 1) are obtained by "pushing" the sigmoid function toward saturation. Note also that there is a difference with the coding in [31] where each output node is associated to a catastrophic fault and not to a component.

Once the net has been trained, it may be used to perform the diagnosis of the circuit. The net must be given the features extracted from the measured test point voltages as input vector $\vec{x}$. The net will produce an output vector $\vec{y}$; a value of $y(k)$ close to 1 will pinpoint a fault of component k; a value close to 0 will denote that the component is correctly functioning.

Although the net has been trained with the results of single fault acquisitions, it is potentially able to diagnose multiple faults. In this case, two or more elements of $\vec{y}$ will be close to 1.

6.2 Network structure and training

The basic architecture we used consists of a three layers backpropagation network. Since the input patterns have been preprocessed to eliminate undistinguishable faults, and thus they are separable, we are sure that eventually there will be a network capable of correctly learning all patterns.

We use *early-stopping* [4] to avoid overfitting. This consists in measuring, during the training, the error with respect to an independent set of patterns, called *validation set*, and in stopping the training when this error reaches a minimum.

Caruana [5] has shown that if early-stopping is used the number of nodes in the hidden layer may vary without appreciably affecting the performance of a neural network, provided it is sufficiently large. The results of our simulations, not reported in this paper, seem to confirm this general rule.

The validation set used for the stopping is independent from the training set. We construct it by performing a new set of PSpice simulations (one for each fault) randomly changing the parameter values of the components within their tolerance range and by adding to the voltage signals of the test points a noise whose magnitude is equivalent to the measurement noise observed during the p fault-free acquisitions.

7 Combining different diagnosis

In the diagnosis of circuits, we have underlined the importance of using more than one power supply. In fact, it is often the case that a given supply can only lead to the detection of a particular subset of all possible faults. The use of different supplies leads to the use of several neural networks N_i, each of which produces its own candidate set A_i. The final diagnosis must be computed combining these sets of candidates.

The combination of neural networks is a problem that has been discussed in the literature and is reviewed in [32]. In particular, since we use neural networks that are all trained on the same task, our approach falls into the *ensemble* (or *committee*) framework [4, 32].

It is clear that the "union" of two sets of candidates magnifies the influence of false alarms, while the "intersection" can be used to filter false alarms at the risk of removing some faulty components from the diagnosis. Keeping this in mind, we propose two different ensemble algorithms.

Let us first give the following definitions. For each candidate k let $v(k)$ be the number of votes it receives, i.e., the number of nets that consider k malfunctioning, and let $\bar{v} = \max_k v(k)$. We consider all non-empty intersections of $\bar{v}$ candidate sets; assume there are α of such intersections and denote them $\mathcal{R}_u$, with $u = 1, \cdots, \alpha$. We also define $\bar{u}$ the index of the intersection $\mathcal{R}_u$ with the smallest cardinality (should there be more than one such intersection we randomly pick one).

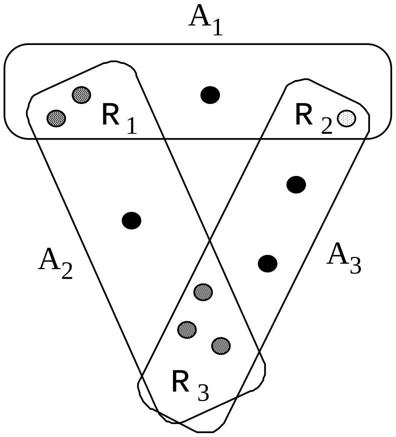


Figure 3: Example of diagnosis combination.

Algorithm 1

The first algorithm considers as faulty all those candidates that have received the highest number of votes. The corresponding diagnosis is:

$$\mathcal{D}_1 := \bigcup_{u=1}^{\alpha} \mathcal{R}_u$$

Algorithm 2

The second algorithm considers as faulty all those candidates that have received the highest number of votes and that belong to the intersection with the smallest cardinality. By considering only the smallest intersection we hope to filter out some false alarms. The corresponding diagnosis is:

$$\mathcal{D}_2 := \mathcal{R}_{\bar{u}}$$

An example is shown in Figure 3. Here $\mathcal{D}_1 = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3$, while $\mathcal{D}_2 = \mathcal{R}_2$.

Note that these algorithms do not give different weight to the candidate sets of each network, but simply perform boolean operations on these sets. We are currently investigating the possibility of associating to each candidate set a different weight, depending on how the network has learned to recognize the single fault on each component that belongs to the candidate set.

8 Experimental results

We discuss the results obtained by the different diagnostic strategies presented in this paper. Two circuits are studied: a DC motor drive board, and an astable multivibrator.

Training

As discussed above, we use early stopping, hence we need both a training and a validation set of patterns.

