http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (5): 944-954, 2011
ISSN 1812-5638 / DOI: 10.3923/itj.2011.944.954

A Hybrid Memetic Algorithm with Back-propagation Classifier for Fish Classification Based on Robust Features Extraction from PLGF and Shape Measurements

M. Alsmadi, K.B. Omar, S.A. Noah and I. Almarashdeh
Center for Artificial Intelligence, Faculty of Information Science and Technology,
Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

Abstract: The aim of this study is a novel hybrid approach for optimizing the performance of back-propagation classifier (BPC) by utilizing the ability of Memetic algorithm (genetic algorithm and great deluge algorithm) to optimize the parameters (weight) of the PBC for fish classification problem. To recognize an isolated pattern of fish in the image based on the combination between robust features extraction which extracted based on Potential Local Geometric Features (PLGF) and shape measurements, which are extracted by measuring the edge detection method, distance and angle measurements. Typical the BPC has such disadvantage as slow practice speed and easy for running into local minimum. We presented a system prototype for dealing with such problem. The process started by acquiring an image-containing pattern of fish, then the image features extraction is performed relying on PLGF and shape measurements. The hybrid Memetic Algorithm (genetic algorithm and great deluge algorithm) with BPC (HGAGD-BPC) has outperformed BPC method and previous methodologies by obtaining better quality results but with a high cost of computational time compared to the BPC method. Where the overall accuracy obtained using the traditional BPC was 86%, while the overall accuracy obtained by the HGAGD-BPC was 96% on the test dataset used. We developed a classifier for fish images classification. Eventually, the classifier is able to classify the given fish into poison and non-poison fish and classify the poison and non-poison fish into its family.

Key words: Back-propagation classifier, memetic algorithm, potential local geometric feature, feature extraction, edge detection method, digital fish images, poison and non-poison fish

INTRODUCTION

Recently, a lot of study was done by depending on the computer; In order to let the processing time to be reduced and to provide more results that are accurate, for example, depending on different types of data, such as digital image and characters and digits. In order to automate systems that deal with numbers such as Fingerprint verification, face recognition, iris discrimination, chromosome shape discrimination, optical character recognition, texture discrimination and speech recognition. And an automatic fish image recognition system is proposed in this study. Digital image recognition has been extremely found and studied. Various approaches in image processing and pattern recognition have been developed by scientists and engineers to solve this problem (Al-Omari *et al*., 2009). That is because it has an importance in several fields. In this study, system for recognized of fish image is built which may benefit various fields, the system concerning on isolated pattern of interest, the input is considered to be an image of specific size and format, the image is processed and then recognized the given fish into its cluster and categorize the clustered fish into poison or non-poison fish and categorizes the non-poison fish into its family. The proposed system recognizes isolated pattern of fish as the system acquire an image consisting pattern of fish, then the image will be processed into several phases such as pre processing and feature extraction before recognizing the pattern of fish. The BPC used for the recognition phase.

Problem statement: The problem statement of this study extracted from the previous studies, several efforts have been devoted to the recognition of digital image but so far it is still an unresolved problem. Due to distortion, noise, segmentation errors, overlap and occlusion of objects in color images (Bai *et al*., 2008; Kim and Hong, 2009). Recognition and classification as a technique gained a lot of attention in the last years wherever many scientists utilize these techniques in order to enhance the scientific fields. Fish recognition and classification still active area

Corresponding Author: M. Alsmadi, Center for Artificial Intelligence, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

in the agriculture domain and considered as a potential research in utilizing the existing technology for encouraging and pushing the agriculture researches a head. Although advancements have been made in the areas of developing real time data collection and on improving range resolutions (Patrick *et al*., 1991; Nery *et al*., 2005) existing systems are still limited in their ability to detect or classify fish. Despite the widespread development in the world of computers and software, the people dies every day because they do not have the ability to distinguish between poison fish and non-poison. Object classification problem lies at the core of the task of estimating the prevalence of each fish species. Solution to the automatic classification of the fish should address the following issues as appropriate:

- Arbitrary fish size and orientation; fish size and orientation are unknown a priori and can be totally arbitrary
- Feature variability; some features may present large differences among different fish species
- Environmental changes; variations in illumination parameters such as power and color and water characteristics, such as turbidity, temperature, not uncommon. The environment can be either outdoor or indoor
- Poor image quality; image acquisition process can be affected by noise from various sources as well as by distortions and aberrations in the optical system
- Segmentation failures; due to its inherent difficulty, segmentation may become unreliable or fail completely

And the vast majority of research-based classification of fish points out that the basic problem in the classification of fish. They typically use small groups of features without previous thorough analysis of the individual impacts of each factor in the classification accuracy (Lee *et al*., 2003; Lee *et al*., 2008; Alsmadi *et al*., 2010b; Alsmadi *et al*., 2010a).

