

16

Medical Imaging and Diagnosis Using Genetic Algorithms

16.1	Introduction.....	16-235
16.2	Preliminaries on Genetic Algorithms.....	16-236
	Basic Principles and Features • Encoding Strategy and Population • Evaluation Technique • Genetic Operators • Parameters of GA	
16.3	Genetic Algorithms for Equipment Design for Medical Image Acquisition	16-239
16.4	Genetic Algorithms for Medical Diagnosis	16-240
	Image Driven Diagnosis • Data Driven Diagnosis	
16.5	Discussion and Conclusions	16-247
	Acknowledgment.....	16-247
	References	16-247

Ujjwal Maulik
Sanghamitra Bandyopadhyay
Sajal K. Das

16.1 Introduction

The last half of the twentieth century has seen a vigorous growth in the field of digital image processing (DIP) and its potential applications. DIP deals with the manipulation and analysis of images that are generated by discretizing the continuous signals. One important area of application that has evolved from the 1970s is that of medical images. Rapid development in different areas of image processing, computer vision, pattern recognition, and imaging technology, and the transfer of technology from these areas to the medical domain has changed the entire way of looking at clinical routine, diagnosis, and therapy. Also, the need for more effective and less (or non) invasive treatment has led to a large amount of research for developing what may be called *computer aided medicine*.

Most modern medical data are expressed as images or other types of digital signals. The explosion in computer technology in recent years introduced new imaging modalities such as x-rays, magnetic resonance imaging (MRI), computer tomography (CT), positron emission tomography (PET), single photon emission computed tomography (SPECT), electrical impedance tomography (EIT), ultrasound, and so on. These images are noninvasive and offer high spatial resolution. Thus the acquisition of a large number of such sophisticated image data has given rise to the development of quantitative and automatic processing and analysis of medical images (as opposed to the manual qualitative assessment done earlier). Moreover, the use of new, enhanced, and efficient computational models and techniques has also become necessary.

A large amount of research is being devoted to the various domains of medical image processing, and some surveys are already published [1,2]. However, in view of the vastness of the field, it has become necessary to specialize any further survey work that is undertaken in this area, so that it can become manageable and can be of more benefit to researchers/users. Some such attempts have already been made, for example, specialization in terms of period of publication [3], image segmentation [4], registration [5], virtual reality, and surgical simulation [6,7].

The area of designing equipments for better imaging and hence improvement in subsequent processing tasks has also received the attention of researchers. The design problem has been viewed as one of optimization, and therefore the use of efficient search strategies has been studied. The application of genetic algorithms, a well-known class of search and optimization strategies, is also one of the important areas that has been investigated in this regard.

Genetic Algorithms (GAs) [8,9] are randomized search and optimization techniques guided by the principles of evolution and natural genetics, and have a large amount of implicit parallelism. They provide near optimal solutions of an objective or fitness function in complex, large, and multimodal landscapes. In GAs, the parameters of the search space are encoded in the form of strings called *chromosomes*. A *fitness function* is associated with each string that represents the degree of *goodness* of the solution encoded in it. Biologically-inspired operators such as *selection*, *crossover*, and *mutation* are used over a number of evolutions (generations) for generating potentially better strings.

The important fallout of (semi-) automated medical image processing tasks is enhanced diagnosis. Several tasks in the area of medical diagnosis have also been modeled as an optimization problem, and researchers have used GAs for solving them. In this chapter, we attempt to provide a state-of-the-art survey in the application of the principles of GAs, an important component of *evolutionary computation*, for improving medical imaging and diagnosis tasks. Section 16.2 describes the basic principles of GAs. Thereafter, the use of GAs in improving equipment design has been studied. Finally, the application of GAs for computer aided diagnosis, including schemes driven by both image and data (consisting of information not derived from images), is provided.

16.2 Preliminaries on Genetic Algorithms

Genetic algorithms [8–10] are efficient, adaptive, and robust search and optimization processes that are usually applied to very large, complex, and multimodal search spaces. They are modeled on the principles of natural genetic systems, in which the genetic information of each individual or potential solution is encoded in structures called chromosomes. They use some domain or problem-dependent knowledge for directing the search in more promising areas of the solution space; this is known as the fitness function. Each individual or chromosome has an associated fitness function, which indicates its degree of goodness with respect to the solution it represents. Various biologically-inspired operators such as selection, crossover, and mutation are applied on the chromosomes to yield potentially better solutions.

16.2.1 Basic Principles and Features

Genetic algorithms emulate biological principles to solve complex optimization problems. It essentially comprises a set of individual solutions or chromosomes (called the population), and some biologically-inspired operators that create a new (and potentially better) population from an old one. According to the theory of evolution, only those individuals in a population who are better suited to the environment are likely to survive and generate offspring, thereby transmitting their superior genetic information to new generations.

The essential components of GAs are the following:

- A representation strategy that determines the way in which potential solutions will be encoded to form string like structures called chromosomes.
- A population of chromosomes.

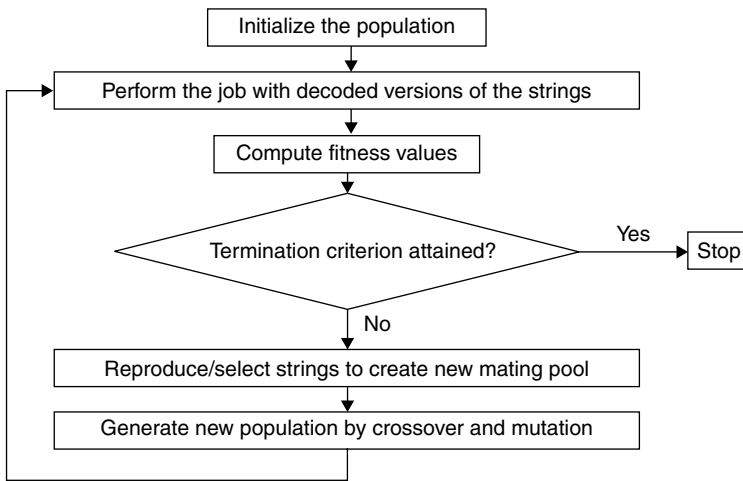


FIGURE 16.1 Basic steps of a genetic algorithm.

- Mechanism for evaluating each string (fitness function).
- Selection/reproduction procedure.
- Genetic operators (crossover and mutation).
- Probabilities to perform genetic operations.

A schematic diagram of the basic structure of a GA is shown in Figure 16.1. The components of GAs are described in the following sections.

16.2.2 Encoding Strategy and Population

To solve an optimization problem, GAs start with the chromosomal representation of a parameter set, which is to be encoded as a finite size string over an alphabet of finite length. Usually, the chromosomes are strings of 0s and 1s. For example, the string

1 0 0 1 1 0 1 0

is a binary chromosome of length 8. It is evident that the number of different chromosomes (or strings) is 2^l , where l is the string length. Each chromosome actually refers to a coded possible solution. A set of such chromosomes in a generation is called a population, the size of which may be constant or may vary from one generation to another. A common practice is to choose the initial population randomly.

16.2.3 Evaluation Technique

The fitness/objective function is chosen depending on the problem to be solved, in such a way that the strings (possible solutions) representing good points in the search space have high fitness values. This is the only information (also known as the payoff information) that GAs use while searching for possible solutions.