The training patterns corresponding to each single fault condition are constructed using a PSpice model of the circuit. This choice gives patterns corresponding to a circuit where the component parameters have nominal values and the voltage signals in each test point are noise free.

The validation set is constructed by performing a new set of PSpice simulations where parameter tolerance and measurement noise are introduced.

Testing

During the test phase, we consider a real circuit and the different faults are implemented by manually shortcircuiting or opening each component terminals. The circuit measurements are collected through a National Instrument Corporation AT-MIO-16E-1 acquisition board.

Thus the test patterns are determined independently of the training patterns. Furthermore, the test patterns are affected by measurement noise and by the error due to the parameter tolerance of the circuit components.

When diagnosing a circuit, we observe the network output corresponding to the input pattern derived from the measurements. Let us recall that the network output layer has as many nodes as there are components. During the training phase we have coded a fault on component k assigning a value 1 to the corresponding output node, while a value 0 was assigned to the output node of a fault-free component.

In general, during the test phase the value of each output node may take any value between 0 and 1. A value close to 0 (1) of an output node will be interpreted as the

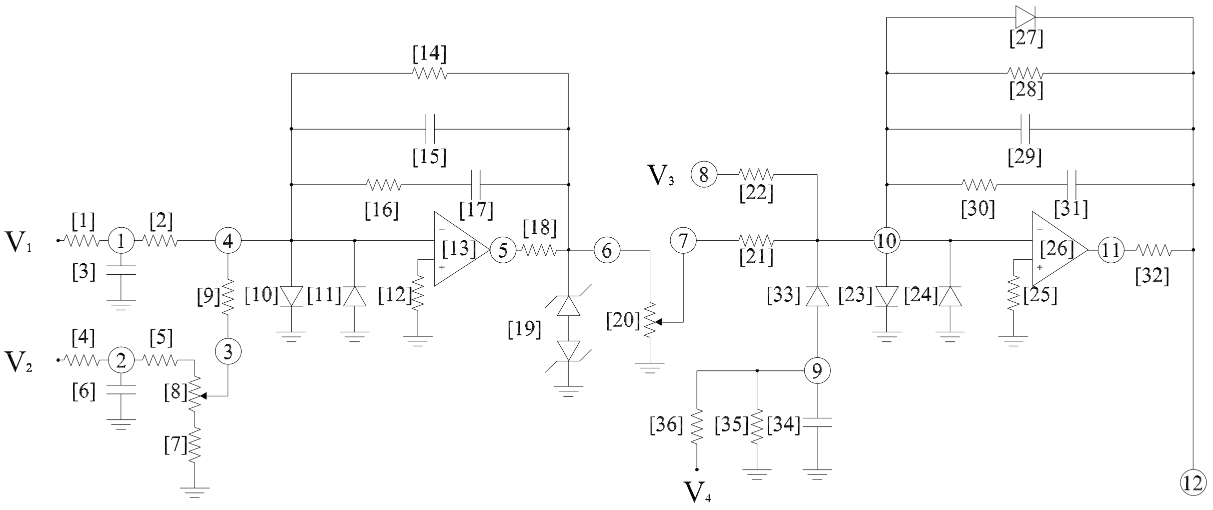


Figure 4: DC motor drive board.

absence (presence) of a fault on the corresponding component. Threshold values need to be set to discriminate between these two cases.

Let v_{max} be the maximum value of all output nodes. If $v_{max} < 0.2$ we consider the circuit as fault-free and the candidate set will be empty. If $v_{max} \geq 0.2$ we consider the circuit as faulty, and the candidate set will contains all components whose corresponding output node has a value greater than $0.5v_{max}$.

8.1 DC motor drive board

We present the results obtained diagnosing the circuit in Figure 4, part of a DC motor drive. The same circuit has also been diagnosed in [8, 13, 14, 15, 16]. In the figure, the $m = 12$ test points are marked by numbers within circles, while the $n = 36$ components are labeled by numbers in square brackets.

Training

There are 70 single faults to consider on this circuit. In fact, the circuit is composed of 36 components but only one fault is considered for each of the two operational amplifiers.

Thus, the overall training set should consist of 76 training patterns — the additional six being obtained by acquisitions of the circuit behavior in absence of fault.

The following sets contain topologically undistinguishable faults: 10/11s = {10s,11s}, 14/15s = {14s,15s}, 23/24s = {23s,24s}, 27/28/29s = {27s, 28s, 29s}, 34/35s = {34s, 35s}, 16/17o = {16o, 17o}, 30/31o ={30o, 31o}. Here 10s represents a short circuit fault on component 10, 16o represents an open circuit fault on component 16, etc. Thus, the training set is reduced to 62+6 patterns by combining the conflicting patterns as discussed in Section 6.1.