Previous work: Selecting suitable variables is a critical step for a successful implementation of image classification. Many potential variables may be used in image classification such as shapes and texture and the feature extraction process can do it. The purpose of feature extraction is to determine the most relevant and the least amount of data representation of the image characteristics in order to minimize the within-class pattern variability, whilst, enhancing the between-class pattern variability. There are two categories of features: Statistic features and structural features. Feature extraction from an image is a major process in image analysis. An image feature is an attribute of an image. Image features can be classified into two types: Natural and artificial ones. The natural features are defined by the visual appearance of an image such as luminance of a region (Wang *et al*., 2005), whilst artificial features are obtained from some manipulations of an image such as image amplitude histogram and filters (Petrou and Kadyrov, 2001). Image analysis requires the use of image features that capture the characteristics of the objects depicted so that they are invariant to the way the objects are presented in the image. Historically, the process of extracting image features has been anthropocentric: The features calculated are defined in a way that captures the attributes the human vision system would recognize in the image. Thus, features like compactness, brightness are features, which have some physical and perceptual meaning. It is not however necessary for the features to have a meaning to the human perception in order to characterize well an object. Indeed, features which broaden the human perception may prove to be more appropriate for the characterization of complex structures, like the objects often one wishes to identify in an image (Sze *et al*., 1999). Zion *et al*. (1999) proposed a classifier based on color and shape features of fish to deal with the shape-based retrieval problem. They mentioned about the necessity of using shape and color of fish to search the fish database of Taiwan. The developed technique is able to perform scale and rotation invariant matching between two fishes. A target object selected by a bounding rectangle has to be processed by a foreground/background separation step. The target object (foreground part) is then converted into a Curvature Scale Space (CSS) map. In order for performing rotation invariant matching. The authors further converts the CSS map into a Circular Vector (CV) map and then find its representative vector based on the concept of force equilibrium. After rotating the representative vector into the canonical orientation, every unknown object can be compared with the model objects efficiently. An image-processing algorithm developed by Shutler and Nixon (2001) has been used for discrimination between images of three fish species for use on freshwater fish farms. Zernike velocity moments were developed by Shutler and Nixon (2001) , to describe an object using not only its shape, but also its motion throughout an image. Classification is the final stage of any image-processing system where each unknown pattern is assigned to a category. The degree of difficulty of the classification problem depends on the variability in feature values for objects in the same category, relative to the difference between feature values for objects in different categories. Lee *et al*. (2008) proposed shape analysis of images of fish to deal with the fish classification problem. A new shape analysis algorithm was developed for removing edge noise and redundant data point such as short straight line. A curvature function analysis was used to locate critical landmark points. The fish contour segments of interest patterns were then extracted based on landmark points for species classification, which were done by comparing individual contour segments to the curves in the database. Regarding the feature extraction process, the authors tackled in their research the following features: Fish contour extraction; fish detection and tracking; shape measurement and descriptions (i.e., shape characters (features), anal and caudal fin and size); data reduction; landmark points; landmark points statistics (i.e., curve segment of interest). In their study, they have chosen nine species of fishes that have similar shape characters and the total of features was nine features. Also, they recommended that the decision tree is considered as a suitable method to obtain high accurate results of fish images based on the common characters used, such as: Caudal, anal and adipose fin. Furthermore, the authors claimed that the number of shape characters needed to be used and how to use them depending on the number of species and what kind of species are required by the system to be classified. Their experiments conducted 22 fish images that belong to 9 species, where the detection percentage of the classification process was 90%.

EXPERIMENTAL

This study had focused on 610 images of fish which collected from Global Information System (GIS) on Fishes (fish-base) and department of fisheries Malaysia Ministry of Agricultural and Agro-based industry in putrajaya, Malaysia region currently, the database contains 610 of fish images. Data acquired on 22th August 2008 are used. Our system has been applied on 20 different fish families, each family has a different number of fish types and our sample consists of distinct 610 of fish images. These images are divided into two datasets: 400 training images and 210 testing images.

The feature selection approach: Feature extraction refers to a process by which fish attributes are computed and collected from PLGF and shape measurements through the edge detection method, distance and angle measurements. The goal of a feature extraction determines a largest set of features.

Anchor/landmark points location detection: In the PLGF and shape measurements, a number of anchor/landmark points are required to be determined as labeled in Fig. 1a, b and Fig. 2. Anchor/landmark points detection is the goal in several works during the last few years. The aim of point detection is to detect a relevant set of point to get the anchor point for patterns of interest. The goal of anchor point detection in our study is to determine nineteen labeled points that will give the location of each features determined for fishes recognition. Then it will be used to calculate the features geometry (distance and angle measurements) for the recognition purpose.

After detecting the anchor/landmark points over the image, we can extract the features from the PLGF and shape measurements.

Shape measurements: This group of measurements consists of planar measurements on the fish's area and the length of external contour of fish will be computed by edge detection method (Mokti and Salam, 2008). These features are not invariant under translation, scales and rotation; they are fundamentally role in computing other relevant features.

Potential local geometric features: PLGF are to determine the significant similarity part, such as the tail shape. Furthermore, through the usage of distance and angle tools, the following features can be determined: The length and width of fish, size of mouth, angle of head,

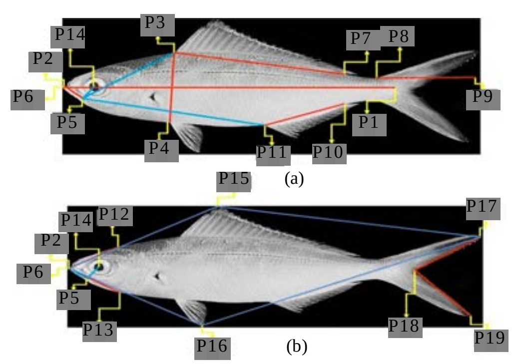

Fig. 1: (a, b) Anchor/landmark point locations

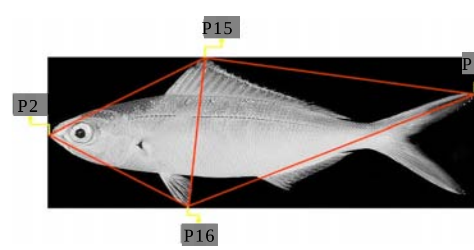

Fig. 2: The two triangles

caudal fin length, dorsal fin length, caudal angle and the angle between the mouth and the eye. Besides, by dividing the fish into two parts it can be a significant step in obtaining a high accuracy of fish classification. According to Fig. 2, two different triangles are drawn based on the maximum and minimum points on the x-axes as well as y-axes, finalizing the triangle drawing process by connecting lines between the maximum and the minimum points on x-axes with the maximum and minimum points on y-axes. This will lead to the classification process through the calculation of triangle's angles between three points and the triangles area.