16.2.4 Genetic Operators

The frequently used genetic operators are selection, crossover, and mutation operators. These are applied to a population of chromosomes to yield potentially new offspring. The operators are described in Sections 16.2.4.1 to 16.2.4.3.

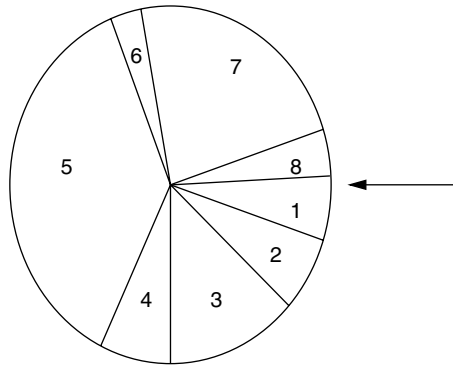


FIGURE 16.2 Roulette wheel selection.

16.2.4.1 Selection

The selection/reproduction process copies individual strings (called parent chromosomes) into a tentative new population (known as mating pool) for genetic operations. The number of copies that an individual receives for the next generation is usually taken to be directly proportional to its fitness value; thereby mimicking the natural selection procedure to some extent. This scheme is commonly called the *proportional selection scheme*. *Roulette wheel parent selection*, *stochastic universal selection*, and *binary tournament selection* [8,10] are some of the most frequently used selection procedures. Figure 16.2 demonstrates the roulette wheel selection. The wheel has as many slots as the population size P , where the size of a slot is proportional to the relative fitness of the corresponding chromosome in the population. An individual is selected by spinning the roulette, and noting the position of the marker when the roulette stops. Therefore, the number of times that an individual will be selected is proportional to its fitness (or, the size of the slot) in the population. In the commonly used *elitist* model of GAs, thereby providing what is called an *elitist* GA (EGA), the best chromosome seen up to the present generation is retained either in the population, or in a location outside it.

16.2.4.2 Crossover

The main purpose of crossover is to exchange information between randomly selected parent chromosomes by recombining parts of their genetic information. It combines parts of two parent chromosomes to produce offspring for the next generation. *Single-point crossover* is one of the most commonly used schemes. Here, first of all, the members of the selected strings in the mating pool are paired at random. Then each pair of chromosomes is subjected to crossover with a probability μ_c where an integer position k (known as the crossover point) is selected uniformly at random between 1 and $l - 1$ ($l > 1$ is the string length). Two new strings are created by swapping all characters from position $(k + 1)$ to l . For example, let the two parents and the crossover points be as shown below.

```

1 0 0 1 1 | 0 1 0
0 0 1 0 1 | 1 0 0
    
```

Then, after crossover the offspring will be the following:

```

1 0 0 1 1 1 0 0
0 0 1 0 1 0 1 0
    
```

Other common crossover techniques are two-point crossover, multiple point crossover, shuffle-exchange crossover, and uniform crossover [9].

The successful operation of GAs depends, to a great extent, on the coding technique used to represent the problem variables [11,12]. The *building block hypothesis* indicates that GAs work by identifying good building blocks, and by eventually combining them to get larger building blocks [8,13,14]. Unless good building blocks are coded tightly, the crossover operation cannot combine them [15,16]. Thus coding–crossover interaction is important for the successful operation of GAs. The problem of tight or loose coding of problem variables is largely known as the *linkage problem* [17]. Recent work on linkage learning GAs that exploits the concept of gene expression can be found in References 18 to 20.

16.2.4.3 Mutation

Mutation is the process by which a random alteration in the genetic structure of a chromosome takes place. Its main objective is to introduce genetic diversity into the population. It may so happen that the optimal solution resides in a portion of the search space that is not represented in the population's genetic structure. The process will therefore be unable to attain the global optima. In such situations, only mutation can possibly direct the population to the optimal section of the search space by randomly altering the information in a chromosome. Mutating a binary gene involves simple negation of the bit, whereas that for real coded genes are defined in a variety of ways [10,21]. Here, we discuss the binary bit-by-bit mutation, where every bit in a chromosome is subject to mutation with a probability μ_m . The result of applying the bit-by-bit mutation on positions 3 and 7 of a chromosome is shown here.

```

1 0 0 1 1 0 1 0
1 0 1 1 1 0 0 0

```

16.2.5 Parameters of GA

There are several parameters in GAs that have to be manually tuned and fixed by the programmer. Among these are the population size, probabilities of performing crossover and mutation, and the termination criteria. Several other things must also be determined by the programmer. For example, one must decide whether to use the generational replacement strategy, in which the entire population is replaced by a new population, or the steady state replacement policy where only the less fit individuals are replaced. Most such parameters in GAs are problem dependent, and no guidelines for their choice exist in the literature. Therefore, several researchers have also kept some of the GA parameters variable and/or adaptive [22–24].

As shown in [Figure 16.1](#), the cycle of selection, crossover, and mutation is repeated a number of times till one of the following occurs:

1. The average fitness value of a population becomes more or less constant over a specified number of generations.
2. A desired objective function value is attained by at least one string in the population.
3. The number of generations (or iterations) is greater than some threshold.

16.3 Genetic Algorithms for Equipment Design for Medical Image Acquisition

Magnetic resonance imaging (MRI) is one of the most commonly used medical imaging techniques used for the diagnosis of several ailments including multiple sclerosis, strokes, tumors, and other infections of the brain, bones, spine, or joints, and for visualizing torn ligaments, soft tissues of the body, etc. Magnetic resonance imaging is based on the phenomenon of nuclear magnetic resonance (NMR) discovered by Felix Bloch and Edward Powell, for which they received the Nobel Prize in 1952. In an MRI scan, the patient is placed inside the bore of a magnet, which is usually cubic in shape. The body part to be scanned

is placed at the center (or, isocenter) of the magnetic field inside the bore. In this noninvasive procedure, strong magnetic fields along with radio waves are used to visualize the structure of a particular body part.

The design of appropriate equipments for the purpose of good imaging may be considered as the first step in medical image processing. For example, large superconducting solenoids with apertures of typically 1 m, highly uniform (20 ppm) central fields of 1–2T, and low fringe fields (5 Gauss at 5 m) are required in clinical MRI. However, these magnets, which are now available in the clinical market, have deep bores, typically between 1.8 and 2 m length, and have a number of disadvantages such as patient claustrophobia and limited access for intervention. In order to overcome the limitations and evolve good designs for the magnets, researchers have mapped the design problem to one of optimization and have investigated the use of computational methods [25,26], including GAs [27–29], for this purpose.