We have used three different voltage supplies and thus three different networks.

1. The first network N_1 is trained with patterns acquired when the circuit has close to nominal voltage supplies: $V_1(t) = 3\sin(2\pi t) + 1.25$ (V); $V_2(t) = -3$ (V); $V_3(t) = 0.25$ (V); $V_4(t) = -12$ (V).

   Fourier The number of significant frequency components is $q = 50$. This gives rise to $m \cdot q = 600$ columns in the input matrix X', that are reduced to $r = 14$ in the matrix X.

   The set of behaviorally undetectable faults for this net is: {3o, 6o, 11o,12s, 15o, 23o, 24o, 25s, 27o, 29o, 31s, 32s, 33o, 34o}.

   The sets of behaviorally undistinguishable faults are: {4o, 6s}, {21o, 27/28/29s, 30s}, {34/35s, 36o}, { 35o, 36s}.

   Wavelets The number of significant wavelets is $q = 8$. This gives rise to $m \cdot q = 96$ columns in the input matrix X', that are reduced to $r = 15$ in the matrix X.

   The set of behaviorally undetectable faults for this net is: {3o, 6o, 11o, 12s, 15o, 23o, 24o, 25s, 27o, 29o, 31s, 32s, 33o, 34o}.

   The sets of behaviorally undistinguishable faults are: {21o, 27/28/29s, 30s}, {34/35s, 36o}, {35o, 36s}.

   PCA Assuming a threshold $c = 0.999$, the number of dominant eigenvalues (i.e., the number of columns of the X matrix) is $r = 25$.

   The set of behaviorally undetectable faults for this net is: {3o, 6o, 11o, 12s, 15o, 23o, 24o, 25s, 27o, 29o, 31s, 32s, 33o, 34o}.

   The sets of behaviorally undistinguishable faults are: {4o, 6s}, {21o, 27/28/29s, 30s}, {34/35s, 36o}, {35o, 36s}.

2. The second network N_2 is trained with patterns acquired when the circuit has far from nominal periodic voltage supplies: $V_1(t), V_2(t)$, and $V_4(t)$ are zero-mean square waves with 160 Hz frequency, 4 V peak-to-peak amplitude; $V_3(t) = 0.25$ (V).

   Fourier The number of significant frequency components is $q = 8$. This gives rise to $m \cdot q = 96$ columns in the input matrix X', that are reduced to $r = 15$ in the matrix X.

   The set of behaviorally undetectable faults for this net is: {10o, 11o, 12s, 21o, 22o, 23o, 24o, 25s, 27/28/29s, 28o, 29o, 30s, 31s, 32s}.

   The sets of behaviorally undistinguishable faults are: {1o, 3s}, {4o, 6s}, {34/35s, 36o}.

   Wavelets The number of significant wavelets is $q = 9$. This gives rise to $m \cdot q = 108$ columns in the input matrix X', that are reduced to $r = 16$ in the matrix X.

   The set of behaviorally undetectable faults for this net is: {8_{up}, 10o, 11o, 12s, 14o, 17s, 21o, 22o, 23o, 24o, 25s, 27/28/29s, 28o, 29o, 30s, 31s, 32s}.

   The sets of behaviorally undistinguishable faults are: {4o, 6s}, {14/15s, 16s}, {18o, 19o}, {34/35s, 36o}.

   PCA Assuming a threshold $c = 0.999$, the number of dominant eigenvalues is $r = 31$.

   The set of behaviorally undetectable faults for this net is: {10o, 11o, 12s, 21o, 22o, 23o, 24o, 25s, 27/28/29s, 28o, 29o, 30s, 31s, 32s, 34/35s, 36o}.

   The set of behaviorally undistinguishable faults is: {1o, 3s}.

3. The third network N_3 is trained with patterns acquired when the circuit has step voltage supplies: $V_1(t) = V_4(t) = \delta_{-1}(t)$ (V); $V_2(t) = V_3(t) = -2\delta_{-1}(t) + 4\delta_{-1}(t - 0.01)$ (V).

   Fourier The number of significant frequency components is $q = 8$. This gives rise to $m \cdot q = 96$ columns in the input matrix X', that are reduced to $r = 15$ in the matrix X.

   The set of behaviorally undetectable faults for this net is: {10o, 12s, 14o, 21o, 23o, 25s, 27o, 28o, 29o, 31s, 32s, 34o}.

   The sets of behaviorally undistinguishable faults are: {1o, 3s}, {4o, 6s}, {22o, 27/28/29s, 30s}.

Wavelets The number of significant wavelets is $q = 32$. This gives rise to $m{\cdot}q = 384$ columns in the input matrix X', that are reduced to $r = 17$ in the matrix X.