Calculation of extracted features: For the PLGF and shape measurements we used the distance and angle measurements to calculate their features. The distance measurements are the length of twelve landmark points (P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11 and P14) as shown in Fig. 1a. While the angle measurements are the angles between three landmark points (P13, P2 and P12), (P14, P5 and P6), (P15, P2 and P16), (P15, P17 and P16) and (P17, P18 and P19) as shown in Fig.1b and two triangle (P15, P2 and P16) and (P15, P17 and P16) illustrated in Fig. 2. The selected anchor/landmark points are explained in Table 1 and 2. This calculation of the distance and the angle will be explained in the next subsection.

Distance measurements: Distance is a numerical description of how far apart objects are at any given moment in the time in physics or everyday discussion, distance may refer to a physical length, a period time or estimation based on other criteria (e.g., two counties over). In mathematics, distance must meet more rigorous criteria. In neutral geometry, the minimum distance between two points is the length of the line segment between them.

Table 1: Distance measurements

Id No.	Feature name	Feature description
D1	Distance between the front of fish and the start of the caudal fin	Dist (P1, P2)
D2	Fish Width excluding the upper and lower fins	Dist (P3, P4)
D3	Fish Mouth length	Dist (P5, P6)
D4	Dorsal Fin Length	Dist (P3, P7)
D5	Caudal Fin Length	Dist (P8, P9)
D6	Distance between the right-end of mouth and the eye center	Dist (P5, P14)
D7	Distance between the right-end of mouth and the start of dorsal fin	Dist (P5, P3)
D8	Distance between anal fin and the right-end of mouth	Dist (P11, P5)
D9	anal fin length	Dist (P10, P11)

Table 2: Angle measurements

Id No.	Feature name	Anchor points
Angle 1	Caudal fin angle	P17, P18, P19
Angle 2	Fish head angle	P13, P2, P12
Angle 3	Eye-end mouth angle	P14, P5, P6
Triangle 1	Front-triangle angle	P15, P2, P16

In algebraic geometry, the distance d between the points A = (x_1, y_1) and B = (x_2, y_2) is given by the formula:

$$\sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

Similarly, given points (x_1, y_1, z_1) and (x_2, y_2, z_2) the distances between them, are given by the formula:

$$\begin{aligned} d = d &= \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \\ &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \end{aligned} \quad (2)$$

The distance calculation can be seen in Table 1 and referred to the ten landmark points as in Fig. 1a shows the distance between mass points as in Table 1. There are ten features produced from this distance measurement category.

Calculate the angles: An angle can be defined as two rays or two line segments having a common end point. The endpoint becomes known as the vertex. An angle occurs when two rays meet or unite at the same endpoint. The angles between two vectors, as we show in Fig. 3 below can be identified as $\angle ABC$ or $\angle CBA$. You can also write this angle as $\angle B$ which names the vertex (common endpoint of the two rays).

The distance formula as mentioned previously can used to find the distance between two points (A, B and C). Once the two side measurements are known, the internal angles 'θ' can be found as well. When the angle (θ) is unknown, the cosine rule is the only option to find the angle. This is represented by an angular separation formula that represents cosine angle between two vectors. Basically, from vector algebra we remember that cosine angle between two vectors can be represented as dot product divided by length of the two vectors as shown in Fig. 3:

$$\cos\theta = \frac{a.b}{|a|.|b|} \quad (3)$$

The length of a vector (also known as modulus) is the root of square of its coordinate:

$$|a| = \sqrt{a_1^2 + a_2^2 + \ldots + a_x^2} \quad (4)$$

Putting the two together, we get:

$$\cos\theta = \frac{a_1 b_1 + a_2 b_2 + \ldots + a_x b_x}{\sqrt{a_1^2 + a_1^2 + \ldots + a_x^2} + \sqrt{b_1^2 + b_1^2 + \ldots + b_x^2}} \quad (5)$$