Analytical techniques have been the preferred approach to design such magnets and gradient sets for MRI. Such technique are computationally efficient but are approximate, particularly away from the axis of symmetry. In Reference 30, an attempt has been made, which uses GA running on massively parallel computers to design an actively shielded whole-body MRI solenoid magnet with a bore of 1 m. The task is to optimize a cost function based on the magnetic field generated by a set of upto 20 circular coils, each having upto 2000 turns. The coils are constrained to be concentric with the magnet, and are arranged in symmetric pairs. A single coil is described by five parameters:

- Two parameters describing its position in the $X-Z$ plane
- Depth of the coil
- Width of the coil
- The direction of the current

A chromosome encodes these five parameters per coil, for upto 20 coils. Since the coils are arranged in pairs that are symmetric about the central $X-Y$ plane of the magnet, only 10 of the coils are independent. Thus a chromosome encodes upto 50 parameters that have to be optimally tuned. The magnetic field is computed using the Biot–Savart law. The fitness function incorporates terms for uniformity of field in the region of interest (ROI), and the smallness of the fringe field. The field is calculated by summing over the contributions from the turns in each coil. Recombination is performed as a two-stage process. In the first stage, a parent subset, which is of half the size of the population, is created by performing binary tournament selection in the top eighth of the population. In the second stage, pairs of chromosomes from the parent subset are picked at random and mated to produce a single offspring. The parents are replaced in the subset, and this process is continued till the next generation is completed. Three types of mutations are considered, which takes care of small perturbations, perturbations to add or remove a coil from the design, and a drastic perturbation for adding extra diversity to the population. The initial result provided in Reference 30 demonstrates the effectiveness of GAs for producing shorter whole-body magnet designs, which are original and innovative.

Two-dimensional ultrasonic arrays provide the possibility of three-dimensional electronic focusing and beam-steering and thus three-dimensional imaging. In simulation studies, it has been demonstrated [31] that reducing the number of elements of a two-dimensional matrix array down to order eight, keeps resolution and leads still to sufficient contrast. These random arrays are usually obtained by generating a random pattern with the desired number of elements. In Reference 32, simulation is presented to show that the imaging quality of a sparse tree can be improved by optimizing the random choice of elements. The optimization is done using GA.

16.4 Genetic Algorithms for Medical Diagnosis

For diagnosis of diseases, several tests are performed on the patient, some of which may involve taking various images such as x-rays, MRI scan, CT scan, etc., and some others that involve pathological and other tests. Data from these tests may become quite large and conflicting, and manual interpretation by collecting all such information may become difficult. This has therefore given rise to the development of

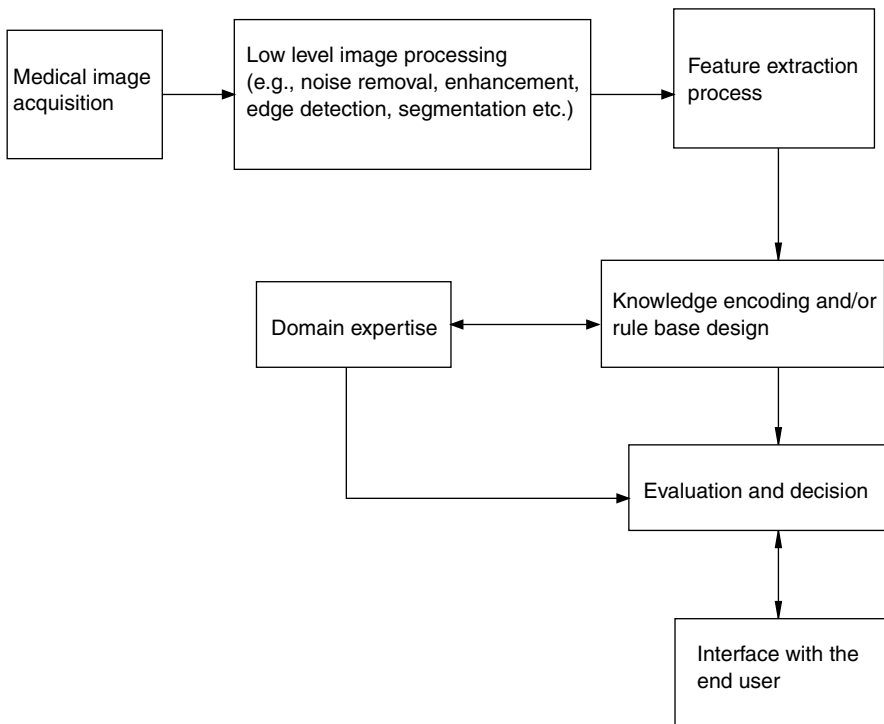


FIGURE 16.3 Block diagram of a CAD system.

computer assisted diagnostic tools, which are intended to help the medical practitioners make sense out of a large amount of data, and make diagnosis and therapy decisions. Figure 16.3 shows a block diagram of a computer aided diagnosis system.

16.4.1 Image Driven Diagnosis

Computer aided detection and classification schemes have the potential of increasing diagnostic accuracy in medical imaging. This may be done, for instance, by alerting radiologists to lesions that they initially overlooked, or assisting in the classification of detected lesions [33]. As with any complicated pattern recognition system, these schemes, generally referred to as computer aided diagnosis (CAD) schemes, typically employ multiple parameters such as threshold values, filter weights, and ROI sizes to arrive at a detection or classification decision. For the system to have a high performance, the values of these parameters need to be set optimally. In general, the optimal set of parameters may change when a component of the imaging chain is modified or changed. When the number of parameters is large, it becomes very difficult to manually determine the optimal choice of parameter values, as some of the values may be correlated in some unknown manner. However, conventional optimizing techniques, including GAs, are designed to optimize a scalar objective function, and the task of optimizing the performance of a diagnostic decision-making system is clearly multiobjective. Therefore, two objectives, increasing the sensitivity and reducing the false-positive rate of the system, are described by a single objective function. This objective function is defined as the weighted sum of the sensitivity and false-positive rate or using area under the receiver operating characteristic (ROC) curve. The use of a niched pareto genetic algorithm (NP-GA) [34] in training two popular diagnostic classifiers for optimizing their performance, has been studied in Reference 35. Unlike conventional classifier techniques that formulate the problem as the solution to a scalar optimization, the NP-GA explicitly addresses the multiobjective nature of the training task. Traditional techniques of classifier training attempt to combine several objective functions into one,

so that conventional scalar optimization technique can be utilized. This involves incorporating a priori information into the aggregation method so that the resulting performance of the classifier is satisfactory for the task at hand. It has been shown in Reference 35 that the multiobjective genetic approach removes the ambiguity associated with defining a scalar measure of classifier performance and that it returns a set of optimal solutions that are equivalent in the absence of any information regarding the preference of the objectives, that is, sensitivity and specificity. The a priori knowledge that is used for aggregating the objective functions in conventional classifier training may instead be applied for post-optimization to select from one of the series of solutions returned from multiobjective genetic optimization. This technique is applied in Reference 35 to train a linear classifier and an artificial neural network, using simulated datasets. The performance of the solutions returned from the multiobjective genetic optimization represents a series of optimal pairs, sensitivity and specificity, which can be thought of as operating points on an ROC curve. It is observed that all possible ROC curves for a given dataset and classifier are less than or equal to the ROC curve generated by the NP-GA optimization. In Reference 36, a multiobjective approach for optimizing the performance of two rule-based CAD schemes has been proposed. One of these CAD schemes is designed to detect clustered microcalcifications in digitized mammograms, while the other scheme is developed to detect the breast masses.