The set of behaviorally undetectable faults for this net is:{10o, 12s, 21o, 23o, 25s, 27o, 28o, 29o, 31s, 32s, 34o}.

The sets of behaviorally undistinguishable faults are: {1o, 3s}, {4o, 6s}, {22o, 27/28/29s, 30s}.

PCA Assuming a threshold $c = 0.999$, the number of dominant eigenvalues is $r = 32$.

The set of behaviorally undetectable faults for this net is: {10o, 12s, 21o, 23o, 25s, 27o, 28o, 29o, 31s, 32s}.

The sets of behaviorally undistinguishable faults are: {1o, 3s}, {4o, 6s}, {22o, 27/28/29s, 30s}.

Testing

We are now ready to study the performance of the neural diagnostic systems previously constructed.

In the initial phase, we test the systems on a fault-free circuit. We observed that when diagnosing a real circuit in absence of faults, all networks correctly identify this behavior, in the sense that all output nodes have a value less than the assigned threshold of 0.2 and thus the candidate set is always empty.

In a second phase, we consider faulty circuits. Table 1 compares the performance (in percent) of the different systems. The first columns of the table shows the diagnosis of N_1, N_2, and N_3 and the diagnosis obtained combining the candidate sets of the three nets with Algorithm 1 and Algorithm 2, using Fourier, Wavelets, and PCA, respectively. The last two columns show the results obtained combining the candidate sets of the nine nets (three for each feature extraction technique) with Algorithm 1 and Algorithm 2.

We consider a fault correctly diagnosed if the candidate set of the net contains a subset of the components associated to this fault, taking into account topologically undistinguishable fault classes. Let us consider some examples in the circuit of Figure 4. The fault 16o belongs to the topologically indistinguishable fault class 16/17o; we say that it is correctly identified if the candidate set is either {16} or {17} or {16, 17}. The fault 16s is correctly identified if the candidate set is {16}.

Single faults

The first row block of Table 1 shows the diagnosis of the 62 possible single faults. There are three different classes of diagnosis.

Class $\mathbf{A}_1$ Faults correctly diagnosed.

Class $\mathbf{B}_1$ Undistinguishable faults: these are the faults that we have classified as behaviorally undistinguishable during the training. As an example, in net N_1 with Fourier, we have identified 9 undistinguishable faults, i.e., 14% of the total 62 faults.

Class $\mathbf{C}_1$ Undetected faults: these are the faults that we have classified as behaviorally undetectable during the training.

Double faults

The second row block of Table 1 shows the performance of the different systems when diagnosing double faults. Each double fault consists in the simultaneous presence of two faulty components. Note that not all possible pairs of single faults constitute a double fault: e.g., a bipolar component cannot be simultaneously open- and short-circuited. We have considered a sample of 168 different double faults randomly chosen from the total population. This sample was large enough to satisfy the χ^2 test for the six different classes of diagnosis.

These are the classes considered.

Class $\mathbf{A}_2$ Both faults correctly diagnosed.

Class $\mathbf{B}_2$ Only one fault correctly diagnosed.

Class $\mathbf{C}_2$ At least one fault correctly diagnosed with one or two false alarms.

Class $\mathbf{D}_2$ At least one fault correctly diagnosed with more than two false alarms.

Class $\mathbf{E}_2$ Empty candidate set.

Class $\mathbf{F}_2$ Only false alarms.

Triple faults

The third row block of Table 1 shows the performance of the different systems when diagnosing triple faults. We have considered a sample of 181 different faults randomly chosen out of the total population.

These are the classes considered.

Class $\mathbf{A}_3$ All three faults correctly diagnosed.

Class $\mathbf{B}_3$ Only one or two faults correctly diagnosed.

Class $\mathbf{C}_3$ At least one fault correctly diagnosed with one or two false alarms.

Class $\mathbf{D}_3$ At least one fault correctly diagnosed with more than two false alarms.

Class $\mathbf{E}_3$ Empty candidate set.

Class $\mathbf{F}_3$ Only false alarms.

Discussion

In the case of multiple faults, we consider correct all diagnoses in class A and in class B. In fact, starting from class B we may use an incremental repair procedure, substituting the faulty components one by one. Diagnosis in class C may also be useful.

From the table it can be seen that the use of several networks improves the system performance provided that a good procedure is used to combine the results of the networks. In particular, Algorithm 1 and Algorithm 2 give the same results when diagnosing: (a) single faults; (b) multiple faults using a system composed of many nets in parallel. When diagnosis multiple faults, if the system is composed by a small number of neural nets Algorithm 2 performs better because it exalts the filtering effect of the intersection operator, reducing the number of diagnoses in class A but increasing the total number of diagnoses in class A+B.