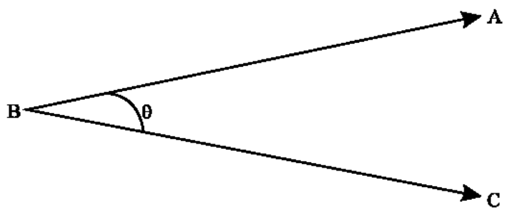

Fig. 3: The angle between two vectors

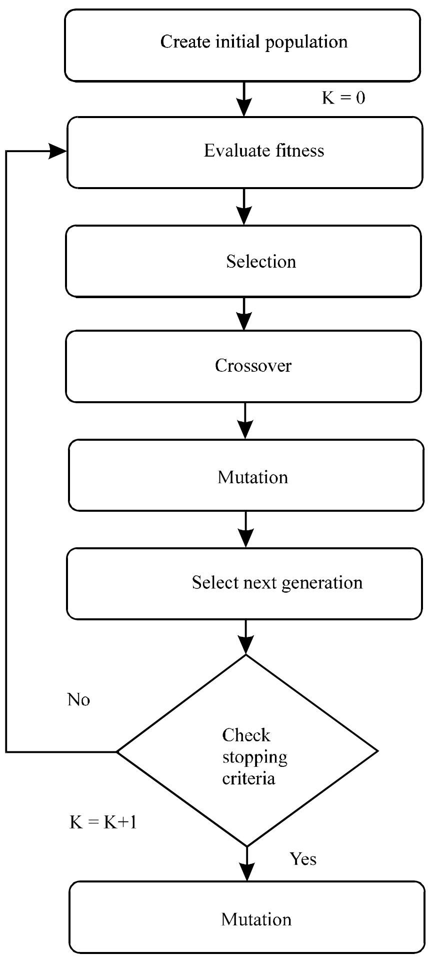

Fig. 4: Flowchart of a simple GAs process

Finally, the obtained angle is converted into angle degrees as follows:

$$angledegree = \theta * (180 / \Pi) \tag{6}$$

Table 2 shows the five angle and triangles features calculated from the angle category calculation based on the anchor/landmark points in Fig. 1b and Fig. 2.

GENETIC ALGORITHM

Genetic Algorithm (GA) optimization is an evolutionary, directed search technique that evaluates hundreds of thousands of possible solutions as it converges on the best solution alternatives. Also GAs are randomized parallel search algorithms that model natural selection, the process of evolution. Over time, natural selection has produced a wide range of robust structures (life forms) that efficiently perform a broad range of functions. The success of natural selection on earth provides an existence proof of the viability of an evolutionary process as a model for constrained optimization problems. Like natural selection, GAs are simple, significantly, faster and robust search method requiring little information to search effectively in large, poorly-understood spaces (Turabieh, 2010).

As the process of implementing GA became better understood and modern computing power became increasingly available, the ability to simulate evolutionary systems on the computer soon began to attract growing interest from the scientific community. In addition to generating further insight into the process of natural evolution, artificial evolutionary techniques also revealed characteristics, which would be particularly useful for applications optimization, engineering and computer science, among other fields.

Recently, many people are using Genetic Algorithms in an easy manner, as long as he/she know how to encode solutions of a given problem to chromosomes in GA and compare the fitness of solutions. The secret behind the success in GA is the representation and meaningful fitness evaluation, GA power come from discovering good solutions rapidly for difficult high-dimensional and complex problems (Turabieh, 2010).

GA EVOLUTIONARY PROCESS

There are two elements are required for any problem before a genetic algorithm can be used to search for a solution and they are:

- There must be a method of representing a solution in a manner that can be manipulated by the algorithm. Traditionally, a solution can be represented by a string of bits, numbers or characters
- There must be some method of measuring the quality of any proposed solution, using a fitness function

The evolutionary processes of GAs start by the computation of the fitness of the each individual in the initial population. A flowchart for a simple GA process is given in Fig. 4. To summarize how genetic algorithms

```
begin
 k=0;
 initialize Pop(k);
 evaluate Pop(k);
 while (termination not reached) do
 recombine Pop(k) to generate Offspring(k);
 evaluate Offspring;
 Select Pop(k + 1) from Pop(k) and
 Offspring(k);
 k = k + 1
 end while
end
```

Fig. 5: Simple Genetic Algorithms (Turabieh, 2010)

work, assume that Pop (k) and Offspring (k) are the parents and offspring in current generation t; the general structure of a genetic algorithm procedure can be described by the simple pseudo code as shown in Fig. 5.

GREAT DELUGE ALGORITHM

The great deluge algorithm is a local search procedure that was introduced by Dueck (1993). The idea of great deluge comes from the analogy that a person climbing a hill and try to move to any direction of finding a way up to keep his feet dry as the water level rises during a great deluge.

Inserting a great deluge algorithm within a genetic algorithm is considered an effective way to produce a high quality solution rather than using a genetic algorithm alone (Nahas *et al*., 2008; Al-Milli, 2010). This research applied a great deluge algorithm to improve the solution quality (weight) by increasing the number of fitness cost. This helps to enhance the exploitation process during the searching process.

Neural network model: The multilayer feed forward neural network model with BPC for training is employed for classification task as shows in Fig. 6. Which illustrates our implemented neural network contains three layers which are the input layer, the hidden layer and the output layer (Alsmadi *et al*., 2009a, b). The number of neurons is varied from layer to another (except The output layer consist of 20 neurons since we need to classify 20 fish families [1, 2,..., 20] each of which correspond to one of the possible family's that might be considered) in order to determine the suitable number of neurons for both input and hidden layers, therefore, obtaining high accurate results. In Fig. 6 the back-propagation classifier is a three-

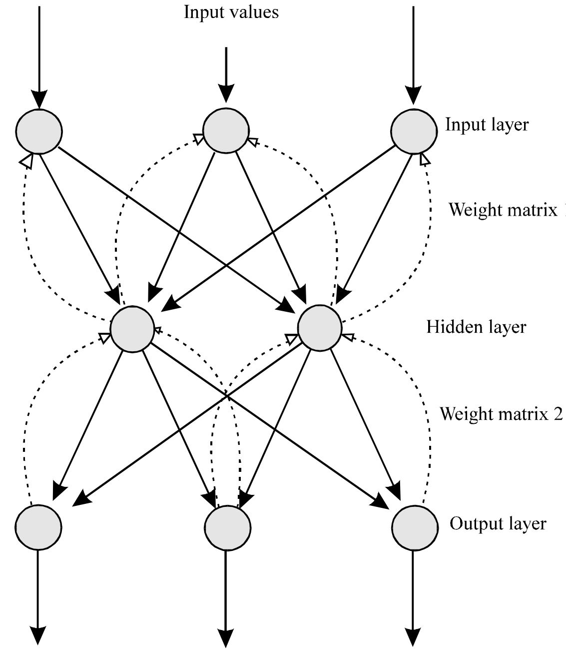

Fig. 6: Topology structure of a three-layer feed-forward NN

Table 3: Number input features and neurons for each layer

Classifier	No. of input features	No. of neurons in layers: Layer 1	Layer 2	Layer 3
BPC	20	25	35	20
HGAGD-BPC	20	27	38	20