Diagnosis and follow up of pigmented skin lesions is an important step toward early diagnosis of skin cancer [37]. For this purpose, digitized epiluminescence microscope (ELM) [38] images of pigmented skin lesions is used. Epiluminescence microscopy is a noninvasive technique that uses an oil immersion to render the skin translucent and make pigmented structure visible. During clinical diagnosis of pigmented skin lesions, one of the main features is the lesion symmetry, which should be evaluated according to its shape, color, and texture. This may be evaluated by drawing two orthogonal axes that maximize the perceived symmetry [39]. The evaluation is binary, that is, the lesion is either symmetrical or asymmetrical. In addition to the small number of possible outcomes, the evaluation is highly subjective and depends on the physicians' experience. As a result, the development of automatic techniques for the quantification of symmetry and the detection of symmetry axes is necessary. Other methods based on principal component technique for computing the axes may be found in References 40 and 41.

Reference 42 has proposed a GA-based technique and an optimization scheme derived from the self-organizing maps theory for the detection of symmetry axes. The notion of symmetry map has been introduced, which allows an object to be mapped to a symmetry space where its symmetry properties can be analyzed. The objective function (ψ) used in Reference 42 is based on a given symmetry measure, which is a function of the mean-square error (MSE) between the original and the reflected images. The MSE is defined as follows:

$$\text{MSE} = E[|\Gamma(x, y) - \Gamma(x', y')|^2] \quad (16.1)$$

where $\Gamma(x, y) : [0, c - 1] \times [0, r - 1] \rightarrow IR^n$ is the vector valued $r \times c$ input image, where the pixel values are in an n dimensional space, (x, y) and (x', y') represent the pixel coordinates and the symmetry coordinates respectively. The input image can be decomposed into symmetric ($\Gamma_s(x, y)$) and asymmetric ($\Gamma_a(x, y)$) components. Therefore, $\Gamma(x, y) = \Gamma_s(x, y) + \Gamma_a(x, y)$. The asymmetric component can be considered as symmetric noise, and the MSE is proportional to its energy. The distortion due to noise can be measured through the peak signal-to-noise ratio (PSNR), and is given by

$$\text{PSNR} = 10 \log_{10} \left(\frac{(N_q - 1)^2}{\text{MSE}} \right), \quad (16.2)$$

where N_q is the number of quantization levels in the image. The fitness function ψ is defined as

$$\psi = 1 - \frac{1}{1 + \text{PSNR}}. \quad (16.3)$$

Real encoding of chromosomes has been used in Reference 42 along with a modified version of simulated binary crossover [43]. The proposed technique is applied for detection and diagnosis of malignant melanoma.

The development of computer supported systems for melanoma diagnosis is of great importance to dermatologists due to its clinical accuracy in identifying malignant melanomas [44,45]. Several techniques for computerized melanoma diagnosis are based on color images making use of image analysis methods to quantify visual features as described by the "ABCD" rule (Asymmetry, irregular Border, varying Color, Diameter) [37,46,47]. Laser profilometry opens up new possibilities to improve tumor diagnostics in dermatology [48].

The recognition task is to classify surface profiles of melanomas and nevi also called moles. Due to the fact that all profiles contain regions with a structure similar to normal skin, each profile is subdivided into 16 non-overlapping quadratic subprofiles and image analysis algorithms are applied to each sub-profile separately. Subsequently, feature selection algorithms are applied to optimize the classification performance of the recognition system.

An efficient computer supported technique that uses GAs for diagnosis of skin tumors in dermatology is presented in Reference 49. High resolution skin surface profiles are analyzed to recognize malignant melanomas and nevocytic nevi. In the initial phase, several types of features are extracted by two-dimensional image analysis techniques characterizing the structure of skin surface profiles: texture features based on co-occurrence matrices [50], Fourier features [51], and fractal features [37,52,53]. Subsequently, several feature selection algorithms based on heuristic strategies, greedy technique, and GA are applied to determine suitable feature subsets for the recognition process by describing it as an optimization problem. As a quality measure for feature subsets, the classification rate of the nearest neighbor classifier computed with the leaving-one-out method is used. Among the different techniques used, GAs show the best result. Finally, neural networks with error back-propagation as learning paradigm are trained using the selected feature sets. Different network topologies, learning parameters, and pruning algorithms are investigated to optimize the classification performance of the neural classifier. With the optimized recognition system, a classification performance of 97.7% is achieved.

In the 1980s, microwave imaging was thought to have great potential in developing medical diagnostic tools. However, because of the problem of inverse scattering, good reconstruction of images was found to be difficult. Carosi et al. [54] have used the approach of focused imaging, where only a part of the body is subjected to the investigation, combined with the capabilities of global search techniques like GAs for accurate reconstruction. For experimental purpose, a human abdomen is considered and different electromagnetic sources operating at the working frequency of 433 MHz and 2.54 GHz are used. Numerical investigations are performed to define the optimal dimensions of the reduced investigation domain. To quantitatively evaluate the effects of the reduction of the original investigation domain on the inversion data, suitable relative errors are defined. Once the reduced domain is defined, preliminary reconstructions are performed aiming to evaluate the imaging capability of GAs when a focussed approach is used for tomographic application.

One of the most important facial paralysis diagnosis techniques is quantitative assessment of patient's facial expression motion. Johnson et al. [55] proposed a technique for achieving maximal regional facial motion while the rest of the face is held at rest based on Maximal Static Response Assay of facial nerve function. This requires the placement of removable adhesive dots and a small adhesive ruler on the face at predefined locations. However, many facts, such as misplacement of facial dots, misplacement of the grid, and reading and entry errors, will cause an error in the assay. Helling et al. [56] used region-specific, subtracted, digitized image light reflectance as a two-dimensional marker for the complex three-dimensional surface deformations of the face during expression. The velocity of region-specific facial motion is estimated from the facial motion image sequences using the optimal flow (OF) technique. The computation of the OF field requires estimates of both the spatial gradient and spatial time derivative at each pixel, and this time-consuming process often limits its use, especially in medical application. To overcome this problem, an OF technique based on GA is proposed in Reference 57 to detect facial motions from dynamic image sequences. Experimental results demonstrate that the proposed technique is very

useful to diagnose the site of facial paralysis and assess progression or recovery profiles of patients when combined with other diagnosis techniques.

16.4.2 Data Driven Diagnosis

The previous section reported some research on development of diagnostic tools in which one of the inputs considered is an image. In this section, we deal with some other diagnostic tools that use other types of input data, mostly numeric.

Electromyography (EMG) is the recording and study of the electrical activity of voluntary contracting muscles. Clinical EMG findings provide useful information in the electrodiagnostic examination of peripheral nerves and skeletal muscle, and in deciding the level of the lesion in patients suffering from neuromuscular disorders. It is also useful in deciding whether the symptom of muscle weakness in the assessment of neuromuscular disorders is myopathic or neurogenetic in origin. The advantages of automated EMG diagnostic systems can be found in Reference 58. Different approaches have been used to address the problem of automated EMG diagnosis. The utility of artificial neural networks in classifying EMG data trained with back propagation or Kohonen's self-organizing feature map algorithm has recently been demonstrated in References 59 and 60. In Reference 61, a study has been made to investigate how genetics-based machine learning (GBML) can be applied for diagnosing certain neuro muscular disorder based on EMG data. The effect of GBML control parameters on diagnostic performance is also examined. Subsequently, a hybrid diagnostic system is introduced that combines both the neural network and GBML.