All three feature extraction techniques give comparable results. PCA performs better than the other two when diagnosing single and double faults, but seems to be less robust when diagnosing three simultaneous faults. Unlike Fourier and Wavelets, PCA requires less data preprocessing in the feature extraction phase, as discussed in Section 5.

	Fourier					Wavelets					PCA					F+W+P	
	N_1	N_2	N_3	$\mathcal{D}_1$	$\mathcal{D}_2$	N_1	N_2	N_3	$\mathcal{D}_1$	$\mathcal{D}_2$	N_1	N_2	N_3	$\mathcal{D}_1$	$\mathcal{D}_2$	$\mathcal{D}_1$	$\mathcal{D}_2$
Single faults (62 fault cases)																	
A_1	63	68	69	82	82	66	60	71	84	84	63	71	73	87	87	90	90
B_1	14	10	12	8	8	11	13	11	6	6	14	3	11	5	5	2	2
C_1	23	22	19	10	10	23	27	18	10	10	23	26	16	8	8	8	8
Double faults (168 fault cases)																	
A_2	1	9	6	5	3	1	6	3	5	2	5	10	11	8	4	5	5
B_2	51	51	50	68	74	52	44	24	74	81	45	50	41	75	83	83	83
C_2	19	19	23	22	10	19	25	28	14	5	26	18	24	10	4	6	6
D_2	5	4	7	0	0	5	5	5	0	0	8	3	11	0	0	0	0
E_2	6	7	4	1	2	5	9	1	1	2	5	7	2	1	1	1	1
F_2	18	10	10	4	11	18	11	39	6	10	11	12	11	6	8	5	5
Triple faults (181 fault cases)																	
A_3	0	3	2	6	2	1	2	3	4	3	2	2	2	6	3	7	7
B_3	28	19	21	62	74	23	21	23	60	71	19	16	13	58	67	76	76
C_3	19	31	31	25	9	16	30	28	27	6	32	26	32	22	8	9	9
D_3	5	10	5	0	0	3	8	5	0	0	10	7	15	0	0	0	0
E_3	2	3	2	0	1	3	3	1	0	1	2	3	1	0	1	0	0
F_3	46	34	39	7	14	54	36	40	9	19	35	46	36	14	21	8	8

Table 1: Diagnosis of the circuit in Figure 4 (in percent).

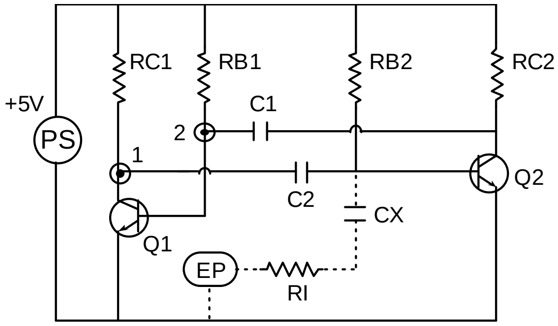


Figure 5: Astable multivibrator. PS is the nominal power supply; EP generates the stimulation voltage pulse.

8.2 Astable multivibrator

Dague et al. in [10] remarked that oscillators are difficult to diagnose because most faults cause the same type of symptoms. This is exactly the case in which a proper choice of the power supplies can improve the diagnosability of the circuit. They proposed using an external “stimulation” and showed the results obtained using their diagnostic expert system on the astable multivibrator shown in Figure 5. In this section we present results obtained using our diagnostic system on the same circuit.

Training

We chose $m = 2$ test-points in the nodes labeled 1 and 2 in the figure. The number of components is $n = 8$ and 24 single faults have been considered. In fact, we consider 6 faults for each transistor: short circuit between base and emitter (QBEs), base and collector (QBCs),collector and emitter (QCEs); open circuit on the base (QBo), collector (QCo), and emitter (QEo). The circuit does not contain topologically undistinguishable faults.

To be able to compare the results of our diagnostic system with the system developed by Dague, we used the same voltage supply proposed in [10]. It consists of the superposition of the nominal supply PS (a continuous voltage signal of +5V) and of an external stimulation EP (a voltage pulse of $10V$ amplitude, applied in $t = 1\mu s$ and lasting $1\mu s$). Since we consider a unique supply, we use a single neural network for each feature extraction.

	Fourier	Wavelets	PCA	Sampling
Single faults (24 fault cases)				
A_1	83	79	83	75
B_1	17	21	17	25
C_1	0	0	0	0
Double faults (140 fault cases)				
A_2	1	0	5	6
B_2	28	8	23	41
C_2	18	50	36	26
D_2	0	0	0	0
E_2	4	1	0	1
F_2	49	41	36	26

Table 2: Diagnosis of the circuit in Figure 5 (in percent).

We used a PSpice model of the oscillator to collect the training and validation patterns for all faulty conditions, as previously described.