layer feed forward neural networks. The sold arrows represent the forward operation. Whilst, the dotted arrows represent the backward operation. The circles represent the neurons. Each connection between layers has a specific weight that is stored in weight matrix. The final output obtained from the summation of the inputs in the final stage for each neuron located in the output layer.

In this research, the input used as a features matrix, that obtained from the input image (fish images). The result from the output layer compared to the original results in the training stage. The feed forward neural networks adjust the weights in order to reduce the error between desired outputs and feed forward neural networks results.

Table 3 shows the number of input features and number of neurons for each layer that determined experimentally.

A hybrid memetic algorithm with back propagation classifier: The MA is used in this work to optimize the

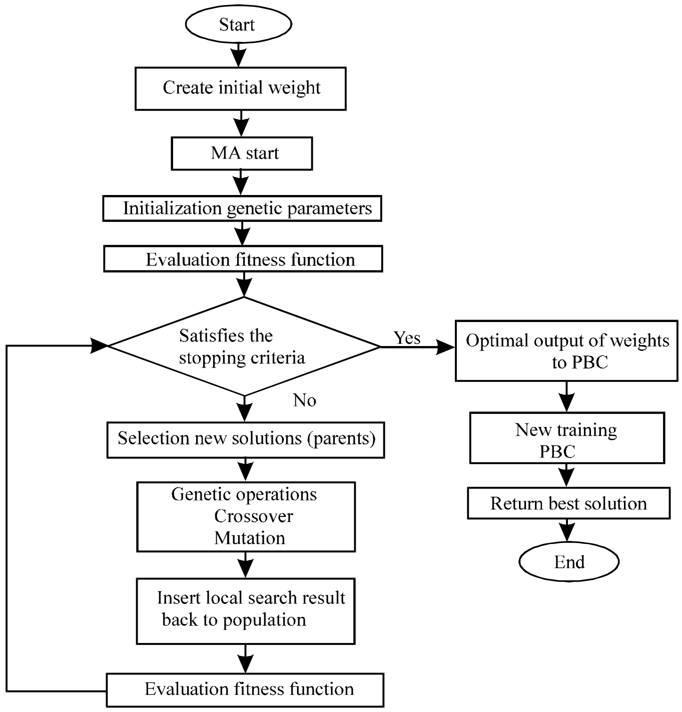

Fig. 7: Flowchart of the hybrid learning algorithm method

weights required by the BPC by initialize a population of diverse weights covering large possibilities of determining the best suited weight for the algorithm's learning process. The parameter learning process, based on MA technique and BPC, involve a two-step learning process, in the first step, the initial parameters of the neural network are tuned by the MA .in the second step, the BPC is introduced to train the initial NN to yield optimal values of weight and biases in the NN.

Basically, the MA is a derivative-free stochastic optimization method based on the features of natural selection and biological evolution. It has several advantages over other optimization algorithms. It can be applied to both continuous and discrete optimization problems. Compared with the BPC, the MA is less likely to get trapped in local optima (Maniezzo, 1994; Tang *et al*., 1995; Siddique and Tokhi, 2001; Leung *et al*., 2003; Gang *et al*., 2006; Kattan *et al*., 2010). This can be avoided by implementing some genetic operators and mechanisms, such as producing new population using solutions crossover and/or mutation. It is a computational model inspired by population genetics. It has been used mainly as function optimizers and it has been demonstrated to be an effective global optimization tool, especially for multi-model and non-continuous functions.

The flowchart of proposed hybrid learning algorithm is illustrated Fig. 7. This model describes the hybridization between MA and BPC by using the MA to optimize the parameters of the BPC. All the parameters of the BPC are encoded to form a long chromosome and tuned by the MA. Then, as a result of the MA process, the BPC is used to train the network..

INITIALIZATION

Usually, generating weights for feed-forward artificial neural network is randomly process. However, Chromosome representation plays an important factor of successful genetic algorithm. This research used a simple presentation for chromosome that is based on briary representation for each solution. Figure 8 represents a pictorial diagram for representation of gene, chromosome and population.

Each chromosome represents a possible solution by a set of parameters. The population size depends on the nature of the problem but typically contains several hundreds or thousands of possible solutions.