In References 62 and 63, a methodology based on GAs for the automatic induction of Bayesian networks from a file containing cases and variables related to the problem of predicting survival in malignant skin melanoma is described. The structure is learned by applying three different techniques: the Cooper and Herskovits metric for a general Bayesian network [64], the Markov blanket approach, and the relaxed Markov blanket method. The methodologies are applied to the problem of predicting survival of people after 1, 3, and 5 years of being diagnosed as having malignant skin melanoma. The induced Bayesian network is used for classifying patients according to their prognosis of survival. These results are compared to those obtained by the Naive-Bayes paradigm. An empirical comparison of Bayesian networks learned using GAs, rule induction, and logistic regression is carried out in Reference 65 where the task is to predict the survival of women suffering from breast cancer. In a more recent attempt, Blanco et al. [66] have studied the problem of learning Bayesian networks using two stochastic, population-based search algorithms: the univariate marginal distribution algorithm and population-based incremental learning. Comparison with the GA-based scheme is also carried out.

The problem of Wisconsin breast cancer diagnosis (WBCD) has been considered in References 67 and 68, by combining fuzzy systems and evolutionary algorithms to design an automatic diagnosis system. The proposed fuzzy-genetic approach produces systems exhibiting two prime characteristics. They attain high classification performance with the possibility of attributing a new confidence measure to the output diagnosis. Moreover, the resulting systems involve a few simple rules, and are, therefore, human interpretable. Another approach for diagnosis of breast cancer by Bayesian networks can be found in Reference 69.

In medicine, prognostic models may be used to assess the likely prognosis of a patient, which, in turn, may determine the treatment of that patient. In Reference 70, a prognostic model based on diffusion GAs is sought to determine whether or not patients suffering from an uncommon form of cancer will survive. The problem considered is a multiobjective one, in which the three objectives are:

- Maximize the correct number of survival predictions
- Maximize the correct number of death predictions
- Minimize the number of factors used

The motivation behind this study is to accurately predict the outcome so that patients who are more likely to die can be identified at diagnosis, and subjected to high dose aggressive chemotherapy (which has several negative side effects). On the other hand, those with a high chance of survival can be spared this treatment and its side effects. Given a set of case histories, a technique is proposed in Reference 70

TABLE 16.1 Different Criteria and Their Ranges Used in Reference 70

Number	Name	Ranges
1	Brain metastases	Yes, No
2	Liver metastases	Yes, No
3	Placental site trophoblastic tumors (PSTT)	Yes, No
4	Prior chemotherapy	Yes, No
5	Pregnancy to term	Yes, No
6	Age	<26, 26–30, 31–41, >41
7	Serum hCG level	<800, 801–29570, 29571–181000, >181000
8	Interval between pregnancy and diagnosis	<4, 5–9, 10–33, >33

that attempts to find the relative weights of the different factors that are used to describe the cases. The eight factors that are used in the prognostic model of Reference 70 and their possible ranges are provided in Table 16.1.

A diffusion GA is used for building the prognostic model that will simultaneously optimize the three objectives given here. A model is represented as a chromosome by encoding the weights associated with each factor. Boolean criteria (1–5 in Table 16.1) have single associated weights, while criteria having a range of values (6–8 in Table 16.1) have one weight per subrange. In addition a combination weight is used in Reference 70, which can incorporate the possibility that a combination of factors might be important. Thus a total of 18 weights are used, 5 for the Boolean criteria, 12 for the subranged criteria, and the combination weight. In the diffusion GA, the individuals in a population are arranged along the vertices of a square lattice. During crossover, each individual chooses one of its four neighbors randomly as the mate. Mutation can either increase or decrease a weight by 10%, or set it to zero. Only if the result of crossover and mutation at a position is better, in the Pareto optimal sense, than the original, the latter will be replaced. Marvin et al. [70] experimented with a population size of 169 placed on a 13 × 13 grid. The weights are randomly set in the range [−2000, 2000], and the algorithm is executed for 1000 generations. It is found to predict 90% of the survivals and 87% of the deaths, while using 6 of the 8 factors, and 13 of the possible 18 weights when the entire data set is used. For training using 90% data and testing using the remaining 10%, several prognostic models are obtained each coming up with a different compromise among the three objective values. Significantly, the method in Reference 70 enables a simple model to be evolved, one that produces well-balanced predictions and one that is relatively easy for clinicians to use.

Ngan et al. [71] employ evolutionary programming (EP) and genetic programming (GP) in the domain of knowledge discovery in medical systems. Evolutionary Programming is used to learn Bayesian networks, which is known to be an intractable problem. Minimum description length principle is used to measure the goodness of solution in EP followed by the use of GP to learn rules. The entire knowledge discovery system is applied on limb fracture and scoliosis data, where it is able to detect many interesting patterns/rules that were uncovered earlier. Ngan et al. had earlier used GP for discovering comprehensible rules in the medical domain; they used grammar to restrict the search space, and to ensure the syntactical correctness of the rules [72]. The discovered rules were evaluated within the framework of support confidence proposed for association rule mining. Here, a major limitation was that the grammar was application dependent, and had to be written for each application domain.

Genetic programming is also used to discover comprehensible rules for predicting 12 different diseases using 189 predicting attributes, or measurements [73]. The 12 diseases considered here are stable angina, unstable angina, acute myocardial infarction, aortic dissection, cardiac tamponade, pulmonary embolism, pneumothorax, acute pericarditis, peptic ulcer, esophageal pain, musculoskeletal disorders, and psychogenic chest pain, the characteristic of all of which was chest pain. All the 189 attributes are binary. Genetic programming is used to learn rules expressed in a kind of first-order logic of the form $\langle Att_i Op Att_j \rangle$, where Att_i and Att_j are the predicting attributes, and Op is some relational operator. Genetic programming evolves a population of “programs” candidate to the solution of a specific problem. Here, a program is represented in the form of a tree, in which the internal nodes are functions (operators) and the leaf nodes

are terminal symbols. In the GP formulation of the problem in Reference 73, the terminal set consists of the 189 attributes, and the function set consists of {AND, OR, NOT}. The GP is executed once for each class, with the appropriate rule for the i th class being evolved in the i th GP run. Thus, each run consists of learning a two-class classification problem, in which the goal is to predict whether a patient has a particular disease (class i) or not (NOT class i). For computing the fitness of a candidate (program or rule), a (labeled) training set is used on which the rule is tested. The size of the following sets are computed:

- True positives (tp): the rule predicts that the patient has a given disease and the patient does have it.
- False positives (fp): the rule predicts that the patient has a given disease and the patient does not have it.
- True negative (tn): the rule predicts that the patient does not have a given disease and the patient actually does not have it.
- False negative (fn): the rule predicts that the patient does not have a given disease but the patient actually has it.