The circuit has been studied in transient behaviour and the voltage signals in each test point have been collected in the interval $0.9 \div 40\mu s$ with a sampling interval of $t_s = 0.1\mu s$. This gives rise to a circuit image before feature extraction composed of $t = 392$ samples for each test point.

We have used all different feature extraction techniques described in Section 5.

Fourier The number of significant frequency components is $q = t/2 = 196$. This gives rise to $m \cdot q = 392$ columns in the input matrix X', that are reduced to $r = 8$ in the matrix X.

There are no behaviorally undetectable faults. The sets of behaviorally undistinguishable faults are: {RC2s, Q2CEs}, {RB2s, Q2BEs}.

Wavelets The number of significant wavelets is $q = 128$. This gives rise to $m \cdot q = 256$ columns in the input matrix X', that are reduced to $r = 12$ in the matrix X.

There are no behaviorally undetectable faults. The sets of behaviorally undistinguishable faults are: {C1o, RC2s, RB2s, Q2CEs, Q2BEs}.

PCA Assuming a threshold $c = 0.999$, the number of dominant eigenvalues (i.e., the number of columns of the X matrix) is $r = 16$.

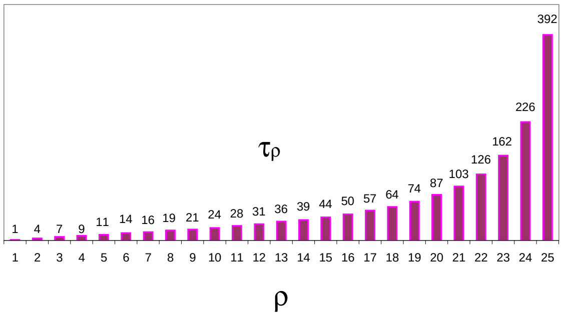


Figure 6: Sampling of the total measurements.

There are no behaviorally undetectable faults. The sets of behaviorally undistinguishable faults are: {RC2s, Q2CEs}, {RB2s, Q2BEs}.

Sampling The Sampling feature extraction retains $q = 25$ samples (out of a total of $t = 392$) for each test point spaced with an hyperbolic law so as to have more samples during the initial phase of the transient and only a few as the circuit reaches the steady state. For $\rho = 1, \cdots q$, let

$$\tau_\rho = 1 + (t-1)\left[\frac{1}{q-1}\left(\frac{q}{q+1-\rho} - 1\right)\right]^{0.75}$$

Then the ρ-th retained sample is the τ_ρ-th sample, as shown in Figure 6.

There are no behaviorally undetectable faults. The sets of behaviorally undistinguishable faults are: {C1o, RC2o}, {RC2s, Q2CEs}, {RB2s, Q2BEs}.

Testing

In a first phase, we test the systems on a fault-free circuit. We observed that when diagnosing a real circuit in absence of faults, all networks correctly identify this behavior.

In a second phase, we consider faulty circuits. Table 2 compares the performances (in percent) of the different neural diagnostic systems. The classes of diagnosis are the same defined in the previous example.

The first row block of Table 2 shows the diagnosis of the 24 single faults considered.

The second row block of Table 2 shows the diagnosis of the 140 double faults considered. Note that in this case the total number of possible double faults is 240.

We can see that in this particular case Sampling appears to be the most effective feature extraction technique (and it is also the easiest to implement). As remarked before, however, it is a viable solution only because we have a small number of test points (two in this example) and the circuit has a short transient.

Comparison with Dague's Expert System

In Table 3 we compare the results obtained using our diagnostic system with sampling feature extraction (Neural Network with Sampling) with the results obtained by Dague's expert system. Note that although only 12 faults are considered by Dague, our diagnostic system has been trained to recognize all 24 single faults.

We observe that the neural network has been able to correctly classify 11 faults in class A and only 1 fault (C1o) in class B. The expert system, on the contrary, can very rarely correctly identify the faulty component (only 3 diagnosis in class A including the diagnosis of RB1s).

The results show that the neural network performs much better than the expert system. This is a consequence of its ability to exploit the information on each single-fault behavior of that particular circuit and to generalize. This information is not taken into account by the expert system, that reasons on more abstract principles.

9 Conclusions

We have shown how a neural network, trained to recognize catastrophic single faults, may be used to diagnose multiple faults on analog circuits.

In general we observe that the network is almost always able to learn and recall the single fault patterns presented during the training. Multiple faults on two and three components may also be diagnosed, although less sharply than in the single fault case, due to the presence of false alarms. In most cases, however, the network is able to detect at least one of the malfunctioning components. Thus one may use an incremental repair procedure, substituting the faulty components one by one.