ROULETTE WHEEL SELECTION

Roulette Wheel Selection is the simplest selection schema; Baker (1987) developed it. This is a stochastic algorithm and involves the following technique:

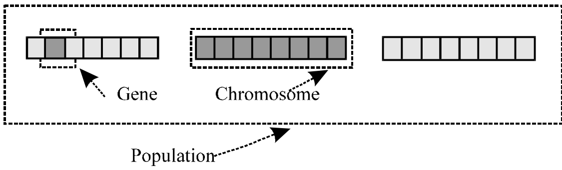

Fig. 8: Representation of Gene, Chromosome and Population

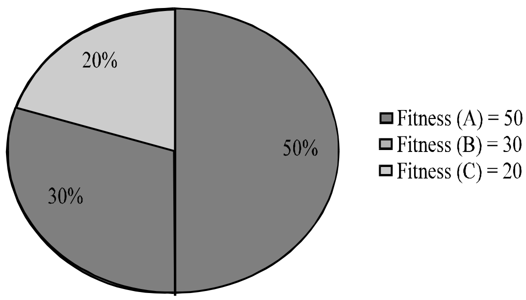

Fig. 9: Roulette wheel technique

- The individuals are mapped to contiguous segments of a line, such that each individual's segment is equal in size to its fitness
- A random number is generated and the individual whose segment spans the random number is selected
- The process is repeated until the desired number of individuals is obtained (called mating population)

Two chromosomes in the population will be selected to undergo genetic operations for reproduction by the method of spinning the roulette wheel (Tang *et al*., 1995). It is believed that high potential parents will produce better offspring (survival of the best ones). The chromosome having a higher fitness value should therefore have a higher chance to be selected as parent as shows in Fig. 9.

From the Fig. 9, three segment areas, for example, represent three solutions (A, B and C) in which each segment area has its size determined by the fitness of the solution. The higher fitness value (of a solution) has the highest probability to be selected in successive iterations.

The basic part of the selection process is to stochastically select from one generation to create the basis of the next generation. The requirement is that the fittest individuals have a greater chance of survival than weaker ones. This replicates nature in that fitter individuals will tend to have a better probability of survival and will go forward to form the mating pool for the next generation. Weaker individuals are not without a chance. In nature such individuals may have genetic coding that may prove useful to future generations.

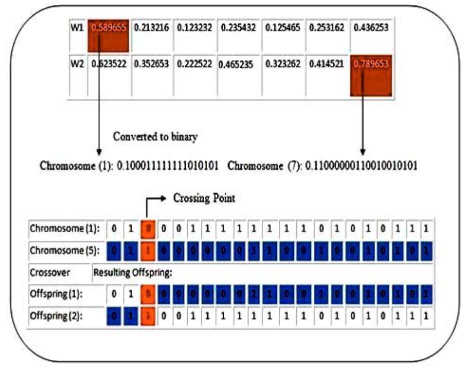

Fig. 10: Binary representation and single point crossover

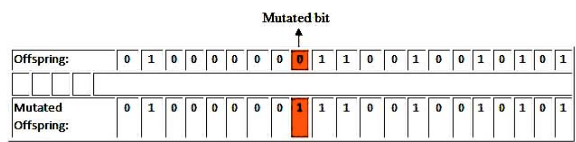

Fig. 11: Mutation process

SINGLE POINT CROSSOVER

This is the simplest method of crossover, which generate one or two child string by random selection of crossover site within the length of pattern string. Single point crossover is done by choosing a randomly point, the chromosomes of the parents will be cut from that point, and the resulting sub-chromosomes will be swapped as shown in Fig. 10.

Mutation: The purpose of mutation in MAs is to allow the algorithm to avoid local minima by preventing the population of chromosomes from becoming too similar to each other, thus slowing or even stopping evolution. Where, a variable is selected with a certain probability and its value is modified by a random value. This reasoning also explains the fact that most MA systems avoid only taking the fittest of the population in generating the next but rather a random (or semi-random) selection with a weighting toward those that are more fit. A simple mutation example shown in Fig. 11.

Here, this study chooses non-uniform mutation method. Non-uniform mutation changes one of the genes of the parent based on a non-uniform probability distribution.

FITNESS FUNCTION

The fitness function is dependent on problem and is used to evaluate the performance of each individual. The fitness function of calculating the performance of each individual is by computing the percentage Variance Account Function (VAF) between two signals. The VAF is calculated as follows:

$$V = 1 - \frac{\text{Variance}(y - y\_est)}{\text{Variance}(y)} \times 100\%$$

where, y is the real output, y_est is the estimated output of a model and VAF is computed for the two signals to result the output v. The VAF of two signals that are the same is 100%. If they differ, the VAF will be lower. When y and y_est have multiple columns, the VAF is calculated for every column in y and y_est. The VAF is often used to verify the correctness of a model, by comparing the real output with the estimated output of the model.

STOPPING CRITERION

A generation consists of the production of a new population in a generational algorithm. A similar definition is used for a steady-state algorithm. A maximum number of generations usually defines the stopping criterion in genetic algorithms. However; when it is possible to achieve an ideal fitness (i.e., optimal weight), this can also serve as the stopping criterion. In this work, a maximum generation number is used, regardless of whether the ideal fitness is achieved or not. Other criteria are possible, such as a measure of diversity loss or a lack of fitness improvement.

Experimental result: As we shows in Table 4, the accuracy of recognition test results for each fish family (20 families) based on the PLGF and Shape features uses BPC and HGAGD-BPC, which are vary from a family to another. The results indicated a high percentage of accuracy for each fish family's recognition, where these results lie between 79% as minimum percentage of accuracy and 96% as a maximum percentage of accuracy by BPC. While, the HGAGD-BPC has obtained the maximum percentage of accuracy is equal to 97% and the minimum percentage of accuracy is equal to 93%.