Thereafter, two measures are computed, the sensitivity (Se) and specificity (Sp):

$$Se = \frac{tp}{tp + fn} \quad (16.4)$$

$$Sp = \frac{tn}{tn + fp}. \quad (16.5)$$

The fitness function is taken to be the product of the two, namely,

$$\text{fitness} = Se \times Sp. \quad (16.6)$$

For the experiments, the data set consisted of 138 samples (patients), which were partitioned into a training set with 90 samples, and a test set with 48 samples. The GP achieved an accuracy of 77.08% on the test set. Other related methods that used GP for classification problem can be found in References 74 to 78. On similar lines, in Reference 79, GAs were used to discover comprehensible IF-THEN rules for the diagnosis of dermatological diseases and prediction of the recurrence of breast cancer. Here, a chromosome is of length n , where n is the number of attributes. The i , the gene corresponding to the i th attribute, is divided into three fields: weight (W_i), operator (O_i), and value (V_i). Each gene corresponds to one condition in the IF part of the rule. The GA is executed once for each class, and therefore the THEN part (indicating the class for which the GA was run) is not required to be encoded in the chromosome. The weight value indicated whether the i th attribute, A_i , is at all present in the rule (if $W_i > \text{Limit}$) or not (if $W_i \leq \text{Limit}$). Limit was set to 0.3. The operator O_i could take values from {=, ≠}, if the corresponding attribute is categorical, and from {≥, <}, if the corresponding attribute is continuous. The value V_i could take values from the domain of the attribute A_i . Normal selection and crossover operators are used. Three mutation operators are defined, namely weight mutation, operator mutation, and value mutation. The fitness function of a chromosome was defined as $\text{fitness} = Se * Sp$ as in Reference 73. The dermatology data set consists of the differential diagnosis of erythematosquamous. There are six different diagnoses (six classes): psoriasis, seboreic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis, and pityriasis rubra pilaris. The data set consists of 366 records with 34 attributes. The breast cancer data consists of 286 records with 9 attributes and 2 classes (recurrence and nonrecurrence of cancer). The accuracy rates achieved in Reference 79 were 95% for the dermatological data and 67% for the cancer data. The resultant rules were also found to be comprehensible, with one rule obtained per class.

16.5 Discussion and Conclusions

This chapter provides a comprehensive survey of the application of GAs to the domain of designing of equipments for medical image acquisition and medical diagnosis. In recent times, with the advent of a variety of sophisticated imaging techniques, use of medical images in clinical diagnosis and therapy has increased manifold. Therefore, attempts at increasing the resolution and quality of such images is an important area of research. This, in turn, leads to research for designing equipments in such a way that the imaging modality becomes faster and more informative. Computer aided diagnosis has also been necessitated due to the large amount of data that is routinely collected for the patients. These problems often turn out to be those of search and optimizations, requiring the use of good optimization tools such as GAs.

The research works reviewed in this chapter have been reported in diverse journals and proceedings like *IEEE Transactions on Medical Imaging*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Neural Networks*, *Artificial Intelligence in Medicine*, *Medical Image Analysis*, *Magnetic Resonance Imaging*, *Evolutionary Computation*, *International Conference on Evolutionary Computation*, *International Conference on Genetic Algorithms*, *ACM SIGMOD Conference on Management of Data*, and so on. Encouraging results have been reported by researchers in this regard. This chapter presented a methodical way in which a large number of such research activities have been compiled and reported within a common platform, namely GA-based techniques.

It may be noted that the main challenges and issues in integrating GAs for solving optimization problems in medical imaging and diagnosis are manifold. First, the encoding strategy must be suitably defined so that it conforms to the building block hypothesis. According to this hypothesis, short low order above average schema, or building blocks, should combine to yield potentially better solutions. Any adhoc encoding strategy may not follow this hypothesis, and hence encoding strategy may not follow this hypothesis, implying GAs may often be found to yield poor results in such situations. Second, the fitness function must be adequately designed. Since fitness computation is often computation intensive, in the medical domain one may need to go for parallel GAs working on massively parallel systems to get real time response. This in turn incorporates an added level of difficulty to the problem. Finally, a still open unsolved issue is the appropriate selection of the operator probabilities and the termination criterion, so as to ensure good performance of the GA-based systems for medical imaging and diagnosis problems.

Acknowledgment

A part of this work was carried out when Dr. U. Maulik visited the University of Texas at Arlington, USA, with the BOYSCAST fellowship provided by Department of Science and Technology, Government of India, during 2001. Dr. S. Bandyopadhyay would like to acknowledge Indian National Science Academy, Government of India sponsored project *Soft computing for medical image segmentation and classification* No. BS/YSP/36/887 for providing partial support to carry out this work.

References

- [1] G. Gerig, T. Pun, and O. Ratib. Image analysis and computer vision in medicine. *Computerized Medical Imaging and Graphics*, 18, 85–96, 1994.
- [2] N. Ayache. Medical computer vision, virtual reality and robotics. *Image and Computer Vision*, 13, 295–313, 1995.
- [3] J.S. Duncan and N. Ayache. Medical image analysis: Progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 85–108, 2000.
- [4] T. McInerney and D. Terzopolous. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1, 91–108, 1996.
- [5] J.B.A. Maintz and M.A. Viergever. A survey of medical image restoration. *Medical Image Analysis*, 2, 1–37, 1998.

- [6] R. Shahidi, R. Tombropoulos, and R.P.A. Grzeszczuk. Clinical applications of three-dimensional rendering of medical data sets. *Proceedings of IEEE*, 86, 555–568, 1998.
- [7] S Lavallee, *Registration for Computer Integrated Surgery: Methodology, and State of Art*. MIT Press, Cambridge, MA, 1996, pp. 77–97.
- [8] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, New York, 1989.
- [9] L. Davis, Ed., *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, New York, 1991.
- [10] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, New York, 1992.
- [11] H. Kargupta, K. Deb, and D.E. Goldberg. Ordering genetic algorithms and deception. In *Proceedings of the Parallel Problem Solving from Nature* (R. Manner and B. Manderick, Eds.). North-Holland, Amsterdam, 1992, pp. 47–56.
- [12] N.J. Radcliffe. Genetic set recombination. In *Foundations of Genetic Algorithms 2* (L.D. Whitley, Ed.). Morgan Kaufmann, San Mateo, CA, 1993, pp. 203–219.
- [13] J.J. Grefenstette, R. Gopal, B. Rosmaita, and D. Van Gucht. Genetic algorithms for the traveling salesman problem. In *Proceedings of the 1st International Conference on Genetic Algorithms* (J.J. Grefenstette, Ed.). Lawrence Erlbaum Associates, Hillsdale, 1985, pp. 160–168.
- [14] J.H. Holland, *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor, MI, 1975.
- [15] D.E. Goldberg, K. Deb, and B. Korb. Messy genetic algorithms: Motivation, analysis, and first results. *Complex Systems*, 3, 493–530, 1989.
- [16] D.E. Goldberg, K. Deb, and B. Korb. Do not worry, be messy. In *Proceedings of the 4th International Conference on Genetic Algorithms* (R.K. Belew and J.B. Booker, Eds.). San Mateo, CA, Morgan Kaufmann, 1991, pp. 24–30.
- [17] D.E. Goldberg, K. Deb, H. Kargupta, and G. Harik. Rapid, accurate optimization of difficult problems using fast messy genetic algorithms. In *Proceedings of the 5th International Conference on Genetic Algorithms* (S. Forrest, Ed.), San Mateo, CA, Morgan Kaufmann, 1993, pp. 56–64.
- [18] H. Kargupta and S. Bandyopadhyay. Further experimentations on the scalability of the GEMGA. In *Proceedings of the V Parallel Problem Solving from Nature (PPSN V), Lecture Notes in Computer Science* (T. Baeck, A. Eiben, M. Schoenauer, and H. Schwefel, Eds.), Vol. 1498. Springer-Verlag, Amsterdam, The Netherlands, 1998, pp. 315–324.
- [19] S. Bandyopadhyay, H. Kargupta, and G. Wang. Revisiting the GEMGA: Scalable evolutionary optimization through linkage learning. In *Proceedings of International Conference on Evolutionary Computation*, 1998, pp. 603–608.
- [20] H. Kargupta and S. Bandyopadhyay. A perspective on the foundation and evolution of the linkage learning genetic algorithms. *The Journal of Computer Methods in Applied Mechanics and Engineering, Special issue on Genetic Algorithms*, 186, 266–294, 2000.
- [21] L.J. Eshelman and J.D. Schaffer. Real-coded genetic algorithms and interval schemata. In *Foundations of Genetic Algorithms 2* (L. Whitley, Ed.). Morgan Kaufmann, San Mateo, CA, 1993, pp. 187–202.
- [22] J.E. Baker. Adaptive selection methods for genetic algorithms. In *Proceedings of the 1st International Conference on Genetic Algorithms* (J.J. Grefenstette, Ed.). Hillsdale: Lawrence Erlbaum Associates, 1985, pp. 101–111.
- [23] M. Srinivas and L.M. Patnaik. Adaptive probabilities of crossover and mutation in genetic algorithm. *IEEE Transaction on System, Man, Cybernetics*, 24, 656–667, 1994.
- [24] S. Bandyopadhyay, *Pattern Classification using Genetic Algorithms*. Ph.D. thesis, Machine Intelligence Unit, Indian Statistical Institute, Calcutta, India, 1998.
- [25] S. Crosier and D.M. Doddrell. Gradient-coil design by simulated annealing. *Journal of Magnetic Resonance*, 103, 354–357, 1993.
- [26] S. Crosier and D.M. Doddrell. Compact MRI magnet design by stochastic optimization. *Journal of Magnetic Resonance*, 127, 233–237, 1997.