We consider several different power supplies in order to detect those faults that do not modify the circuit behavior under nominal supplies. We use several neural networks "in parallel", one for each different supply configuration. Each network is specialized in

Defect	Expert System (Dague et al.)	Neural Network with Sampling
RC1s	RC1, Q1, {C2} × {PS, CX, C1, EP, Q2, RB1, RB2, RC2, RI}	RC1
RB1s (*)	Q1	RB1
C1s	C1 + 10 double candidates	C1
Q1CEs	Q1, C2, {RC1} × {CX, C1, EP, Q2, RB2, RC2, RI}	Q1
Q1BEs	Q1, C1, RB1	Q1
Q1BCs	Q1, C1, RB1, RC1, {C2} × {PS, EP, Q2, RB2, RC2, RI}	Q1
RC1o	RC1, Q1, C2	RC1
RB1o	RB1, Q1, C1	RB1
C1o	C1, Q2, RC2	{C1, RC2}
Q1Eo	Q1	Q1
Q1Bo	Q1	Q1
Q1Co	Q1, C1, RB1, C2, RC1	Q1

(*) Note that a short-circuit on RB1 induces destruction of Q1.

Table 3: A comparison between the diagnosis of the circuit in Figure 5 done with Dague's Expert System and with Neural Network with Sampling.

detecting a given set of faults. Thus, it is not necessary to force a network to recognize a fault that is more easily detected by another one.

The use of different networks, leads to the problem of composing different sets of candidates into a single diagnosis. We showed that a suitable choice of the composition algorithm may dramatically improve the system performance, especially when diagnosing multiple faults.

We compared the results obtained by our system when using different feature extraction techniques. In fact, the performance of the diagnostic system is noticeably affected by the choice of features that we consider as representative of the device behavior.

Although we have only presented two simple examples of diagnosis, extensive experiments convinced us that this approach is fairly general and that it gives better results than other diagnostic systems, such as expert systems, whenever it can be applied.

References

[1] Azimi-Sadjadi M.R., Ghaloum S., Zoughi R. Terrain Classification in SAR Images Using Principal Component Analysis and Neural Networks. IEEE Trans. on Geoscience and Remote Sensing, March 1993, 31(2):511–515.

[2] Bandler J.W., Salama A.E. Fault Diagnosis of Analog Circuits. Proc. IEEE 1985; 73(8):1279–1325.

[3] Bernieri A., D'Apuzzo M., Sansone L., Savastano M. A Neural Network Approach for Identification and Fault Diagnosis of Dynamic Systems. In: Conf. Rec IMTC93 (Irvine, California), May 1993, pp 564–569.

[4] Bishop C.M. Neural Networks for Pattern Recognition. Oxford, Clarendon Press, 1995.

[5] Caruana R. Extra capacity rarely hurts generalization if yo u use early stopping. Presented at NIPS 96 workshop on Model Complexity. December, 1996.

[6] Ceccarelli M., Farina A., Petrosino A., Vaccaro R., Vinelli F. SAR Image Segmentation Using Textural Information and Neural Classifiers. L'Onde Electronique. May-June 1994, 74(3):511–515.

[7] Console L., Torasso P. A Spectrum of Logical Definitions of Model-Based Diagnosis. Computational Intelligence 1991; 7(3):133–141.

[8] Contu S., Fanni A., Marchesi M., Montisci A., Serri A. Wavelet Analysis for Diagnostic Problems. In: Proc. 8th MELECON, Bari, Italy, May 1996, pp 1571–1574.

[9] Craven M.W. Extracting Comprehensible Models from Trained Neural Networks. Ph.D. Thesis, Dept. of Computer Sciences, University of Wisconsin-Madison, 1996.

[10] Dague Ph., Jehl O., Devès Ph. Luciani P., Taillibert P. When Oscillators Stop Oscillating. In: Int. Joint Conf. on Artificial Intelligence, Sydney, Australia, August 1991, pp 1109–1115.

[11] deKleer J., Williams B.B. Diagnosing Multiple Faults. Artificial Intelligence 1987; 32:97–129.

[12] Duda R.O., Hart P.E. Pattern Classification and Scene Analysis. John Wiley & Sons, 1973.

[13] Fanni A., Diana P., Giua A., Perezzani M. Qualitative Dynamic Diagnosis of Circuits. Artificial Intelligence. for Engineering Design, Analysis and Manufacturing 1993; 7(1):53–64

[14] Fanni A., Giua A., Micheli F., Montisci A. A Multiple Neural Network Diagnostic System for Analog Circuits Based on Fourier Transforms. In: Proc. 5th Int. Work. on Principles of Diagnosis, New Paltz, New York, October 1994, pp 98–105.

[15] Fanni A., Giua A., Montisci A. Diagnosis of Electrical Circuits Using Neural Networks and Principal Components Analysis. In: Proc. Int. Conf. on Engineering Applications of Neural Networks, Helsinki, Finland, August 1995, pp 629–632.