In the BPC some of the results that are close to the minimum percentage (e.g., Sigamidae) because of sharing some common features with each other (e.g., Megalopidae) which causes a noise identification interruption to the neural network. However, in the other hand, some families share the same features with each other, but each one has its own species-specific traits. This gives the ability of the neural network to recognize the respected family easier. For example, some of the poison fishes have the same angle tail with other non-poison fishes but with some dissimilarity such as length of dorsal fin and the distance between the pelvic fin and the right-end of the mouth. The same situation goes with the non-poison fishes, for example, the size of mouth, anal fin length, the distance between the right-end of mouth and the dorsal fin, are usually vary from family to another.

Table 4: The accuracy of recognition test results for each fish family based on the PLGF and shape features

Family name	BPC	HGAGD-BPC
Acestrorhynchidae	91	97
Acropomaatidae	92	97
Albulidae	85	95
Anomalopidae	87	97
Caesionidae	89	95
Drepanidae	80	93
Istiophoridae	90	97
Leiognathidae	84	94
Megalopidae	80	95
Platycephalidae	80	96
Priacanthidae	84	96
Scombridae	86	96
Siganidae	79	97
Sillaginidae	80	97
Stromateidae	90	94
Triacanthidae	80	95
poison/Red Snapper	96	97
poison/Trigger	91	96
poison/Porcupine	90	94
poison/Thorn	92	96

As shown in the Table 4, the poison fish families are recognized with high accurate results, due to their species-specific traits unlike to the non-poison fish families. The obtained results of the poison fish families are within 91 and 94%.

RESULTS AND DISCUSSION

The methods have been implemented in MATLAB programming language on a CPU Core 2 Duo 2.33 GHZ. We have considered different fish images families, obtained from Global Information System (GIS) on Fishes (fish-base) and department of fisheries. For experimentation purpose 610 hundred fish images families are considered, 400 fish images for training and the rest 210 for testing. Table 5 describes the overall training and testing accuracy obtained based on robust features extracted from PLGF and shape measurements using BPC. In addition, the problem in fish recognition is to find meaningful features based on the features extraction. An efficient classifier that produce better fish images recognition accuracy rate is also required. As we show in Table 5 the overall training accuracy equals to 89% and the overall testing accuracy equals to 86%.

Table 6 describes the fitness cost and the overall accuracy of training and testing for the PLGF and shape features. The results shown in the Table are the overall

Table 5: Description of the overall accuracy of training and testing

Description	Results (%)
Overall training accuracy	89
Overall testing accuracy	86

Table 6: The overall accuracy outcome for both training and testing accuracy obtained from the trained the HGAGD-BPC

Description	Results (%)
Fitness cost	98
Overall training accuracy	97
Overall testing accuracy	96

accuracy outcome for both training and testing accuracy obtained from the trained the HGAGD-BPC. The fitness cost and the overall training and testing accuracy was 98%, 97% and 96% respectively.

The feature extraction is done based on PLGF and shape measurements, utilizing local geometric approach that uses distance and angles measurements. This is to obtain 12 features that rely on distance measurements, 5 features that rely on angles measurements and 2 features obtained by edge detection. We determined 19 anchor/landmark points on the shape of pattern of interest (fish), where 20 landmark/anchor points were extracted manually. Only one fish-based study is reported in the literature that extracted the features using the distance measurements, while in our work, we increased the number of features extracted using the distance measurements. In addition, we added (for the first time in the fish classification) the angles measurements and dividing the pattern of interest (fish) into two triangles. The main advantage of the local geometric approach that is less affected by global changes in the appearance of fish images including fish expression. Nevertheless, this approach has received little attention due to the fact that it requires an additional step of reliably locating fish landmarks/anchor points, which may affect their overall performance (Gupta *et al*., 2007; Lee *et al*., 2008).

CONCLUSION

Nineteen features representation have been extracted from nineteen detected landmark points as shown in the second section of the paper. All features were obtained from PLGF and shape measurements of fish images, through angle and distance measurements. Our experimental results suggest that our feature selection methodology can be successfully used to significantly improve the performance of fish classification systems. Unlike previous approaches which propose feature descriptors and do not analyze their impact in the classification task as a whole. We propose a general set of 19 features and their corresponding weights which may be used as a priori information by the classifier. Moreover, our study presents a novel set of features extracted from PLGF and shape measurements. The performance of the BPC has been improved significantly by the hybridization of the MA with the BPC. Where proved to be much better than the BPC. The experiments showed the effectiveness and robustness of the MA incorporated by the BPC. The HGAGD-BPC has outperformed BPC method and previous methodologies by obtaining better quality results but with a high cost of computational time compared to the BPC method. Where the overall accuracy obtained using the traditional BPC was 86%, while the overall accuracy obtained by the HGAGD-BPC was 96% on the test dataset used.

REFERENCES

Al-Milli, N.R., 2010. Hybrid genetic algorithms with great deluge for course timetabling. Int. J. Comput. Sci. Network Secur., 10: 283-288.

Al-Omari, S.A.K., P. Sumari, S.A. Al-Taweel and A.J.N. Husain, 2009. Digital recognition using neural network. J. Comput. Sci., 5: 427-434.

Alsmadi, M.K., K.B. Omar, S.A. Noah and I. Almarashdah, 2009a. Performance comparison of multi-layer. Perceptron (Back Propagation, Delta Rule and Perceptron) algorithms in neural networks. Proceedings of the IEEE International Advance Computing Conference, March 6-7, Patiala, pp: 296-299.

Alsmadi, M.K.S., K.B. Omar and S.A. Noah, 2009b. Back propagation algorithm the best algorithm among the multi-layer perceptron algorithm. Int. J. Comput. Sci. Network Secur., 9: 378-383.