- [27] B.J. Fisher, N. Dillon, T.A. Carpenter, and L.D. Hall. Design by genetic algorithm of a z gradient set for magnetic resonance imaging of the human brain. *Measurement Science and Technology*, 6, 904–909, 1995.
- [28] B.J. Fisher, N. Dillon, T.A. Carpenter, and L.D. Hall. Design of a biplanar gradient coil using a genetic algorithm. *Magnetic Resonance Imaging*, 15, 369–376, 1997.
- [29] G.B. Williams, B.J. Fisher, C.L.-H. Huang, T.A. Carpenter, and L.D. Hall. Design of biplanar gradient coils for magnetic resonance imaging of the human torso and limbs. *Magnetic Resonance Imaging*, 17, 739–754, 1999.
- [30] R.E. Ansoorge, T.A. Carpenter, L.D. Hall, N.R. Shaw, and G.B. Williams. Use of parallel supercomputer to design magnetic resonance systems. *IEEE Transactions on Applied Superconductivity*, 10, 1368–1371, 2000.
- [31] D.H. Turnbull, A.T. Kerr, and F.S. Foster. Simulation of B-Scan images from two dimensional transducer arrays. In *Proceedings of the 1990 IEEE Ultrasonic Symposium*, 1990, pp. 769–773.
- [32] P.K. Weber, R.M. Schmitt, B.D. Tylkowski, and J. Steak. Optimization of random sparse 2-D transducer arrays for 3-D electronic beam steering and focusing. In *Proceedings of the 1994 IEEE, IEEE Ultrasonics Symposium*, 1994, pp. 1503–1506.
- [33] M.A. Anastasio, H. Yoshida, R. Nagel, R.M. Nishikawa, and K. Doi. A genetic algorithm-based method for optimizing the performance of a computer-aided diagnosis scheme for detection of clustered microcalcifications in mammograms. *Medical Physics*, 25, 1613–1620, 1998.
- [34] J. Horn and N. Nafpliotis. Multiobjective Optimization using the Niche Pareto Genetic Algorithms. *illiGAL report no. 93005*. University of Illinois at Urbana-Champaign, 1993.
- [35] M.A. Kupinski and M.A. Anastasio. Multiobjective genetic optimization of diagnostic classifiers with implications for generating receiver operating characteristic curves. *IEEE Transactions on Medical Imaging*, 18, 675–685, 1999.
- [36] M.A. Anastasio, M.A. Kupinski, R.M. Nishikawa, and M.L. Giger. A multiobjective approach to optimizing computerized detection schemes. *IEEE Nuclear Science Symposium*, 3, 1879–1883, 1998.
- [37] A. Green, N. Martin, J. Pfitzner, M. O'Rourke, and N. Knight. Computer image analysis in the diagnosis of melanoma. *Journal of American Academies of Dermatology*, 31, 958–964, 1994.
- [38] Z.B. Argyenyi. Dermatoscopy (epiluminescence microscopy) of pigmented skin lesions. *Dermatology Clinical*, 15, 79–95, 1997.
- [39] W. Stolz, O. Braun-Falco, M. Landthaler, P. Bilek, and A.B. Cognetta, *Color Atlas of Dermatoscopy*. Blackwell Science, Oxford, 1994.
- [40] W.V. Stoecker, W.W. Li, and R.H. Moss. Automatic detection of asymmetry in skin tumors. *Computerized Medical Imagings and Graphics*, 16, 191–197, 1992.
- [41] D. Gutkowitz-Krusin, M. Elbaum, P. Szwaykowski, and A.W. Kopf. Can early malignant melanoma be differentiated from a typical melanocytic nerves by in vivo techniques?, Part II, automatic machine vision classification. *Skin Research and Technology*, Vol. 3, pp. 15–22, 1997.
- [42] P. Schmid-Saugeon. Symmetry axis computation for almost-symmetrical and asymmetrical objects: Application to pigmented skin lesions. *Medical Image Analysis*, 4, 269–282, 2000.
- [43] K. Deb and A. Kumar. Simulated binary crossover for continuous search space. *Complex Systems*, 9, 115–148, 1995.
- [44] C.M. Balch and G.W. Milton, *Hautmelanome* Springer, Berlin, 1988.
- [45] C.M. Grin, A.W. Kopf, B. Welkovich, R.S. Bart, and M.J. Levenstein. Accuracy in the clinical diagnosis of malignant melanoma. *Archives of Dermatology*, 126, 763–766, 1990.
- [46] J.E. Golston, W.V. Stoecker, R.H. Moss, and I.P.S. Dhillon. Automatic detection of irregular borders in melanoma and other skin tumors. *Computer Medical Imaging Graphics*, 16, 163–177, 1992.
- [47] A.J. Sober and J.M. Burstein. Computerized digital image analysis: An aid for melanoma diagnosis preliminary investigations and brief review. *Journal of Dermatology*, 21, 885–890, 1994.
- [48] K.P. Wilhelm, P. Elsner, E. Beradesca, and H. Baibach, *Bioengineering of the Skin: Skin Surface Imaging and Analysis*. CRC Press, Boca Raton, FL, 1997.