[16] Fanni A., Giua A., Sandoli E., Neural Networks for Multiple Fault Diagnosis in Analog Circuits. In: Proc. IEEE Int. Work. on Defect and Fault Tolerance in VLSI Systems, Venezia, Italy, October 1993, pp 303–310.

[17] Gu X.P., Yang Y.H., Zang W.Q., Gao S. Integration of Artificial Neural Networks and Expert Systems for Power System Fault Diagnosis. In: Proc. IPEC '95, Singapore, February 1995.

[18] Hertz J., Krogh A., Palmer R.G. Introduction to the Theory of Neural Computation. Santa Fe Institute Studies in the Science of Complexity, Lecture Notes, Vol. 1, Addison-Wesley, 1991.

[19] Hochwald W., Bastian J.D. A DC Approach for Analogue Fault Dictionary Determination. IEEE Trans. Circuits and Systems 1979; 26(7):523–529.

[20] Jain A. K. Fundamentals of Digital Image Processing. Englewood Cliffs, NJ, Prentice-Hall, 1989.

[21] Keagle B.J., Murphy J.H., Koos L.J., Reeder J.R. Multi-Fault Diagnosis of Electronic Circuit Boards Using Neural Networks. In: Proc. Int. Joint Conf. on Neural Networks, San Diego, California, June, 1990, pp 197–202.

[22] Kirkland L.V., Dean J.S. Monitoring Power Supply Current and Using a Neural Network Routine to Diagnose Circuit Faults. In: Proc. AUTOTESTCON '94, Anaheim, California, September 1994, pp 649–651.

[23] Liu R. (Ed.) Selected Papers on Analog Fault Diagnosis. Advances in Circuits and Systems, IEEE Circuits and Systems Society, IEEE Press, 1987.

[24] Liu R. (Ed.) Testing and Diagnosis of Analog Circuits and Systems. Van Nostrand Reinhold, 1991.

[25] Mallat S.G. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Analysis and Machine Intelligence, July 1989, 11(7):674-693.

[26] Meador J., Wu A., Tseng C.T., Lin T.S. Fast Diagnosis of Integrated Circuit Faults Using Feedforward Neural Networks. In: Proc. Int. Joint Conf. on Neural Networks, Seattle, Washington, July 1991, pp 269–273.

[27] Meyer Y. Wavelets - Algorithms and Applications. SIAM, 1993.

[28] Milne R. Stategy for Diagnosis. IEEE Trans. on Systems, Man, and Cybernetics 1987; 17(3):333–339.

[29] Parten C.R., Saeks R., Pap R. Fault Diagnosis and Neural Networks. In: Proc. IEEE Int. Conf. on Systems, Man and Cybernetics, Charlottesville, Virginia, October 1991, pp 1517–1521.

[30] Rioul O., Vetterli M. Wavelet and signal processing. IEEE SP Magazine, October 1991, pp 14-38.

[31] Rutkowsky G. A Neural Network Approach to Fault Location in Non Linear DC Circuits. In: Proc. 3rd Int. Conf. on Artificial Neural Networks, Brighton, England, September 1992, pp 1123–1126.

[32] Sharkey A.J.C. On combining Artificial Neural Nets. Connection Science, Special Issue: Combining Artificial Neural Nets: Ensemble Approaches, December 1996, 8(3/4):299–313.

[33] Simpson P.K. Artificial Neural Systems. Pergamon Press Inc., 1990.

[34] Spence H.F., Burris D.P., Lopez J., Houston R.A. An Artificial Neural Network Printed Circuit Board Diagnostic System Based on Infrared Energy Emission. In: Proc. AUTOTESTCON 91, Anheim, California, September 1991, pp 41–45.

[35] Spence H.F. Printed Circuit Board Diagnosis Using Artificial Neural Networks and Circuit Magnetic Fields. IEEE Aerospace Systems Magazine 1994; 20–24.

[36] Spina R., Upadhyaya S. Linear Circuit Fault Diagnosis Using Neuromorphic Analyzers. IEEE Trans. on Circuits and Systems—II , March 1997, 44(3):188–196.

[37] Tarassenko L. A Guide to Neural Computing Applications. Arnold, 1998.

[38] Thompson A.B., Sutton J.C., Nagle H.T. Diagnosis of Telephony Line Card Component Failures Using an Artificial Neural Network. In: Proc. SOUTHEASTCON 91, Raleigh, North Carolina, April 1991, pp 229–233.

[39] Totton K.A.E., Limb P.R. Experience in Using Neural Networks for Electronic Diagnosis. In: Proc. 2nd Int. Conf. on Artificial Neural Networks, Bournemouth, England, November 1991, pp 115-118.