Alsmadi, M.K., K.B. Omar, S.A. Noah and I. Almarashdeh, 2010a. Fish recognition based on robust features extraction from color texture measurements using back-propagation classifier. J. Theor. Applied Inform. Technol., 6: 1088-1094.

Alsmadi, M.K., K.B. Omar, S.A. Noah and I. Almarashdeh, 2010b. Fish recognition based on robust features extraction from size and shape measurements using neural network. J. Comput. Sci., 6: 1059-1065.

Bai, X., X. Yang and L.J. Latecki, 2008. Detection and recognition of contour parts based on shape similarity. Patt. Recog., 41: 2189-2199.

Baker, J.E., 1987. Reducing bias and inefficiency in the selection algorithm. Proceedings of the 2nd International Conference on Genetic Algorithms on Genetic Algorithms and their Application, July 1987, Cambridge, Massachusetts, United States, pp: 14-21.

Dueck, G., 1993. New optimization heuristics: The great deluge algorithm and the record-to-record travel. J. Comput. Phys., 104: 86-92.

Gang, L., T.M. McGinnity and G. Prasad, 2006. Design for self-organizing fuzzy neural networks based on genetic algorithms. IEEE Trans. Fuzzy Syst., 14: 755-766.

Gupta, S., M.K. Markey and A.C. Bovik, 2007. Advances and Challenges in 3D and 2D+3D Human Face Recognition. In: Pattern Recognition Research Horizons, Zoeller, E.A. (Ed.). Nova Science Publishers, New York, pp: 1-41.

Kattan, A.R.M., R. Abdullah and R.A. Salam, 2010. Training feed-forward neural networks using a parallel genetic algorithm with the best must survive strategy. Proceedings of the International Conference on Intelligent Systems, Modelling and Simulation, Jan. 27-29, Liverpool, pp: 96-99.

Kim, J.S. and K. S. Hong, 2009. Color-texture segmentation using unsupervised graph cuts. Patt. Recog., 42: 735-750.

Lee, D.J., S. Redd, R. Schoenberger, X. Xiaoqian and Z. Pengcheng, 2003. An automated fish species classification and migration monitoring system. Proceedings of the 29th Annual Conference of the IEEE Industrial Electronics Society, Nov. 2-6, Provo, UT, USA., pp: 1080-1085.

Lee, D.J., J.K. Archibald, R.B. Schoenberger, W.A. Dennis and D.K. Shiozawa, 2008. Contour matching for fish species recognition and migration monitoring. Appl. Comput. Intell. Biol., 122: 183-207.

Leung, F.H.F., H.K. Lam, S.H. Ling and P.K.S. Tam, 2003. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. IEEE Trans. Neural Networks, 14: 79-88.

Maniezzo, V., 1994. Genetic evolution of the topology and weight distribution of neural networks. IEEE Trans. Neural Networks, 5: 39-53.

Mokti, M.N. and R.A. Salam, 2008. Hybrid of Mean-shift and median-cut algorithm for fish segmentation. Proceedings of International Conference on Electronic Design, Dec. 1-3, Penang, pp: 1-5.

Nahas, N., A. Khatab, D. Ait-Kadi and M. Nourelfath, 2008. Extended great deluge algorithm for the imperfect preventive maintenance optimization of multi-state systems. Reliability Eng. Syst. Safety, 93: 1658-1672.

Nery, M.S., A.M. Machado, M.F.M. Campos, F.L.C. Padua, R. Carceroni and J.P. Queiroz-Neto, 2005. Determining the appropriate feature set for fish classification tasks. Proceedings of the 18th Brazilian Symposium on Computer Graphics and Image Processing, Oct. 09-12, Brazil, pp: 173-180.

Patrick, P.H., N. Ramani, W.G. Hanson and H. Anderson, 1991. The potential of a neural network based sonar system in classifying fish. Proceedings of the IEEE Conference on Neural Networks for Ocean Engineering, Aug. 15-17, IEEE Xplore Press, Washington DC., USA., pp: 207-213.

Petrou, M. and A. Kadyrov, 2001. Features invariant to affine distortions from the trace transform. Proceedings of the International Conference on Image Processing, Oct. 7-10, Thessaloniki, Greece, pp: 852-855.

Shutler, J.D. and M.S. Nixon, 2001. Zernike velocity moments for description and recognition of moving shapes. Proceedings of the British Machine Vision Conference, September 2001, EPrints, USA., pp: 705-714.

Siddique, M.N.H. and M.O. Tokhi, 2001. Training neural networks: Backpropagation vs. genetic algorithms. Proceedings of the International Joint Conference on Neural Networks, July 15–19, Washington, DC, USA., pp: 2673-2678.

Sze, C.J., H.R. Tyan, H.Y.M. Liao, C.S. Lu and S.K. Huang, 1999. Shape-based retrieval on a fish database of Taiwan. Tamkang J. Sci. Eng., 2: 163-173.

Tang, K.S., C.Y. Chan, K.F. Man and S. Kwong, 1995. Genetic structure for NN topology and weights optimization. Proceedings of the 1st International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications, (Conf. Publ. No. 414), Sept. 12-14, Sheffield, UK., pp: 250-255.

Turabieh, H.I., 2010. Single and multiobjective evolutionary algorithms for university timetabling problems. Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia.

Wang, X., X. Ding and C. Liu, 2005. Gabor filters-based feature extraction for character recognition. Pattern Recogn., 38: 369-379.

Zion, B., A. Shklyar and I. Karplus, 1999. Sorting fish by computer vision. Comput. Electron. Agric., 23: 175-187.