- [49] H. Handels, T. Rob, J. Kreuzsch, H.H. Wolff, and S.J. Poppl. Feature selection for optimized skin tumor recognition using genetic algorithms. *Artificial Intelligence in Medicine*, 16, 283–297, 1999.
- [50] R.M. Haralick, K. Shanmugam, and I. Dinstein. Texture features for image classification. *IEEE Transactions on System, Man, & Cybernetics*, 3, 610–621, 1973.
- [51] D.H. Ballard and M.B. Brown. *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ, 1982.
- [52] K.J. Falconer. *Fractal Geometry, Mathematical Foundations and Applications*. Wiley, Chichester, 1990.
- [53] H.O. Peitgen and D. Saupe. *The Science of Fractal Images*. Springer, Berlin, 1988.
- [54] S. Carosi, A. Massa, and M. Pastorino. Numerical assessment concerning a focussed microwave diagnostic method for medical application. *IEEE Transactions of Microwave Theory and Techniques*, 48, 1815–1830, 2000.
- [55] P.C. Johnson, H. Brown, W.M. Kuzon, J.R. Ballit, J.L. Garrison, and J. Campbell. Simultaneous quantitation of facial movements: The maximal static response assay of facial nerve function. *Annals of Plastic Surgery*, 32, 171–179, 1994.
- [56] T.D. Helling and M.J.G. Neely. Validation of objective measures for facial paralysis. *Laryngoscope*, 107, 1345–1349, 1997.
- [57] Y. Cui, M. Wan, and J. Li. A new quantitative assessment method of facial paralysis based on motion estimation. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 3. IEEE Press, Hong Kong, 1998, pp. 1412–1413.
- [58] L.J. Dorfman and K.C. McGill. AAEE minimonograph #29: Automatic quantitative electromyography. *Muscle Nerve*, 11, 804–818, 1988.
- [59] C.N. Schizas, C.S. Pattichis, I.S. Schofield, P.R. Fawcett, and L.T. Middleton. Artificial neural nets in computer-aided macro motor unit potential classification. *IEEE Engineering Medicine and Biology Magazine*, 9, 31–38, 1990.
- [60] C.S. Pattichis. *Artificial Neural Networks in Clinical Electromyography*. Ph.D. thesis, Queen Mary and Westfield College, University of London, UK, 1992.
- [61] C.S. Pattichis and C.N. Schizas. Genetic-based machine learning for the assessment of certain neuromuscular disorders. *IEEE Transactions on Neural Networks*, 7, 427–439, 1996.
- [62] P. Larranaga, B.S.M.Y. Gallego, M.J. Michelena, and J.M. Pikaza. Learning bayesian networks by genetic algorithms. A case study in the prediction of survival in malignant skin melanoma. In *Lecture Notes in Artificial Intelligence*, Vol. 1211. Springer-Verlag, Heidelberg, 1997, pp. 261–272.
- [63] B. Sierra and P. Larranaga. Predicting survival in malignant skin melanoma using Bayesian networks automatically induced by genetic algorithms. An empirical comparison between different approaches. *Artificial Intelligence in Medicine*, 14, 215–230, 1998.
- [64] F.V. Jensen, *Introduction to Bayesian Networks*. University College of London, UK, 1996.
- [65] P. Larranaga, M.Y. Gallego, B. Sierra, L. Urkola, and M.J. Michelena. Bayesian networks, rule induction and logistic regression in the prediction of women survival suffering from breast cancer. In *Lecture Notes in Artificial Intelligence*, Vol. 1323. Springer-Verlag, Heidelberg, 1997, pp. 303–308.
- [66] R. Blanco, I. Inza, and P. Larranaga. Learning bayesian networks in the space of structures by estimation of distribution algorithms. *International Journal of Intelligent Systems*, 18, 205–220, 2003.
- [67] C.A. Pena-Reyes and M. Sipper. Evolving fuzzy rules for breast cancer diagnosis. In *Proceedings of 1998 International Symposium on Nonlinear Theory and Applications (NOLTA '98)*, Vol. 2. Lausanne, Presses Polytechniques et Universitaires Romandes, pp. 369–372, 1998.
- [68] C.A. Pena-Reyes and M. Sipper. A fuzzy-genetic approach to breast cancer diagnosis. *Artificial Intelligence in Medicine*, 17, 131–155, 1999.
- [69] C.J. Kahn, L. Roberts, K. Shaffer, and P. Haddawy. Construction of a bayesian network for mammographic diagnosis of breast cancer. *Computers in Biology and Medicine*, 27, 19–29, 1997.
- [70] N. Marvin, M. Bower, and J.E. Rowe. An evolutionary approach to constructing prognostic models. *Artificial Intelligence in Medicine*, 15, 155–165, 1999.

- [71] S. Ngan, M.L. Wong, W. Lam, K.S. Leung, and J.C.Y. Cheng. Medical data mining using evolutionary computation. *Artificial Intelligence in Medicine*, 16, 73–96, 1999.
- [72] S. Ngan, M.L. Wong, and K.S. Leung. Using grammar based genetic programming for data mining of medical knowledge. In *Genetic Programming 1998: Proceedings of 3rd Annual Conference*, Morgan Kaufmann, San Mateo, CA, 1998, pp. 254–259.
- [73] C.C. Bojarczuk, H.S. Lopes, and A.A. Freitas. Discovering comprehensible classification rules by using genetic programming: A case study in a medical domain. In *Proceedings of the Genetic and Evolutionary Computation Conference* (W. Banzhaf, J. Daida, A.E. Eiben, M.H. Garzon, V. Honavar, M. Jakiela, and R.E. Smith, Eds.), Vol. 2, (Orlando, FL, USA), pp. 953–958, Morgan Kaufmann, San Mateo, CA, 1999, pp. 13–17.
- [74] A. Teller and M. Veloso. Program evolution for data mining. *International Journal of Expert Systems*, 8(3), 213–236, 1995.
- [75] B. Marchesi, A.L. Stelle, and H.S. Lopes. Detection of epileptic events using genetic programming. In *Proceedings of the 19th International Conference IEEE/EMBS*, pp. 1198–1201. IEEE Press, Washington, 1997.
- [76] J.R. Sherrah, R.E. Bogner, and A. Bouzerdoum. The evolutionary pre-processor: Automatic feature extraction. In *Genetic Programming: Proceedings of 2nd Annual Conference*, Morgan Kaufmann, San Mateo, CA, 1997, pp. 304–312.
- [77] N.I. Nikolaev and V. Slavov. Inductive genetic programming with decision trees. In *Proceedings of the 1997 European Conference on Machine Learning (ECML)*, 1997.
- [78] L. Martin, F. Moal, and C. Vrain. A relational data mining tool based on genetic programming. In *Principles of Data Mining and Knowledge Discovery: Proceedings of 2nd European Symposium (LNAI)*, Vol. 1510 Springer-Verlag, Heidelberg, 1998, pp. 130–138.
- [79] M.V. Fidelis, H.S. Lopes, and A.A. Freitas. Discovering comprehensible classification rules a genetic algorithm. In *Proceedings of the 2000 Congress on Evolutionary Computation CEC00*, La Jolla Marriott Hotel La Jolla, CA, USA, pp. 805–810, IEEE Press, Washington 2000, pp. 6–9.